Improved Methods for the study of Hadronic Physics from Lattice QCD

Kostas Orginos^{1,2}, David Richards¹

¹Jefferson Lab, 12000 Jefferson Avenue, Newport News, VA 23606, USA

² Department of Physics, College of William and Mary, Williamsburg, Virginia 23187, USA

E-mail: kostas@jlab.org, dgr@jlab.org

Abstract. The solution of QCD on a lattice provides a first-principles method for understanding QCD in the low-energy regime, and is thus an essential tool for nuclear physics. The generation of gauge configurations, the starting point for lattice calculations, requires the most powerful leadership-class computers available. However, to fully exploit such leadership-class computing requires increasingly sophisticated methods for obtaining physics observables from the underlying gauge ensembles. In this paper, we describe a variety of recent methods that have been used to advance our understanding of the spectrum and structure of hadrons through lattice QCD.

1. Introduction

Lattice gauge calculations have assumed a central rôle in the toolkit of theoretical methods in nuclear and particle physics by enabling Quantum Chromodynamics (QCD), the theory of the strong interactions, to be solved, rather than modeled, for many quantities in the strong-coupling regime. Such quantities span nuclear physics, from the confinement of the fundamental quarks and gluons of QCD into the observed hadrons, the distribution of charge, currents and mass in those hadrons, to the binding of those hadrons to form nuclei.

Lattice QCD (LQCD) calculations proceed through a Monte Carlo procedure, in which the basic building blocks of the lattice calculations, ensembles of gauge configurations, are generated, and then a variety of physics quantities are calculated on those ensembles. The generation of an equilibrated ensemble of gauge configurations is a “capability” problem, with jobs typically employing tens of thousand of cores running for tens of hours, to which the most powerful computational resources available, leadership-class computers such as the IBM BlueGene and Cray XK/XE series of machines, must be devoted. The calculation of physical quantities on these configurations is then a “capacity” computing task, with a very large number of jobs each typically employing tens or hundreds of cores running for several hours, in which quantities are computed on each gauge configuration as an independent calculation.

Whilst, historically, the emphasis in the lattice community has been on the optimization of the “capability” stage, there is an increasing effort on the “capacity” stage of lattice calculations, with a variety of drivers. Firstly, it is important, given the investment in generating an ensemble, that a given quantity be calculated with the highest possible precision on that ensemble. Secondly, as the field has developed, an increasingly ambitious range of quantities becomes accessible to lattice computation: the excited-state spectrum of hadrons, the numerous quark and gluon distributions in nucleons, and the binding between nucleons, quantities which were inaccessible to previous lattice calculations. Finally, there is the need to attenuate the degrading signal-to-noise ratio on lattice calculations as the quark masses approach their physical values, and as higher energies in the spectrum are explored. Indeed, for many calculations in nuclear physics the aggregate computational resources devoted to the “capacity” phase exceed those devoted to the “capability” phase, with the important distinction that those resources can be satisfied with clusters, albeit of very large size, rather than requiring the highest-performance interconnect that characterizes leadership-class facilities.

In this contribution, we describe progress at tackling the challenges encountered in the “capacity” phase. We begin by introducing methods that enable the complete four-dimensional volume of a gauge configuration to be sampled, exemplified by the “distillation” method, and algorithmic improvements for solvers that enable distillation and its variants to be implemented efficiently. We then discuss a promising approach, so-called “all-mode averaging”, that provides a highly efficient variance reduction method exploiting the various symmetries, such as translation, of the lattice. Finally, we describe

a recent extension to the applicaiton of the variational method enabling us to further refine the extraction of physical information.

2. Physical observables in lattice calculations

Lattice QCD is formulated in Euclidean space with a strong analogy to calculations in statistical physics. The generating functional for lattice QCD for N_f mass-degenerate quark flavors is given by

$$\mathcal{Z} = \int \mathcal{D}U \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{-S_G(U) + \sum_{f,x,y} \bar{\psi}_f(x) M(x,y,U) \psi_f(y)} \quad (1)$$

where $U_\mu(x)$, $\mu = 1, \dots, 4$ are SU(3) matrices representing the gauge fields, $S_G(U)$ is the gauge action, ψ_f represent the fermion fields, and M is the Dirac operator encapsulating the fermion part of the action. Since the fermion fields are Grassman variables, these can be integrated out to yield

$$\mathcal{Z} = \int \mathcal{D}U [\det M(U)]^{N_f} e^{-S_G(U)}. \quad (2)$$

Lattice calculations now proceed through the Monte Carlo method, in which an ensemble of gauge configurations $\{U\}$ are generated with probability distribution

$$P(U) \propto [\det M(U)]^{N_f} e^{-S_G(U)}. \quad (3)$$

The formulation in Euclidean space ensures that this distribution is *real* and *positive*. The generation of ensembles is the starting point of all lattice calculations, and for the largest, state-of-the-art ensembles this stage is performed on leadership-scale facilities. Observables, in general explicit functions both of the gauge fields and of the fermion fields, are computed as

$$\begin{aligned} \langle \mathcal{O}(U, \psi, \bar{\psi}) \rangle &= \frac{1}{\mathcal{Z}} \int \mathcal{D}U \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{O}(U, \psi, \bar{\psi}) \\ &\equiv \frac{1}{N_{\text{ens}}} \sum_{i=1}^{N_{\text{ens}}} \mathcal{O}(U_i, \psi, \bar{\psi}), \end{aligned} \quad (4)$$

where the $\{U_i : i = 1, \dots, N_{\text{ens}}\}$ are distributed according to eqn. 3. The calculation of this ensemble average is a *capacity* computing problem; each term in the sum of eqn. 4 can be evaluated independently as a stand-alone computation.

In nuclear physics, the underlying quantities computed are energies derived from examining the exponential decay in Euclidean time of two-point correlation functions, and the matrix elements of operators computed from two- or three-point correlation functions. Focusing on two-point functions, we construct correlators in the form

$$C_{AB}(t, \vec{z}, t_0) = \sum_{\vec{x}} \langle 0 | \mathcal{O}_A(\vec{x}, t) \mathcal{O}_B(\vec{z}, t_0) | 0 \rangle e^{-i\vec{p} \cdot (\vec{x} - \vec{z})}, \quad (5)$$

where $\mathcal{O}_A, \mathcal{O}_B$ are interpolating operators having the quantum numbers of the states of interest. We then use completeness of the states to express this correlation function as

$$C_{AB}(t, t_0) = \sum_n \frac{1}{2E_n(\vec{p})} Z_A^n Z_B^{n\dagger} e^{-E_n(\vec{p})(t-t_0)}, \quad (6)$$

where E_n is the energy of the n^{th} state in the spectrum, and $Z_A^n = \langle 0 | \mathcal{O}_A(0) | n \rangle$.

To explore the computational tasks involved in computing a correlator of eqn. 5, we will focus on the isovector-meson correlation function. We can express the interpolating operators as

$$\mathcal{O}_A(\vec{x}, t) = \bar{u}(\vec{x}, t) \Gamma^A d(\vec{x}, t) \quad (7)$$

where, for the case of strictly local interpolating operators, Γ^A is one of the sixteen Dirac matrices, yielding

$$C_{AB}(t; \vec{z}, t_0) = \sum_{\vec{x}} \langle \bar{u}(\vec{x}, t) \Gamma^A d(\vec{x}, t) \bar{d}(\vec{x}, t_0) \bar{\Gamma}^B u(\vec{z}, t_0) \rangle. \quad (8)$$

The building blocks of the correlators of eqn. 5 are so-called quark propagators, the solution of the Dirac equation for a source at a fixed point z :

$$M_{\alpha\beta}^{pq}(x, y) G_{\beta\gamma}^{qr}(y, z) = \delta^{pr} \delta_{\alpha\gamma} \delta_{x,z}. \quad (9)$$

where α, β, γ are spinor indices, and p, q, r are colour indices. Quark propagators computed in this manner enable the calculation of the correlation function from a single point $z \equiv (\vec{z}, t_0)$. Spatial translational invariance enables us to can gain greater precision by sampling the correlator of eqn. 5 over *all* equivalent points z , and that has been an important quest in recent large-scale calculations. In particular, summing over all equivalent spatial points $\vec{z}$ in eqn. 5 enables a momentum projection to be made both at source and at sink. Invariance in time implies we can average over all t_0 that leave $t - t_0$ fixed. Naively, the cost of doing so the that we must compute solutions to eqn. 9 for each z on the lattice.

A important means of decreasing the systematic uncertainties on the energies E_n and residues Z_A of eqn. 6, and of increasing the information we can derive from the correlators, is through the use of the variational method, in which we calculate all elements of the correlation matrix C_{AB} for some judiciously chosen basis of operators $\{\mathcal{O}_A : A = 1, \dots, N\}$. Crucial to the efficacy of this procedure is the use of non-local interpolating operators in eqn. 7, where the Γ^A now encode (gauge-covariant) displacements, so as resolve the spatial structure of the states in the spectrum. Indeed, this has been crucial to investigations of the excited-state spectrum of QCD, and to the extraction of the momentum-dependent phase shifts that characterize resonances in the theory. However, such a procedure comes at a potentially heavy computational cost since the use of non-local operators at the source t_0 of the correlation function of necessity requires that propagators be computed from several sources, not only from a single point z .

3. Improved Sampling of the Gauge Configurations

A computational method that satisfies both these requirements above, namely to efficiently sample the whole of the lattice, and to enable the efficient computation of correlation functions between spatially extended operators at both source and at sink,

is “distillation”[1], a variant of the “smearing” procedure. The use of “smearing” in LQCD calculations, in which the quark fields are replaced by extended objects through some gauge-invariant, spatially symmetric operation:

$$\tilde{\psi}(\vec{x}, t) = \sum_{\vec{y}} L(\vec{x}, \vec{y}) \psi(\vec{y}, t) \quad (10)$$

has long been used as a means of improving the projection onto the low-lying states in correlation functions. Distillation and its descendants rely on the observation that we can write such as smearing matrix, for example that of three-dimensional Jacobi-smearing, as

$$L^{(J)} \equiv (1 - \frac{\kappa}{n} \Delta)^n = \sum_{i=1} f(\lambda_i) \xi^{(i)} \otimes \xi^{*(i)}, \quad (11)$$

where $\xi^{(i)}$ is the i^{th} (three-dimensional) eigenvector of Δ with eigenvalue λ_i , and $f(\lambda_i) = (1 - \frac{\kappa \lambda}{n})^n$. We now truncate the sum in eqn. 11 at some sufficient N_{vec} that the low-energy-scale physics is faithfully captured. We will take $f(\lambda) = 1$, and obtain the so-called distillation function

$$\square(t) = \sum_{i=1}^{N_{\text{vec}}} \xi^{(i)}(t) \xi^{(i)\dagger}(t). \quad (12)$$

which we now use as the kernel for our quark smearing in eqn. 10. Focusing on the mesons, we write the “smeared” two-point correlation function of eqn. 8 as

$$C_M(t, t') = \text{Tr} \langle \phi^A(t') \tau(t', t) \Phi^B(t) \tau^\dagger(t', t) \rangle, \quad (13)$$

where

$$\Phi_{\alpha\beta}^{A,ij} = \xi^{*(i)}(t) [\Gamma^A(t) \gamma_5]_{\alpha\beta} \xi^{(j)}(t') \quad (14)$$

$$\tau_{\alpha\beta}^{ij}(t, t') = \xi^{*(i)}(t') M_{\alpha\beta}^{-1}(t', t) \xi^{(j)}(t). \quad (15)$$

Here $i, j = 1, \dots, N_{\text{vec}}$, α, β are spinor indices, and t, t' the source and sink timeslices respectively.

The crucial observation in the above construction is that the Φ 's of eqn. 14 encode the structure of the operator, with, for example, derivatives now acting on the (quark-mass-independent) ξ 's, whilst the parallel transport of the quark fields is encoded in the “perambulators” of eqn. 15. This factorization enables operators to be constructed not only at the sink but also at the source with essentially unrestricted spatial structures, for any set of perambulators.

The Φ 's and the τ 's are the building blocks of our calculation, and analogous extensions allow the computation of baryon and multi-hadron operators, with annihilation diagrams at source and/or sink. The calculation of quark propagators is replaced by the calculation of the perambulators for some set of eigenvectors $\{\xi^i; i = 1, \dots, N_{\text{vec}}\}$; the calculation of the correlation functions is performed as in eqn. 13. Distillation has been used by numerous groups to compute the low-lying “single-hadron” meson spectrum[2, 3, 4, 5], the baryon spectrum[6, 7, 8], and scattering amplitudes in both resonant[9, 10, 11] and non-resonant channels[12, 13].

The number of eigenvectors N_{vec} needed to capture the low-lying physics is expected to scale as the (physical) spatial volume, and the cost of forming the resulting correlation functions scales as N_{vec}^3 or faster. Thus, as the volume is increased, not only does the total number of “perambulators” to be computed increase, but so does the cost of constructing the correlation functions; most of the calculations employing distillation have been performed on spatial volumes of 24^3 or less.

There are two recent approaches that have been used to overcome the challenge of proceeding to larger volumes. The first method is based on the stochastic sampling of the eigenvectors, the so-called “Stochastic LapH Method”[14]; whilst the number of eigenvectors required still scales with the volume, the number of *samples* appears to be largely insensitive to the volume. The use of this method has enabled the hadronic correlation functions needed for investigations of the excited-meson spectrum to be computed on spatial lattices as large as 32^3 [15]. The second approach relies on introducing a smearing kernel in eqn. 12 that dramatically reduces the cost of computing the correlation functions, namely through the use of a so-called color-wave basis, rather than overcoming the volume scaling[16]. This method has been applied successfully both to the study of LQCD at non-zero chemical potential[17], and to the investigation of tetra-quark bound states in heavy-light systems[16].

3.1. Improved Solvers

The need to extract as much information as possible from an ensemble of configurations typically requires that we obtain many solutions of the Dirac equation using the same background gauge configuration, but for many right-hand sides. Such a requirement is clear for distillation where the Dirac equation has to be solved for each of the eigenvectors of eqn. 15.

There are several numerical linear algebra methods that accelerate the solution of linear systems when the same coefficient matrix is used on multiple right hand sides. Because of the size and sparsity pattern of the Dirac matrix, dense-matrix methods that reuse the matrix factorization cannot be used. However, Krylov subspace algorithms that exploit the fact that solutions for multiple right hand sides are required have been proposed in the numerical linear algebra literature. These include Block methods [18]–[19], Seed methods [20]–[21], and deflation methods [22]–[23]. A combination of these methods has also been suggested (see for example [24]). The Block methods that solve simultaneously the linear system for many right hand sides achieve their efficiency primarily from the increased ratio of flops to memory access. These methods have already been tried in LQCD applications (for example see [25, 26]). Seed methods do not seem to work well for LQCD because the right hand sides that appear in these calculations are orthogonal, hence seeding subsequent solves with past solutions gain very little in efficiency. Deflation methods have been the most successful in accelerating the linear system solves with multiple right hand sides in LQCD. These methods gain their efficiency by solving exactly the linear system in a low eigenvalue subspace

hence effectively reducing the condition number of the system. Although, deflation can be implemented using eigenvectors produced by a stand alone eigen-solver, one can also produce estimates of low eigenvectors while solving the first few right hand sides and hiding this way the cost of eigenvector calculation. Such methods include GMRES-DR [22]–[23], the eigCG [27] for solving the normal equation problem, and the eigBiCG [28] that extends eigCG to the case of a non-hermitian coefficient matrix.

The main problem that deflation algorithms encounter is scaling with volume. In order to achieve constant reduction of the condition number as the volume of the system grows, one needs to grow the number of vectors used in the deflation proportionally with the volume. The result is that, at large volumes, the algorithm becomes inefficient, due to the large memory required to store the deflation vectors as well as the large number of inner products to achieve deflation. A solution to this problem was proposed by Luscher [29, 30], where a large basis of the low eigenvector subspace was constructed using geometric blocking of vectors designed to have strong overlap with this low eigenvector subspace. The locality of the Dirac operator, as well as the blocked nature of the basis used for deflation, allows for efficient deflation with a small memory footprint and small number of global inner products. Such a deflation strategy works well even on very large volumes.

Recently, Algebraic Multi-Grid (AMG) methods have been shown to be very efficient for LQCD applications [31, 32, 33, 34, 35, 36]. Lattices have now reached such extents[‡] that such methods represent an essential state-of-the-art solver technology that give orders of magnitude increase in performance over traditional methods, especially for calculations at physical quark masses.

3.2. All-mode Averaging

A recently proposed method of reducing the uncertainty in the extraction of physics information from a lattice calculation that exploits the ability to compute the low-eigenvalue spectrum is all-mode averaging (AMA), within the context of covariant approximation averaging[37, 38]. The idea is that for an observable or correlation function $\mathcal{C}$ one can perform a covariant transformation $g \in G$ such that the expectation value is invariant

$$\langle \mathcal{C}^g(U, \psi, \bar{\psi}) \rangle = \langle \mathcal{C}(U^g, \psi, \bar{\psi}) \rangle. \quad (16)$$

For the example considered in eqn. 5, G would be simply the translations of the two-point functions on a time slice, i.e. all possible $\vec{z}$, and all possible t_0 . Then we have

$$\langle \mathcal{C}_G[U] \rangle \equiv \frac{1}{N_G} \sum_{g \in G} \mathcal{C}^g[U] = \frac{1}{N_G} \sum_{g \in G} \mathcal{C}[U^g].$$

As we have seen, performing the sum over all $g \in G$ is in general prohibitively expensive. A far more efficient way of obtaining an improved estimate of $\langle \mathcal{C} \rangle$ is to introduce an approximation $\mathcal{C}^{(\text{approx})}$ that is less computationally demanding but still invariant under

[‡] Lattice sizes of $64^3 \times 128$ are now not uncommon.

$g \in G$, compute $\langle \mathcal{C} \rangle$ exactly for a single, or at least a few, $g \in G$, and then perform the average over all $g \in G$ on the cheaper approximation, whence

$$\langle \mathcal{C}^{(\text{imp})} \rangle = \{ \langle \mathcal{C} \rangle - \langle \mathcal{C}^{(\text{approx})} \rangle \} + \langle \mathcal{C}_G^{(\text{approx})} \rangle. \quad (17)$$

One way of constructing the approximate correlator is through replacing the quark propagators by some good approximation. The starting point is the eigenvectors of the (Hermitian) Dirac operator D_H :

$$D_H v_i = \lambda_i v_i, i = 1, \dots, N_{\text{tot}}, \quad (18)$$

where N_{tot} is the dimension of the Dirac matrix, and we assume the eigenvalues are ordered in λ_i . The spectral decomposition enables the propagator to be written as

$$G(x, y) = \sum_{i=1}^{N_{\text{tot}}} v_i(x) \frac{1}{\lambda_i} v_i(y).$$

In the case of low-mode averaging (LMA), only those eigenvectors $v_i, i = 1, \dots, N_{\text{cut}}$ corresponding to $|\lambda| < \lambda_{\text{cut}}$ are included, yielding

$$G_{\text{LMA}}(x, y) = \sum_{i=1}^{N_{\text{cut}}} v_i(x) \frac{1}{\lambda_i} v_i(y). \quad (19)$$

The calculation of correlation functions using G_{LMA} would be biased, but eqn. 17 removes that bias. However, it is likely to provide a poor variance-reduction method for correlation functions not dominated by the low modes, such as those in excited-state spectroscopy where we require the extraction of higher energies in the spectrum. An alternative, and better, method is provided by All-Mode Averaging (AMA), in which the lowest eigenmodes are treated exactly as in eqn. 19, and the contribution of the higher modes is treated approximately:

$$G_{\text{AMA}}(x, y) = \sum_{i=1}^{N_{\text{cut}}} v_i(x) \frac{1}{\lambda_i} v_i^\dagger(y) + P_n(D_H) \left(1 - \sum_{i=1}^{N_{\text{cut}}} v_i v_i^\dagger \right), \quad (20)$$

where P_n is some approximation to the propagator in the space orthogonal to the lowest N_{cut} modes.

All-mode averaging, despite its recent introduction, has already become an important member of the standard repertoire of lattice QCD methods, applied to such calculations in nuclear physics as that of the proton and neutron electric-dipole moments[39], nucleon electromagnetic form factors[40], with applications to other areas in progress. Whilst we have described AMA as applied to the computation of the quark propagators of eqn. 9, the application to “distillation” should be straightforward.

4. Generalized Eigenvalue Methods

The methods outlined above enable the construction of correlation functions that are key to exploring the properties of hadronic systems, notably the excited hadron spectrum through the use of the variational method. In this final section, we outline

the variational method through the solution of generalized eigenvalue equations, and provide an extension to the case where the matrix of correlations of eqn. 5 is not square, and indeed the basis of operators at source and sink can be distinct.

We have seen that a meson correlation function, such as that of eqn. 5 at zero momentum, behaves at large Euclidean times as

$$C_M(t, t') = Z e^{-m_M(t-t')} + \dots, \quad (21)$$

where m_M is the ground state mass of the hadron; the additional terms contain contributions from the excited states. These contributions are suppressed at large time separations, and that suppression is controlled by the gap between the first excited state and the ground state. However, at large Euclidean time separations the statistical noise from the Monte Carlo evaluation of the correlation functions grows exponentially relative to the signal[41]. Thus one means of improving the efficiency of the calculation is by developing methods that enable the energy spectrum to be obtained at small Euclidean time separations.

In the LQCD spectrum calculations cited previously, the variational method has been exploited to extract the key physics information from matrices of two-point correlation functions; these can be used not only to obtain excited state energy levels, but also to control the systematic errors and improve the statistical precision of the spectrum and properties of the lower-lying states; a recent review and explanation of how systematic errors are controlled in the energy spectrum calculations is provided in ref. [42]. In this latter work, a Hermitian matrix of correlation functions is used and the systematic effects on the computed energies are discussed leading to a robust methodology for such calculations. In the following we will describe how matrices of correlation functions lead to the extraction of the energy spectrum of hadrons in the slightly more general case of non-hermitian matrices; the hermitian case is contained as a special case of our discussion.

Let us assume that we have a non-symmetric matrix C_{ij} of correlation functions constructed from two sets of interpolating operators: $\mathcal{O}_i$ are the annihilation operators and $\tilde{\mathcal{O}}_j$ are the creation operators. Then we have

$$C_{ij}(t) = \langle \mathcal{O}_i(t) \tilde{\mathcal{O}}_j^\dagger(0) \rangle = \langle 0 | \mathcal{O}_i e^{-\hat{H}t} \tilde{\mathcal{O}}_j^\dagger | 0 \rangle, \quad (22)$$

where $\hat{H}$ is the Hamiltonian. This matrix can be decomposed as a sum over all energy eigenstates of the Hamiltonian $\hat{H}$

$$C_{ij}(t) = \sum_{n=1}^{\infty} Z_i^n \tilde{Z}_j^{n*} e^{-E_n t}, \quad (23)$$

where the eigenstates of the Hamiltonian $|n\rangle$ are

$$\hat{H}|n\rangle = E_n|n\rangle, \quad Z_i^n = \langle 0 | \mathcal{O}_i | n \rangle, \quad \tilde{Z}_j^{n*} = \langle n | \tilde{\mathcal{O}}_j^\dagger | 0 \rangle. \quad (24)$$

In matrix form, the spectral representation of the correlation function matrix can be written as

$$C(t) = Z \Lambda \tilde{Z}^\dagger, \quad (25)$$

with

$$\Lambda = \text{diag}(e^{-E_1 t}, e^{-E_2 t}, \dots) \equiv \text{diag}(\lambda_1(t), \lambda_2(t), \dots). \quad (26)$$

$$\lambda_i(t) \equiv e^{-E_i t}. \quad (27)$$

In the remainder of this section, we will assume that the energy levels are ordered

$$E_1 < E_2 < E_3 \dots; \quad (28)$$

degenerate energy levels introduce complexities that we will avoid for now.

If we now assume that the system has N energy levels and we use a set of N linearly independent interpolating fields $\mathcal{O}_A$ and N linearly independent interpolating fields $\tilde{\mathcal{O}}_A$ we can compute all the energy levels from the correlator matrix C_{AB} . One way of doing this is by constructing two matrices Q and $\tilde{Q}$ so that

$$Q^\dagger = (Z)^{-1}, \quad \tilde{Q} = (\tilde{Z}^\dagger)^{-1}. \quad (29)$$

This is always possible since Z and $\tilde{Z}$ are $N \times N$ matrices with a non-vanishing determinants. The fact that $\mathcal{O}_A$ and $\tilde{\mathcal{O}}_A$ are each linearly independent guarantees that $\det(Z) \neq 0$ and $\det(\tilde{Z}) \neq 0$. Using this property we can now extract the energies by

$$Q^\dagger C(t) \tilde{Q} = \Lambda(t). \quad (30)$$

It turns out that $\tilde{Q}$ is a solution to a generalized eigenvalue problem:

$$\begin{aligned} C(t) \tilde{Q} &= Q^\dagger \Lambda(t) = Z \Lambda(t_0) \Lambda(t - t_0) = \\ &= Z \Lambda(t_0) \tilde{Z}^\dagger \tilde{Q} \Lambda(t - t_0) = C(t_0) \tilde{Q} \Lambda(t - t_0) \end{aligned} \quad (31)$$

Note that the factorization of $\Lambda(t) = \Lambda(t_0) \Lambda(t - t_0)$ will not hold for the case where periodic images need to be included due to boundary conditions in the finite time direction. Equivalently the columns $\tilde{q}_i$ of the matrix $\tilde{Q}$ are the generalized right eigenvectors that satisfy

$$C(t) \tilde{q}_i = C(t_0) \tilde{q}_i \lambda_i(t - t_0); \quad (32)$$

Q can be similarly determined as the solution of a left generalized eigenvalue problem.

Given the above discussion the algorithm for determining the energy levels contributing to the correlation function is simple. We first solve the generalized eigenvalue problem of eqn. 31 and then determine the energy levels by taking the logarithm of the eigenvalues λ_i :

$$E_i = \frac{1}{\tau} \log \left(\frac{\lambda_i(t - t_0)}{\lambda_i(t + \tau - t_0)} \right). \quad (33)$$

In this simple example, the result is independent of t_0 and t and is exactly the eigenvalues of the Hamiltonian that contribute to the correlator. It can be shown that deviations from the ideal case where the basis of interpolating fields has dimension equal to the dimension of the space the Hamiltonian acts on, behaves similarly to the case of symmetric correlation matrices.

If the matrix of correlation functions is not sufficiently large, one can further increase the basis of interpolating fields using time translation invariance. In Euclidean space,

we translate the annihilation interpolating operators forward in time, and the creation interpolating fields backward in time as follows:

$$\mathcal{O}_A(t + m\tau) = e^{m\tau H} \mathcal{O}_A e^{-m\tau H} \quad (34)$$

$$\tilde{\mathcal{O}}_A(t - m\tau) = e^{-m\tau H} \mathcal{O}_A e^{m\tau H} \quad (35)$$

where τ is a choice of time step and m is an integer. Using time translation one can construct several new interpolating fields by increasing the integer m . With this in mind, we can always take a correlator matrix with unequal values for the row and column dimensions, and convert it to a square matrix by a sufficient number of time translations of the existing interpolating fields. To do so, we only need to translate in time the previously computed matrix of correlators since

$$\langle 0 | \mathcal{O}_A(t + m\tau) \tilde{\mathcal{O}}_B(t - n\tau) | 0 \rangle = C_{AB}(t + n\tau + m\tau) . \quad (36)$$

Using time translation, we can now work with a new square matrix of correlators K_{ij} .

With this enlarged time-translated basis of interpolating fields we now have a square matrix of correlators from which energy levels can be extracted as described above by solving the generalized eigenvalue problem

$$K(t) \tilde{q}_i = K(t_0) \tilde{q}_i \lambda_i(t - t_0) \quad , \quad q_i^\dagger K(t) = \lambda_i(t - t_0) q_i^\dagger K(t_0) \quad (37)$$

One complication that arises is that the interpolating fields are now no longer linearly independent. For example, if in the time range we are exploring, the size of the resulting basis is larger than the number of energy levels significantly contributing to the correlators, then the resulting linearly-dependent interpolating fields yields an ill-conditioned eigenvalue problem. This is avoided by use of a singular value decomposition (SVD). If we consider the SVD of $K(t_0)$ then we write

$$K(t_0) = U \Sigma V^\dagger, \quad (38)$$

with $S = \text{diag}(\sigma_i)$ a diagonal matrix with the singular values of $K(t_0)$ in the diagonal and U and V unitary matrices. We now define S' as the sub-matrix of S containing only the non-vanishing singular values; in practice we include the singular values σ_i that are larger than some predetermined tolerance. We also define U' and V' to be the columns of U and V that correspond to the selected non-vanishing singular values. We can now project the creation and annihilation operators to new sets that are linearly independent using

$$\tilde{\Omega}_i = \sum_k [\Sigma'^{-1/2} V'^\dagger]_{ik} \tilde{\mathcal{O}}_k \quad , \quad \Omega_i = \sum_k [\Sigma'^{-1/2} U'^\dagger]_{ik} \mathcal{O}_k \quad , \quad (39)$$

With the above projection the new correlation matrix $K'(t)$ becomes

$$K'(t) = \Sigma'^{-1/2} U'^\dagger K(t) V' \Sigma'^{-1/2} \quad (40)$$

§ As will be clear from the subsequent analysis, the matrix K could remain non-square. The resulting generalized eigenvalue problem can be solved with using SVD.

The dimension of $K'(t)$ is now equal to the number of non-vanishing singular values, and furthermore. $K'(t_0) = \mathbf{1}$. We can now re-write the generalized eigenvalue problem as regular eigenvalue problem using

$$K'(t)\tilde{\omega}_i = \tilde{\omega}_i\lambda_i(t - t_0) \quad , \quad \omega_i^\dagger K'(t) = \lambda_i(t - t_0)\omega_i^\dagger \quad (41)$$

This is now a regular non-symmetric matrix eigenvalue problem that can be solved for the left (ω) and right ($\tilde{\omega}$) eigenvectors and the eigenvalues $\lambda_i(t - t_0)$. The eigenvalues can now be used to extract the desired energy levels using eqn. 33, or through fits to single exponentials.

The idea of using multiple time shifts of a signal to recast the problem of estimating the contributing frequencies into a generalized eigenvalue problem has been used before in signal processing, for example refs. [43, 44]. In [43, 44], SVD was used to obtain the number as well as the values of the contributing frequencies. In the above, we have generalized the method to the case of a matrix of correlators as well as established the connection to the creation and annihilation operators of the underlying quantum field theory. This latter connection enables the use of the obtained eigenvectors and projection operators in the analysis of three point functions for extracting hadronic matrix elements [45].

5. Conclusions

We have described a variety of methods that have been introduced to enable increasingly sophisticated physical information to be obtained from lattice calculations. These methods have facilitated an understanding of the excited state spectrum of QCD, and are being employed to reveal the quark and gluon structure of hadrons. With the advent of exascale computing, the development and implementation of such methods is essential if the full physics potential of the emerging generation of leadership-class computers is to be realized.

Acknowledgement

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics under contract DE-AC05-06OR23177. In addition, KO was supported by the U.S. Department of Energy through Grant Number DE-FD02-04ER41302.

References

- [1] Michael Peardon et al. A Novel quark-field creation operator construction for hadronic physics in lattice QCD. *Phys.Rev.*, D80:054506, 2009.
- [2] Jozef J. Dudek, Robert G. Edwards, Michael J. Peardon, David G. Richards, and Christopher E. Thomas. Highly excited and exotic meson spectrum from dynamical lattice QCD. *Phys.Rev.Lett.*, 103:262001, 2009.

- [3] Jozef J. Dudek, Robert G. Edwards, Michael J. Peardon, David G. Richards, and Christopher E. Thomas. Toward the excited meson spectrum of dynamical QCD. *Phys.Rev.*, D82:034508, 2010.
- [4] Jozef J. Dudek, Robert G. Edwards, Balint Joo, Michael J. Peardon, David G. Richards, et al. Isoscalar meson spectroscopy from lattice QCD. *Phys.Rev.*, D83:111502, 2011.
- [5] Jozef J. Dudek, Robert G. Edwards, Peng Guo, and Christopher E. Thomas. Toward the excited isoscalar meson spectrum from lattice QCD. *Phys.Rev.*, D88(9):094505, 2013.
- [6] Robert G. Edwards, Jozef J. Dudek, David G. Richards, and Stephen J. Wallace. Excited state baryon spectroscopy from lattice QCD. *Phys.Rev.*, D84:074508, 2011.
- [7] Jozef J. Dudek and Robert G. Edwards. Hybrid Baryons in QCD. *Phys.Rev.*, D85:054016, 2012.
- [8] R.G. Edwards, J.J. Dudek, D.G. Richards, and S.J. Wallace. Baryon spectroscopy in lattice QCD with spin identification. *AIP Conf.Proc.*, 1432:33–38, 2012.
- [9] Jozef J. Dudek, Robert G. Edwards, and Christopher E. Thomas. Energy dependence of the ρ resonance in $\pi\pi$ elastic scattering from lattice QCD. *Phys.Rev.*, D87(3):034505, 2013.
- [10] C.B. Lang, Luka Leskovec, Daniel Mohler, and Sasa Prelovsek. Axial resonances $a_1(1260)$, $b_1(1235)$ and their decays from the lattice. *JHEP*, 04:162, 2014.
- [11] Jozef J. Dudek, Robert G. Edwards, Christopher E. Thomas, and David J. Wilson. Resonances in coupled $\pi K, \eta K$ scattering from quantum chromodynamics. 2014.
- [12] Jozef J. Dudek, Robert G. Edwards, Michael J. Peardon, David G. Richards, and Christopher E. Thomas. The phase-shift of isospin-2 pi-pi scattering from lattice QCD. *Phys.Rev.*, D83:071504, 2011.
- [13] Jozef J. Dudek, Robert G. Edwards, and Christopher E. Thomas. S and D-wave phase shifts in isospin-2 pi pi scattering from lattice QCD. *Phys.Rev.*, D86:034031, 2012.
- [14] Colin Morningstar, John Bulava, Justin Foley, Keisuke J. Juge, David Lenkner, et al. Improved stochastic estimation of quark propagation with Laplacian Heaviside smearing in lattice QCD. *Phys.Rev.*, D83:114505, 2011.
- [15] C. Morningstar, J. Bulava, B. Fahy, J. Foley, Y.C. Jhang, et al. Extended hadron and two-hadron operators of definite momentum for spectrum calculations in lattice QCD. *Phys.Rev.*, D88(1):014511, 2013.
- [16] Zachary S. Brown and Kostas Orginos. Tetraquark bound states in the heavy-light heavy-light system. *Phys.Rev.*, D86:114506, 2012.
- [17] William Detmold, Kostas Orginos, and Zhifeng Shi. Lattice QCD at non-zero isospin chemical potential. *Phys.Rev.*, D86:054507, 2012.
- [18] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, PA, USA, 2003.
- [19] A. El Guennouni, K. Jbilou, and H. Sadok. A block version of BiCGStab for linear systems with multiple right-hand sides. *Electronic Transactions on Numerical Analysis*, 16:129–142, 2003.
- [20] C. F. Smith. *The Performance of preconditioned iterative methods in computational electromagnetics*. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, IL, 1987.
- [21] Y. Saad. On the Lanczos method for solving symmetric linear systems with several right hand sides. *Math. Comp.*, 48:651–662, 1987.
- [22] R. B. Morgan. GMRES with deflated restarting. *SIAM J. Sci. Comput.*, 24:20–37, 2002.
- [23] Abdou M. Abdel-Rehim, Ronald B. Morgan, Dywayne A. Nicely, and Walter Wilcox. Deflated and restarted symmetric lanczos methods for eigenvalues and linear equations with multiple right-hand sides. *SIAM J. Sci. Comput.*, 32:129–149, 2010.
- [24] Ronald B. Morgan. Restarted block-GMRES with deflation of eigenvalues. *Applied Numerical Mathematics*, 54:222–236, 2005.
- [25] Y. Nakamura, K.-I. Ishikawa, Y. Kuramashi, T. Sakurai, and H. Tadano. Modified Block BiCGSTAB for Lattice QCD. *Comput.Phys.Commun.*, 183:34–37, 2012.
- [26] Y. Nakamura, K.I. Ishikawa, Y. Kuramashi, T. Sakurai, and H. Tadano. Modified block BiCGSTAB for lattice QCD. *PoS, LATTICE2011*:042, 2011.
- [27] Andreas Stathopoulos and Kostas Orginos. Computing and deflating eigenvalues while solving multiple right hand side linear systems in quantum chromodynamics. *SIAM J.Sci.Comput.*,

- 32:439–462, 2010.
- [28] A.M. Abdel-Rehim, Andreas Stathopoulos, and Kostas Orginos. Extending the eigCG algorithm to nonsymmetric Lanczos for linear systems with multiple right-hand sides. *Linear Algebra Appl.*, 21:473–493, 2014.
 - [29] Martin Luscher. Deflation acceleration of lattice QCD simulations. *JHEP*, 0712:011, 2007.
 - [30] Martin Luscher. Local coherence and deflation of the low quark modes in lattice QCD. *JHEP*, 0707:081, 2007.
 - [31] A. Frommer, K. Kahl, S. Krieg, B. Leder, and M. Rottmann. An adaptive aggregation based domain decomposition multilevel method for the lattice wilson dirac operator: multilevel results. 2013.
 - [32] Andreas Frommer, Karsten Kahl, Stefan Krieg, Bjrn Leder, and Matthias Rottmann. Adaptive Aggregation Based Domain Decomposition Multigrid for the Lattice Wilson Dirac Operator. 2013.
 - [33] James Brannick and Karsten Kahl. Bootstrap Algebraic Multigrid for the 2D Wilson Dirac system. 2013.
 - [34] J.C. Osborn, R. Babich, J. Brannick, R.C. Brower, M.A. Clark, et al. Multigrid solver for clover fermions. *PoS, LATTICE2010:037*, 2010.
 - [35] R. Babich, J. Brannick, R.C. Brower, M.A. Clark, T.A. Manteuffel, et al. Adaptive multigrid algorithm for the lattice Wilson-Dirac operator. *Phys.Rev.Lett.*, 105:201602, 2010.
 - [36] J. Brannick, R.C. Brower, M.A. Clark, J.C. Osborn, and C. Rebbi. Adaptive Multigrid Algorithm for Lattice QCD. *Phys.Rev.Lett.*, 100:041601, 2008.
 - [37] Thomas Blum, Taku Izubuchi, and Eigo Shintani. New class of variance-reduction techniques using lattice symmetries. *Phys.Rev.*, D88(9):094503, 2013.
 - [38] Eigo Shintani, Rudy Arthur, Thomas Blum, Taku Izubuchi, Chulwoo Jung, et al. Covariant approximation averaging. 2014.
 - [39] E. Shintani. Neutron and proton EDM with $N_f=2+1$ domain-wall fermion. *PoS, LATTICE2013:298*, 2013.
 - [40] Meifeng Lin. Nucleon Form Factors with $2+1$ Flavors of Domain Wall Fermions and All-Mode-Averaging. *PoS, LATTICE2013:275*, 2014.
 - [41] G. P. Lepage. The analysis of algorithms for lattice field theory. *Boulder TASI’89*, page 97, 1989.
 - [42] B. Blossier, M. Della Morte, G. von Hippel, T. Mendes, and R. Sommer. On the generalized eigenvalue method for energies and matrix elements in lattice field theory. *JHEP*, 0904:094, 2009.
 - [43] Y. Hua and T. K. Sarkar. Generalized pencil-of-function method for extracting poles of an em system from its transient response. *Antennas and Propagation*, 37:229, February 1989.
 - [44] T. Sarkar and O. Pereira. Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. *IEEE Antennas and Propagation Magazine*, 37:48, February 1995.
 - [45] C. Aubin and K. Orginos. A new approach for Delta form factors. 2010.