

Block Lanczos algorithm for lattice QCD spectroscopy and matrix elements

Daniel C. Hackett[✉] and Michael L. Wagman[✉]

Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA



(Received 27 February 2025; accepted 19 May 2025; published 25 July 2025)

Recent work introduced a new framework for analyzing correlation functions with improved convergence and signal-to-noise properties, as well as rigorous quantification of excited-state effects, based on the Lanczos algorithm and spurious eigenvalue filtering with the Cullum-Willoughby test. Here, we extend this framework to the analysis of correlation-function matrices built from multiple interpolating operators in lattice quantum chromodynamics (QCD) by constructing an oblique generalization of the block Lanczos algorithm, as well as a new physically motivated reformulation of the Cullum-Willoughby test that generalizes to block Lanczos straightforwardly. The resulting block Lanczos method directly extends generalized eigenvalue problem (GEVP) methods, which can be viewed as applying a single iteration of block Lanczos. Block Lanczos provides qualitative and quantitative advantages over GEVP methods analogous to the benefits of Lanczos over the standard effective mass, including faster convergence to ground- and excited-state energies, explicitly computable two-sided error bounds, straightforward extraction of matrix elements of external currents, and asymptotically constant signal-to-noise. No fits or statistical inference are required. Proof-of-principle calculations are performed for noiseless mock-data examples as well as two-by-two proton correlation-function matrices in lattice QCD.

DOI: [10.1103/fp74-q35q](https://doi.org/10.1103/fp74-q35q)

I. INTRODUCTION

Numerical lattice quantum chromodynamics (QCD) is an integral part of the particle and nuclear physics toolkit. In lattice QCD, Euclidean-time path integrals are discretized and evaluated stochastically with Monte Carlo methods to estimate QCD correlation functions, allowing first-principles calculation of hadronic energies and matrix elements. This approach has been employed with great success to study the dynamics of QCD, including its spectrum of states, scattering amplitudes, and various aspects of hadron structure, as well as to constrain the Standard Model. However, in practice, the mutually compounding issues of exponentially decaying signal-to-noise and excited-state contamination limit what is possible. This has led to ongoing research into improved methods to alleviate these issues.

Challenges in reliably extracting the spectrum of QCD from single correlation functions (“correlators”) have led to the development of methods which analyze correlation-function matrices (“correlator matrices”) defined from sets

of multiple interpolating operators (“interpolators”), each providing complementary information. This includes the variational method based on solving generalized eigenvalue problems (henceforth, GEVP methods or GEVP) [1–6], the block generalization [5] of Prony’s method [7–10], and variations on generalized pencil of functions (GPOF) [6,11–15]. GEVP methods are the present standard and state of the art for spectroscopy. In practice these methods provide better control over excited-state contamination than that available with single-correlator analyses, in both application to spectroscopy as well as in their generalizations to the extraction of matrix elements. Importantly, GEVP energy levels admit rigorous interpretation as variational upper bounds, allowing determination of excited-state energies with controlled uncertainties. However, these bounds are one-sided, leading to various well-known difficulties in interpretation.

Recent work [16,17] has explored a new analysis framework based on the Lanczos algorithm. In this framework, bounds on numerical Lanczos convergence properties can be reframed to provide rigorous two-sided bounds on energy levels in correlator analyses. Separately, tools developed to treat numerical noise from finite-precision arithmetic can be repurposed to discard spurious states that arise in application to noisy correlator data. The resulting Lanczos estimators have improved signal-to-noise properties, without the exponential degradation in Euclidean time separation present in “effective mass” estimators.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP³.

However, the method has thus far been limited to analysis of single (scalar) correlation functions.

In this work, we extend the Lanczos analysis formalism to treat correlator matrices. The framework is built around the block Lanczos algorithm [18–22], which generalizes the standard (scalar) Lanczos algorithm to use a basis of multiple vectors. We construct its oblique generalization [23–26], which uses distinct left and right bases, in the form of recursion relations that can be applied to treat noisy correlator data [16,17]. This method yields simple block generalizations of the recursion relations and change-of-basis matrices that arise for oblique Lanczos in the scalar case, allowing similar constructions of estimators for energies, overlap factors, and operator matrix elements.

The key element to the improved signal-to-noise properties of the Lanczos framework is filtering of spurious states that arise as noise artifacts. We reformulate the methods introduced in Refs. [16,17]—the Cullum-Willoughby and Hermitian subspace tests—to filter spurious noise-artifact states and generalize them to the block case. In the process, we find that the Cullum-Willoughby test used in previous examples admits a simple physical interpretation as identifying states that have zero overlap with the entire set of interpolating operators and therefore have “wrong quantum numbers” that should not arise from application of the physical transfer matrix. With this interpretation, the entire state filtering scheme can be understood in physical terms as restricting to the subspace of Hilbert space of physical interest.

We test the oblique block Lanczos method on noiseless mock data and a noisy lattice quantum chromodynamics (LQCD) 2×2 nucleon correlator matrix. We demonstrate that the improved convergence and signal-to-noise properties observed in the scalar case extend to the block case as well. We find that the block analysis provides nontrivial advantages over separate applications of the scalar analysis, including cleaner state resolution and access to relative phases of transition matrix elements.

We find that GEVP is the one-step limit of block Lanczos, and thus that block Lanczos is a direct extension and superset of GEVP methods. In comparisons on noiseless and noisy data, we find that block Lanczos offers similar advantages over GEVP as does scalar Lanczos over effective energies and other effective estimators.

The remainder of this paper proceeds as follows. In Sec. II, we construct an oblique block Lanczos algorithm and corresponding recursion relations for analyzing correlator matrices. In Sec. III, we apply the recursive block Lanczos analysis formalism to extraction of the spectrum, overlap factors, and matrix elements in noiseless mock-data examples, comparing its performance with both GEVP and scalar Lanczos. In Sec. IV, we discuss the convergence properties of this algorithm and construct two-sided bounds allowing rigorous determination of energies. In Sec. V, we discuss the physics of the Cullum-Willoughby test and

spurious eigenvalue filtering and demonstrate how block Lanczos can be applied to noisy data, extracting the spectrum and overlap factors. In Sec. VI, we review the GEVP formalism and discuss how it coincides with single-iteration block Lanczos. We conclude in Sec. VII.

Several appendixes complement the main text. Appendix A presents further details of the derivation of the oblique block Lanczos algorithm. Appendix B translates notation from Ref. [21] and derives the form of the Kaniel-Paige-Saad bound presented in Sec. IV. Appendix C presents additional analysis of noiseless mock data and addresses the potential for “pseudoplateau” behavior. Appendix D proves that correlator matrices constructed with most commonly used interpolating operators are real in expectation. Appendix E discusses the left GEVP and differences with the standard right GEVP that can arise for noisy data applications. Appendix F discusses Gaussianity tests and uncertainty assumptions for bootstrap-median and sample-mean estimators. Appendix G proves that block Lanczos is not equivalent to the correlator matrix analysis method introduced in Ref. [6] based on applying Prony’s method to GEVP results.

II. BLOCK LANCZOS FORMALISM

This section presents the oblique block Lanczos algorithm for rank- r correlator matrices of the form

$$C_{ab}(t) = \langle \chi_a(t) \psi_b^\dagger(0) \rangle = \langle \chi_a | T^t | \psi_b \rangle, \quad (1)$$

where $t \in \{0, \dots, N_t - 1\}$, T is the transfer matrix, and $\chi_a(t)$ and $\psi_a(t)$ are interpolating operators chosen to excite states

$$|\psi_b\rangle = \psi_b^\dagger |\Omega\rangle \quad \text{and} \quad |\chi_a\rangle = \chi_a^\dagger |\Omega\rangle \quad (2)$$

with the quantum numbers of physical interest from the vacuum $|\Omega\rangle$. For a zero-temperature lattice quantum field theory, the transfer matrix is typically Hermitian and positive definite.¹ For lattice gauge theories, T is infinite dimensional [28], in which case objects like $|\psi_a\rangle$ and $\langle \chi_a |$ are vectors in the infinite-dimensional Hilbert space of states. The (nonthermal)² eigenvalues of T are denoted λ_n

¹This implies real, positive eigenvalues. Complex transfer-matrix eigenvalues can arise in theories with improved actions, but they are typically exponentially suppressed in the continuum limit compared to a sector of (strictly positive) physical eigenvalues [27]. Negative eigenvalues can arise for theories with staggered fermions, where they can be treated analogously to thermal eigenvalues [16].

²For a finite-temperature theory, such as Euclidean LQCD with a finite temporal extent, T should be augmented to include “backwards-propagating” thermal states as described in Appendix A of Ref. [16]. The resulting thermal eigenvalues are eigenvalues of the augmented transfer matrix with magnitudes larger than unity [16]. Their determination is discussed in Sec. V F.

and ordered such that $1 \geq \lambda_0 \geq \lambda_1 \geq \dots \geq 0$. The spectrum of LQCD energies is related to the transfer matrix elements by $E_n = -\ln \lambda_n$ where units in which the lattice spacing is set to unity are used here and below.

In this section, Hermiticity is *not* assumed for either T or $C_{ab}(t)$. Restrictions to Hermitian T and to real, symmetric $C_{ab}(t)$ will be made in Secs. IV and V, respectively. To simplify the discussion of applications to noisy LQCD data it is convenient to begin by constructing the oblique block Lanczos algorithm for a general infinite-dimensional operator T and time series of $r \times r$ invertible matrices $C_{ab}(t)$. For $r = 1$, oblique block Lanczos reduces to the (scalar) oblique Lanczos algorithm [29].

To avoid proliferation of indices below, we use boldface matrix notation for block indices $a, b, c, d \in \{1, \dots, r\}$ while keeping all other indices explicit. When a symbol like $C_{ab}(t)$ is bolded, $\mathbf{C}(t)$, it indicates the same quantity as a matrix in the missing indices a, b . Brackets with subscripts around a matrix expression indicate taking the indices of the result, e.g.,

$$[\mathbf{a}_j]_{ab} \equiv \alpha_{jab} \quad \text{or} \quad [\mathbf{AB}]_{ab} \equiv \sum_c A_{ac} B_{cb}. \quad (3)$$

Bolded symbols are always matrices and never vectors. All sums are taken explicitly, with repeated indices never indicating summation.

The practical steps required to carry out a Lanczos correlator matrix analysis are summarized here for convenience:

- (1) Apply oblique block Lanczos recursion relations to $\mathbf{C}(t)$ to compute the elements of the block-tridiagonal matrix $\mathbf{T}_{ij}^{(m)}$ (Sec. II B).
- (2) Diagonalize $\mathbf{T}_{ij}^{(m)}$ to obtain Ritz values $\lambda_k^{(m)}$ and the eigenvectors $\omega_{iak}^{(m)}$ (Sec. II C). These provide all information necessary to estimate energies, overlaps (Sec. II F), and for spurious state filtering in the noisy case (Sec. V).
- (3) If residual bounds (Sec. IV) or operator matrix elements are desired, evaluate an auxiliary recursion to compute the Krylov coefficients $\mathbf{K}_{ij}^{R/L}$ (Sec. II D).
- (4) If operator matrix elements are desired, compute the Ritz coefficients $\mathbf{P}_{kta}^{R/L(m)}$ (Sec. II E), normalize the Ritz vectors by computing overlap factors (Sec. II F), and evaluate the matrix element estimators (Sec. II G).
- (5) For noisy data, identify and filter spurious states using the Hermitian-subspace test (Sec. V B) and the “block Z-factor Cullum-Willoughby (ZCW) test” (Sec. V C).
- (6) Estimate uncertainties idiomatically using (nested) bootstrapping (Sec. V E).
- (7) Extract thermal modes separately using a thermal ZCW test if desired (Sec. V F).

A. Oblique block Lanczos

An oblique block Lanczos algorithm iteratively constructs a sequence of left- and right-Lanczos vectors $\langle v_{ia}^L |$ and $|v_{jb}^R \rangle$ indexed by $i, j = 1 \dots m$ where m is the iteration count. The Lanczos vectors are defined to satisfy the generalized biorthogonality condition $\langle v_{ia}^L | v_{jb}^R \rangle = \delta_{ij} \delta_{ab}$ along with the three-term recurrence relations

$$\begin{aligned} T |v_{ja}^R \rangle &= \sum_b \left(|v_{jb}^R \rangle \alpha_{jba} + |v_{(j-1)b}^R \rangle \beta_{jba} \right. \\ &\quad \left. + |v_{(j+1)b}^R \rangle \gamma_{(j+1)ba} \right), \\ \langle v_{ja}^L | T &= \sum_b \left(\alpha_{jab} \langle v_{jb}^L | + \gamma_{jab} \langle v_{(j-1)b}^L | \right. \\ &\quad \left. + \beta_{(j+1)ab} \langle v_{(j+1)b}^L | \right), \end{aligned} \quad (4)$$

where the block-tridiagonal matrix elements α_{jab} , β_{jab} , and γ_{jab} are defined as

$$\begin{aligned} \alpha_{jab} &\equiv \langle v_{ja}^L | T | v_{jb}^R \rangle, \\ \beta_{jab} &\equiv \langle v_{(j-1)a}^L | T | v_{jb}^R \rangle, \\ \gamma_{jab} &\equiv \langle v_{ja}^L | T | v_{(j-1)b}^R \rangle. \end{aligned} \quad (5)$$

The consistency of these definitions with each other and with the biorthonormality condition $\langle v_{ia}^L | v_{jb}^R \rangle = \delta_{ij} \delta_{ab}$ is derived in Appendix A.

To construct each new set of Lanczos vectors, the algorithm 1) applies T and orthogonalizes with previous iterations to make a new set of left and right residual vectors, then 2) biorthonormalizes the new left and right residual vectors to obtain the Lanczos vectors. How to initialize the iteration is discussed below. Step 1 defines the residual vectors for $j > 1$ as

$$\begin{aligned} |r_{(j+1)a}^R \rangle &\equiv T |v_{ja}^R \rangle - \sum_b \left(|v_{jb}^R \rangle \alpha_{jba} + |v_{(j-1)b}^R \rangle \beta_{jba} \right), \\ \langle r_{(j+1)a}^L | &\equiv \langle v_{ja}^L | T - \sum_b \left(\alpha_{jab} \langle v_{jb}^L | + \gamma_{jab} \langle v_{(j-1)b}^L | \right), \end{aligned} \quad (6)$$

where terms involving $j - 1$ are omitted for the case $j = 1$. In each equation, the first term applies T to expand the Krylov space, and the remaining terms implement orthogonality with all previous Lanczos vectors. However, the new sets of residual vectors are not biorthonormal, as quantified by the residual norm,

$$\Delta_{jab} \equiv \langle r_{ja}^L | r_{jb}^R \rangle. \quad (7)$$

Step 2 thus decomposes the residual norm as

$$\Delta_{jab} = \sum_c \beta_{jac} \gamma_{jcb} \quad (\text{i.e., } \Delta_j = \beta_j \gamma_j) \quad (8)$$

to define the new set of Lanczos vectors as

$$\begin{aligned} |v_{ja}^R\rangle &\equiv \sum_b |r_{jb}^R\rangle \gamma_{jba}^{-1}, \\ \langle v_{ja}^L| &\equiv \sum_b \beta_{jab}^{-1} \langle r_{jb}^L|, \end{aligned} \quad (9)$$

for which $\langle v_{ja}^L | v_{jb}^R \rangle = \delta_{ab}$. As in the case of scalar oblique Lanczos, there is no unique correct choice for the decomposition $\Delta_j = \beta_j \gamma_j$. The same feature³ leading to a three-term recurrence is achieved for any definition of $|v_{(j+1)a}^R\rangle$ with $\text{span}\{|v_{(j+1)a}^R\rangle\} = \text{span}\{|r_{(j+1)a}^R\rangle\}$. Thus, any choice for which β_j and γ_j are both invertible is a valid convention.

We present several options for block oblique conventions for concreteness, but first emphasize that all physical results are necessarily independent of convention; other than insight, different choices can only offer advantages in numerical precision and efficiency. One simple choice is to take $\beta_j = \Delta_j$ such that $\gamma_j = \mathbf{1}$; this is invoked as a simple way to connect block Lanczos to GEVP methods in Sec. VI. Roots of matrices offer another set of options, as well as any factorization applicable for complex, non-Hermitian matrices; this includes lower-upper (LU) and QR, but not Cholesky. Another class of conventions follows from the eigendecomposition

$$\Delta_{jab} = \sum_c \Omega_{jac}^{-1} \Lambda_{jc} \Omega_{jcb}, \quad (10)$$

which for any choice of $\rho_{ja} \tau_{ja} = \Lambda_{ja}$ can be used to define

$$\gamma_j = \rho_j \Omega_j^{-1}, \quad \beta_j = \Omega_j \tau_j, \quad (11)$$

where ρ_j and τ_j are diagonal matrices with entries $\rho_a \delta_{ab}$ and $\tau_a \delta_{ab}$, respectively. This convention intuitively separates orthogonalization, implemented by Ω and Ω^{-1} , and normalization, implemented by ρ and τ ; the remaining ambiguity in the choice of ρ , τ is the same ambiguity as arises in the scalar case, which is recovered identically for $r = 1$ where $\Omega_j = \mathbf{1}$. This convention is used to produce the numerical results below.

³The left and right Lanczos vectors $\langle v_{(j+1)a}^L|$ and $|v_{(j+1)a}^R\rangle$ are defined so that their spans are identical to those of the residuals $\langle r_{(j+1)a}^L|$ and $|r_{(j+1)a}^R\rangle$. This ensures that the action of T on the j -th Lanczos vectors results only in terms proportional to the $(j+1)$ -th Lanczos vectors in addition to the terms proportional to the $(j-1)$ -th Lanczos vectors explicitly appearing in Eq. (6). This feature is what leads to the three-term recurrence in Eq. (4).

In practice, the initial blocks of right and left states, $|\psi_a\rangle$ and $\langle\chi_a|$, will typically not satisfy biorthonormality and thus cannot be used immediately as the initial Lanczos vectors $|v_{1a}^{R/L}\rangle$. They instead must be treated as with the residual vectors $|r_{ja}^{R/L}\rangle$ above and orthonormalized before proceeding. It is cleanest to define this similarly to subsequent biorthonormalizations, in terms of the decomposition

$$\langle\chi_a|\psi_b\rangle \equiv [\Delta_1]_{ab} = [\beta_1 \gamma_1]_{ab}, \quad (12)$$

with a similar choice of convention required. This then defines the initial Lanczos vectors as

$$\begin{aligned} |v_{1a}^R\rangle &= \sum_b |\psi_b\rangle \gamma_{1ba}^{-1}, \\ \langle v_{1a}^L| &= \sum_b \beta_{1ab}^{-1} \langle\chi_b|. \end{aligned} \quad (13)$$

These relations are used ubiquitously below to relate expressions in terms of Lanczos vectors to correlator data via

$$\begin{aligned} \langle v_{1a}^L | T^t | v_{1b}^R \rangle &= \sum_{cd} \beta_{1ac}^{-1} \langle\chi_c | T^t | \psi_d \rangle \gamma_{1db}^{-1} \\ &= [\beta_1^{-1} C(t) \gamma_1^{-1}]_{ab}, \end{aligned} \quad (14)$$

and similar expressions for three-point functions and quantities defined with, e.g., right Lanczos vectors on both sides, etc. However, we emphasize that β_1 and γ_1 do not appear explicitly in the tridiagonal matrix $T^{(m)}$ defined below, and instead generalize the normalization factors $|\psi|$ and $|\chi|$ from the scalar case.

B. Recursion relations

To apply oblique block Lanczos to analyze LQCD correlator matrices, the iterative procedure on infinite-dimensional Hilbert-space vectors defined above must be used to derive recursion relations on finite-dimensional matrix elements—i.e., correlation function data of the form Eq. (1).

To begin, we first compute β_1 and γ_1 by decomposing (with whatever choice of convention)

$$C_{ab}(0) = \langle\chi_a|\psi_b\rangle = [\beta_1 \gamma_1]_{ab}, \quad (15)$$

per Eq. (12) above, recognizing the definition Eq. (1). When $r = 1$, these reduce to the ubiquitous factors of $C(0)$ in the scalar formalism. The first block of the Lanczos approximation to the transfer matrix is then given by

$$\alpha_{1ab} \equiv \langle v_{1a}^L | T | v_{1b}^R \rangle = \sum_{cd} \beta_{1ac}^{-1} C_{cd}(1) \gamma_{1db}^{-1}, \quad (16)$$

or in matrix form

$$\alpha_1 = \beta_1^{-1} C(1) \gamma_1^{-1}. \quad (17)$$

As discussed further in Sec. VI A, the eigenvalues of α_1 are the same as the generalized eigenvalues provided by the GEVP method, given appropriately aligned definitions.

Further iterations allow improving this initial approximation of the transfer matrix. The recursions are most naturally defined in terms of the residuals

$$\Delta_{jab} = \langle r_{ja}^L | r_{jb}^R \rangle, \quad (18)$$

and the generalized correlators

$$\begin{aligned} A_{jab}(t) &\equiv \langle v_{ja}^L | T^t | v_{jb}^R \rangle, \\ G_{jab}(t) &\equiv \langle v_{ja}^L | T^t | v_{(j-1)b}^R \rangle, \\ B_{jab}(t) &\equiv \langle v_{(j-1)a}^L | T^t | v_{jb}^R \rangle. \end{aligned} \quad (19)$$

The $j = 1$ and $j = 2$ terms form the recursion base case. For $j = 1$, the matrix elements defining G_1 and B_1 are not well defined and only

$$A_1(t) = \beta_1^{-1} C(t) \gamma_1^{-1}. \quad (20)$$

is required. The residual norm after the first iteration can then be computed and decomposed as

$$\Delta_2 = A_1(2) - \alpha_1 \alpha_1 \equiv \beta_2 \gamma_2, \quad (21)$$

to define β_2 and γ_2 using some convention for decomposing Δ_2 into a product of matrices as discussed above. The next step proceeds by computing

$$G_2(t) = \beta_2^{-1} [A_1(t+1) - \alpha_1 A_1(t)], \quad (22)$$

$$B_2(t) = [A_1(t+1) - A_1(t) \alpha_1] \gamma_2^{-1}, \quad (23)$$

$$\begin{aligned} A_2(t) &= \beta_2^{-1} [A_1(t+1) + \alpha_1 A_1(t) \alpha_1 \\ &\quad - (\alpha_1 A_1(t+1) + A_1(t+1) \alpha_1)] \gamma_2^{-1}. \end{aligned} \quad (24)$$

Once these initial computations are completed to seed the recursion, the calculation of quantities for subsequent $j \geq 3$ proceeds regularly. In each step, first, β_{j+1} and γ_{j+1} are obtained by decomposing the residual norm

$$\Delta_{j+1} = A_j(2) - \alpha_j \alpha_j - \gamma_j \beta_j \equiv \beta_{j+1} \gamma_{j+1}. \quad (25)$$

Then, the next set of generalized correlators can be computed as

$$G_{j+1}(t) = \beta_{j+1}^{-1} [A_j(t+1) - \alpha_j A_j(t) - \gamma_j B_j(t)], \quad (26)$$

$$B_{j+1}(t) = [A_j(t+1) - A_j(t) \alpha_j - G_j(t) \beta_j] \gamma_{j+1}^{-1}, \quad (27)$$

$$\begin{aligned} A_{j+1}(t) &= \beta_{j+1}^{-1} [A_j(t+2) - (\alpha_j A_j(t+1) + A_j(t+1) \alpha_j) \\ &\quad + \alpha_j A_j(t) \alpha_j + \gamma_j A_{(j-1)}(t) \beta_j \\ &\quad - (\gamma_j B_j(t+1) + G_j(t+1) \beta_j) \\ &\quad + \gamma_j B_j(t) \alpha_j + \alpha_j G_j(t) \beta_j] \gamma_{j+1}^{-1}. \end{aligned} \quad (28)$$

The next diagonal block of the Lanczos transfer matrix approximation is then given by

$$\alpha_{j+1} = A_{j+1}(1). \quad (29)$$

Although β_j and γ_j are computed otherwise, the off-diagonal blocks satisfy

$$\gamma_{j+1} = G_{j+1}(1) \quad \text{and} \quad \beta_{j+1} = B_{j+1}(1), \quad (30)$$

which can be useful for consistency checks.

C. Ritz values and vectors

After m iterations of oblique block Lanczos, the block-tridiagonal matrix

$$T_{iajb}^{(m)} \equiv \langle v_{ia}^L | T | v_{jb}^R \rangle, \quad (31)$$

expressing matrix elements of T in the Lanczos-vector basis, is given by

$$T_{iajb}^{(m)} = \begin{pmatrix} \alpha_{1ab} & \beta_{2ab} & & & 0 \\ \gamma_{2ab} & \alpha_{2ab} & \beta_{3ab} & & \\ & \gamma_{3ab} & \alpha_{3ab} & \ddots & \\ & & \ddots & \ddots & \beta_{(m-1)ab} \\ 0 & & & \gamma_{(m-1)ab} & \alpha_{(m-1)ab} & \beta_{mab} \\ & & & & \gamma_{mab} & \alpha_{mab} \end{pmatrix}_{ij}. \quad (32)$$

This matrix may be understood as the matrix elements of the true transfer matrix T in the Krylov subspace of Hilbert space, i.e.,

$$\langle v_{ia}^L | T^{(m)} | v_{jb}^R \rangle \equiv \begin{cases} T_{iajb}^{(m)}, & i, j \leq m \\ 0, & \text{otherwise} \end{cases}. \quad (33)$$

It can be used to define a Hilbert-space operator

$$\begin{aligned} T^{(m)} &= \sum_{i,j=1}^m \sum_{ab} |v_{ia}^R\rangle \langle v_{ia}^L | T | v_{jb}^R\rangle \langle v_{jb}^L| \\ &= \sum_{ijab} |v_{ia}^R\rangle T_{iajb}^{(m)} \langle v_{jb}^L|, \end{aligned} \quad (34)$$

which can be identified as T multiplied by projection operators $\sum_{ia} |v_{ia}^R\rangle\langle v_{ia}^L|$ that restrict its left- and right-action to left- and right-Krylov spaces defined by

$$\begin{aligned}\mathcal{K}^{L(m)} &= \text{span}\{|v_{ia}^L\rangle | i = 1, \dots, m, a = 1, \dots, r\}, \\ \mathcal{K}^{R(m)} &= \text{span}\{|v_{ia}^R\rangle | i = 1, \dots, m, a = 1, \dots, r\}.\end{aligned}\quad (35)$$

Thus, $T^{(m)}$ provides a Krylov-space approximation to the Hilbert-space operator T .

In this sense, the eigenvalues and eigenvectors of the Hilbert-space operator $T^{(m)}$ provide the Lanczos algorithm's optimal approximations of the true transfer matrix eigenvalues and eigenvectors in the Krylov subspace. Viewing (ia) and (jb) as composite indices, the matrix may be diagonalized as

$$T_{iajb}^{(m)} = \sum_k \omega_{iak}^{(m)} \lambda_k^{(m)} (\omega^{-1})_{kjb}^{(m)}, \quad (36)$$

where ω^{-1} is the matrix inverse of $\omega_{k(jb)}$. The eigenvalues $\lambda_k^{(m)}$ are the Ritz values, from which Lanczos energy estimators are obtained via

$$E_k^{(m)} \equiv -\ln \lambda_k^{(m)}. \quad (37)$$

These are the primary quantities required for LQCD spectroscopy analyses. Calculations of matrix elements and residual bounds require the Ritz vectors, $|y_k^{R/L(m)}\rangle$, which decompose the Hilbert-space $T^{(m)}$ as

$$T^{(m)} = \sum_k |y_k^{R(m)}\rangle \lambda_k^{(m)} \langle y_k^{L(m)}|. \quad (38)$$

Inserting Eq. (36) into Eq. (34) and comparing with Eq. (38), we may identify the right and left eigenvectors $\omega_k^{(m)}$ and $(\omega^{-1})_k^{(m)}$ of $T_{iajb}^{(m)}$ as change-of-basis matrices relating Lanczos and Ritz vectors,

$$\begin{aligned}|y_k^{R(m)}\rangle &\equiv \mathcal{N}_k^{(m)} \sum_{i=1}^m \sum_a |v_{ia}^R\rangle \omega_{iak}^{(m)}, \\ \langle y_k^{L(m)}| &\equiv \frac{1}{\mathcal{N}_k^{(m)}} \sum_{i=1}^m \sum_a (\omega^{-1})_{kia}^{(m)} \langle v_{ia}^L|,\end{aligned}\quad (39)$$

where the $\mathcal{N}_k^{(m)}$ are constants used to normalize the Ritz vectors as discussed further below. We may verify this construction by applying the Hilbert-space $T^{(m)}$ to both sides of Eq. (39). We first use Eq. (34) and $\langle v_{ia}^L | v_{jb}^R \rangle = \delta_{ij} \delta_{ab}$ to find the action of $T^{(m)}$ on the Lanczos vectors,

$$\begin{aligned}T^{(m)} |v_{jb}^R\rangle &= \sum_{i=1}^m \sum_b |v_{ia}^R\rangle T_{iajb}^{(m)}, \\ \langle v_{ia}^L | T^{(m)} &= \sum_{j=1}^m \sum_b T_{iajb}^{(m)} \langle v_{jb}^L|.\end{aligned}\quad (40)$$

It is then straightforward to show that

$$\begin{aligned}T^{(m)} |y_k^{R(m)}\rangle &= \mathcal{N}_k^{(m)} \sum_{j=1}^m \sum_b T^{(m)} |v_{jb}^R\rangle \omega_{jbk}^{(m)} \\ &= \mathcal{N}_k^{(m)} \sum_{j=1}^m \sum_{ab} |v_{ia}^R\rangle T_{iajb}^{(m)} \omega_{jbk}^{(m)} \\ &= \mathcal{N}_k^{(m)} \sum_{i=1}^m \sum_a |v_{ia}^R\rangle \omega_{iak}^{(m)} \lambda_k^{(m)} \\ &= |y_k^{R(m)}\rangle \lambda_k^{(m)},\end{aligned}\quad (41)$$

i.e., that $|y_k^{R(m)}\rangle$ is a right eigenvector of $T^{(m)}$. An identical proof applies to $\langle y_k^{L(m)}|$. The properties and uses of Ritz vectors are discussed further in the remainder of this section and in Sec. IV.

D. Krylov coefficients

Matrix element calculations essentially require performing changes of basis from the ‘‘Krylov basis’’ in which correlation functions are naturally defined to the Lanczos- and Ritz-vector bases [17]. The block Lanczos change-of-basis matrices are straightforward generalizations of their scalar Lanczos counterparts and explicitly constructed in this and the next subsection.

Krylov-basis vectors are defined by

$$|k_{ia}^R\rangle \equiv T^i |v_{1a}^R\rangle, \quad \langle k_{ia}^L| \equiv \langle v_{1a}^L | T^i. \quad (42)$$

The Krylov coefficients K_{ij}^R and K_{ji}^L are defined to relate the Lanczos and Krylov bases as

$$\begin{aligned}|v_{ja}^R\rangle &= \sum_{i=0}^{j-1} \sum_b |k_{ib}^R\rangle K_{ijba}^R, \\ \langle v_{ja}^L| &= \sum_{i=0}^{j-1} \sum_b K_{jtab}^L \langle k_{ib}^L|.\end{aligned}\quad (43)$$

As in the scalar case [17], these may be computed from the elements of $T_{iajb}^{(m)}$ by deriving an auxiliary recursion. The $j = 1$ Krylov coefficients follow directly from these definitions,

$$K_{i1ab}^R = K_{1tab}^L = \delta_{i0} \delta_{ab}. \quad (44)$$

The $j = 2$ case can be computed by inserting these definitions into the recurrence Eq. (4), obtaining

$$\begin{aligned} K_{t2}^R &= \delta_{t1}\gamma_2^{-1} - \delta_{t0}\alpha_1\gamma_2^{-1}, \\ K_{2t}^L &= \delta_{t1}\beta_2^{-1} - \delta_{t0}\beta_2^{-1}\alpha_1. \end{aligned} \quad (45)$$

Krylov coefficients with $j > 2$ can be computed from these $j \in \{1, 2\}$ results using the recursion relations

$$\begin{aligned} K_{t(j+1)}^R &= \left(K_{(t-1)j}^R - K_{tj}^R \alpha_j - K_{t(j-1)}^R \beta_j \right) \gamma_{j+1}^{-1}, \\ K_{(j+1)t}^L &= \beta_{j+1}^{-1} \left(K_{j(t-1)}^L - \alpha_j K_{jt}^L - \gamma_j K_{(j-1)t}^L \right). \end{aligned} \quad (46)$$

Krylov coefficients can be used to compute Lanczos vector norms as

$$\begin{aligned} \langle v_{ia}^R | v_{jb}^R \rangle &= \sum_{s=0}^{i-1} \sum_{t=0}^{j-1} \sum_{cd} K_{sica}^{R*} \langle v_{1c}^R | (T^\dagger)^s T^t | v_{1d}^R \rangle K_{tjdb}^R \\ &= \sum_{st} [[K_{si}^R]^\dagger [\gamma_1^{-1}]^\dagger C(s+t) \gamma_1^{-1} K_{tj}^R]_{ab}, \\ \langle v_{ia}^L | v_{jb}^L \rangle &= \sum_{s=0}^{i-1} \sum_{t=0}^{j-1} \sum_{cd} K_{tsac}^L \langle v_{1c}^L | T^s (T^\dagger)^t | v_{1d}^L \rangle K_{jtb}^{L*} \\ &= \sum_{st} [K_{is}^L \beta_1^{-1} C(s+t) [\beta_1^{-1}]^\dagger [K_{jt}^L]^\dagger]_{ab}. \end{aligned} \quad (47)$$

Both of these expressions reduce to $\delta_{ij}\delta_{ab}$ in noiseless applications where $T^{(m)}$ is Hermitian. They are useful in particular for computing the residual bounds described below.

E. Ritz coefficients

The block-tridiagonal matrix eigenvectors and Krylov coefficients may be combined to compute the Ritz coefficients

$$\begin{aligned} P_{kta}^{R(m)} &\equiv \mathcal{N}_k^{(m)} \sum_{i=1}^m \sum_b K_{taib}^R \omega_{ibk}^{(m)}, \\ P_{kta}^{L(m)} &\equiv \frac{1}{\mathcal{N}_k^{(m)}} \sum_{i=1}^m \sum_b (\omega^{-1})_{kib}^{(m)} K_{ibta}^L, \end{aligned} \quad (48)$$

which directly relate the Ritz and Krylov vectors as

$$\begin{aligned} |y_k^{R(m)}\rangle &= \sum_{t=0}^{m-1} \sum_a P_{kta}^{R(m)} |k_{ta}^R\rangle, \\ \langle y_k^{L(m)}| &= \sum_{t=0}^{m-1} \sum_a \langle k_{ta}^L | P_{kta}^{L(m)}. \end{aligned} \quad (49)$$

The factors $\mathcal{N}_k^{(m)}$ are arbitrary but can be set to enforce unit normalization using overlap factors as described in the next subsection.

It can be useful to define Hilbert-space operators called Ritz rotators, polynomials in the transfer matrix T built from the Ritz coefficients as

$$P_{ka}^{R/L(m)} \equiv \sum_{t=0}^{m-1} P_{kta}^{R/L(m)} T^t, \quad (50)$$

which excite the Ritz vectors from the initial block of Lanczos vectors as

$$\begin{aligned} |y_k^{R(m)}\rangle &= \sum_a P_{ka}^{R(m)} |v_{1a}^R\rangle, \\ \langle y_k^{L(m)}| &= \sum_a \langle v_{1a}^L | P_{ka}^{L(m)}. \end{aligned} \quad (51)$$

Inserting Eq. (13) obtains versions which can be applied to the starting states prior to biorthonormalization,

$$\begin{aligned} |y_k^{R(m)}\rangle &= \sum_{ab} [\gamma_{1ba}^{-1} P_{ka}^{R(m)}] |\psi_b\rangle, \\ \langle y_k^{L(m)}| &= \sum_{ab} \langle \chi_b | [P_{ka}^{L(m)} \beta_{1ab}^{-1}]. \end{aligned} \quad (52)$$

Because the Ritz vectors as well as $|\psi_a\rangle$ and $|\chi_a\rangle$ are convention-independent, it follows that the combinations

$$\sum_a \gamma_{1ba}^{-1} P_{kta}^{R(m)} \quad \text{and} \quad \sum_a P_{kta}^{L(m)} \beta_{1ab}^{-1}, \quad (53)$$

must be convention-independent as well, which can be useful for consistency checks.

F. Overlap factors

Overlap factors between Ritz vectors and the starting interpolating operators can be computed as

$$\begin{aligned} [Z_{ka}^{R(m)}]^* &\equiv \langle \chi_a | y_k^{R(m)} \rangle \\ &= \mathcal{N}_k^{(m)} \sum_{i=1}^m \sum_{bc} \beta_{1ab} \langle v_{1b}^L | v_{ic}^R \rangle \omega_{ick}^{(m)} \\ &= \mathcal{N}_k^{(m)} \sum_b \beta_{1ab} \omega_{1bk}^{(m)}, \\ Z_{ka}^{L(m)} &\equiv \langle y_k^{L(m)} | \psi_a \rangle \\ &= \frac{1}{\mathcal{N}_k^{(m)}} \sum_{j=1}^m \sum_{bc} (\omega^{-1})_{kjc}^{(m)} \langle v_{jc}^L | v_{1b}^R \rangle \gamma_{1ba} \\ &= \frac{1}{\mathcal{N}_k^{(m)}} \sum_b (\omega^{-1})_{k1b}^{(m)} \gamma_{1ba}. \end{aligned} \quad (54)$$

As in the scalar case [17], the overlap factors provide a convenient means of computing the Ritz vector normalization constants $\mathcal{N}_k^{(m)}$ in the diagonal case where

$|\chi_a\rangle = |\psi_a\rangle$. The true transfer matrix is Hermitian, with degenerate left and right eigenvectors. Physical interpretability of Lanczos outputs requires that this holds also for the Lanczos approximation to these eigenvectors, i.e., that right and left Ritz vectors coincide as $|y_k^{R(m)}\rangle = |y_k^{L(m)}\rangle$. When $|\chi_a\rangle = |\psi_a\rangle$, this requires $Z_{ka}^{R(m)} = Z_{ka}^{L(m)}$, which in turn implies the $\mathcal{N}_k^{(m)}$ can be computed as

$$|\mathcal{N}_k^{(m)}|^2 = \frac{\sum_b (\omega^{-1})_{k1b}^{(m)} \gamma_{1ba}}{\left[\sum_c \beta_{1ac} \omega_{1ck}^{(m)} \right]^*} \quad \forall a, \quad (55)$$

where there is no summation over a implied, and the right-hand side must be independent of a so that the equality is true for all a as indicated. For noiseless data and a symmetric oblique convention $\beta_j = \gamma_j$, these normalization factors will equal unity automatically. For asymmetric conventions and in applications to noisy LQCD data, the $|\mathcal{N}_k^{(m)}|^2$ can be computed directly using Eq. (55). In the off-diagonal case $|\chi_a\rangle \neq |\psi_a\rangle$, $Z_{ka}^{R(m)} \neq Z_{ka}^{L(m)}$ and moreover right and left Ritz vectors can only be approximately equal, so (at least this version of) the norm trick does not apply for spurious state filtering and may distort results if applied naively.

The overlap factors together with the Ritz values provide all information necessary to diagnose and discard spurious states which arise in applications to noisy LQCD data. As in the scalar case, Eq. (55) provides a means to diagnose noise-artifact states which arise due to non-Hermiticity of $T^{(m)}$: when the computation yields a nonreal or nonpositive $|\mathcal{N}_k^{(m)}|^2$, it signals an inconsistency, specifically that $|y_k^{R(m)}\rangle \neq |y_k^{L(m)}\rangle$. The a independence of Eq. (55) provides a separate, nontrivial constraint not present in the scalar case that only holds for physical states. Finally, we find the Cullum-Willoughby test as employed in Refs. [16,17] can be reformulated as a cut on anomalously small overlap factors. These points are discussed in further detail in Sec. V where they may be explored with example data.

Once $|\mathcal{N}_k^{(m)}|^2$ is determined, normalized Ritz coefficients provide an alternate means to compute the overlap factors as

$$\begin{aligned} [Z_{ka}^{R(m)}]^* &= \sum_{t=0}^{m-1} \sum_b \langle \chi_a | T^t | v_{1b}^R \rangle P_{ktb}^{R(m)} \\ &= \sum_{t=0}^{m-1} \sum_{bc} C_{ac}(t) \gamma_{1cb}^{-1} P_{ktb}^{R(m)}, \\ Z_{ka}^{L(m)} &= \sum_{t=0}^{m-1} \sum_b P_{ktb}^{L(m)} \langle v_{1b}^L | T^t | \psi_a \rangle \\ &= \sum_{t=0}^{m-1} \sum_{bc} P_{ktb}^{L(m)} \beta_{1bc}^{-1} C_{ca}(t). \end{aligned} \quad (56)$$

This can be useful for consistency checks.

G. Operator matrix elements

Matrix elements of generic (local or nonlocal) operators J can be computed from three-point correlation function matrices,

$$C_{ab}^{3pt}(\sigma, \tau) \equiv \langle \psi'_a | T^\sigma J T^\tau | \psi_b \rangle, \quad (57)$$

where primes denote quantities for the sector of final states, which could be distinct from those of the initial states (e.g., in the case of off-forward matrix elements with different initial and final momenta). As in Ref. [17], the strategy is simply to compute $\langle y_f^{L(m)} | J | y_i^{R(m)} \rangle$, i.e., the matrix elements in the basis of Ritz vectors. The block generalization adds no formal complications over the scalar case. Inserting the definition of the Ritz rotators, Eq. (51), yields the prescription

$$\begin{aligned} \langle y_f^{L(m)} | J | y_i^{R(m)} \rangle &= \sum_{\sigma\tau ab} P_{f\sigma a}^{L(m)} \langle v_{1a}^{L'} | T^\sigma J T^\tau | v_{1b}^R \rangle P_{itb}^{R(m)} \\ &= \sum_{\sigma\tau ab} P_{f\sigma a}^{L(m)} [\beta_1^{-1} C^{3pt}(\sigma, \tau) \gamma_1^{-1}]_{ab} P_{itb}^{R(m)}, \end{aligned} \quad (58)$$

where unprimed and primed quantities are computed from separate diagonal two-point correlator matrices with initial- and final-state quantum numbers,

$$C_{ab}(t) = \langle \psi_a | T^t | \psi_b \rangle \quad \text{and} \quad C'_{ab}(t) = \langle \psi'_a | T^t | \psi'_b \rangle. \quad (59)$$

Thus, as in the scalar case, once two-point data has been analyzed, oblique block Lanczos allows explicit computation of matrix elements for (approximate) energy eigenstates by simple matrix multiplication of three-point correlator matrices with no additional analysis choices. Assuming the initial and final states of interest are in the physical subspace of Krylov space, the choice of L vs R labels in Eq. (58) is irrelevant.

Note that, in certain cases, block Lanczos matrix-element estimators are numerically identical to GEVP estimators (see Sec. VI) as well as estimators based on GPOF methods [11–15]. These and other coincidences between methods will be explored more deeply in future work [30].

H. Correlator decomposition

Lanczos energies and overlap factors provide an exact representation of correlation functions, even for noisy Monte Carlo data. Specifically,

$$\begin{aligned} C_{ab}(t) &= \langle \chi_a | T^t | \psi_b \rangle = \langle \chi_a | [T^{(m)}]^t | \psi_b \rangle \\ &= \sum_k Z_{ka}^{R(m)*} Z_{kb}^{L(m)} [\lambda_k^{(m)}]^t, \end{aligned} \quad (60)$$

for all $t \leq 2m - 1$. The first equality is the definition of $C(t)$. The third can be shown by inserting

$$[T^{(m)}]^t = \sum_k |y_k^{R(m)}\rangle [\lambda_k^{(m)}]^t \langle y_k^{L(m)}|, \quad (61)$$

which follows from the eigendecomposition Eq. (38) and $\langle y_k^{L(m)} | y_l^{R(m)} \rangle = \delta_{kl}$, and recognizing the definitions of the overlap estimators $Z_{ka}^{L/R(m)}$, Eq. (54). We prove the remaining second equality below. The result and proof thereof generalize the scalar result found in Ref. [17].

Some preliminary results are necessary before demonstrating the second equality of Eq. (60). Comparing the action of T , following from the three-term recurrence Eq. (4), with the action of its Krylov-space approximation $T^{(m)}$, following from Eqs. (40) and (32), shows that

$$T^{(m)}|v_{ja}^R\rangle = \begin{cases} T|v_{ja}^R\rangle, & j < m, \\ \sum_b (|v_{jb}^R\rangle \alpha_{jba} + |v_{(j-1)b}^R\rangle \beta_{jba}), & j = m. \end{cases} \quad (62)$$

The difference between these operators therefore acts on Lanczos vectors as

$$\begin{aligned} [T - T^{(m)}]|v_{ja}^R\rangle &= \delta_{jm} \sum_b |v_{(m+1)b}^R\rangle \gamma_{(m+1)ba}, \\ \langle v_{ja}^L|[T - T^{(m)}] &= \delta_{jm} \sum_b \beta_{(m+1)ab} \langle v_{jb}^L|. \end{aligned} \quad (63)$$

These results are also used in the explicit construction of residual bounds below. Equation (63) implies that

$$T^t|v_{1a}^R\rangle = [T^{(m)}]^t|v_{1a}^R\rangle \quad \text{and} \quad \langle v_{1a}^L|T^t = \langle v_{1a}^L|[T^{(m)}]^t, \quad (64)$$

for all $t < m$, and

$$\begin{aligned} T^m|v_{1a}^R\rangle &= [T^{(m)}]^m|v_{1a}^R\rangle + \sum_b |v_{(m+1)b}^R\rangle \gamma_{(m+1)ba}, \\ \langle v_{1a}^L|T^m &= \langle v_{1a}^L|[T^{(m)}]^m + \sum_b \beta_{(m+1)ab} \langle v_{1b}^L|, \end{aligned} \quad (65)$$

for $t = m$, with the second term on each right-hand side representing a contribution outside the rm -dimensional Krylov space. Finally, note that the three-term recurrence Eq. (4) implies

$$[T^{(m)}]^{m-1}|v_{1a}^R\rangle = \sum_{j=1}^m \sum_b |v_{jb}^R\rangle c_{jba}^{R(m)}, \quad (66)$$

where $c_{jab}^{R(m)}$ are some (in-principle computable) coefficients. Their values are not important, but rather that the

vector only has support inside the rm -dimensional Krylov space (i.e., the sum runs over $j \in [1, m]$). An analogous left expression holds as well.

We are now equipped to prove the second equality of Eq. (60). Note first that

$$\langle \chi_a | T^t | \psi_b \rangle = \sum_{cd} \beta_{1ac} \langle v_{1c}^L | T^t | v_{1d}^R \rangle \gamma_{1db}, \quad (67)$$

by Eq. (13), so it is sufficient to show $\langle v_{1c}^L | T^t | v_{1d}^R \rangle = \langle v_{1c}^L | [T^{(m)}]^t | v_{1d}^R \rangle$ for all $t \leq 2m - 1$. To do so, we must consider two cases, in each decomposing $T^t = T^{t_L} T^{t_R}$ with some choice of $t = t_L + t_R$. In the case $t \leq 2m - 2$, we may choose both of $t_L, t_R \leq m - 1$. It follows immediately from Eq. (64) that

$$\begin{aligned} \langle v_{1c}^L | T^{t_L} T^{t_R} | v_{1d}^R \rangle &= \langle v_{1c}^L | [T^{(m)}]^{t_L} [T^{(m)}]^{t_R} | v_{1d}^R \rangle \\ &= \langle v_{1c}^L | [T^{(m)}]^t | v_{1d}^R \rangle. \end{aligned} \quad (68)$$

For the remaining case $t = 2m - 1$, consider $t_L = m$ and $t_R = m - 1$, such that by Eq. (65)

$$\begin{aligned} \langle v_{1c}^L | T^m T^{m-1} | v_{1d}^R \rangle &= \langle v_{1c}^L | [T^{(m)}]^{2m-1} | v_{1d}^R \rangle \\ &\quad + \sum_b \beta_{(m+1)cb} \langle v_{(m+1)b}^L | [T^{(m)}]^{m-1} | v_{1d}^R \rangle. \end{aligned} \quad (69)$$

The second term is zero by Eq. (66), i.e.,

$$\begin{aligned} \langle v_{(m+1)b}^L | [T^{(m)}]^{m-1} | v_{1d}^R \rangle &= \sum_{j=1}^m \sum_c \langle v_{(m+1)b}^L | v_{jc}^R \rangle c_{jcd}^{R(m)} \\ &= \sum_{j=1}^m \sum_b \delta_{j,(m+1)} \delta_{bc} c_{jcd}^{R(m)} = 0, \end{aligned} \quad (70)$$

using the biorthonormality of the Lanczos vectors, so the desired equality holds for $t = 2m - 1$ as well. This completes the proof of Eq. (60).

III. NOISELESS DEMONSTRATION

This section works through an application of the block Lanczos algorithm to noiseless mock-data examples to assess its performance in the infinite-statistics limit. We verify the validity of the formalism presented in Sec. II and compare block Lanczos extractions of energies, overlaps, and matrix elements with those computed using standard effective estimators and scalar Lanczos. Comparison with GEVP methods is deferred to Sec. VI. The same mock-data examples defined in this section are also used in Secs. IV and VI.

For the demonstration of spectroscopy and overlap factor extractions, we construct a correlator of length $N_t = 32$ with 16 states,

$$C_{ab}(t) = \sum_{k=0}^{15} Z_{ka}^* Z_{kb} e^{-E_k t}, \quad (71)$$

where $a, b \in \{0, 1\}$ and

$$\begin{aligned} E_k &= 0.1(k+1), \\ \tilde{Z}_{k0} &= \begin{cases} 1, & k \text{ even} \\ 0.1, & k \text{ odd} \end{cases}, \\ \tilde{Z}_{k1} &= \begin{cases} -0.1, & k = 0 \\ 1, & k \text{ odd} \\ 0.1, & k > 0 \text{ even} \end{cases}, \\ Z_{ka} &= \frac{\tilde{Z}_{ka}}{\sqrt{2E_k}}. \end{aligned} \quad (72)$$

The overall factor of $1/\sqrt{2E_k}$ on the overlaps mimics the single-particle relativistic normalization of states. These definitions are engineered to produce common pathologies that arise in analysis of correlator matrix, visible in the effective energies

$$E_{ab}^{\text{eff}}(t) \equiv -\ln \frac{C_{ab}(t)}{C_{ab}(t-1)}, \quad (73)$$

shown in Fig. 1. The overlaps Z_{k0} for $|\psi_0\rangle$ are defined to have large overlap with even- k states including the ground state, resulting in the clear asymptote of E_{00}^{eff} towards the true $E_0 = 0.1$. Meanwhile, overlaps Z_{k1} for $|\psi_1\rangle$ are suppressed for the ground state but large for odd- k states, including the first excited state. This results in an approximate “pseudoplateau” in E_{11}^{eff} near the true $E_1 = 0.2$ with slow but visible convergence toward its asymptotic value E_0 . The ground-state overlap Z_{01} for $|\psi_1\rangle$ is negative, while all other Z_{ka} are positive. This gives the observed

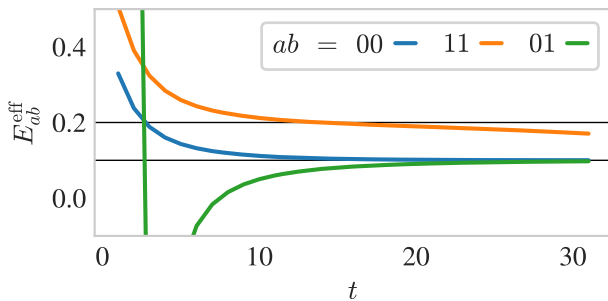


FIG. 1. Effective energies $E_{ab}^{\text{eff}} = -\ln C_{ab}(t)/C_{ab}(t-1)$ for the elements of the correlator matrix for noiseless example as defined in Eq. (72). Not shown is E_{10}^{eff} , which is identical to E_{01}^{eff} . Black lines show the exact E_0 and E_1 .

nonmonotonic behavior (i.e., large oscillations) for the off-diagonal $E_{01}^{\text{eff}} = E_{10}^{\text{eff}}$ at early times and convergence toward E_0 from below. Section III C extends this example to resemble the case of different initial- and final-state quantum numbers for the matrix element demonstration.

As discussed in Refs. [16,17], applying Lanczos methods in the noiseless case requires high-precision arithmetic to avoid numerical instabilities; see Appendix A of Ref. [17] for a detailed discussion. As in that work, we use the MPMATH Python package [31]; 100 decimal digits of precision is sufficient for these demonstrations. In practice, the same precision concerns do not apply in the noisy case as explored in Sec. V. More specifically, applications to noisy data below require Lanczos recursions to be performed in high precision while $T^{(m)}$ eigensolves can be safely performed in double precision, apart from the case of thermal modes discussion in Sec. V F.

In this and following sections, we compare block Lanczos with scalar Lanczos applied only to the diagonal elements of the correlator matrix. As explored in Ref. [16], scalar Lanczos may in principle be applied to the off-diagonal elements without issue for spectroscopy, but in this noiseless example it produces nonpositive Ritz values that cannot be shown straightforwardly. Properly treating overlap factors and matrix elements with off-diagonal Lanczos requires nontrivial extensions of the formalism presented in Ref. [17], which are left for future work.

A. Spectrum

Computing the various quantities in Sec. II, it is straightforward to verify that all R and L quantities coincide for any symmetric convention $\beta_j = \gamma_j$ or any other convention that reduces to a symmetric one in the noiseless limit. As in the scalar case, all physical quantities—Ritz values as well as overlap factor and matrix element estimators—are insensitive to choice of oblique Lanczos convention, but basis-dependent quantities like Krylov coefficients may vary.

Figure 2 shows the Ritz values computed by block Lanczos, compared with those extracted by scalar Lanczos applied to the diagonal correlators C_{00} and C_{11} . The improvements offered by block Lanczos are immediately apparent. Each block iteration produces $r = 2$ new Ritz values, in contrast to scalar Lanczos, which yields only one. Consequently, block Lanczos exactly solves the spectrum of this 16-state system exactly after $m = 16/r = 8$ steps, i.e., twice as quickly as scalar Lanczos, which requires $m = 16$ steps. Roughly speaking, this may be characterized as block Lanczos achieving similar convergence to scalar Lanczos in $1/r$ steps, albeit while consuming matrices of r^2 values per iteration.

More quantitatively, we observe that the low-lying states in the block Lanczos spectrum converges at least as well as the union of the two scalar Lanczos spectra. As can be observed in Fig. 2, scalar Lanczos applied to C_{00} quickly

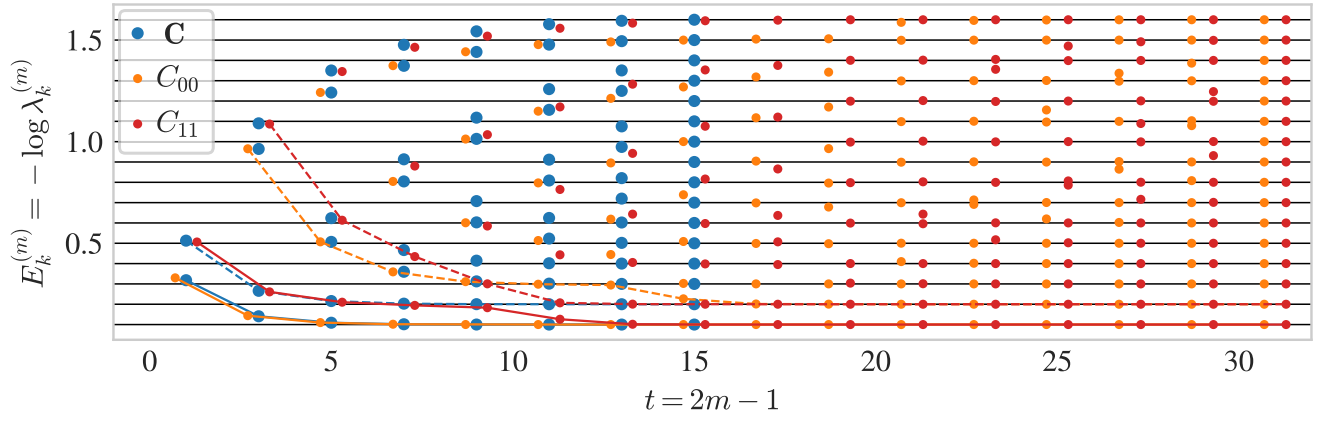


FIG. 2. Spectrum extracted by block Lanczos (blue) applied to the noiseless correlator matrix Eq. (72) compared with the spectra extracted by scalar Lanczos (orange, red) applied to each diagonal correlator, as a function of the number of Lanczos iterations m . Solid lines connect the lowest energies in each extraction, and dashed lines the second-lowest. Block Lanczos solves for the true eigenvalues exactly after 8 steps, and scalar Lanczos after 16 steps. The black horizontal lines indicate the true energies $E_n = 0.1(n+1)$.

extracts E_0 and E_2 , for which the corresponding overlaps are large, but does not give a Ritz value for the suppressed state E_1 until its 8th iteration. The emergence of a Ritz value for E_1 is not associated with any disruption of convergence of its neighboring Ritz values toward E_0 or E_2 . Meanwhile, the lowest-lying Ritz value for scalar Lanczos applied to C_{11} rapidly approaches E_1 ; due to the suppressed overlap associated with E_0 , the lowest-lying Ritz value does not begin converging toward E_0 until $m \geq 5$.⁴ In contrast, block Lanczos yields a sequence of Ritz values that quickly and smoothly converge to the true E_0 and E_1 . Similar comparisons hold for all other pairs of energies in the spectrum, as expected given the alternating pattern of overlap suppression in the example. Further quantitative comparisons are presented in Sec. IV, where they can be better understood in the context of the rigorous bounds on convergence presented therein.

Block Lanczos Ritz values appear to arise in comoving pairs. In this example, the first such pair appears⁵ to eventually converge to the ground and first excited state, while the second such pair apparently converges to the two highest-lying excited states. This is expected behavior for Lanczos algorithms, which tend to first extract the most extremal eigenvalues [22,32–34].

B. Overlap factors

As presented in Sec. IIF, block Lanczos provides an estimator for overlap factors $Z_{ka} = \langle k | \psi_a \rangle$, generalizing the scalar construction. As with the spectrum, block Lanczos solves for the full set of overlap factors exactly after $m = 8$

steps, and scalar does after $m = 16$ up to the sign ambiguity discussed further below. Figure 3 compares block and scalar extractions of the overlap factors in this example. It is clear from the comparison that the relative sizes of the overlaps govern which order states are resolved in. Scalar Lanczos applied to either diagonal correlator preferentially resolves the states with larger overlaps in that correlator, not yielding values for the suppressed states until the latter half of iterations. Block Lanczos, which sees at least one interpolator with unsuppressed overlap with each state, shows no similar structure—as with the spectrum, block overlap estimates converge at least as well as the union of scalar estimates. This indicates that block Lanczos is able to cleanly resolve spectra which must be accessed using multiple interpolators with nearly disjoint support in the space of states. This is exactly the circumstance that arises in lattice studies of multiparticle spectroscopy.

Importantly, while the overall phase of overlap factors is a matter of definition,⁶ relative signs between different overlaps with the same state are physical. Block Lanczos is able to correctly extract these relative signs (in this case, between Z_{00} and Z_{01}). In contrast, scalar Lanczos—which does not simultaneously process data involving $|\psi_0\rangle$ and $|\psi_1\rangle$ —is fundamentally incapable of doing so. In this example, when scalar Lanczos applied to C_{11} solves this noiseless example exactly at $m = 16$, our definitions necessarily return a positive $|Z_{01}|$. The ability to resolve such signs represents a qualitative improvement in capabilities of block over scalar Lanczos.

⁴We caution against interpreting these behaviors as “pseudo-plateaus” like those that arise in effective energies, as inspection of the residual bounds indicates true convergence to states but in a nonmonotonic order; see Appendix C.

⁵This association of Ritz values between different Lanczos iterations is purely visual and not quantitative.

⁶The true transfer matrix eigenstates are eigenvectors, which are defined only up to an overall phase. Conventionally, overlap factors for one interpolator are defined to be real to remove this ambiguity. Typically, this is sufficient to guarantee all overlaps are real—see the discussion of correlator reality in Sec. V and Appendix D.

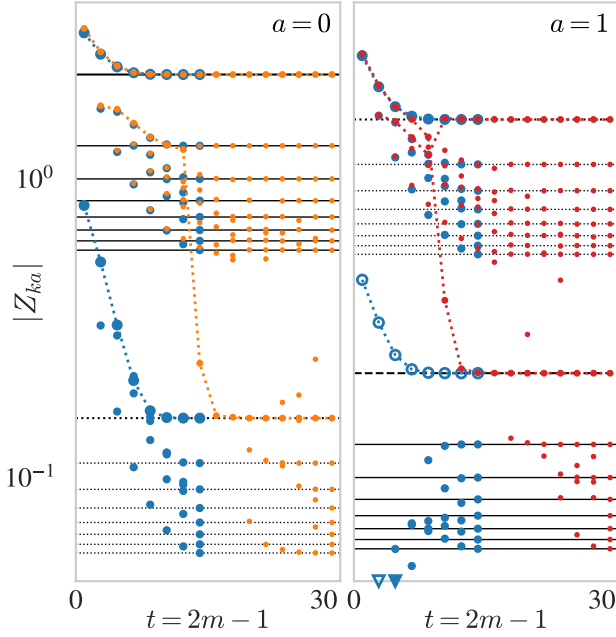


FIG. 3. Overlap factors estimated in the noiseless example Eq. (72) by block Lanczos (large blue) versus scalar Lanczos applied to C_{00} (small orange at left) and C_{11} (small red at right), with Z_{k0} at left and Z_{k1} at right. Black horizontal lines represent the true overlaps Z_{ka} , with solid lines for even k , dotted for odd k , and a dashed line in the right panel for the absolute value of the single negative overlap Z_{01} . Lines and symbols for the two lowest-lying states are bolded, and to guide the eye, dotted lines connect estimates corresponding to the two largest Ritz values in each extraction. No visual distinction is made between values for different k otherwise, but note that, restricting to even or odd k , the true Z_{ka} decrease monotonically in k . Hollow symbols represent negative values whose absolute value is shown instead; this includes all block estimates of the negative overlap Z_{01} . The triangles at the bottom of the right panel indicate block estimates outside the plot range. Block Lanczos solves for all overlaps exactly after 8 steps, while each scalar Lanczos requires 16 and finds the wrong sign for Z_{01} ; see the discussion in the text.

C. Matrix elements

For the demonstration of matrix elements, we extend the example defined in Eq. (72) to resemble the case of an off-forward three-point function, with different initial- and final-state momenta. Specifically, we additionally define a final-state two-point function and an off-diagonal three-point function

$$C'_{ab}(t) = \sum_k Z_{ka}^* Z'_{kb} e^{-E_k t},$$

$$C_{ab}^{\text{3pt}}(\sigma, \tau) = \sum_{fi} Z_{fa}^* J_{fi} Z_{ib} e^{-E_f \sigma - E_i \tau}, \quad (74)$$

respectively, with parameters

$$E_k = 0.1(k+1), \quad E'_k = \sqrt{E_k^2 + 0.1^2},$$

$$Z_{ka} = \frac{\tilde{Z}_{ka}}{\sqrt{2E_k}}, \quad Z'_{ka} = \frac{\tilde{Z}_{ka}}{\sqrt{2E'_k}},$$

$$J_{fi} = \sqrt{\frac{4E'_0 E_0}{4E'_f E_i}} \tilde{J}_{fi}, \quad (75)$$

where $\tilde{Z}_{ka}$ is as in Eq. (72), $\tilde{J}_{00} = 1$, and all other elements of $\tilde{J}_{fi}$ are randomly drawn from a unit normal distribution.⁷ The initial-state spectrum E_k and overlaps Z_{ak} are the same as Eq. (72). The final-state spectrum E'_k and overlaps Z'_{ak} are defined to resemble those of the initial state in a boosted reference frame. The matrix elements J_{fi} are defined to impose the expected energy scaling in C^{3pt} for single-particle states while maintaining $J_{00} = 1$. The block and scalar Lanczos treatments of the final-state spectrum are similar to the initial-state results presented in Sec. III A and thus not shown.

The example is sufficiently pathological that standard ratio plots are not useful to provide. The data are displayed below in Fig. 4, which shows effective estimators for J_{00} defined for the different elements of C_{ab}^{3pt} separately. Specifically, we use the same two definitions as in Ref. [17] following from the standard ratio used in the scalar case⁸

$$R(\sigma, \tau) = \frac{C^{\text{3pt}}(\sigma, \tau)}{C'(\sigma + \tau)} \sqrt{\frac{C(\sigma) C'(\sigma + \tau) C'(\tau)}{C'(\sigma) C(\sigma + \tau) C(\tau)}}. \quad (76)$$

From this ratio, power iteration effective matrix elements can be constructed as

$$J_{00}^{\text{PI}}(t) = \begin{cases} R\left(\frac{t}{2}, \frac{t}{2}\right), & t \text{ even} \\ \frac{1}{2} \left[R\left(\frac{t+1}{2}, \frac{t-1}{2}\right) + R\left(\frac{t-1}{2}, \frac{t+1}{2}\right) \right], & t \text{ odd} \end{cases} \quad (77)$$

and summation-method [36–39] effective matrix elements as

$$\Sigma_{\Delta_\tau}(t_f) = \sum_{\tau=\Delta_\tau}^{t_f-\Delta_\tau} R(t_f - \tau, \tau),$$

$$J_{00, \Delta_\tau}^\Sigma(t_f) = \Sigma_{\Delta_\tau}(t_f + 1) - \Sigma_{\Delta_\tau}(t_f)$$

$$= J_{00} + (\text{excited states}). \quad (78)$$

⁷The precise values used are provided in the Supplemental Material [35].

⁸As noted in Ref. [17], this is not the precise ratio prescribed by the power iteration method; we employ the standard ratio Eq. (76) for ease of comparison with other works.

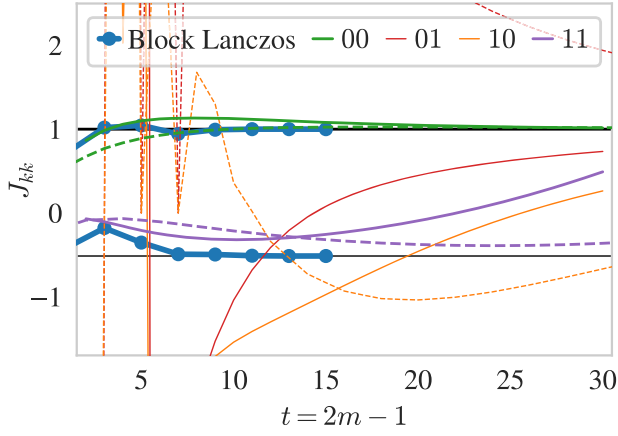


FIG. 4. For the noiseless example Eq. (75), estimates of diagonal matrix elements for the ground state J_{00} (top) and first excited state J_{11} as computed by block Lanczos (blue), as compared with the same summation [Eq. (78), solid] and power iteration [Eq. (77), dashed] estimators. The horizontal lines indicate the true values of $J_{00} = 1$ and J_{11} . Block Lanczos solves for the true J_{00} and J_{11} exactly after $m = N_i/4 = 8$ steps.

Each may be applied to any individual element a, b of a three-point correlator matrix $C_{ab}^{3pt}(t)$. As visible in Fig. 4, while analysis of C_{00}^{3pt} provides estimators which converge near J_{00} quickly, those for C_{11}^{3pt} neither converge to J_{00} —each initially “anticonverges” away from its asymptotic value—nor provide an accurate estimator of J_{11} with which the overlap is larger. The estimators for off-diagonal C_{ab}^{3pt} show large oscillations at early t and poor convergence in the available range $0 \leq t < 32$. Considered collectively, these estimators do not provide any convincing estimate of any matrix element.

Figure 4 compares the same estimators against the block Lanczos extractions of the diagonal matrix elements for the ground and first excited state, J_{00} and J_{11} . In contrast to the confused and incoherent behavior of the effective estimators, block Lanczos estimates converge quickly and reliably to the true values, providing a clear picture. This improvement is just as observed for scalar Lanczos over effective estimators in Ref. [17]. Comparisons with improved GEVP estimators for J in Sec. VI are similar. While we have not repeated the adversarial exercise of Ref. [17] in the block case, we expect block Lanczos to inherit the same robustness against pathological cases observed for scalar Lanczos.

Figure 5 compares the convergence of block and scalar Lanczos estimates for low-lying diagonal and transition matrix elements. The improvement is similar as with energies and overlaps: block Lanczos provides a better estimator than scalar Lanczos applied to either diagonal correlator, even allowing for state misidentification in scalar Lanczos where the spectrum is extracted nonmonotonically. However, the advantage is not as obvious as over the effective estimators.

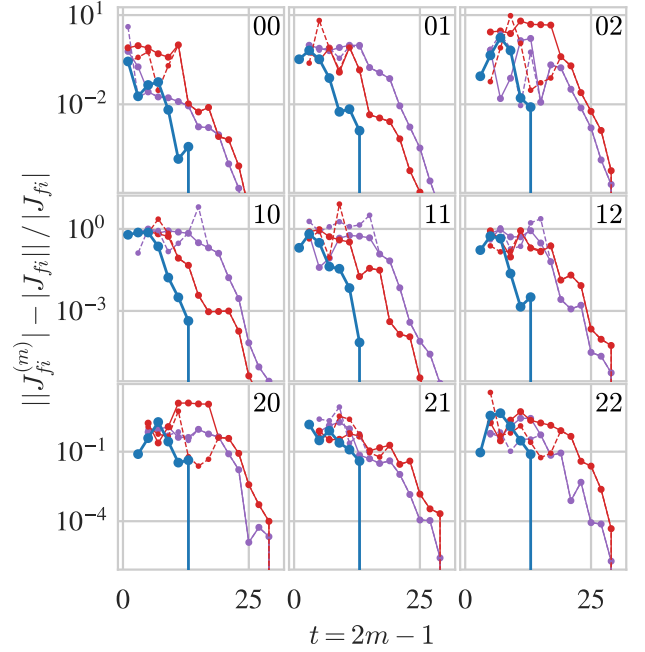


FIG. 5. For the noiseless example Eq. (75), relative errors in (absolute values of) estimates of operator matrix elements J_{fi} involving the lowest-lying three states, as computed by block Lanczos (blue) and scalar Lanczos applied to C_{00} (purple) and C_{11} (red). Solid lines connect points for relative errors defined with respect to i th and f th largest Ritz values. To check for the effects of state misidentification, dashed lines connect points for the scalar results where the error is defined with respect to matrix elements for the states with true energies nearest to the initial- and final-state Ritz values. In most cases, these coincide with the other points. Block Lanczos solves for all matrix elements exactly after $m = N_i/4 = 8$ steps. Scalar Lanczos solves after $m = N_i/2 = 16$ steps, but in some cases is off by an overall sign, as discussed in the text.

A critical improvement over scalar Lanczos is not reflected in Fig. 5: block Lanczos allows self-consistent determination of the relative phases between transition matrix elements for states probed by different interpolators. The formalism presented in both Sec. II and Ref. [17] provides self-consistent determinations of relative signs between transition matrix elements. In the scalar case, this means that the phases of operator matrix elements are defined relative to the overlap factors—for which it produces real and positive values by convention. In this example, this results in scalar Lanczos applied to C_{11} yielding ground-excited matrix elements off by a sign from the definitions in the problem. In contrast, because block Lanczos is able to resolve the relative phases of overlaps as discussed above, it is thereby able to consistently determine relative phases of other transition matrix elements as well. This may be useful for, e.g., lattice studies of electroweak transitions involving resonances [40,41].

IV. CONVERGENCE AND BOUNDS

The same formal frameworks use to quantify and bound convergence in scalar Lanczos can be extended to the block case as well. In this section, we first review the block extension of Kaniel-Paige-Saad (KPS) convergence theory [21,42,43] and discuss its implications for correlator matrix analyses. We then derive the block generalization of the oblique Lanczos residual bounds originally presented in Ref. [16]. In each case, we demonstrate the validity of the bounds using the noiseless examples of Sec. III.

A. KPS bound

In finite-dimensional linear algebra applications of the Lanczos algorithm, the convergence of Ritz values and vectors to the eigenvalues and eigenvectors of a matrix is described by KPS convergence theory. The KPS bound not only quantifies the asymptotic converge rates of Ritz values and vectors, it also provides two-sided bounds on differences between Ritz values and true eigenvalues after finitely many iterations. The KPS bound for applications of scalar Lanczos to infinite-dimensional transfer matrices in LQCD is discussed in Refs. [16,17].

Two important limitations arise for LQCD applications of the KPS bound. First, the KPS bound depends on the full spectrum of T and cannot be directly computed only from the matrix elements of $T^{(m)}$ available after finitely many steps of Lanczos. Second, the KPS bound only applies to nonoblique Lanczos and therefore only to LQCD results for diagonal correlator matrices in the infinite-statistics limit. For these reasons, the KPS bound should be thought of as a theoretical bound on the convergence rate of Lanczos applications to LQCD at infinite statistics, rather than as a directly computable measure of finite-iteration effects in a given calculation. In the remainder of this subsection, L/R superscripts will therefore be omitted.

The block Lanczos extension of Kaniel-Paige-Saad convergence theory [21,42,43] was developed by Saad and reported along with results for standard Lanczos in Ref. [21],

$$0 \leq \frac{\lambda_n - \lambda_n^{(m)}}{\lambda_n - \lambda_\infty} \leq \left[\frac{K_n^{(m)} \tan \theta_n}{T_{m-n-1}(\Gamma_n^r)} \right]^2, \quad (79)$$

where $r = \dim[C_{ab}(0)]$ is the size of the initial Lanczos block, the $T_k(x)$ are Chebyshev polynomials of the first kind defined by $T_k(\cos x) = \cos(kx)$,

$$\Gamma_n^r \equiv 1 + \frac{2(\lambda_n - \lambda_{n+r})}{\lambda_{n+r} - \lambda_\infty} = 2e^{E_{n+r} - E_n} - 1, \quad (80)$$

and

$$K_n^{(m)} \equiv \prod_{l=0}^{n-1} \frac{\lambda_l^{(m)} - \lambda_\infty}{\lambda_l^{(m)} - \lambda_n}, \quad n > 0, \quad (81)$$

with $K_0^{(m)} \equiv 1$ and λ_∞ the smallest eigenvalue of T —for a bounded infinite-dimensional operator T it must be that $\lambda_\infty = 0$. The remaining ingredient in the KPS bound is the angle θ_n between the vector $|n\rangle$ and the subspace $\mathcal{K}_1$ spanned by the initial Lanczos vectors. It can be expressed as

$$\tan^2 \theta_n = |||n\rangle - |\hat{x}_n\rangle||^2, \quad (82)$$

where $|\hat{x}_n\rangle$ is the vector whose orthogonal projection onto $\mathcal{K}_1$ is equal to $|n\rangle$. The explicit construction of this vector is discussed in Ref. [21] and summarized in Appendix B, where it is shown that this can be expressed as

$$\tan^2 \theta_n = [X_{(n,r)}]_{nn}^{-1} - 1, \quad (83)$$

where $X_{(n,r)}$ is the $r \times r$ block of the matrix

$$[X]_{n',n''} \equiv \sum_{ab} Z_{n'a} C_{ab}^{-1}(0) Z_{n''b}^*, \quad (84)$$

where $n', n'' \in [n, n+r-1]$. For the ground state this simplifies to

$$0 \leq \frac{\lambda_0 - \lambda_0^{(m)}}{\lambda_0} \leq \frac{\tan^2 \theta_0}{T_{m-1}(2e^{\delta_r} - 1)^2}, \quad (85)$$

where $\delta_r = E_r - E_0$ and, with $\mathbf{Z}$ the $r \times r$ matrix defined by Z_{na} with $n \in [0, r-1]$,

$$\tan^2 \theta_0 = [\mathbf{Z}\mathbf{C}(0)^{-1}\mathbf{Z}^\dagger]_{00}^{-1} - 1. \quad (86)$$

For large m , $T_m(x) \approx \frac{1}{2}(x + \sqrt{x^2 - 1})^m$, so this further simplifies to

$$0 \leq \frac{\lambda_0 - \lambda_0^{(m)}}{\lambda_0} \lesssim 4 \tan^2 \theta_0 \times \begin{cases} e^{-2(m-1)\delta_r} & \delta_r \gg 1 \\ e^{-4(m-1)\sqrt{\delta_r}} & \delta_r \ll 1 \end{cases}. \quad (87)$$

For $r = 1$ these results reduce to scalar Lanczos results, while for $r > 1$ block Lanczos converges exponentially faster than scalar Lanczos because $\delta_r > \delta_1$.

The convergence of block and scalar Lanczos and the corresponding KPS bounds are compared in Fig. 6. For the ground state, block Lanczos shows nearly identical convergence as for scalar Lanczos applied to C_{00} until the last few iterations before solving the system exactly. Convergence for scalar Lanczos applied to C_{11} parallels that of C_{00} after an initial pause due to the delay in resolving the true ground state. Figure 6 further demonstrates the validity of the block and scalar KPS bounds in comparisons with relative errors of eigenvalue estimates

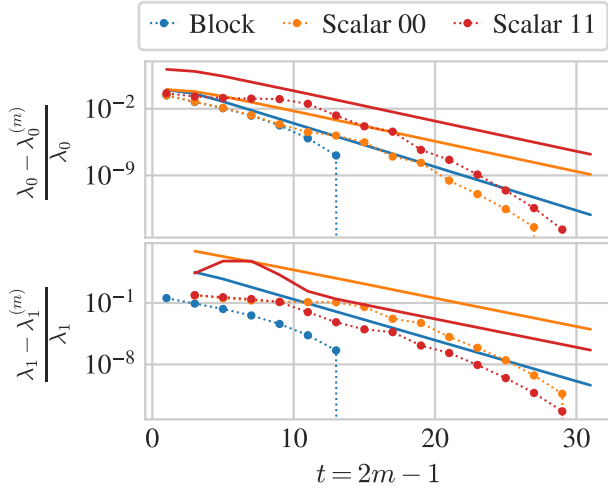


FIG. 6. Convergence of block (blue) and scalar (orange, red) extractions of the two largest eigenvalues for the noiseless example of Sec. III (Eq. (72)), alongside corresponding KPS bounds. Markers connected by dashed lines indicate the true relative error, while solid lines indicate the bound. The block KPS curve beyond $m = 8$ is an extrapolation computed by taking $K_n^{(m)} = K_n^{(8)}$.

computed for the noiseless example of Sec. III. Each bound is satisfied, although the bound is not well saturated in any case and especially not for the excited state; it becomes increasingly less saturated for higher states not shown. More positively, observed convergence is substantially faster than guaranteed. The same bounds predict block Lanczos will converge exponentially faster than scalar Lanczos. This is clearly reflected in empirical convergence for the excited state, for which block Lanczos's improved extraction is especially apparent, as expected for a method which analyzes correlator matrices rather than individual correlators.

B. Residual bounds

The residual bound provides a distinct and complementary two-sided bound on the minimum distance between a Ritz value at finite m and a true eigenvalue of T for scalar Lanczos [22,43–45]. Unlike the KPS bound, it can be calculated directly from matrix elements of $T^{(m)}$ and therefore provides a practically computable bound on the size of excited-state effects at finite m . However, also unlike the KPS bound, it is noteworthy that the residual bound applies to the difference between a given Ritz value and the *closest true eigenvalue*, as opposed to a particular true eigenvalue such as λ_0 . It was shown in Ref. [16] that the residual bound applies to oblique Lanczos assuming $T = T^\dagger$, even when $T^{(m)} \neq [T^{(m)}]^\dagger$. We discuss the block generalization next.

The derivation proceeds by comparing the matrix-element expression for a Ritz-vector residual norm with

its spectral expansion. Right and left Ritz-vector residual norms are defined by

$$R_k^{R/L(m)} \equiv \langle y_k^{R/L(m)} | |T - T^{(m)}|^2 | y_k^{R/L(m)} \rangle, \quad (88)$$

and quantify the difference between the action of the exact transfer matrix T and the Krylov-space approximation $T^{(m)}$ on Ritz vectors. For oblique block Lanczos, the right residual norms can be computed by noting that Eq. (63) implies

$$[T - T^{(m)}] | y_k^{R(m)} \rangle = \mathcal{N}_k^{(m)} \sum_{ab} | v_{(m+1)b}^R \rangle \gamma_{(m+1)ba} \omega_{mak}^{(m)}, \quad (89)$$

i.e., that the action of T on a Ritz vector is a rescaling within rm -dimensional Krylov space by its Ritz value plus the addition of a term orthogonal to this space. It is then straightforward to derive

$$R_k^{R(m)} = |\mathcal{N}_k^{(m)}|^2 \sum_{abcd} \omega_{mak}^{(m)*} \gamma_{(m+1)ba} \times \langle v_{(m+1)b}^R | v_{(m+1)c}^R \rangle \gamma_{(m+1)cd} \omega_{mdk}^{(m)}. \quad (90)$$

An analogous calculation for left residual norms gives

$$R_k^{L(m)} = \frac{1}{|\mathcal{N}_k^{(m)}|^2} \sum_{abcd} (\omega^{-1})_{kma}^{(m)} \beta_{(m+1)ab} \times \langle v_{(m+1)b}^L | v_{(m+1)c}^L \rangle \beta_{(m+1)dc}^* (\omega^{-1})_{kmd}^{(m)*}. \quad (91)$$

These expressions differ from the scalar Lanczos analogs only by the presence of block indices.

The derivation of the spectral representation of the residual bound in Ref. [16] applies here without modification, under the same assumption that T is Hermitian—the Ritz vectors $| y_k^{R/L(m)} \rangle$ can be expanded as linear combinations of true energy eigenstates regardless of whether they are obtained from scalar or block Lanczos. This provides the inequalities

$$\min_{\lambda \in \{\lambda_n\}} |\lambda_k^{(m)} - \lambda|^2 \leq B_k^{R/L(m)}, \quad (92)$$

where

$$B_k^{R/L(m)} \equiv \left| \frac{R_k^{R/L(m)}}{\langle y_k^{R/L(m)} | y_k^{R/L(m)} \rangle} \right| \quad (93)$$

can be conveniently expressed as

$$\begin{aligned}
B_k^{R(m)} &= \left| \sum_{ab} \omega_{mak}^{(m)*} [\gamma_{(m+1)}^\dagger V_k^{R(m)} \gamma_{(m+1)}]_{ab} \omega_{mbk}^{(m)} \right|, \\
B_k^{L(m)} &= \left| \sum_{ab} (\omega^{-1})_{kma}^{(m)} [\beta_{(m+1)} V_k^{L(m)} \beta_{(m+1)}^\dagger]_{ab} (\omega^{-1})_{kmb}^{(m)*} \right|,
\end{aligned} \tag{94}$$

in terms of the matrices

$$\begin{aligned}
V_{kab}^{R(m)} &\equiv \frac{|\mathcal{N}_k^{(m)}|^2 \langle v_{(m+1)a}^R | v_{(m+1)b}^R \rangle}{\langle y_k^{R(m)} | y_k^{R(m)} \rangle} \\
&= \frac{\langle v_{(m+1)a}^R | v_{(m+1)b}^R \rangle}{\sum_{ijcd} \omega_{ick}^{(m)*} \langle v_{ic}^R | v_{jd}^R \rangle \omega_{jdc}^{(m)}}, \\
V_{kab}^{L(m)} &\equiv \frac{\langle v_{(m+1)a}^L | v_{(m+1)b}^L \rangle}{|\mathcal{N}_k^{(m)}|^2 \langle y_k^{L(m)} | y_k^{L(m)} \rangle} \\
&= \frac{\langle v_{(m+1)a}^L | v_{(m+1)b}^L \rangle}{\sum_{ijcd} (\omega^{-1})_{kic}^{(m)} \langle v_{ic}^L | v_{jd}^L \rangle (\omega^{-1})_{kjdc}^{(m)*}}. \tag{95}
\end{aligned}$$

Given a symmetric choice of oblique convention $\beta_j = \gamma_j$, the matrices $V_{kab}^{R/L(m)}$ are equal to δ_{ab} in the noiseless limit where $|v_j^R\rangle = |v_j^L\rangle$ and the oblique algorithm reduces to the standard symmetric one. However, they are nontrivial and must be computed explicitly in applications to noisy data where the oblique formalism is required. This may be accomplished straightforwardly using Eq. (47) once Krylov coefficients have been computed. Thus, two-sided Ritz value error bounds can be computed in practical applications for all m where Lanczos results with $m+1$ iterations are available using Eqs. (92)–(95) and Lanczos-vector norms in Eq. (47), i.e., for all iterations but the last one, $m = N_t/2$. These bounds hold for oblique block Lanczos with Hermitian T , regardless of whether $T^{(m)}$ is Hermitian. The explicit absolute values in Eqs. (93) and (95) are not required for noiseless applications but are helpful for regulating negative values that can arise in applications to noisy data.

Figure 7 demonstrates the validity of the residual bound for Ritz values extracted by applying block Lanczos to the noiseless example of Sec. III; Appendix C examines the residual bounds for scalar Lanczos applied to the same example. In the figure, each Ritz value is shown with an error bar representing the extent of the window of values allowed by the bound. For all points, the error bar crosses at least one line representing a true eigenvalue, indicating the bound is satisfied for all values. A separate question is whether the bounds are sufficiently tight to be useful in practice. As can be observed in Fig. 7, as a Ritz value converges, its residual bound shrinks as well, with each Ritz value quickly becoming consistent with only a single

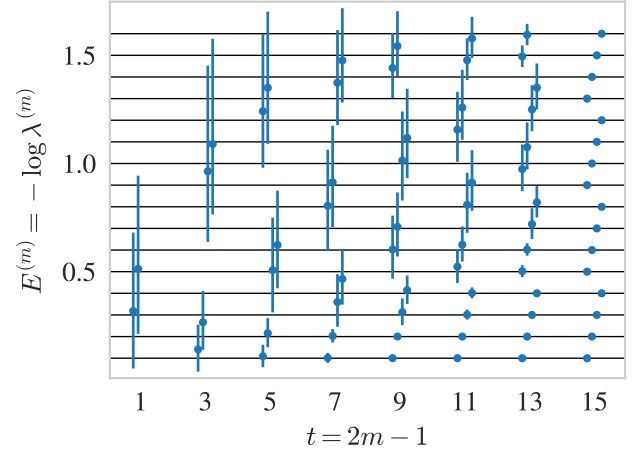


FIG. 7. Demonstration of block Lanczos residual bounds for the noiseless example Eq. (72). Error bars are not statistical, but rather represent the extent of values allowed by the residual bound Eq. (92). The window of allowed values is computed for Ritz values and mapped through the logarithm, and thus asymmetric for energies.

true value. For the lower-lying states, the bounds rapidly become smaller than visible on the plot.

V. NOISE

Block Lanczos applied to noisy correlator data exhibits an analogous version of the “Lanczos phenomenon” [46,47] as the scalar case: some Ritz vectors consistently describe physical states that have all the properties expected at infinite statistics, while other Ritz vectors describe states that have manifestly unphysical properties and must arise from noise. This section describes techniques for identifying and removing such spurious states from Lanczos results, including the construction of a simple physical picture that is equivalent to the Cullum-Willoughby (CW) test for scalar Lanczos. This picture, based on the identification of Ritz vectors that have pathologically small overlaps with all of the initial states ($\sim$ interpolating operators) considered, is used to define a physically motivated block Lanczos generalization of the CW test. This is the first generalization of the CW test to block Lanczos, as far as we are aware.

Proof-of-principle results are demonstrated for the extraction of energies and overlap factors from a 2×2 matrix of nucleon correlation functions. Comparisons are made between block and scalar results for this dataset, as well as scalar Lanczos results from the higher-statistics dataset with identical physical parameters studied in Ref. [17]. Comparisons with GEVP are deferred to Sec. VI. We focus on exploring how block versus scalar Lanczos affects the spurious eigenvalue filtering and state identification that arise during spectroscopy; once Ritz coefficients have been computed for a given state, the subsequent evaluation of matrix elements for external

currents only involves matrix multiplication of three-point functions and proceeds identically for block and scalar Lanczos.

A. Problem setup and data

For the demonstration, we use data computed for the same ensemble with $a \approx 0.091$ fm and $m_\pi \approx 170$ MeV [48–50] as used for demonstrations in Refs. [16,17]. Configurations were generated by the JLab/LANL/MIT/WM groups [51] using the tadpole-improved Lüscher-Weisz gauge action [52] and $N_f = 2 + 1$ flavors of clover fermions [53] defined with stout smeared [54] links on a $48^3 \times 96$ lattice volume. Quark propagators and nucleon correlator matrices were computed by the NPLQCD Collaboration.

The data for the block Lanczos example are a 2×2 matrix of nucleon two-point functions projected to zero momentum, computed on $N_{\text{cfg}} = 80$ configurations independently from and on a (partially) different subensemble than the data used in Refs. [16,17]. The two interpolators differ only in the quark smearing, and are labeled by $a \in \{T, W\}$ where T is for “thin” and W for “wide.” Each is of the form

$$\psi_a(x) = \epsilon^{ijk} [u_j^S(x)^T C \gamma_5 d_k^S(x)] u_i^S(x), \quad (96)$$

where C is the charge conjugation matrix, and $u^S(x)$ and $d^S(x)$ are up- and down-quark fields smeared using gauge-invariant Gaussian smearing, with radius 3.0 for ψ_T and 4.5 for ψ_W . The measurement for each configuration is obtained by evaluating the correlator

$$C_{ab}(t; x_0) = \sum_{\vec{x}} \text{Tr}[\Gamma_\pm \langle \psi_a(\vec{x}, t + t_0) \bar{\psi}_b(x_0) \rangle], \quad (97)$$

where

$$\Gamma_\pm = P_+(1 \pm i\gamma_x \gamma_y) \quad \text{with} \quad P_+ = \frac{1}{2}(1 + \gamma_t) \quad (98)$$

on an 8^3 grid of $N_{\text{src}} = 512$ source positions all on a single random timeslice, then summing over sources and averaging both signs of $\Gamma_\pm$, i.e., over the up-up and down-down channels. This source- and spin-averaged correlator matrix is denoted $\bar{C}_{ab}(t)$. The ψ_a will also sometimes be denoted N_a^+ below in order to distinguish them from interpolating operators N_a^- obtained by replacing $P_+ = \frac{1}{2}(1 + \gamma_t)$ with $P_- = \frac{1}{2}(1 - \gamma_t)$ in all expressions above.

It is essential for the spurious state filtering methods described below to explicitly take the real and symmetric part of the correlator matrix, denoted

$$C'_{ab}(t) = \frac{1}{2} \text{Re}[\bar{C}_{ab}(t) + \bar{C}_{ba}(t)]. \quad (99)$$

As with the imaginary part of a scalar correlator, the imaginary and antisymmetric parts of the correlator matrix are zero in expectation, and we are free to set them to their known infinite-statistics values. Hermiticity follows immediately for any correlator matrix built from the same source and sink operators, $C_{ab}(t) = \langle \psi_a | T^\dagger | \psi_b \rangle$, from the underlying Hermiticity of the transfer matrix. The stronger assumption of a real, symmetric correlator matrix is valid for any interpolating operators that are complex conjugated by a CP transformation and a 2π rotation, as shown in Appendix D. This includes the interpolators used here as well as all commonly used single and multihadron interpolating operators with plane-wave spatial wave functions built from products of (momentum-)smeared [55–57] quark fields.

To compare overlap factors between different operators, it is convenient to work directly with interpolating operators $\hat{\psi}_a$ creating unit-normalized states such that $\langle |\hat{\psi}_a|^2 \rangle = 1$. The corresponding correlator matrix built from unit-normalized interpolating operators is

$$C_{ab}(t) = \frac{C'_{ab}(t)}{\sqrt{C'_{aa}(0)C'_{bb}(0)}}, \quad (100)$$

and satisfies $C_{aa}(0) = C_{bb}(0) = 1$. This is not a necessary precursor to applying the Lanczos analysis, which may treat correlators with arbitrary norms, but instead simply to render the overlap factors for the different smearings of comparable magnitude. This normalization is not included in the error propagation.⁹ This furthermore does not affect the value of the ZCW cut quantity $\Delta_k^{\text{ZCW}(m)}$ defined and discussed in Sec. V C below.

We employ bootstrap resampling to study the effects of noise on the analysis and propagate uncertainties. In particular, we construct 200 bootstrap ensembles (bootstraps) by drawing 80 samples with replacement from the original dataset of 80 and apply Lanczos to the ensemble average within each bootstrap. We further study the effects of a nested bootstrap procedure in which these 200 (outer) bootstrap ensembles are resampled again, generating 200 (inner) bootstrap ensembles for each outer ensemble by again drawing 80 samples with replacement from the 80 (nonunique) samples of the outer ensemble. We then apply Lanczos to the ensemble average of each inner bootstrap. As discussed below, this enables the calculation of an outlier-robust bootstrap-median based estimator and its uncertainties.

For some tests below, we also use the same high-statistics dataset of $N_{\text{cfg}} = 1381$ configurations employed in Ref. [17], generated in the course of the studies in Refs. [39,58]. This scalar nucleon correlator is defined similarly to the WW channel of the 2×2 matrix, using the

⁹That is, it is computed from the mean of $C_{ab}(0)$ and this common value is applied for all bootstrap draws.

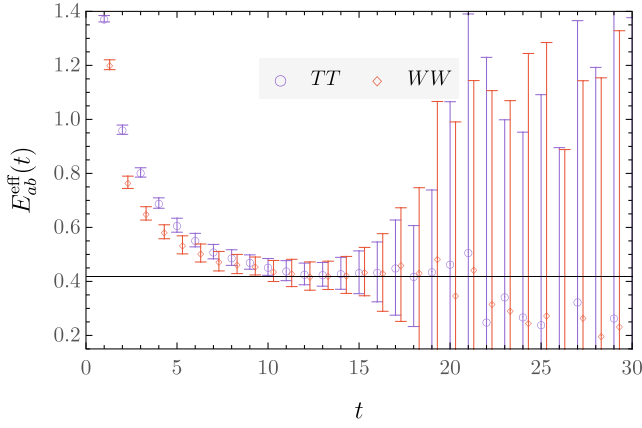


FIG. 8. Effective energies $E_{ab}^{\text{eff}} = -\ln C_{ab}(t)/C_{ab}(t-1)$ for the diagonal elements of the correlator matrix, $a = b = T$ (purple, TT) and $a = b = W$ (red, WW) with a horizontal offset for clarity. Recall $a = T$ indicates the interpolator with “thin” smearing and $a = W$ the “wide” one. Uncertainties are computed using bootstrap confidence intervals. The black line is the fit from Refs. [39,58], 0.4169(18); its uncertainties are not visible on the scale of the plot.

same interpolator ψ_W with smearing radius 4.5. Two-point functions are computed and summed over two offset $4^3 \times 8$ grids of 512 source positions with an overall random offset, for a total of 1024 positions per configuration. It is similarly projected to zero momentum, averaged over both signs of $\Gamma_{\pm}$, and the real part is taken. The correlator is normalized to equal unity at $t = 0$, which is similarly not included in the error propagation. Effective energies for the diagonal entries of the correlator matrix are shown in Fig. 8.

B. Hermitian subspace filtering

As in the scalar case, a subset of block Lanczos outputs (states) are real Ritz values $\lambda_k^{(m)} = \lambda_k^{(m)*}$ with corresponding degenerate left/right Ritz vectors $|y_k^{R(m)}\rangle = |y_k^{L(m)}\rangle$. These obey the physical properties that follow from Hermiticity of T in the infinite-statistics limit and define a Hermitian subspace of Krylov space in the sense made precise below. We denote¹⁰ this subset of Ritz values and vectors by $k \in \mathcal{H}$ where $\mathcal{H} \subset \{0, 1, 2, \dots, rm-1\}$. The objective of Hermitian subspace filtering is to identify and retain the physically interpretable states in $\mathcal{H}$ and discard those in its complement $\bar{\mathcal{H}}$.

Sorting the Ritz values and vectors into sets labeled by $\mathcal{H}$ and $\bar{\mathcal{H}}$ provides the operator-level decomposition

$$T^{(m)} = T_{\mathcal{H}}^{(m)} + T_{\bar{\mathcal{H}}}^{(m)}, \quad (101)$$

¹⁰The definition of $\mathcal{H}$ necessarily differs for different m and could more explicitly be denoted $\mathcal{H}^{(m)}$. We leave this m dependence implicit to improve readability.

where $T_{\mathcal{H}}^{(m)} = T_{\mathcal{H}}^{(m)\dagger}$ acts on states in the Hermitian subspace as

$$T_{\mathcal{H}}^{(m)} \equiv \sum_{k \in \mathcal{H}} |y_k^{(m)}\rangle \lambda_k^{(m)} \langle y_k^{(m)}|, \quad (102)$$

with $\lambda_k^{(m)} = \lambda_k^{(m)*}$ and L/R labels omitted because $|y_k^{R(m)}\rangle = |y_k^{L(m)}\rangle$. Distinctions between left and right quantities only arise for states in the non-Hermitian subspace where the action of $T^{(m)}$ is governed by

$$T_{\bar{\mathcal{H}}}^{(m)} \equiv \sum_{k \in \bar{\mathcal{H}}} |y_k^{R(m)}\rangle \lambda_k^{(m)} \langle y_k^{L(m)}|. \quad (103)$$

The associated Hermitian and non-Hermitian subspaces of Krylov space can be explicitly defined as

$$\begin{aligned} \mathcal{K}^{(m)} &\equiv \text{span}\{|y_k^{(m)}\rangle, k \in \mathcal{H}\}, \\ \bar{\mathcal{K}}^{L(m)} &\equiv \text{span}\{|y_k^{L(m)}\rangle, k \in \bar{\mathcal{H}}\}, \\ \bar{\mathcal{K}}^{R(m)} &\equiv \text{span}\{|y_k^{R(m)}\rangle, k \in \bar{\mathcal{H}}\}. \end{aligned} \quad (104)$$

Ritz vector orthogonality $\langle y_k^{L(m)} | y_l^{R(m)} \rangle = \delta_{kl}$ is exact even for noisy applications because $\langle v_{ia}^L | v_{jb}^R \rangle = \delta_{ij} \delta_{ab}$ follows identically from the recursion relations. This means that these spaces provide an exact decomposition of Krylov space of the form

$$\begin{aligned} \mathcal{K}^{L(m)} &= \mathcal{K}^{(m)} \oplus \bar{\mathcal{K}}^{L(m)}, \\ \mathcal{K}^{R(m)} &= \mathcal{K}^{(m)} \oplus \bar{\mathcal{K}}^{R(m)}. \end{aligned} \quad (105)$$

We define all nonspurious states to include only elements of the Hermitian subspace $\mathcal{K}^{(m)} \subset (\mathcal{K}^{L(m)} \cap \mathcal{K}^{R(m)})$.

The features of the remaining states with labels $\bar{\mathcal{H}}$ clearly identify them as noise artifacts (or at least noise contaminated). This can be seen by noting their role in the correlator decomposition (cf. Sec. II H),

$$C_{ab}(t) = \sum_k Z_{ka}^{R(m)*} Z_{kb}^{L(m)} [\lambda_k^{(m)}]^t, \quad (106)$$

which holds exactly for all $t \leq 2m-1$. Reproducing noisy data generically requires states with complex and/or negative Ritz values that make oscillatory contributions, since noise does not preserve the existence of a convex spectral representation. Complex-eigenvalue states necessarily come in conjugate pairs when $C(t)$ is real. Other states have real $\lambda_k^{(m)} > 0$ but distinct L/R overlaps $\langle y_k^{R/L(m)} | \psi_a \rangle$, allowing them to make unphysical negative contributions to diagonal correlators.

One may also separately filter on the positivity of Ritz values, i.e., discard any $\lambda_k^{(m)} < 0$, on the grounds that in the

infinite-statistics limit the underlying transfer matrix is positive definite. However, it remains unclear whether this is the best option in practical analyses of noisy data. While such states make oscillatory contributions that are clearly associated with noise, it is not unreasonable for a small eigenvalue to fluctuate negative, and clipping the distribution at zero may complicate error analysis. The results below do not filter explicitly on positivity and instead rely on subsequent filtering by the Cullum-Willoughby test to define which are spurious.

The practical Hermitian subspace filtering prescription proceeds similarly as in the scalar case. States with complex Ritz values may be immediately sorted into $\tilde{\mathcal{H}}$, defining reality at working precision as

$$\left| \frac{\text{Im}[\lambda_k^{(m)}]}{\lambda_k^{(m)}} \right| = \left| \sin \arg \lambda_k^{(m)} \right| < \epsilon^{\text{float}}, \quad (107)$$

where $\epsilon^{\text{float}} \sim 10^{-8}$ is often appropriate.¹¹ Further removal of unphysical k where right and left Ritz vectors do not coincide may be accomplished with the block version of the “norm trick” introduced in Ref. [17], by attempting to compute $|\mathcal{N}_k^{(m)}|^2$ per Eq. (55) as

$$\frac{\sum_b (\omega^{-1})_{k1b}^{(m)} \gamma_{1ba}}{[\sum_c \beta_{1ac} \omega_{1ck}^{(m)*}]}, \quad (108)$$

yielding for each state k a set of r values indexed by a . As in the scalar case, equality with real $|\mathcal{N}_k^{(m)}|^2 > 0$ requires that these values all be real and positive; states are thus sorted into $\tilde{\mathcal{H}}$ for which $\arg[\mathcal{N}_{ka}^{(m)}] > \epsilon^{\text{float}}$ for any a . As advertised in Sec. II F, in the block case we also impose an additional criterion—that the normalization of a Ritz vector $|y_k^{(m)}\rangle$ can be computed consistently from its overlap with any $|\psi_a\rangle$. In practice, this amounts to requiring that the value computed by Eq. (108) is numerically identical for all a at the level of ϵ^{float} . We note that this last condition fails to hold for *any* state if the correlator matrix is not taken to be real, even if it is Hermitian. It is essential to apply oblique block Lanczos to the real, symmetric part of a noisy correlator matrix $\text{Re}[\mathcal{C}(t) + \mathcal{C}(t)^T]/2$ rather than the complex, Hermitian part $(\mathcal{C}(t) + \mathcal{C}(t)^\dagger)/2$ in order for this part of the Hermitian subspace filter to be valid.

¹¹When eigensolves are performed with floating precision arithmetic, lowering ϵ^{float} to significantly smaller values than 10^{-8} leads to all Ritz values being labeled as outside the Hermitian subspace. If eigensolves are performed in higher precision, then ϵ^{float} can be lowered accordingly without changing the number of states in the Hermitian subspace. This suggests that these states have imaginary parts arising solely from the use of floating point precision at some stages of the calculation, motivating the notation ϵ^{float} .

Following this filtering, we find that left and right Ritz coefficients of the surviving states coincide within numerical precision, i.e.,

$$\sum_b \gamma_{1ab}^{-1} P_{ktb}^{R(m)} = \sum_b P_{ktb}^{L(m)*} \beta_{1ba}^{-1}, \quad (109)$$

where the factors of β_1^{-1} and γ_1^{-1} give a convention-independent definition. This is sufficient to guarantee that all left and right estimators of physical quantities will be identical.

C. The Cullum-Willoughby test

In the context of numerical linear algebra, procedures like the Cullum-Willoughby test [46,47] and selective reorthogonalization [44] are motivated by the observation that numerical noise (i.e., errors from finite-precision arithmetic) causes a loss of (bi)orthogonality in the Lanczos vectors after many iterations. In practice, this results in an artificial expansion of Krylov space outside what would be obtained in the noiseless case and thus the seeding of spurious Ritz vectors which must be diagnosed and removed. Intuitively speaking, the effect of noise is to mix new information into the Lanczos vectors which is independent of the initial vectors $|v_1^R\rangle$ and $\langle v_1^L|$. The observation motivates the CW test, which diagnoses spuriousity by examining sensitivity of Ritz values to the removal of the initial vectors from the Krylov space. As discussed in detail below, this same observation motivates an alternative “ZCW test” that diagnoses spuriousity by directly computing the overlaps of Ritz vectors with the initial vectors.

Figure 9 schematically illustrates the line of argumentation of this subsection, whose goal is to provide a block generalization of CW filtering. As discussed in detail below, the original CW test is challenging to generalize to the block case directly; in contrast, the ZCW test generalizes naturally on physical grounds. Separately, we find the CW and ZCW tests are closely related by the eigenvalue-eigenvector identity [59] and functionally

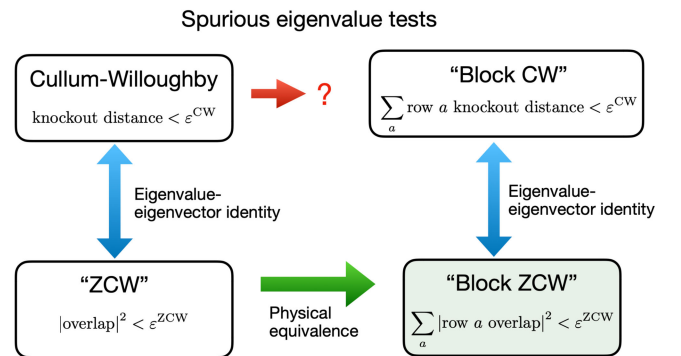


FIG. 9. Schematic depiction of the line of argumentation used to construct block generalizations of the CW test in Sec. V C.

equivalent in practice. This finally allows an indirect derivation of block CW, although we find the resulting prescription is practically equivalent to block ZCW and more expensive.

1. The CW test

As presented in its original formulation [46,47] and employed in Refs. [16,17], the CW test compares the Ritz values with the eigenvalues $\tilde{\lambda}_k^{(m)}$ of the matrix $\tilde{T}^{(m)}$ constructed by removing (“knocking out”) the first row and column from $T^{(m)}$. The matrix $\tilde{T}^{(m)}$ corresponds to the projection of T to a Krylov space with the initial states removed. Thus, by the reasoning above, Ritz values insensitive to this removal must have been seeded by noise and can be identified as spurious, where “insensitive” is defined as

$$\Delta_k^{\text{CW}(m)} \equiv \min_l \left| \lambda_k^{(m)} - \tilde{\lambda}_l^{(m)} \right| < \epsilon^{\text{CW}}, \quad (110)$$

i.e., whether any eigenvalue of $\tilde{T}^{(m)}$ is equal to a given Ritz value within a chosen threshold ϵ^{CW} . In numerical applications where noise is due to finite-precision arithmetic, ϵ^{CW} may be taken to be machine precision. Where noise is statistical, as in analysis of lattice correlator data, ϵ^{CW} must be set accordingly larger. References [16,17] propose different adaptive procedures to set this threshold. Below, we propose a simpler alternative to these procedures which follows naturally from the reformulation of the CW test in more physically intuitive terms.

As already advertised, our new physical perspective presents a straightforward path to a block generalization of the CW test not offered by its original motivating argument. The original argument [46,47] relies on an identity relating the (scalar) residual bound [Eq. (93) with $r = 1$] to the characteristic polynomials of $\tilde{T}$ and $T^{(m-1)}$, defined as $\tilde{a}_m(\mu) \equiv \det(\tilde{T}^{(m)} - \mu \mathbb{1})$ and $a_{m-1}(\mu) \equiv \det(T^{(m-1)} - \mu \mathbb{1})$:

$$B_k^{(m)} = \frac{\prod_{j=2}^{m+1} |\beta_j|^2}{\tilde{a}_m(\lambda_k^{(m)}) a'_{m-1}(\lambda_k^{(m)})} \quad (r = 1), \quad (111)$$

where $a'_m(\mu)$ denotes the derivative of $a_m(\mu)$. This identity was derived for symmetric Lanczos; we have numerically verified that it remains true for oblique Lanczos applications to noisy LQCD data. This identity shows that $B_k^{(m)}$ has a pole whenever $\tilde{a}_m(\lambda_k^{(m)}) = 0$, i.e., whenever $\lambda_k^{(m)}$ is also an eigenvalue of $\tilde{T}^{(m)}$ and thus $\Delta_k^{\text{CW}(m)} = 0$. This directly associates $\Delta_k^{\text{CW}(m)} = 0$ with nonconvergence of $\lambda_k^{(m)}$ to a true eigenvalue, as indicated by a diverging residual bound. As noted in Eq. (111), this identity is only valid for scalar

Lanczos ($r = 1$) and we are not aware of a simple generalization that is valid for block Lanczos with $r > 1$. In the scalar case, it is valid to identify Ritz values with spuriously small $\Delta_k^{\text{CW}(m)}$ and Ritz values with spuriously large residuals; however, this perspective does not generalize straightforwardly to the block case.

2. CW in terms of overlaps (ZCW)

The original motivation for the CW test—independence of the starting vectors as a characteristic property of spurious Ritz vectors—suggests an alternative and more literal procedure: directly compute and examine the overlaps of Ritz vectors with the starting states $\sim \langle v_1 | y_k \rangle$, with a small (i.e., zero up to noise) overlap indicating a spurious state. These are simply the usual overlap factor estimators $Z^{(m)}$, up to normalization (and, in the block case, choice of basis, as discussed further below). Specifically, restricting initially to scalar Lanczos, we consider the normalized overlap factor product as our test quantity,

$$\begin{aligned} \Delta_k^{\text{ZCW}(m)} &= \left| \langle v_1^L | y_k^{R(m)} \rangle \langle y_k^{L(m)} | v_1^R \rangle \right| \\ &= \left| \frac{Z_k^{R(m)*} Z_k^{L(m)}}{C(0)} \right| = \left| \omega_{1k}^{(m)} (\omega^{-1})_{k1}^{(m)} \right|, \end{aligned} \quad (112)$$

and define the ZCW test for spuriousity as

$$\Delta_k^{\text{ZCW}(m)} < \epsilon^{\text{ZCW}}. \quad (113)$$

How to choose the cut ϵ^{ZCW} is discussed in the next subsection.

Figure 10 illustrates a practical comparison of the CW and ZCW tests using results for $\Delta_k^{\text{CW}(m)}$ and $\Delta_k^{\text{ZCW}(m)}$ computed for the high-statistics scalar nucleon correlator of Ref. [17]. The relation between $\Delta_k^{\text{ZCW}(m)}$ and $\Delta_k^{\text{CW}(m)}$ is remarkably linear for the range of $\Delta_k^{\text{CW}(m)}$ corresponding to spurious eigenvalues when ϵ^{CW} is defined using the bootstrap histogram method from Ref. [16]. This demonstrates that an appropriate choice of ϵ^{ZCW} will lead to ZCW test results that are completely equivalent to those from the CW test with this choice of ϵ^{CW} .

The close relationship between the CW and ZCW tests may be understood in terms of the “eigenvalue-eigenvector identity,” first derived by Jacobi in 1834 [60] and rediscovered many times since [59]. For $\tilde{T}^{(m)}$ specifically,¹² it reads

¹²Similarly to how Cullum and Willoughby use Eq. (111) to relate $B_k^{(m)}$ to eigenvalues of a minor of $T^{(m)}$ (i.e., $\tilde{T}^{(m)}$), the general eigenvalue-eigenvector identity relates products of eigenvector elements $\omega_{ik}^{(m)} (\omega^{-1})_{kj}^{(m)}$ to eigenvalues of the minor of $T^{(m)}$ with row i and column j knocked out.

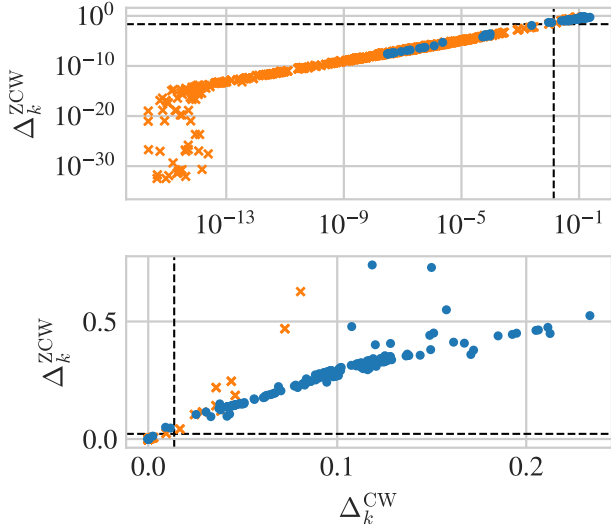


FIG. 10. Comparison of $\Delta_k^{CW(m)}$ and $\Delta_k^{ZCW(m)}$ for scalar Lanczos results using the high-statistics nucleon correlator analysis of Ref. [17]. Each marker corresponds to a pair of values for some m, k , for all $m \in [2, 48]$ and all k , as computed on the central-value correlator. No graphical distinction is made between different m and k . Blue dots (orange Xs) are states inside (outside) the Hermitian subspace. The top and bottom panels show the same data with log and linear axes. The vertical line indicates the cut $\epsilon^{CW} \approx 1.4 \times 10^{-2}$ selected using the bootstrap histogram procedure of Ref. [16], while the horizontal line indicates the cut $\epsilon^{ZCW} \approx 2.2 \times 10^{-2}$ selected by Eq. (117).

$$\prod_{\tilde{l}} [\lambda_k^{(m)} - \tilde{\lambda}_{\tilde{l}}^{(m)}] = \omega_{1k}^{(m)} (\omega^{-1})_{k1}^{(m)} \prod_{l \neq k} [\lambda_k^{(m)} - \lambda_l^{(m)}]. \quad (114)$$

For any state, $\Delta_k^{CW(m)}$ appears in the product on the left-hand side. Dividing the right-hand side by the remaining terms and recognizing $\Delta_k^{ZCW(m)}$ allows $\Delta_k^{CW(m)}$ to be expressed as

$$\Delta_k^{CW(m)} = \Delta_k^{ZCW(m)} \left(\frac{\prod_{l \neq k} |\lambda_k^{(m)} - \lambda_l^{(m)}|}{\prod_{\tilde{l} \neq \tilde{c}} |\lambda_k^{(m)} - \tilde{\lambda}_{\tilde{l}}^{(m)}|} \right), \quad (115)$$

where $\tilde{c} = \text{argmin}_{\tilde{l}} |\lambda_k^{(m)} - \tilde{\lambda}_{\tilde{l}}^{(m)}|$ indexes the closest eigenvalue of $\tilde{T}^{(m)}$ to $\lambda_k^{(m)}$.

Equation (115) implies that if $\Delta_k^{ZCW(m)}$ is small then so must be $\Delta_k^{CW(m)}$, and vice versa, and thus that they are equivalent quantities for filtering. The key feature of Eq. (115) is that the factor in parentheses corresponding to $\Delta_k^{CW(m)} / \Delta_k^{ZCW(m)}$ is necessarily nonzero and finite. To see this, note first that the eigenvalues of any tridiagonal matrix are nondegenerate [22]. This means $\lambda_l^{(m)} \neq \lambda_k^{(m)}$ for $l \neq k$ and therefore the numerator $\prod_{l \neq k} |\lambda_k^{(m)} - \lambda_l^{(m)}|$ has no zeros. It is possible to have $\lambda_k^{(m)} = \tilde{\lambda}_{\tilde{l}}^{(m)}$, but only for one

choice of $\tilde{l}$ because $\tilde{T}_{ij}^{(m)}$ is tridiagonal. This means that it is only possible to have $\lambda_k^{(m)} = \tilde{\lambda}_{\tilde{l}}^{(m)}$ when $\tilde{l} = \tilde{c}$, so the denominator—defined as a product over $\tilde{l} \neq \tilde{c}$, which excludes the one possible zero—has no zeroes. The finiteness of the numerator and denominator follow immediately from the fact that they are finite products of finite factors. It follows that $\Delta_k^{CW(m)} = 0$ can occur if and only if $\Delta_k^{ZCW(m)} = 0$. Since both $\Delta_k^{CW(m)}$ and $\Delta_k^{ZCW(m)}$ are smooth functions of $\lambda_k^{(m)}$, Ritz values within a sufficiently small neighborhood of $\Delta_k^{CW(m)} = 0$ are in one-to-one correspondence with the Ritz values within some small neighborhood of $\Delta_k^{ZCW(m)} = 0$. Therefore, the CW test $\Delta_k^{CW(m)} < \epsilon^{CW}$ for sufficiently small ϵ^{CW} is formally equivalent to a ZCW test $\Delta_k^{ZCW(m)} < \epsilon^{ZCW}$ for some ϵ^{ZCW} .

3. Physical interpretation and choice of ϵ^{ZCW}

Rephrasing the CW test in terms of overlaps suggests a clear physical (re)interpretation. Specifically, filtering with the CW test corresponds to restricting to the physical subspace of states to those with the correct quantum numbers. In practice, for any interpolating operator, the eigenexpansions of the initial states $|v_1^{L/R}\rangle$ have support over all states with the desired quantum numbers, and noiseless time evolution preserves quantum numbers. Thus, any Ritz vector $|y_k^{L/R(m)}\rangle$ corresponding to a physical state must have a finite overlap with the initial state, i.e., $\Delta_k^{ZCW} > 0$. By the same logic, a state with zero overlap cannot have the correct quantum numbers. Such states are necessarily noise artifacts arising from artificial expansion of Krylov space by noise, do not admit a physical interpretation, and should be disregarded.¹³

In practice, the overlaps and thus $\Delta_k^{ZCW(m)}$ can only be measured noisily, and the formal spuriosity criterion $\langle \psi | y_k^{R(m)} \rangle = 0$ must be replaced by the practical ZCW test, $\Delta_k^{ZCW(m)} < \epsilon^{ZCW}$. However, physical intuition about overlaps can be used to guide the selection of ϵ^{ZCW} . For physical states, $\Delta_k^{ZCW(m)}$ corresponds to the absolute, normalized squared overlap factor, such that $0 \leq \Delta_k^{ZCW(m)} \leq 1$. A simple choice of $\epsilon^{ZCW} = 10^{-2}$ or 10^{-3} is sufficient for analyses of typical datasets where available precision allows extraction of states with $\Delta_k^{ZCW(m)} \sim O(0.1)$. Estimates for any physical state with overlaps

¹³In some cases, such as sectors where finite-volume analogs of multiparticle scattering states are present at low energies, there can be physical states that are approximately orthogonal due, e.g., to approximate symmetries that emerge at low energies. Noise contributions then implicitly include contributions from any approximately orthogonal physical states whose overlaps are too small to be resolved from zero at a given level of statistics.

above the cut will converge in the limit of infinite statistics, while physical states below the cut will not be extracted. However, small-overlap states are difficult to resolve from noise; studies targeting such states will necessarily require higher statistical precision, allowing smaller values of ϵ^{ZCW} to be chosen.

Although simple and practical, a fixed choice of ϵ^{ZCW} recovers only a subset of the full spectrum of Ritz values in the infinite-statistics limit. It is thus desirable to define a procedure with a better approach to this limit. The same line of physical reasoning provides a simple and natural definition. First, note that results with small m are least impacted by noise and for at least some range of m will have all Ritz vectors in the Hermitian subspace and all Ritz values satisfying $0 < \lambda_k^{(m)} < 1$. Define $m_{\mathcal{H}}$ to be the largest iteration where this holds, that is

$$m_{\mathcal{H}} \equiv \max \left\{ m \mid [k \in \mathcal{H} \ \& \ 0 < \lambda_k^{(m)} < 1] \ \forall k \right\}. \quad (116)$$

The minimum value of $\Delta_k^{\text{ZCW}(m_{\mathcal{H}})}$ provides a natural cutoff for the smallness of overlap factors that can be unambiguously associated with physical states. For subsequent iterations $m > m_{\mathcal{H}}$, Ritz values can be safely labeled as spurious if the overlap factors are “much smaller” than this minimum value. This provides the cutoff ϵ^{ZCW} needed to define the ZCW test can then be defined as

$$\epsilon^{\text{ZCW}} \equiv \frac{1}{F^{\text{ZCW}}} \min_{k \in \{1, \dots, m_{\mathcal{H}}\}} \Delta_k^{\text{ZCW}(m_{\mathcal{H}})}. \quad (117)$$

For sufficiently noisy data, spurious states can arise as early as $m = 2$ leading to $m_{\mathcal{H}} = 1$; this results in $m_{\mathcal{H}} = 1$ where by definition $\Delta_k^{\text{ZCW}(1)} = 1$. For the $m_{\mathcal{H}} = 1$ case in particular it is preferable to impose a simple cut $\epsilon^{\text{ZCW}} = 10^{-2}$ or 10^{-3} since the data does not provide additional information on the sizes of physical overlaps.

The only hyperparameter that enters Eq. (117) is the factor F^{ZCW} , which simply quantifies the “much” in “much smaller.” All numerical results below apply this procedure with a default value of $F^{\text{ZCW}} = 10$; we find results to be broadly insensitive to other choices in the range $F^{\text{ZCW}} \in [2, 20]$. When applying bootstrap resampling, $m_{\mathcal{H}}$ and ϵ^{ZCW} can be computed independently for each bootstrap sample to provide idiomatic bootstrap estimators.

One important edge case that must be considered separately is the appearance of exponentially growing modes associated with thermal effects on Euclidean lattices with finite temporal extent. Exponentially growing modes have effective overlaps suppressed by $O(e^{-\beta E/2})$, and for practical analyses where this is much smaller than ϵ^{ZCW} their identification requires a separate thermal ZCW test as discussed in Sec. V F below.

4. Block ZCW

The physical perspective of ZCW as filtering on the overlaps between Ritz vectors and the initial state generalizes immediately to block Lanczos. Specifically, Eq. (112) generalizes naturally to define a separate test quantity for each initial state $|v_{1a}\rangle$,

$$\begin{aligned} \Delta_{ka}^{\text{ZCW}(m)} &\equiv \left| \langle v_{1a}^L | y_k^{R(m)} \rangle \langle y_k^{L(m)} | v_{1a}^R \rangle \right| \\ &= \left| \omega_{1ak}^{(m)} (\omega^{-1})_{k1a}^{(m)} \right|. \end{aligned} \quad (118)$$

The same logic as above—states that have vanishingly small overlap with the initial state cannot arise from physical transfer matrix evolution and must be due to noise—applies in the block case to Ritz vectors that have vanishingly small overlap with *all* interpolating operators present. This suggests a particular construction of the “block ZCW test”

$$\begin{aligned} \epsilon^{\text{ZCW}} > \Delta_k^{\text{ZCW}(m)} &\equiv \left| \sum_a \omega_{1ak}^{(m)} (\omega^{-1})_{k1a}^{(m)} \right| \\ &= \left| \sum_{ab} Z_{ka}^{R(m)*} [\mathbf{C}(0)^{-1}]_{ab} Z_{kb}^{L(m)} \right|, \end{aligned} \quad (119)$$

where the sum over a encodes that all overlaps must be small and provides a choice independent of block oblique convention, as made clear by the reexpression in terms of convention-independent quantities in the second equality. The exact same prescription for choosing a physically motivated ϵ^{ZCW} using Eqs. (116) and (117) can be used for the block case. No special treatment is required for oscillating modes in applications to staggered fermions; see Appendix A of Ref. [16].

5. Block CW

The eigenvalue-eigenvector identity can be applied to $T^{(m)}$ exactly as in the scalar case to give

$$\prod_{\tilde{l}} [\lambda_k^{(m)} - \tilde{\lambda}_{\tilde{l}}^{a(m)}] = \omega_{1ak}^{(m)} (\omega^{-1})_{k1a}^{(m)} \prod_{\tilde{l} \neq k} [\lambda_k^{(m)} - \lambda_{\tilde{l}}^{(m)}], \quad (120)$$

where $\tilde{\lambda}_{\tilde{l}}^{a(m)}$ denotes the $\tilde{l}$ th eigenvalues of the matrix obtained by removing the a th row and column of $T^{(m)}$. Similar reasoning as in the scalar case gives the a -dependent relation

$$\Delta_{ka}^{\text{CW}(m)} = \Delta_{ka}^{\text{ZCW}(m)} \left(\frac{\prod_{\tilde{l} \neq k} |\lambda_k^{(m)} - \lambda_{\tilde{l}}^{(m)}|}{\prod_{\tilde{l} \neq \tilde{c}} |\lambda_k^{(m)} - \tilde{\lambda}_{\tilde{l}}^{a(m)}|} \right), \quad (121)$$

with $\tilde{c} = \operatorname{argmin}_l |\lambda_k^{(m)} - \tilde{\lambda}_l^{a(m)}|$ and

$$\Delta_{ka}^{\text{CW}(m)} \equiv \min_l |\lambda_k^{(m)} - \tilde{\lambda}_l^{a(m)}|. \quad (122)$$

The same reasoning leading to Eq. (119) motivates considering

$$\Delta_k^{\text{CW}(m)} \equiv \sum_a \min_l |\lambda_k^{(m)} - \tilde{\lambda}_l^{a(m)}|, \quad (123)$$

i.e., the sum of the minimum distances between a Ritz value and the eigenvalues obtained after knocking out the first r rows and columns. This provides a quantity that vanishes if and only if $\Delta_k^{\text{ZCW}(m)}$ vanishes and so provides a physically motivated starting point for a “block CW” test $\Delta_k^{\text{CW}(m)} < \varepsilon^{\text{CW}}$ based on the sum of minimum knockout distances. Numerical tests on the 2×2 nucleon correlator matrix shown in Figs. 11 and 12 illustrate clear proportionality between $\Delta_{ka}^{\text{CW}(m)}$ and $\Delta_{ka}^{\text{ZCW}(m)}$, as well the same quantities Eqs. (119) and (123) summed over a . Unlike the scalar case, proportionality breaks down for Ritz values where both are $O(1)$.

Several other plausible extensions of the CW test could be imagined, for example based on knocking out the first block of rows/columns collectively. Ones we have explored tend to provide roughly similar levels of spurious

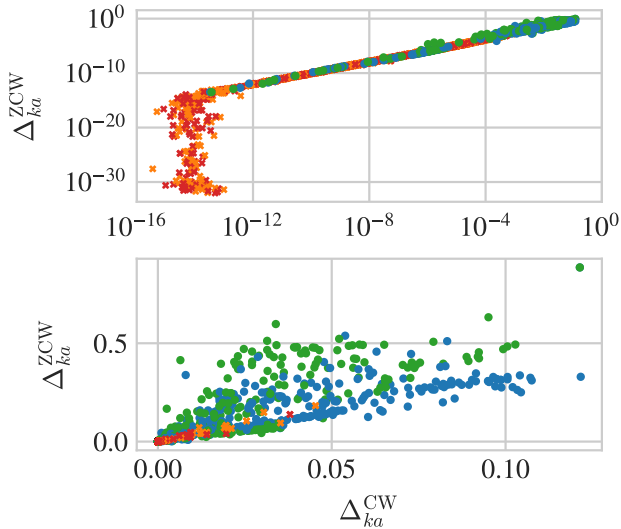


FIG. 11. Comparison of $\Delta_{ka}^{\text{ZCW}(m)}$ and $\Delta_{ka}^{\text{CW}(m)}$ for block Lanczos results, computed on the central-value correlator of the 2×2 nucleon example. Each marker corresponds to a pair of values for some m, k, a , for all smearings $a \in \{T, W\}$ and k for all $m \in [2, 48]$. No graphical distinction is made between different m and k . Blue and orange markers indicate results for $a = T$, while green and red are for $a = W$. Blue and green dots (orange and red Xs) are states inside (outside) the Hermitian subspace. The top and bottom panels show the same data with log and linear axes.

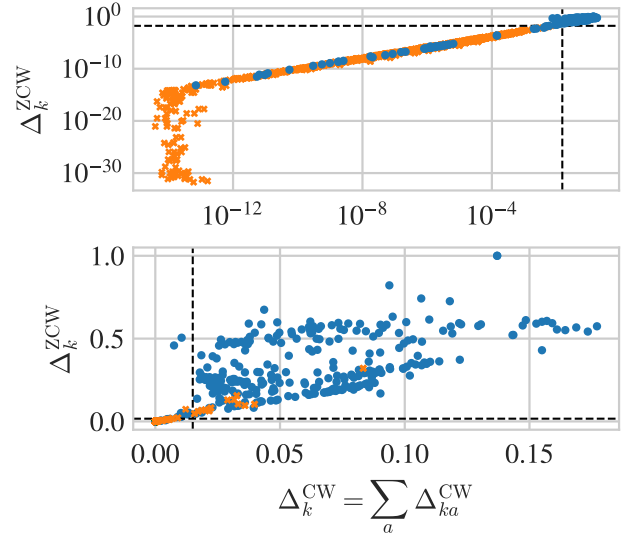


FIG. 12. Comparison of $\Delta_k^{\text{ZCW}(m)}$ and $\Delta_k^{\text{CW}(m)}$ for block Lanczos results. Details are as in Fig. 10, but for the block generalizations Eqs. (119) and (123). The data is the same as shown in Fig. 11 but summed over a . The vertical line indicates the cut $\varepsilon^{\text{CW}} \approx 1.5 \times 10^{-2}$ selected using the bootstrap histogram procedure of Ref. [16], while the horizontal line indicates the cut $\varepsilon^{\text{ZCW}} \approx 1.7 \times 10^{-2}$ selected by Eq. (117).

eigenvalue filtering as the block ZCW/block CW tests defined above. Unsurprisingly, they show slightly worse correspondence with $\Delta_k^{\text{ZCW}(m)}$ than this definition.

At a practical level, the block ZCW test provides almost identical spurious eigenvalue filtering as the block CW test as seen in Fig. 12. The block ZCW test is computationally more efficient, since it only requires calculation of $\omega_{1ak}^{(m)}(\omega_{k1a}^{-1})^{(m)}$, which must already be computed to evaluate residual bounds, overlap factors, and/or matrix elements. In contrast, the block CW test requires r additional eigensolves at every iteration. Eigensolves of $T^{(m)}$ are one of the dominant computational costs of Lanczos analyses, and so the need to perform $r + 1$ times as many amounts to a significant overall increase.

It is interesting to note that, as visible in both the scalar and block cases in Figs. 10–12, states in and out of the Hermitian subspace are well separated such that—although there is no reason not to use the Hermitian subspace test, which introduces no ambiguities or additional hyperparameters—in principle, an appropriate choice of ε^{CW} or ε^{ZCW} is a sufficient state filtering scheme by itself.

6. Results of filtering

Our final prescription for spurious state identification used below is application of the Hermitian subspace test, as described in Sec. V B, followed by application of the ZCW test, Eq. (119). Figure 13 summarizes the results. After the initial few iterations, there are a gradually increasing

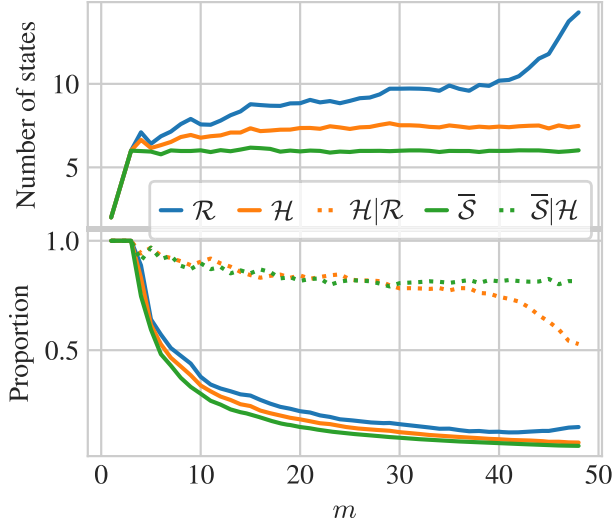


FIG. 13. Census of states surviving the various stages of spuriosity filtering, computed as a mean over $N_{\text{boot}} = 200$ (outer) bootstraps on the 2×2 nucleon correlator matrix example. In the legend, $\mathcal{R}$ denotes states with numerically real Ritz values per Eq. (107), $\mathcal{H}$ denotes states in the Hermitian subspace, and $\bar{\mathcal{S}}$ denotes nonspurious states which survive both Hermitian subspace and block ZCW filtering. $\mathcal{H}|\mathcal{R}$ indicates the fraction of states which are in the Hermitian subspace, given their Ritz values are real. $\bar{\mathcal{S}}|\mathcal{H}$ indicates the fraction of states which are nonspurious, given they are in the Hermitian subspace.

number of real Ritz values, but typically ~ 8 – 10 . Roughly 80% of real Ritz values correspond to states in the Hermitian subspace, with a decreasing fraction at late m , and a similar fraction of Hermitian states survive additional ZCW filtering. The average number of states in the Hermitian subspace is nearly constant at ≈ 7 after $m \sim 5$ – 10 . More sharply, the total number of states identified as physical is nearly constant at ≈ 6 after $m = 4$. However, these figures represent averages, and will vary from bootstrap to bootstrap.

D. Assessing state identification

Providing estimates and uncertainties for energy levels requires state identification, i.e., labeling the Ritz values within each bootstrap as the ground state $n = 0$, first excited state $n = 1$, etc. In this work, we restrict our attention to the maximally simple “filter and sort” method where, after filtering using Hermitian subspace and ZCW tests, we associate the largest surviving eigenvalue with the ground state, the second largest with the first excited state, etc. However, the cautious analysis below finds some ambiguities in this scheme, likely due to the low statistics used for the demonstration. State identification is significantly more challenging for states beyond the lowest-lying two, i.e., those with $n \geq r$, and future work on improved state identification methods may be useful for improving the precision of excited-state energy determinations.

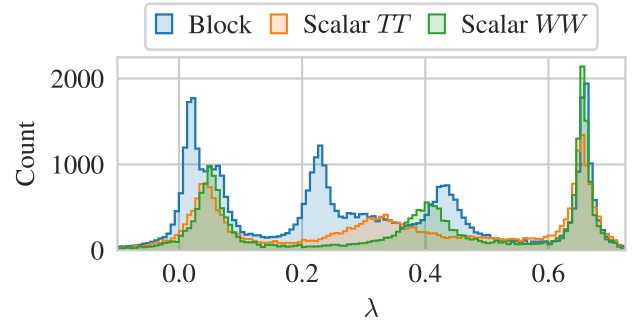


FIG. 14. For the nucleon correlator matrix, distributions of Ritz values for all $m \geq 6$ over $N_{\text{boot}} = 200$ (outer) bootstraps as extracted by block Lanczos or scalar Lanczos applied to either diagonal correlator, after removing spurious eigenvalues with Hermitian subspace and ZCW tests.

However, to begin, we can demonstrate the advantage of block over scalar Lanczos without the need to discuss state identification. Figure 14 shows histograms of the overall distributions of Ritz values, comparing those extracted by block Lanczos with those found by scalar Lanczos applied to the diagonal correlators C_{TT} or C_{WW} . The histograms are restricted to Ritz values that survive spurious state filtering and to $m \geq 6$ to remove preconvergence behavior, but then all remaining values for all m, k over all bootstraps are binned together. All three procedures produce similar-shaped peaks for the ground state at right, as well as a mass of Ritz values near $\lambda = 0$ corresponding to very high-energy states.

The important difference lies in the intermediate part of the spectrum: block Lanczos yields two clearly distinct peaks, while each scalar Lanczos analysis yields only one. Qualitatively, this suggests exactly the improvement in ability to distinguish nearby states expected when analyzing correlator matrices instead of scalar ones. Complicating this interpretation is that, as found in Ref. [17], when scalar Lanczos is applied to a higher-statistics measurement of C_{WW} , it similarly finds two states in the intermediate regime (as reproduced in Fig. 15). This demonstrates that scalar Lanczos with sufficiently large statistics can sometimes achieve similar energy resolution as block Lanczos.

Figure 15 assesses the performance of filter-and-sort state labeling by decomposing the total Ritz value histograms into separate histograms for each state label. The figure shows results for both the block Lanczos analysis, as well as a scalar Lanczos analysis of the high-statistics C_{WW} data for comparison. Interestingly, despite the much lower statistics, the block Lanczos analysis appears to resolve an intermediate state not seen by the high-statistics scalar analysis. The peak for this state is wide and thus not apparent in the total histograms of Fig. 14.

The reduced statistics in the block example cause difficulties with state identification. In the high-statistics setting, each of the four peaks is clearly associated with a single state, up to some minor confusion between the

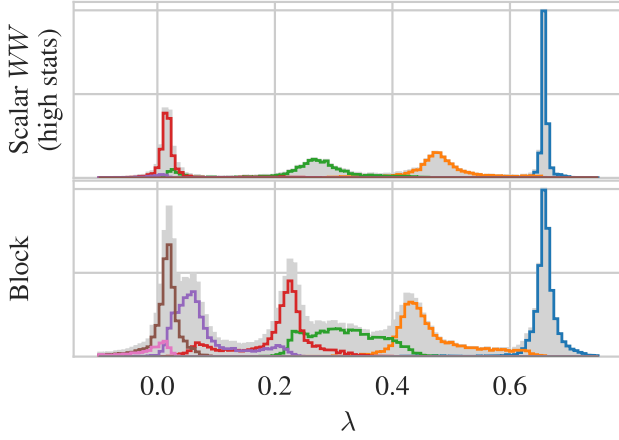


FIG. 15. Distributions of Ritz values for all $m \geq 6$ over N_{boot} (outer) bootstraps after removing spurious eigenvalues with Hermitian subspace and ZCW tests, for scalar Lanczos on the high-statistics dataset of Ref. [17] (top) and block Lanczos on the 2×2 nucleon correlator example (bottom). In each panel, the gray filled histogram indicates all Ritz values, while colored histograms indicate associations of Ritz values with different states, with state identification made simply by sorting Ritz values after filtering. The colored histograms sum to the gray ones.

second and third excited state (i.e., the small green subpeak under the mostly red leftmost peak). In the block example, the validity of the labeling is less obvious. In several instances, the labeling for a single state includes density from multiple peaks (e.g., red, especially). While unambiguous peak-state association in such a plot is not necessary, it does raise suspicion.

Comparison with the noiseless demonstration of Sec. III provides a mechanistic picture of the issue: the spectrum is not extracted in monotonic order, and rather states with large overlaps tend to be resolved earlier. The apparent mislabelings can be explained as a signal-dependent version of this same effect, which gives rise to dislocations in the indexing of the spectrum due to intermediate states being resolved in some bootstraps but not others. These labeling ambiguities motivate our use of outlier-robust estimators based on bootstrap medians discussed in the next subsection.

Before proceeding, we caution against overinterpreting histograms like those of Figs. 14 and 15. These are intended only to give a simple, global view of the distribution of Ritz values extracted by Lanczos analyses. They do not admit a clean statistical interpretation in terms of uncertainties or confidence about the location of true eigenvalues. However, exploring state-identification-agnostic constructions which may admit such interpretations could be an interesting topic for future work.

E. Bootstrap median with nested bootstrap confidence intervals

The state-identification challenges above underscore the fact that noise affects the spectra of Ritz values obtained

from Lanczos in nontrivial ways, even after spurious state filtering. Moreover, filtering is imperfect in practice, introducing additional difficulties. The combination of bootstrap resampling and outlier-robust estimators provides a natural means to work past these difficulties, giving reliable results as long as the imperfect filtering and labeling are not *too* imperfect. In particular, we find that the median of Ritz values over bootstraps provides a more reliable energy estimator than the sample mean. This can be intuitively understood from the fact that state misidentification is a finite-statistics artifact that can lead to large effects in comparison with statistical fluctuations of a particular Ritz value; outlier-robust estimators that remove these large effects can be expected to provide more accurate, as well as more precise, estimators at finite statistics.

The statistical uncertainties of the bootstrap median may be computed by bootstrap resampling and then repeating the entire analysis, including a second (nested) step of bootstrap resampling and then taking the median. To calculate uncertainties, outer and inner bootstrap ensembles are constructed as detailed in Sec. V A. To define a Ritz-value estimator for a given outer ensemble, compute (filtered and labeled) Ritz values for each of the N_{boot} inner ensembles resampled from the outer one, then take the median over the inner estimates. Uncertainties are computed simply from the variance over outer bootstraps: for any quantity X with bootstrap samples $X^{(b)}$ the uncertainty δX is defined as [61–63]

$$\delta X^2 = \frac{1}{N_{\text{boot}}} \sum_b (X^{(b)})^2 - \left(\frac{1}{N_{\text{boot}}} \sum_b X^{(b)} \right)^2. \quad (124)$$

This formula provides an asymptotically unbiased estimator for the variance of any random variable; in particular we apply it to the case where the random variable X is

$$E_n^{(m)} \equiv -\ln \text{median}_{b \in \mathbb{S}_n^{(m)}} \lambda_n^{(b,m)}, \quad (125)$$

where $\mathbb{S}_n^{(m)}$ is the subset of bootstrap draws b that contain at least n nonspurious Ritz values for iteration m —i.e., the median over draws where an estimate is available. The bootstrap samples $X^{(b)}$ are then

$$E_n^{(b,m)} \equiv -\ln \text{median}_{b' \in \mathbb{S}_n^{(b,m)}} \lambda_n^{(b,b',m)}, \quad (126)$$

where $\mathbb{S}_n^{(b,m)}$ is the subset of inner bootstrap draws b' from the outer ensemble b that contain at least n nonspurious Ritz values for iteration m , from which the bootstrap variance $\delta E_n^{(m)}$ can be computed using Eq. (124). Note that in noisy data applications below, $E_n^{(m)}$ will always refer to the bootstrap median estimator, Eq. (125), rather than the sample mean estimator $-\ln \lambda_n^{(m)}$ unless explicitly specified.

To define the overall number of nonspurious states $N_{\max}^{(m)}$ below, we first define $N_{\max}^{(b,m)}$ for each outer bootstrap as the maximum n for which at least n nonspurious Ritz values exist in 95% of the $\{\lambda_n^{(b,b',m)}\}$, or in other words the largest n for which $S_n^{(b,m)}$ contains at least $0.95N_{\text{boot}}$ elements.¹⁴ Applying Eq. (124) requires that $E_n^{(b,m)}$ is defined for each outer bootstrap,¹⁵ and therefore $N_{\max}^{(m)} = \min N_{\max}^{(b,m)}$ is the maximum number of states for which consistent error quantification is possible.

Both $E_n^{(m)}$ and $\delta E_n^{(m)}$ are asymptotically unbiased but include finite-statistics bias suppressed by $O(1/N)$. Bias correction factors that reduce this bias to $O(1/N^2)$ are well known [64]. However, as $N \rightarrow \infty$ the correlation function for $t \leq 2m - 1$ will approach a convex form and (assuming numerical precision is also taken to infinity appropriately) there will be no spurious eigenvalues at iteration m . Outlier-robust estimators will eventually become unimportant in the asymptotic regime where $N \rightarrow \infty$ at fixed m . Iterations where the variance of bootstrap-median estimators is significantly smaller than that of sample-mean estimators are therefore in a qualitatively different regime than the asymptotic regime, and it is not obvious that $1/N$ bias correction will produce a more accurate estimator. Since there are additional $1/\sqrt{N_{\text{boot}}}$ statistical uncertainties introduced by the bias correction to $E_n^{(m)}$ and taking N_{boot} arbitrarily large becomes computationally expensive for nested median estimators, we expect $E_n^{(m)}$ to be a preferable estimator to the bias-corrected analog $2E_n^{(m)} - \frac{1}{N_{\text{boot}}} \sum_b E_n^{(b,m)}$.

Figure 16 illustrates how this scheme regulates the issues with state identification noted in the previous subsection. Note that the median is only defined *after* state labeling, and thus the median estimator necessarily depends on the precise choice of state identification scheme. The distributions of Ritz values for states $n \in \{1, 2, 3\}$ are clearly more well separated in the bootstrap median case and show better resolved peaks. These peaks qualitatively resemble log-normal distributions that would correspond to Gaussian $E_n^{(b,m)}$. As detailed in Appendix F, bootstrap median Lanczos energy estimators are much closer to Gaussian distributed than sample-mean versions of the estimators; however, $E_0^{(b,m)}$ and $E_3^{(b,m)}$ still have observable departures

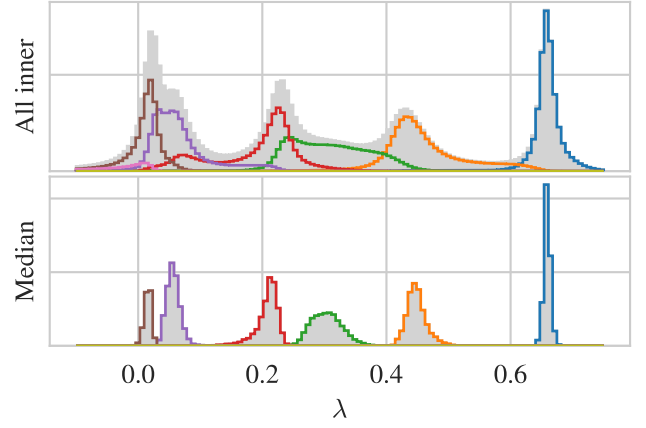


FIG. 16. Distributions of Ritz values estimated over all N_{boot}^2 inner bootstraps (top), and of the N_{boot} outer bootstrap samples of the median estimator $E_n^{(b,m)}$ constructed from them (bottom), as extracted by block Lanczos applied to the 2×2 nucleon correlator matrix, for all $m \geq 6$ after removing spurious eigenvalues with Hermitian subspace and ZCW tests. The colored histograms sum to the gray ones.

from Gaussianity as quantified by the Kolmogorov-Smirnov and Shapiro-Wilk tests.

Equation (124) provides an asymptotically unbiased estimate of the variance of any random variable; however, $s\delta E_n^{(m)}$ can only be interpreted as an $s\sigma$ confidence interval under the stronger assumption that the $E_n^{(b,m)}$ are Gaussian distributed. Given these observed departures from Gaussianity, it is important to test how accurately $s\delta E_n^{(m)}$ approximates $s\sigma$ confidence intervals. These Gaussian approximations are compared with direct confidence-interval calculations using empirical bootstrap confidence intervals in Appendix F. Despite the departures from Gaussianity visible in Shapiro-Wilk test results for E_0 , results for $n \in \{0, 1\}$ at the largest m show that Gaussian confidence interval estimates $s\delta E_n^{(m)}$ are within 10% of the empirical bootstrap confidence intervals for $s\sigma$ with $s \in \{1, 2, 3\}$. For $n \in \{2, 3\}$ the differences are somewhat larger but still less than 20% in all cases. This quantifies the systematic uncertainties associated with identifying $s\sigma \approx s\delta E_n^{(m)}$.

Figure 17 shows the spectrum as extracted by the block analysis using the Hermitian subspace test, ZCW test, and bootstrap-median estimators. Energies $E_n^{(m)}$ of the cleanly identified Ritz values, $n = \{0, \dots, 3\}$, converge after 5–10 iterations and then fluctuate about stable values with roughly constant uncertainties. The corresponding residual bounds $B_n^{(m)}$ are shown in Fig. 18.

Using bootstrap-median estimators allows simply taking the estimates at the last iteration where residual bounds can be computed, $m_{\max} = N_t/2 - 1$, as the final output of the analysis. As observed in the initial application

¹⁴The choice of 95% is arbitrary; any choice $\gtrsim 50\%$ will lead to identical results in this case. Note that demanding 100% inclusion would lead to pathological behavior in the $N_{\text{boot}} \rightarrow \infty$ limit since outliers due to (spurious or nonspurious) state misidentification occur with nonzero probability.

¹⁵The pathological behavior of the $N_{\text{boot}} \rightarrow \infty$ limit for inner bootstraps with a 100% acceptance cut does not apply here because the inner bootstrap median is already an outlier-robust quantity.

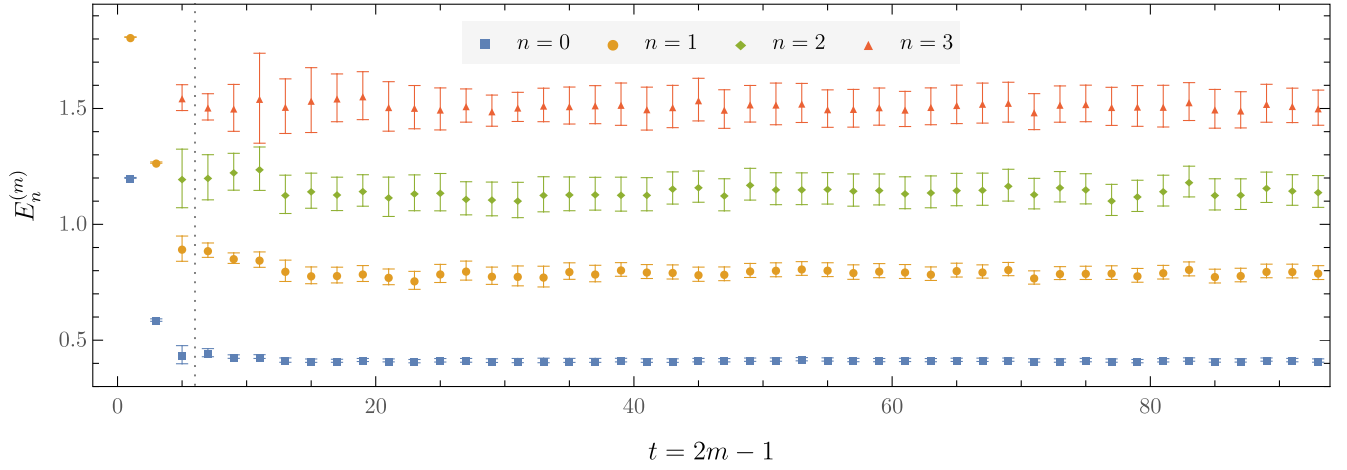


FIG. 17. Lanczos energy estimators for the lowest-lying four states in the nucleon spectrum extracted using block Lanczos. Errors are computed for $E_n^{(m)}$ using bootstrap-median estimators as described in the text. Dotted vertical lines are placed just after $m_{\mathcal{H}}$ as defined in Eq. (116), the last iteration when all Ritz values/vectors satisfy all the physical constraints that arise in the infinite-statistics limit, which is used to determine ϵ^{ZCW} for the ZCW test. For $m = 1$ there are only two Ritz values; for $m = 2$ the higher pair of energies are much above the plot range.

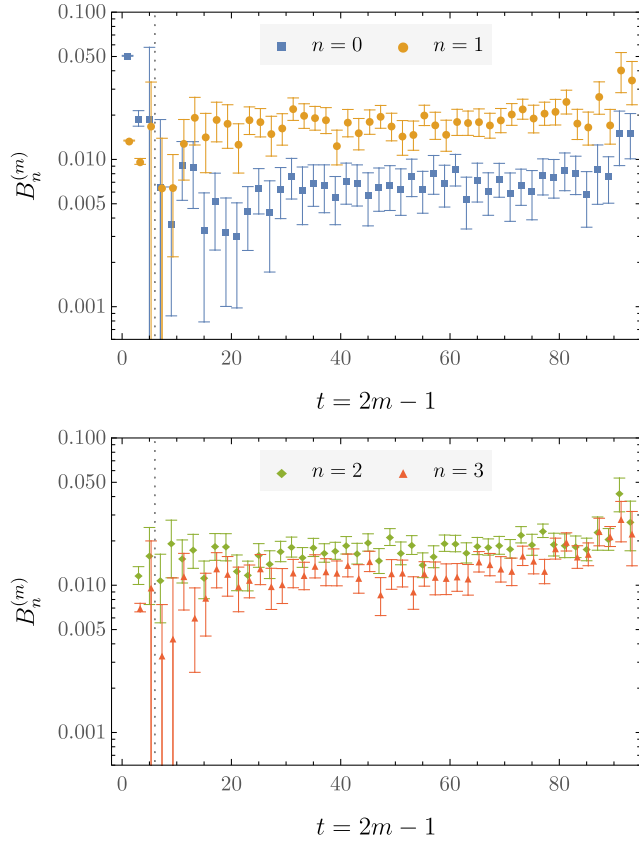


FIG. 18. Residual bounds corresponding to the energy estimators shown in Fig. 17. Note the uncertainties are defined symmetrically but appear asymmetric due to the logarithmic y axes.

of bootstrap-median estimators to scalar Lanczos results in Ref. [16], correlations between Ritz values at large m emerge for bootstrap-median estimators that are obscured by additional uncorrelated noise present in sample-mean estimators.¹⁶ This holds similarly for block Lanczos analyses; as shown in Fig. 19, large correlations ($\gtrsim 0.7$) between energy estimators at large m arise for block Lanczos results using bootstrap-median estimators but not for results using sample-mean estimators. These correlations signal that most of the useful statistical information found in block Lanczos analysis of $\mathbf{C}(t)$ considered over all m is captured by energy estimators at any single large value of m . There is therefore little gain in precision from, e.g., fitting the large- m points to a constant, which moreover comes at the cost of rigor. Such fitting processes require estimating the covariance matrix of many strongly correlated data points, which requires statistical techniques to regulate numerical ill-posedness that often introduce additional subjective analysis choices and/or hyperparameters.

It is interesting to note that, as apparent from comparison of Figs. 19 and 20, correlations $\gtrsim 0.7$ appear at smaller $m \gtrsim 8$ in the block case than for scalar Lanczos applied to either diagonal correlation function, which do not stably display similar correlations until $m \gtrsim 20$ for WW and $m \gtrsim 25$ for TT . The faster approach of $E_0^{(m)}$ to values strongly correlated with the large- m asymptotic result could be related to the faster convergence of block than scalar

¹⁶These differences between sample mean and nested median estimators only appear after 10s of iterations, when many spurious eigenvalues are present. In contrast, the uncertainties of the effective mass are effectively identical when calculated using bootstrap median and sample-mean estimators.

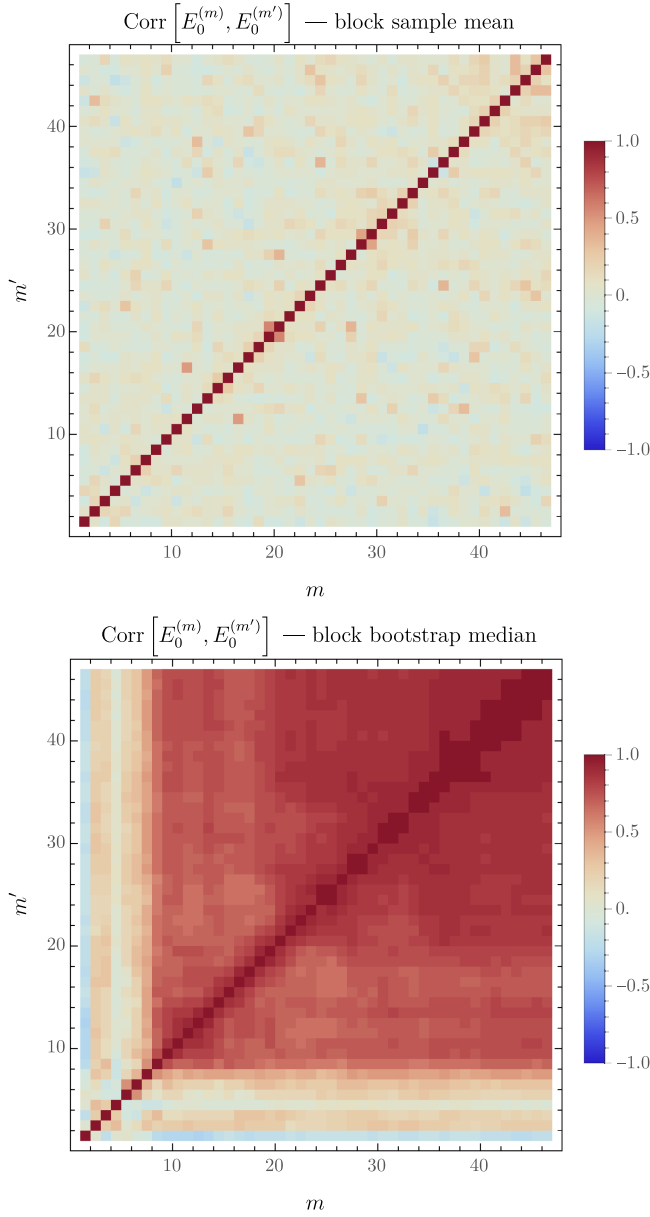


FIG. 19. Correlations between Lanczos energy estimators at different iteration counts for determinations using sample-mean Ritz values (top) and bootstrap-median Ritz values with idiomatic bootstrap uncertainties (bottom). The bootstrap-median case corresponds to the (Pearson) correlation matrix of the “outer bootstrap” samples of the “inner bootstrap” medians.

Lanczos for noiseless data. However, from this example alone it is not clear whether this behavior is generic or what features of the Lanczos algorithm (e.g., loss of biorthogonality after convergence of particular Ritz vectors [22,43]) give rise to these differences.

Figure 21 compares block Lanczos energy estimators with their scalar Lanczos counterparts for both of the diagonal correlators. The strictly faster converge of block than scalar Lanczos is manifest but not large, particularly for the ψ_W operator with larger ground-state overlap.

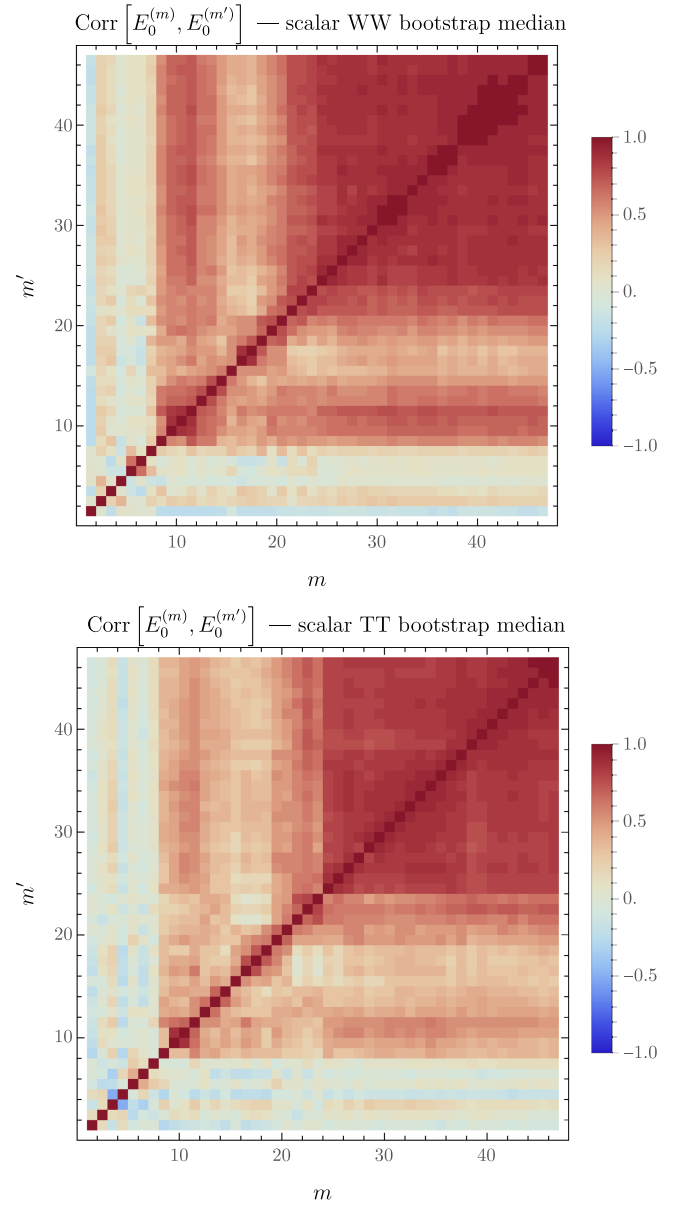


FIG. 20. Correlations between scalar Lanczos energy estimators using bootstrap-median Ritz values estimators; details are as in Fig. 19.

The block Lanczos and WW scalar Lanczos energies both converge within uncertainties for the ground state at $m = 3$ ($t = 5$), while the TT correlation function converges within uncertainties at around $m = 7$ ($t = 13$). Results and uncertainties for the ground state are very similar between block and both scalar analyses.

More significant differences between block and scalar Lanczos are seen in the determination of E_1 . Here, block estimates have significantly smaller uncertainties than either WW or TT scalar estimates. Block Lanczos converges more quickly and stably to a $2 - 3\sigma$ lower value than the E_1 determined from either TT or WW scalar Lanczos analyses.

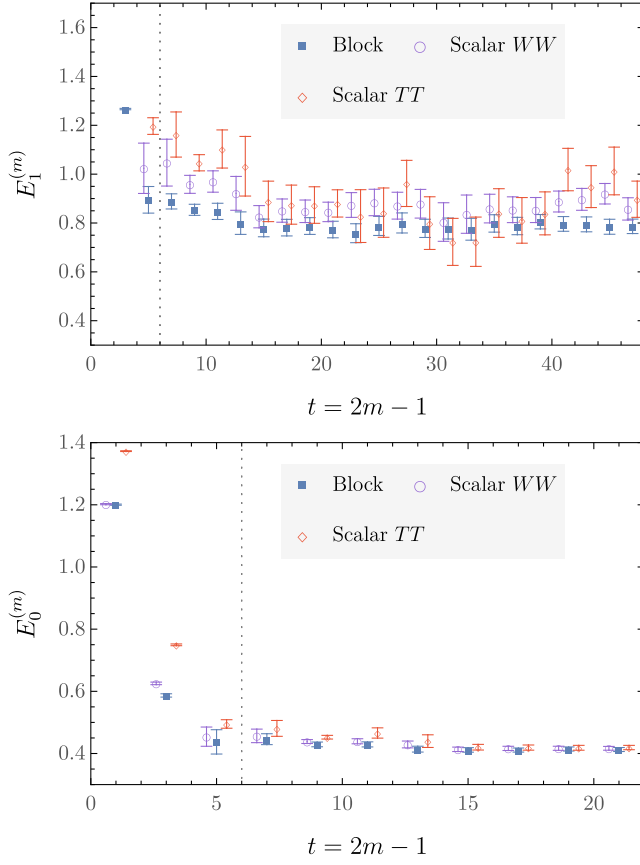


FIG. 21. Comparisons of block Lanczos energy estimators with scalar Lanczos estimators for the corresponding diagonal correlators. Details are as in Fig. 17.

As already discussed, the last iteration where residual bounds can be computed, $m_{\max} = N_t/2 - 1$, provides an estimator $E_n^{(m_{\max})}$ that uses the maximum amount of information in the correlation function to constrain the spectrum. Tables I and II compare these estimates with results from the scalar Lanczos analysis of the high-statistics WW correlator, showing statistical consistency between all methods for E_0 . It is noteworthy that the E_1

TABLE I. Comparison of $E_0^{(m_{\max})}$ and $E_1^{(m_{\max})}$ from block and scalar Lanczos on the 2×2 matrix of $\{\psi_W, \psi_T\}$ operators described in Sec. VA, its diagonals, and the high-statistics superset of this data analyzed in Ref. [17]. The row labeled “High-stats WW [17]” quotes fits results to the estimators without nested bootstrap used in that work.

Lanczos analysis	E_0	E_1	E_2	E_3
Block	0.4122(72)	0.792(30)	1.14(7)	1.50(8)
Scalar TT	0.4240(91)	0.937(67)	...	...
Scalar WW	0.4217(69)	0.891(38)	...	...
High-stats WW	0.4170(22)	0.735(25)	1.31(7)	...
High-stats WW [17]	0.4175(17)	0.736(34)	1.30(8)	...

TABLE II. The corresponding overlap factors, details as in Table I. Note that Ref. [17] uses non-unit-normalized interpolating operators whose overlaps differ from those shown here by a normalization factor $\sqrt{C_{WW}(0)} \approx 6.2801 \times 10^{-4}$.

Lanczos analysis	Z_{0W}	Z_{0T}	Z_{1W}	Z_{1T}
Block	0.366(15)	0.238(11)	0.535(21)	0.437(23)
Scalar TT	...	0.261(15)	...	0.642(58)
Scalar WW	0.392(16)	...	0.650(13)	...
High-stats WW	0.375(4)	...	0.486(15)	...
High-stats WW [17]	0.376(7)	...	0.482(29)	...

obtained from block Lanczos is in 1σ agreement with the high-statistics E_1 result.

The power of block Lanczos to resolve higher-energy excited states is visible in determinations of E_2 and E_3 , both of which have comparable precision to high-statistics scalar results for E_2 . It is noteworthy that the high-statistics E_2 results are in between the two block results and consistent with each within 2σ (and the corresponding residual bounds overlap at 1σ , see Sec. VI below), but it is difficult

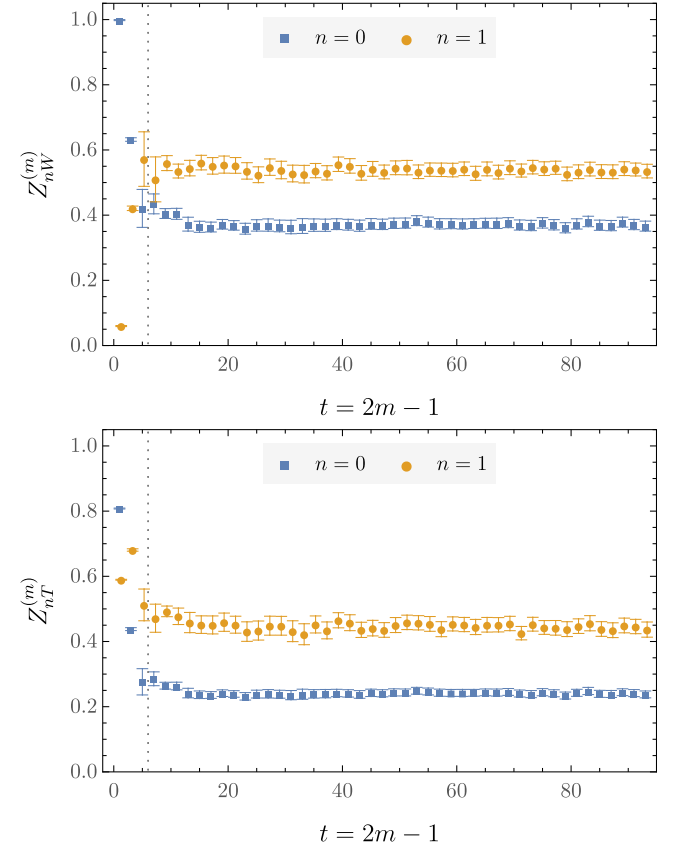


FIG. 22. For the nucleon correlator matrix example, overlap factors $Z_{na}^{(m)}$ as estimated by m steps of block Lanczos for states $n \in \{0, 1\}$ and smearings labeled by $a \in \{W, T\}$. Uncertainties are computed using bootstrap-median estimators with idiomatic bootstrap confidence intervals.

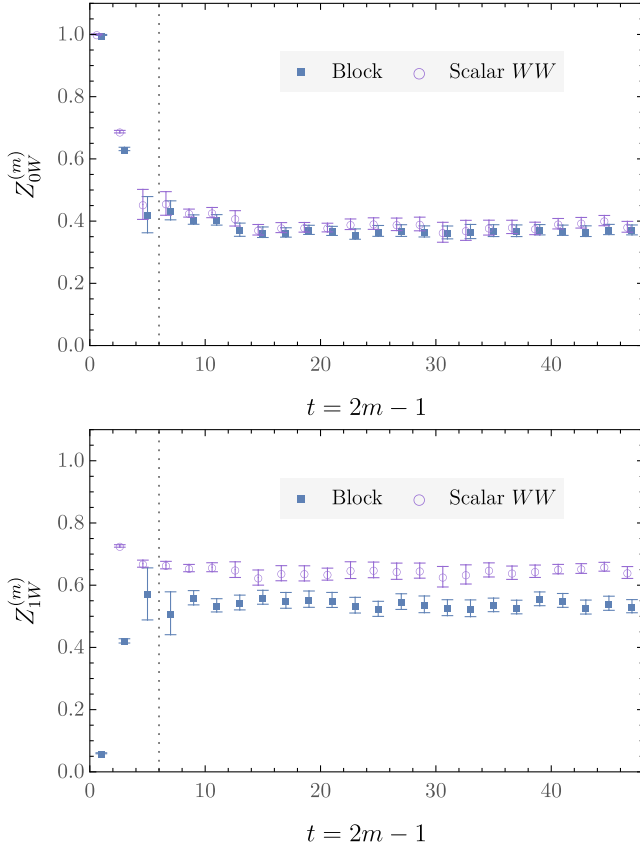


FIG. 23. Comparisons of the block Lanczos overlap factors for the “wide” ψ_W interpolator with those obtained using scalar Lanczos, details as in Fig. 22.

to conclude from this data alone which determination is more accurate.

As presented in Sec. II F, block Lanczos provides an estimator for the overlap factors. Figure 22 shows the resulting estimates, $Z_{na}^{(m)}$,¹⁷ for each interpolator ψ_W and ψ_T and the lowest two states $n \in \{0, 1\}$. As with the spectrum, the overlap estimators converge rapidly and provide clean signals with no signal-to-noise ratio (SNR) degradation with increasing m . Predictably, the signal is cleaner for the ground-state overlaps than for the excited state.

As with the spectrum, the block Lanczos extraction of the overlaps demonstrates clear advantages over those made by scalar Lanczos applied to the diagonal correlators C_{WW} and C_{TT} , as compared in Figs. 23 and 24. While the block and scalar estimators for the ground-state overlaps Z_{0a} are similar in value and uncertainty, the excited-state estimates differ substantially. Each scalar estimate for Z_{1a} is shifted upwards versus its block equivalent; this is likely associated with the downward shift of the corresponding peaks in the scalar-extracted Ritz value histograms of

¹⁷For states in the Hermitian subspace, the R and L estimators necessarily become identical, thus no distinction is necessary.

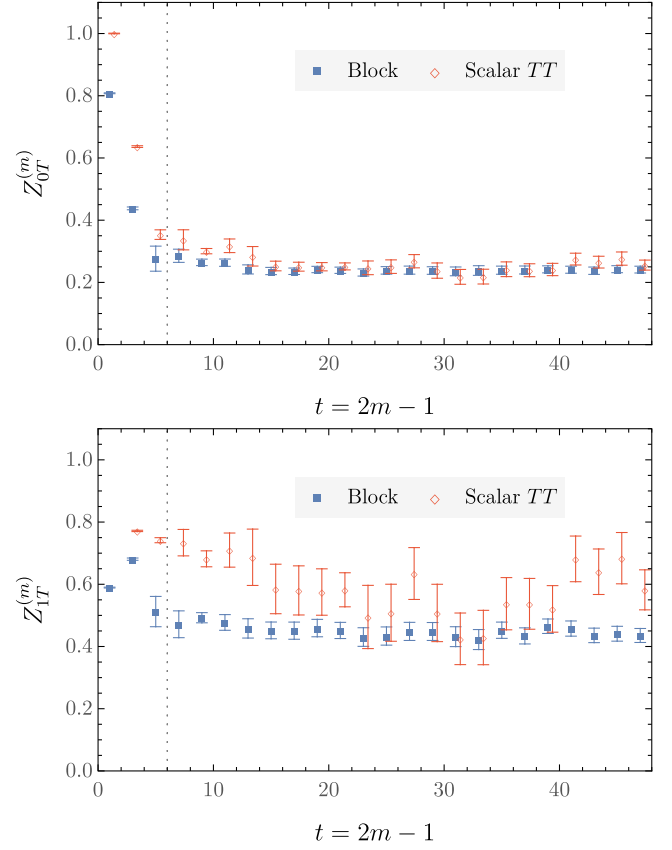


FIG. 24. Comparisons of the block Lanczos overlap factors for the “thin” ψ_T interpolator with those obtained using scalar Lanczos, details as in Fig. 22.

Fig. 14 versus the block-extracted one. The scalar estimator for Z_{1T} is significantly noisier than the block one, which may again be associated with the relatively broader (orange) peak in Fig. 14.

To summarize, in comparison with scalar Lanczos, block appears most advantageous for excited-state energy and overlap determinations. These advantages—in terms of both signal-to-noise and control of excited-state effects—are clear but do not qualitatively change the results for the lowest-energy states. However, it is important to note that the two interpolators ψ_T and ψ_W used in these example are very similar, differing only by the choice of quark smearing radius. This is dissimilar to the most important use case for correlator matrix data, wherein qualitatively different operators are used to probe different single- and multiparticle channels. Comparisons between block and scalar analyses in this setting are a critical topic for future work. Separately but no less importantly, as explored in Sec. VI below, the advantages offered by block Lanczos over GEVP are qualitative and readily apparent even in this example.

F. Thermal states

The ZCW test proposed in Sec. V C and employed to produce the results above is based on the assumption that

states with sufficiently small overlaps with all $|\psi_a\rangle$ must be spurious noise artifacts. However, as already noted, there is a physically motivated counterexample to this assumption: thermal effects lead to a modification of the spectral representation

$$C(t) = \sum_{n=0}^{\infty} \left[|Z_n|^2 e^{-E_n t} + |Z_n^T|^2 e^{-E_n^T(N_t-t)} \right], \quad (127)$$

where the E_n^T are the energies associated with states that have time-reversed quantum numbers compared to those of the $|\psi_a\rangle$. For the (positive-parity) nucleon example, E_n^T corresponds to the energy spectrum of negative-parity nucleon states—i.e., states with baryon-number 2, isospin 1/2, and G_1^- cubic transformation properties, describing negative-parity spin-1/2 finite-volume states. Rearranging this expression gives

$$C(t) = \sum_{n=0}^{\infty} \left[|Z_n|^2 e^{-E_n t} + |\mathcal{Z}_n^T|^2 e^{E_n^T t} \right], \quad (128)$$

defining “apparent overlaps” $\mathcal{Z}_n^T \equiv e^{-E_n^T N_t/2} Z_n^T$. Thermal states therefore enter the spectral expansion of $C(t)$ as growing exponentials with very small overlap factors $\mathcal{Z}_n^T \sim O(e^{-N_t E/2})$.

As discussed in Ref. [16], applying Lanczos to a thermal field theory correlator $C(t)$ provides approximations to both the transfer-matrix eigenvalues $\lambda_n = e^{-E_n}$ and also the thermal modes $\lambda_n^T = e^{E_n^T}$. Thermal states can be expected to have physically reasonable overlaps $Z_n^T > \epsilon^{\text{ZCW}}$ with the time-reversed initial state, but contribute to the correlator with apparent overlaps $\mathcal{Z}_n^T \ll 1$ likely much smaller than ϵ^{ZCW} for realistic lattice sizes and statistical precision. Nevertheless, Lanczos should in principle be able to compute accurate approximations to these modes from correlation functions at large t , where their contributions grow exponentially and statistically precise signals dominated by these modes are visible.

Accessing these exponentially growing modes requires an adapted “thermal ZCW test” analogous to Eq. (113), which defines states as nonspurious if

$$1 < [\lambda_k^{(m)}]^{N_t} \Delta_k^{\text{ZCW}(m)} < \epsilon^{\text{ZCW}}. \quad (129)$$

Here, scaling $\Delta_k^{\text{ZCW}(m)}$ by $[\lambda_k^{(m)}]^{N_t} = e^{N_t E_k^{(m)}}$ is equivalent to applying the ZCW test to the physical overlaps Z_n^T rather than the apparent overlaps $\mathcal{Z}_n^T$.

To study the practical utility of this prescription, we apply the thermal ZCW test to the Ritz values obtained from analyzing the high-statistics nucleon correlator from Ref. [17]. For iterations $m < 44$, no Ritz values pass the thermal ZCW test. For $m \in [44, 46]$, a single Ritz value passes the thermal ZCW test. For $m = m_{\text{max}} = 47$, two Ritz

TABLE III. Comparison of $E_0^{(m_{\text{max}})}$ and $E_1^{(m_{\text{max}})}$ determined with the usual ZCW test with those determined from exponentially growing modes using the thermal ZCW test. The parity of the state corresponds to the (opposite of the) interpolating-operator parity for the ZCW (thermal ZCW) test. The correlator analyzed to produce the results of the second and third rows is the (time-) reverse of the correlator analyzed for the first and fourth rows, $C^{\text{rev}}(t) = C(N_t - t)$.

State parity	Spurious eval test	E_0	E_1
+	ZCW	0.4170(22)	0.735(25)
+	Thermal ZCW	0.4186(23)	0.756(32)
−	ZCW	0.638(10)	1.164(37)
−	Thermal ZCW	0.632(7)	1.128(24)

TABLE IV. The corresponding overlap factors, details as in Table III.

State parity	Spurious eval test	Z_{0W}	Z_{1W}
+	ZCW	0.375(4)	0.486(15)
+	Thermal ZCW	0.379(5)	0.509(21)
−	ZCW	0.284(9)	0.542(6)
−	Thermal ZCW	0.277(7)	0.545(4)

values pass the ZCW test with energies 0.632(7) and 1.128(23).¹⁸ These states, which can be interpreted as part of the spectrum of negative-parity nucleon states, have thermally suppressed overlap factors which after rescaling by $e^{E_n N_t/2}$ predict physical overlaps of 0.277(7) and 0.545(4).

These energies and overlap factors can be compared with the results for negative-parity states obtained by analyzing the time-reversed correlator $C^{\text{rev}}(t) \equiv C(N_t - t)$ with $C(N_t) \equiv C(0)$ using the (nonthermal) ZCW test. As illustrated in Tables III and IV, the negative-parity energies and overlap factors extracted from thermal and nonthermal modes are in excellent agreement. It is noteworthy that almost identical precision is achieved using the time-reversed correlator where the negative-parity modes are visible for small m and the nontime-reversed correlated where negative-parity modes are only visible for the last few iterations.

The thermal ZCW test can also be applied to the Ritz values from the time-reversed (and thus negative-parity) correlator analysis. In this case thermal modes correspond to positive-parity states, and should therefore be consistent with the precise determinations of the positive-parity

¹⁸In practice, finding any thermal modes passing both the ZCW test and the Hermitian subspace test with $\epsilon^{\text{float}} = 10^{-8}$ requires performing $T^{(m)}$ eigensolves using high-precision arithmetic (in addition to the Lanczos recursions). In contrast, all analyses of nonthermal modes presented herein use double-precision eigensolves and find negligible differences with results obtained using high-precision eigensolves.

spectrum from the nonreversed correlator presented above. This is indeed the case; as shown in Tables III and IV the energy and overlap results determined from thermal effects on negative-parity correlators are in 1σ agreement with the standard determination of the positive-parity spectrum. Again, the uncertainties of the thermal and nonthermal determinations are almost identical.

The nearly identical precision of results from reversed and nonreversed correlators is unsurprising from the point of view that, with $m = m_{\max}$, the Krylov spaces appearing in both cases should have almost the same physical content only differing by the inclusion of states for t very close to either end of the correlator. Since the Ritz values are uniquely defined, optimal eigenvalue approximations in Krylov space [22], the same Ritz values ought to be determined from reversed and nonreversed correlators. The fact that this expectation is borne out in practice even in the presence of noise suggests that the spurious eigenvalue filtering described enough is sufficient to provide a physical subspace of states that obey the expected properties of Krylov-space estimators.

Note that because the ZCW test is applied as a post-processing step after all Ritz values are calculated, the nonthermal modes are completely unaffected by whether thermal modes are discarded as spurious or included as a separate category of physical Ritz values via the thermal ZCW test. This means that the nonthermal version of the ZCW test can be safely applied in finite-temperature LQCD calculations and will not bias results for nonthermal modes; it will simply result in thermal modes being labeled as spurious rather than physical. All numerical results outside of this subsection apply the nonthermal version of the ZCW test to avoid the need for high-precision eigensolves.

Although a detailed study of nucleon resonances requires extraction of a range of finite-volume energy levels including those overlapping with multiparticle (e.g., $N\pi$) interpolating operators, it is interesting to note that in the positive parity sector $E_1^+/E_0^- = 1.76(6)$ is nearly consistent with the experimental ratio $M_{N^*(1710)}/M_N \approx 1.823$, while $E_0^-/E_0^+ = 1.53(3)$ is reasonably close to the experimental ratio $M_{N^*(1535)}/M_N \approx 1.636$. The deviations from the physical ratios may easily be accounted for by the slightly heavier quark masses used here ($m_\pi \approx 170$ MeV) than in nature and finite lattice-spacing effects. This underscores the fact that Lanczos converges to genuine energy levels in the spectrum; however, they are not necessarily the lowest energy levels. In particular, it is noteworthy that this Lanczos analysis does not reveal an energy level close to the expected position of the positive-parity Roper resonance, $M_{N^*(1440)}/M_N \approx 1.535$. Further, there are no energy levels obtained by this Lanczos analysis that are close to the energies of noninteracting P -wave $N\pi$ and S -wave $N\pi\pi$ states, which for this volume give $E_{N\pi}/M_N \gtrsim 1.26$ and $E_{N\pi\pi}/M_N \gtrsim 1.24$. Block Lanczos analysis of correlator matrices involving N , $N\pi$, and $N\pi\pi$ are likely

to be required in order to obtain a complete picture of the low-energy spectrum at finite statistics.

VI. COMPARISON WITH GEVP

The block Lanczos formalism and numerical results above can be compared with the state-of-the-art method to date for analyzing LQCD correlator matrices $C_{ab}(t)$, which starts by considering the right¹⁹ GEVP [3],

$$\sum_b C_{ab}(t) G_{bk}(t, t_0) = \lambda_k(t, t_0) \sum_b C_{ab}(t_0) G_{bk}(t, t_0), \quad (130)$$

where $\lambda_k(t, t_0)$ and $G_{bk}(t, t_0)$ are the k th generalized eigenvalues and eigenvectors, respectively. For Hermitian $C(t) = C(t)^\dagger$, the $\lambda_k(t, t_0)$ provide upper bounds on the largest eigenvalue λ_0 of T [1–3]. Further, if the $\lambda_k(t, t_0)$ are ordered such that $\lambda_0(t, t_0) \geq \dots \geq \lambda_{r-1}(t, t_0)$, then Cauchy's interlacing theorem guarantees that there are at least k transfer matrix eigenvalues satisfying $\lambda_0 \geq \dots \geq \lambda_k \geq \lambda_k(t, t_0)$ for all $k \in \{0, \dots, r-1\}$ [5]. That is, $\lambda_k(t, t_0)$ provides a lower bound on the k th true eigenvalue λ_k . Moreover, $\lambda_k(t, t_0)$ and $G_{bk}(t, t_0)$ may be used to construct estimators for energies, overlap factors, and matrix elements as described below.

A. Block Lanczos generalizes GEVP

Before proceeding, we may demonstrate immediately that GEVP coincides with a single step of block Lanczos. Noting that, in numerical applications, correlator matrices $C(t)$ are always invertible for all t permits reworking the GEVP into an equivalent (nongeneralized) eigenproblem,

$$\sum_b [C(t_0)^{-1} C(t)]_{ab} G_{bk}(t, t_0) = G_{ak}(t, t_0) \lambda_k(t, t_0). \quad (131)$$

Thus, the GEVP eigenvalues and eigenvectors are simply those of $C(t_0)^{-1} C(t)$. Meanwhile, the Ritz values after a single step of block Lanczos are obtained by diagonalizing Eq. (17),

$$T_{11}^{(1)} = \alpha_1 = \beta_1^{-1} C(1) \gamma_1^{-1}. \quad (132)$$

Eigenvalues are invariant under conjugation by an arbitrary matrix and its inverse; conjugating by γ_1 and using $C(0) = \beta_1 \gamma_1$ per Eq. (15) gives

$$\begin{aligned} \gamma_1^{-1} T_{11}^{(1)} \gamma_1 &= \gamma_1^{-1} \beta_1^{-1} C(1) \\ &= C(0)^{-1} C(1). \end{aligned} \quad (133)$$

Therefore the Ritz values from one step of block Lanczos and the GEVP eigenvalues for $t_0 = 0$ and $t = 1$ are both

¹⁹See Appendix E for discussion of the corresponding left GEVP.

equal to the eigenvalues of $\mathbf{C}(0)^{-1}\mathbf{C}(1)$ and thus identical to each other,

$$\lambda_k(1, 0) = \lambda_k^{(1)}. \quad (134)$$

Alternately, one may simply observe that a valid choice of oblique convention is $\beta_1 = \mathbf{C}(0)$ and $\gamma_1 = \mathbf{1}$, for which $\mathbf{T}_{11}^{(1)} = \mathbf{C}(0)^{-1}\mathbf{C}(1)$ identically.

The coincidence applies not only for eigenvalues, but also eigenvectors and the problem in abstract. Rearranging Eq. (133) gives $\mathbf{T}_{1ab}^{(1)} = [\gamma_1 \mathbf{C}(0)^{-1} \mathbf{C}(1) \gamma_1^{-1}]_{ab}$ which, in conjunction with the eigenvector relation $\sum_b \mathbf{T}_{1ab}^{(1)} \omega_{k1b}^{(1)} = \omega_{1ak}^{(1)} \lambda_k^{(1)}$ for $m = 1$, gives

$$\sum_b [\gamma_1 \mathbf{C}(0)^{-1} \mathbf{C}(1) \gamma_1^{-1}]_{ab} \omega_{1bk}^{(1)} = \omega_{1ak}^{(1)} \lambda_k^{(1)}. \quad (135)$$

Left multiplying both sides by $\mathbf{C}(0) \gamma_1^{-1}$ yields the equivalent GEVP,

$$\sum_{bc} C_{ab}(1) (\gamma_{1bc}^{-1} \omega_{1ck}^{(1)}) = \lambda_k^{(1)} \sum_{bc} C_{ab}(0) (\gamma_{1bc}^{-1} \omega_{1ck}^{(1)}). \quad (136)$$

This is precisely the form of Eq. (130), allowing the GEVP eigenvectors to be identified as

$$G_{bk}(1, 0) = \sum_c \gamma_{1bc}^{-1} \omega_{1ck}^{(1)}. \quad (137)$$

The exact relation between the two methods holds in full generality, not only for $t_0 = 0$ and $t_d = 1$: a single step of block Lanczos applied to the reindexed correlation function

$$\tilde{C}_{ab}(t) = C_{ab}(t_0 + tt_d) \quad (138)$$

will result in an eigenproblem for $\tilde{\mathbf{C}}(0) = \mathbf{C}(t_0)$ and $\tilde{\mathbf{C}}(1) = \mathbf{C}(t_d)$ equivalent to the GEVP for any t_0, t_d . The Ritz values $\tilde{\lambda}_k^{(1)}$, eigenvectors $\tilde{\omega}_{1bk}^{(1)}$, and $\tilde{\gamma}_1^{-1}$ obtained from this reindexed correlation function are therefore related to the GEVP eigenvalues and eigenvectors by

$$\begin{aligned} \lambda_k(t_d, t_0) &= \tilde{\lambda}_k^{(1)}, \\ G_{ak}(t_d, t_0) &= \sum_b \tilde{\gamma}_{1ab}^{-1} \tilde{\omega}_{1bk}^{(1)}. \end{aligned} \quad (139)$$

Formally, this may be understood as applying a single step of block Lanczos to estimate the eigenvalues of the operator $T^{t_d-t_0}$ with redefined initial states. For even $t_0 \neq 0$, the initial states may be taken as $\langle \chi_a | \rightarrow \langle \chi_a | T^{t_0/2}$ and $|\psi_b\rangle \rightarrow T^{t_0/2} |\psi_b\rangle$. For odd t_0 , it must instead be viewed as applying a single step of oblique block Lanczos with distinct initial and final states, e.g., $\langle \chi_a | \rightarrow \langle \chi_a | T^{(t_0-1)/2}$ and $|\psi_b\rangle \rightarrow T^{(t_0+1)/2} |\psi_b\rangle$.

Block Lanczos can therefore be viewed as a generalization of GEVP that uses an m -times larger-dimensional Krylov space to approximate the transfer matrix after m iterations. This is analogous to how scalar Lanczos generalizes the power-iteration algorithm (i.e., standard effective estimators for energies, overlaps, and matrix elements [16,17]). Since both GEVP and block Lanczos construct optimal Krylov-space approximations to the transfer-matrix eigensystem by explicitly diagonalizing their respective approximations of the transfer matrix, the larger Krylov space explored by block Lanczos necessarily leads to faster convergence than GEVP. Indeed, for systems with small gaps (in lattice units) such that $E_r - E_0 \ll 1$, where achieving fast convergence is most important, block Lanczos converges to the ground state at a rate of $e^{-2t\sqrt{E_r-E_0}}$ after m iterations with $t = 2m - 1$ according to Eq. (87), while GEVP converges at an exponentially slower rate of $e^{-t(E_r-E_0)}$ [3,4].

B. Numerical comparisons

To demonstrate that block Lanczos not only extends but improves upon GEVP methods, we compare the two in applications to both the noiseless mock-data examples of Sec. III and the noisy nucleon example of Sec. V. Comparisons are complicated by the many different variations in precise definitions of GEVP estimators for energies, overlap factors, and matrix elements. For the sake of demonstration, we define and apply two different sets of estimators—specifically, one “moving pivot” scheme and another with a “fixed pivot,” as defined precisely below. In the noiseless applications, in extractions of the spectrum, overlap factors, and matrix elements, we demonstrate the improved convergence properties of block Lanczos over both GEVP and scalar Lanczos. In determinations of the spectrum and overlap factors in the noisy case, we demonstrate not only improved convergence but also an advantage in SNR properties, extending the success of scalar Lanczos to the block case as well. While different variations on the GEVP analyses may perform slightly differently, the demonstrated advantages of block Lanczos should be expected to hold generically.

1. GEVP definitions

Rather than using GEVP eigenvalues directly, it is common practice to define an effective energy estimator [3,4],

$$E_k(t, t_0) \equiv -\ln \left(\frac{\lambda_k(t, t_0)}{\lambda_k(t-1, t_0)} \right), \quad (140)$$

resembling but with distinct properties from the usual effective energy, as discussed further below. Different “moving pivot” (MP) schemes to choose t_0 as a function

of t are employed in practice. For large t , choosing²⁰ $t_0 = \lfloor t/2 \rfloor$ to define

$$E_k^{\text{MP}}(t) \equiv E_k(t, \lfloor t/2 \rfloor) \quad (141)$$

guarantees that $E_0^{\text{MP}}(t)$ converges to E_0 exponentially quickly with excited-state contamination suppressed by $\mathcal{O}(e^{-t(E_r-E_0)})$ [3,4]. At finite t , the inequalities above guarantee that there exist k energies with $E_0 \leq \dots \leq E_k \leq E_k^{\text{MP}}(t)$ for all $k \in \{0, \dots, r-1\}$. This allows GEVP solutions to provide rigorous one-sided bounds on LQCD energy levels as emphasized in Refs. [5,65,66].

A downside of the moving-pivot approach is that $\lambda_k(t, t_0)$ does not possess an exact spectral representation as a sum of exponentials even for fixed t_0 , because the relative contributions of different energy eigenstates to $\lambda_k(t, t_0)$ depend on both t and t_0 . This can sometimes complicate the analysis of excited-state effects. A widely used approach for circumventing this issue are “fixed-pivot” (FP) schemes, wherein one defines fixed-pivot correlation functions [1–3]

$$\hat{C}_k(t; t_d, t_0) \equiv \langle \Psi_k(t; t_d, t_0) \Psi_k^\dagger(0; t_d, t_0) \rangle, \quad (142)$$

using optimized interpolators

$$\Psi_k(t; t_d, t_0) \equiv \sum_a \psi_a(t) G_{ak}^*(t_d, t_0) \quad (143)$$

defined from the generalized eigenvectors or “GEVP weights” $G_{ak}(t_d, t_0)$ at pivot time t_0 and “reference time” t_d . Such fixed-pivot correlation functions possess convex spectral representations as functions of t [67,68],

$$\hat{C}_k(t; t_d, t_0) = \sum_{n=0}^{\infty} |\hat{Z}_{kn}(t_d, t_0)|^2 e^{-tE_n}, \quad (144)$$

where $\hat{Z}_{kn}(t_d, t_0) = \langle n | \Psi_k(0; t_d, t_0)^\dagger | \Omega \rangle$. These can be used to define an effective energy

$$E_k^{\text{FP}}(t; t_d, t_0) \equiv -\ln \left(\frac{\hat{C}_k(t; t_d, t_0)}{\hat{C}_k(t-1; t_d, t_0)} \right), \quad (145)$$

which shares all the properties of the usual effective energy defined on a scalar correlator due to Eq. (144).

Explicit effective estimators can also be defined for overlap factors by generalizing simpler constructions from scalar correlator analyses [68]. For example, the moving-pivot effective overlap employed here is

$$Z_{ka}^{\text{MP}*}(t, t_0) \equiv \frac{\sum_b C_{ab}(t) G_{bk}(t, t_0)}{\lambda_k(t, t_0)^{\frac{1}{2}(t-t_0)} \sqrt{\hat{C}_k(t; t, t_0)}} \quad (146)$$

where $\hat{C}_k(t; t, t_0)$ is as in Eq. (142); we take $t_0 = \lfloor t/2 \rfloor$ in all demonstrations below. We use an analogous definition for the fixed-pivot estimator,

$$Z_{ka}^{\text{FP}*}(t; t_d, t_0) \equiv \frac{\sum_b C_{ab}(t) G_{bk}(t_d, t_0)}{e^{-\frac{1}{2}tE_k^{\text{FP}}(t; t_0, t_d)} \sqrt{\hat{C}_k(t; t_d, t_0)}} \quad (147)$$

where $E_k^{\text{FP}}(t; t_0, t_d)$ is the effective energy estimator of Eq. (145).

Explicit GEVP estimators can be defined for operator matrix elements $J_{fi} = \langle f | J | i \rangle$ in terms of the three-point function matrix $C^{3\text{pt}}(\sigma, \tau)$ and two-point matrices $C(t)$ and $C'(t)$ with initial- and final-state quantum numbers, respectively. In demonstrations below, we use the moving-pivot effective estimator [4]

$$J_{fi}^{\text{MP}}(t, m) = \sum_{ab} G_{af}^*(t, 2m) C_{ab}^{3\text{pt}}(m, m) G_{bi}(t, 2m) \quad (148)$$

where G and G' have been computed from C and C' , respectively. This expression directly generalizes the power-iteration estimator for the effective matrix element derived in Ref. [17]. Note that only a single point of three-point function data at $\sigma = \tau = m$ is incorporated, and varying t only changes how G and G' are derived from the two-point data. An analog of the $t_0 = \lfloor t/2 \rfloor$ MP scheme can be obtained by taking $m = \lfloor t/4 \rfloor$.

Fixed-pivot approaches are more commonly employed in GEVP matrix element analyses. Using GEVP interpolators as in Eq. (143), one may define [69]

$$\begin{aligned} \hat{C}_{fi}^{3\text{pt}}(\sigma, \tau; t_d, t_0) &\equiv \langle \Psi_f'(\sigma + \tau; t_d, t_0) J(\tau) \Psi_i^\dagger(0; t_d, t_0) \rangle \\ &= \sum_{ab} G_{af}^*(t_d, t_0) C_{ab}^{3\text{pt}}(\sigma, \tau) G_{ib}(t_d, t_0). \end{aligned} \quad (149)$$

As with the FP two-point function, the elements of the resulting FP three-point matrix admit good spectral expansions and thus can be analyzed as with any other three-point function. While this is often accomplished by fitting, to reduce the complexity of comparisons we instead restrict to explicit effective estimators applied to the individual correlators in $\hat{C}_{fi}(t)$. Specifically, we employ summation [Eq. (78)] and power-iteration [Eq. (77)] effective matrix elements as in Sec. III C.

Some additional GEVP definitions are required for analysis of noisy data. Rather than filtering and treating the resulting partial data missingness as in a Lanczos analysis, we take the more standard approach of fully discarding GEVP estimates for t_0, t_d where any complex

²⁰Where $\lfloor \cdot \rfloor$ indicates the integer floor operation, i.e., rounding down to the nearest integer.

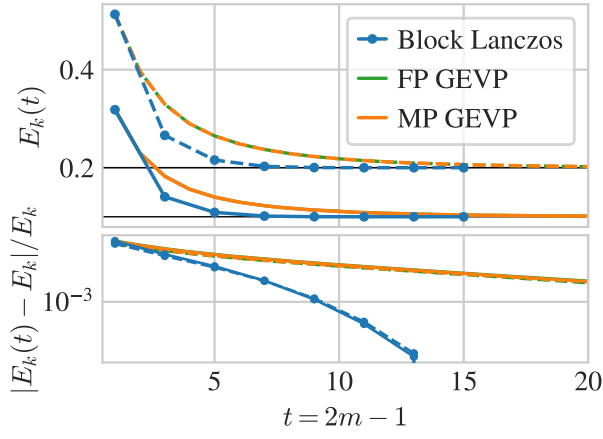


FIG. 25. The lowest two energies in the spectrum of the noiseless example Eq. (72) (top) and their relative errors versus the true value (bottom) as extracted by block Lanczos (blue), GEVP with a moving pivot at $t_0 = \lfloor t/2 \rfloor$ (orange), and GEVP with a fixed pivot at $t_0 = 5$ and $t_d = 10$ (green). For each k , the two GEVP estimates closely coincide at all t . In the bottom panel, the relative errors for both $k = 0$ and 1 collapse into two curves, one for block Lanczos and one for the GEVP estimators. Block Lanczos solves for the true eigenvalues exactly after $m = N_t/4 = 8$ steps.

eigenvalues arise. Separately, the FP scheme requires some choice of t_0 and t_d to define the pivot; here, we use $t_0 = 12$ and $t_d = 19$ as selected using the automated procedure of Ref. [66].²¹ The FP scheme is defined taking a common pivot matrix G_{ak} across all bootstraps, computed from the central-value expectation of $C(t)$ —i.e., the change of operator basis is not included in the error propagation. The same definitions are used below for overlaps and matrix elements.

2. Spectrum

We first compare block Lanczos with GEVP in application to the noiseless mock-data example from Sec. III, Eq. (72). The faster convergence of block Lanczos over GEVP extractions of the spectrum is unambiguous, as shown in Fig. 25, which compares block Lanczos with both the moving-pivot scheme with $t_0 = \lfloor t_d/2 \rfloor$ as well as a fixed-pivot GEVP with $t_0 = 5$ and $t_d = 10$ (note FP results

²¹Contributions to $G_{ak}(t_d, t_0)$ from states with energies above E_r are exponentially suppressed for large t_0 and t_d . For large enough t_0 and t_d where these contributions are negligible, $E_k^{\text{FP}}(t, t_0, t_d)$ will be identical to $E_k^{\text{MP}}(t)$ up to statistical uncertainties. In this regime, the positive features of both moving- and fixed-pivot GEVP effective mass estimators, including monotonicity with t and the interlacing theorem, apply to either definition. However, statistical noise increases with t_0 and t_d and results will become unreliable if t_0 and t_d are taken too large at finite statistics. The algorithm of Ref. [66] selects the largest t_0 and t_d where there is statistical agreement between fixed- and moving-pivot estimators.

are only weakly sensitive to the choice of t_0 and t_d in the noiseless case). With $m = 1$, block Lanczos and both GEVP schemes give identical results as expected,²² and for a few iterations, GEVP and block Lanczos provide comparable extractions. However, with increasing iterations, block Lanczos converges exponentially faster, as predicted by comparison of the block KPS bound of Sec. IV leading to $e^{-2t\sqrt{E_r-E_0}}$ convergence in comparison to the $e^{-t(E_r-E_0)}$ convergence provided by GEVP [3,4]. Besides improving convergence on the lowest $r = 2$ states shown in Fig. 25, block Lanczos yields additional energy estimates for $(m - 1)r$ higher states that are unconstrained by GEVP, as explored in Sec. III.

Proceeding to the noisy nucleon example, Fig. 26 compares the block Lanczos extraction of the ground and first excited states with estimates made with two different GEVP schemes: the FP scheme noted above and the moving-pivot scheme with $t_0 = \lfloor t/2 \rfloor$. The improved convergence and signal-to-noise properties of block Lanczos over either GEVP scheme are apparent. At $t = 1$, the block Lanczos and MP GEVP estimators are (statistically) identical; the FP estimators as defined here differ nontrivially in the noisy case. However, examining $t > 1$, energy estimates by block Lanczos reach their steady-state values earlier than those of either GEVP scheme; this is especially clear for the excited-state energy E_1 . At early times before convergence, the GEVP estimates are more precise than the Lanczos ones, similar as observed in analogous scalar correlator analyses [16,17]. However, signal-to-noise for the Lanczos extractions remains constant (up to fluctuations) after the estimates have converged, while SNR for both GEVP schemes decreases exponentially. The Lanczos estimate of E_0 is consistent with the plateaus observed in each GEVP estimator, but for E_1 , SNR degrades in the GEVP estimates before any convincing plateau is achieved.

Block Lanczos results from the last iteration where residual bounds can be computed, $m_{\text{max}} = 47$, are compared to multistate fits to FP GEVP results as well as other scalar Lanczos results in Figs. 27 and 28. The only block Lanczos hyperparameter is $F^{\text{ZCW}} = 10$, parametrizing how much smaller overlaps must be than ones at small m to be labeled spurious. This amounts to an enormous practical reduction in complexity versus the fits to the FP correlation functions, which are performed using the strategy summarized in Fig. 16 of Ref. [70]: in brief, fits are performed for all possible $t_{\text{min}} \in [2, t_{\text{max}} - 6]$ where t_{max} is chosen as the largest t where the SNR of $E_k^{\text{FP}}(t)$ is greater than 0.2. Linear

²²For GEVP estimators defined as effective energies, Eq. (141) (145), this holds for the MP definition because $\lambda(0, 0) = 1$. This alignment of definitions is intentional, to reflect the deeper underlying coincidence between GEVP eigenvalues and Ritz values discussed above. For FP schemes the alignment is not exact, see noisy results below.

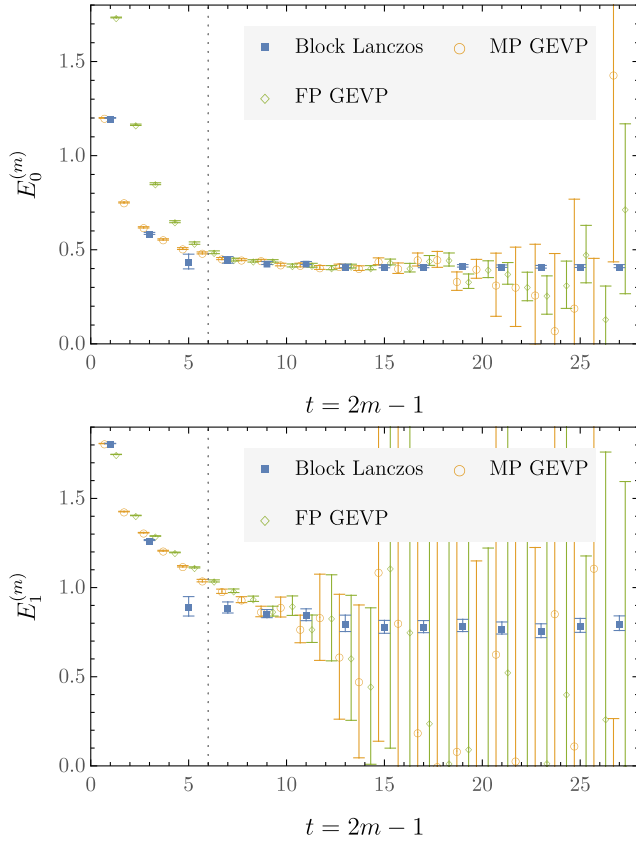


FIG. 26. Energy estimates for the two lowest-lying states in the nucleon spectrum as extracted by block Lanczos (blue), GEVP with a moving pivot $t_0 = \lfloor t/2 \rfloor$ as in Ref. [4] (orange), and GEVP with a fixed pivot with $t_0 = 11$ and $t_d = 18$ chosen using the algorithm of Ref. [66] (green). Bootstrap-median estimators are used for block Lanczos. In all cases, uncertainties are computed using idiomatic bootstrap confidence intervals.

shrinkage [71,72] is used to regulate numerical instabilities in covariance matrix inversion. One-, two-, three-, ... state fits are performed until adding states does not lower the Akaike information criterion [73] (AIC) by more than a threshold of $0.5N_{\text{dof}}$. Various checks on the consistency of numerical optimization detailed in Ref. [70] are then performed. A weighted average of the AIC-preferred fits for each $t_{\min}$ passing these checks is then taken, where the weights are proportional to the p -value divided by the variance of each fit [74]. Hyperparameters enter this scheme through the numerical values specified above and others detailed in Ref. [70], as well as the hyperparameters for choosing t_0 and t_d summarized in Fig. 16 of Ref. [66] that here lead to $t_0 = 11$ and $t_d = 18$. For $n = 0$, the fits with weights above 0.1 that result from this procedure are one-state fits with $t_{\min} \in [9, 12]$ and two-state fits for $t_{\min} = 4$. For $n = 1$, the only fits with weights above 0.1 are two-state fits with $t_{\min} \in [2, 3]$.

Note that this GEVP analysis prescription is far from unique. The GEVP formalism admits many other possible

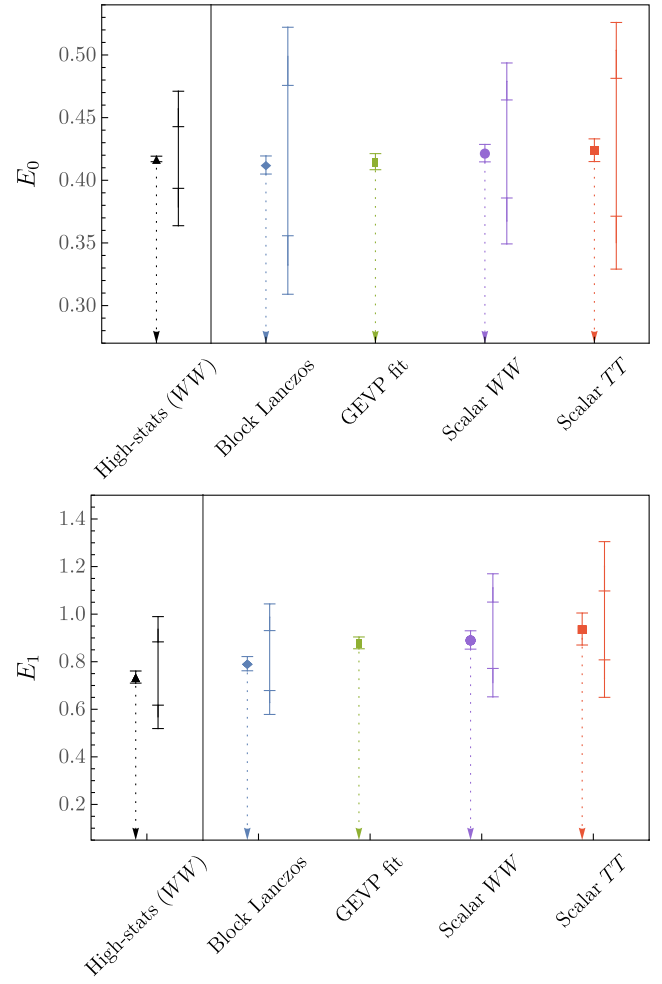


FIG. 27. Block Lanczos results for $E_0^{(m_{\max})}$ (top) and $E_1^{(m_{\max})}$ (bottom) with $m_{\max} = 47$, shown in Fig. 17, are compared with FP GEVP fits performed as described in Refs. [65,66,70] as well as scalar Lanczos results from the corresponding diagonal correlator matrix entries, higher-statistics scalar Lanczos analysis. Results with bootstrap uncertainties are shown as points with error bars in all cases. Dotted arrows indicate variational bounds. Residual bounds for Lanczos results are shown as intervals whose upper and lower edges are marked by error bars showing bootstrap uncertainties on $-\ln(\lambda_k \pm \sqrt{B_k})$.

analyses of the same data, all of which ideally should give statistically compatible results. Tables V and VI tabulate multistate fit results for various $t_{\min}$ and numbers of states included in the fit.

A weighted average of GEVP fit results performed as described above gives $E_0 = 0.4150(59)$. This is in excellent agreement with all Lanczos results, and all analyses of the same data give comparable uncertainty estimates. For E_1 , a weighted average of GEVP fit results gives $0.879(25)$, which differs from the high-statistics scalar Lanczos result $0.735(25)$ by 4σ , while the block Lanczos result, $0.792(32)$, shows only 1.4σ tension (see Table I). Other choices of GEVP fit hyperparameters lead to similar tensions; for

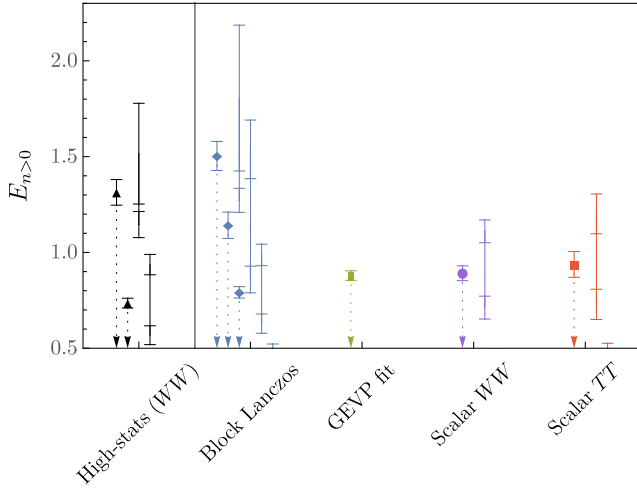


FIG. 28. Results as in Fig. 27 shown for E_2 and E_3 where available along with E_1 .

instance increasing the noise tolerance to 0.5 leads to $t_{\max} = 11$ instead of $t_{\max} = 9$, which changes the weighted average fit for E_1 to 0.879(21). Scalar Lanczos results for the diagonal correlation functions also show $2-3\sigma$ tensions in comparison with high-statistics results. However, the residual bounds $-\ln(\lambda_1 \pm \sqrt{B_1})$ provided by scalar Lanczos results²³ are in 1σ agreement with the corresponding high-statistics results. This residual-bound consistency shows that each result could be converging to the same true energy eigenvalue but with significant excited-state effects still present in one or more determinations. Agreement at the level of Ritz value statistical uncertainties between block Lanczos and high-statistics scalar Lanczos suggests that the low-statistics scalar Lanczos results for E_1 still include statistically significant excited-state effects that are effectively removed by block Lanczos (or by going to higher statistics where they can be resolved more easily). With only GEVP results for guidance, any interpretation of E_1 as more than a variational upper bound is therefore likely to be incorrect at 4σ in this example.

The residual bounds for $n \in \{2, 3\}$ from block Lanczos both overlap the residual bounds for $n = 2$ from the high-statistics scalar analysis, making it difficult to conclude whether block has resolved two distinct physical states or

TABLE V. Multistate fit results for the $n = 0$ FP GEVP correlator with $t_{\max} = 19$; see the main text for details. Fits where χ^2 is not decreased by increasing N_{states} , indicating optimizer convergence failures, are omitted.

N_{states}	$t_{\min}$	E_0	χ^2/N_{dof}	N_{dof}
3	2	0.403(12)	1.0	13
3	3	0.4221(50)	1.7	12
3	5	0.4025(93)	1.3	10
3	8	0.4139(60)	1.1	7
3	9	0.4157(40)	1.4	6
3	10	0.4143(42)	1.6	5
3	11	0.4138(71)	1.9	4
3	12	0.416(10)	2.5	3
3	13	0.42(21)	3.7	2
3	14	0.422(12)	7.1	1
2	2	0.4293(30)	2.0	15
2	4	0.4146(59)	1.2	13
2	5	0.4448(26)	7.2	12
2	6	0.4331(50)	3.4	11
2	7	0.410(11)	1.0	10
2	8	0.4233(49)	2.3	9
2	9	0.4157(40)	1.0	8
2	10	0.4143(54)	1.1	7
2	11	0.4139(62)	1.3	6
2	12	0.4159(69)	1.5	5
2	13	0.4173(94)	1.9	4
2	14	0.42(21)	2.4	3
2	15	0.41(21)	3.1	2
1	2	0.8147(84)	260	17
1	3	0.5213(51)	110	16
1	4	0.4658(30)	21	15
1	5	0.4448(26)	6.2	14
1	6	0.4331(30)	2.9	13
1	7	0.4289(34)	2.7	12
1	8	0.4233(36)	1.9	11
1	9	0.4157(40)	0.83	10
1	10	0.4143(42)	0.87	9
1	11	0.4138(61)	0.97	8
1	12	0.4158(69)	1.1	7
1	13	0.417(10)	1.2	6
1	14	0.422(11)	1.4	5
1	15	0.411(16)	1.5	4

encountered an artificial degeneracy due to the appearance of multiple eigenvalues, which can generically occur for block Lanczos states with $n \geq r$ [22].

3. Overlap factors

For the noiseless example of Sec. III, Fig. 29 compares convergence of different Lanczos and GEVP estimators of overlaps with ψ_T and ψ_W for the lowest-lying two states. As with the spectrum, while GEVP provides a comparable-quality estimate at early times, block Lanczos estimators converge exponentially more rapidly. In practice, it is necessary to manually enforce the sign convention $Z_{k0} > 0$ on the GEVP estimates.

²³It is clear from Fig. 18 that $B_k^{(m)}$ shows more nontrivial m dependence than $E_k^{(m)}$. For simplicity, we have computed the widths of the residual bounds using the m where the central value of the bootstrap median estimator $B_k^{(m)}$ is minimized. The m dependence of residual bounds, which must be a monotonic decrease for symmetric (block) Lanczos [22] but can be nontrivial for oblique (block) Lanczos, will be investigated in more detail in future work. It will also be interesting to study whether the increase in the widths of the residual bounds between $r = 1$ and $r = 2$ seen in Fig. 27 appears in other noisy examples.

TABLE VI. Multistate fit results for the $n = 1$ FP GEVP correlator with $t_{\max} = 9$; see the main text and Table V for details.

N_{states}	$t_{\min}$	E_1	χ^2/N_{dof}	N_{dof}
3	2	0.86(10)	0.60	3
3	3	0.862(25)	0.92	2
2	2	0.892(12)	1.1	5
2	3	0.862(20)	0.46	4
2	4	0.84(14)	0.44	3
2	5	1.01(10)	36	2
1	2	1.2561(17)	500	7
1	3	1.1608(23)	170	6
1	4	1.0852(33)	57	5
1	5	1.0131(56)	18	4

For the noisy nucleon data of Sec. V, Fig. 30 compares block Lanczos overlap estimators with the GEVP estimators, demonstrating an even more clear advantage than for the spectrum. Note that we find that the sign convention $Z_{k0} > 0$ must be enforced by hand for the GEVP results, achieved by applying an overall sign to Z_{ka} at each t and within each bootstrap. For all GEVP overlaps, the signal breaks down more rapidly than for the spectrum, with SNR quickly degrading before each estimator begins producing complex values and thus becoming unreliable. While the FP GEVP produces consistent results for the ground state overlaps Z_{0a} , the MP estimator does not provide a clear plateau and passes through the value indicated by the Lanczos estimator in each case. For the excited state overlaps Z_{1a} , neither GEVP scheme produces

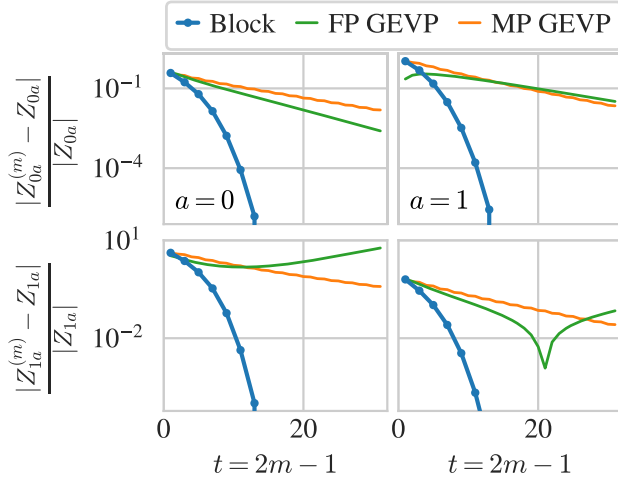


FIG. 29. Relative error in ground-state overlap factors estimated in the noiseless example Eq. (72) by block Lanczos (blue), the moving-pivot GEVP estimator Eq. (146) with $t_0 = \lfloor t/2 \rfloor$ (orange), and the fixed-pivot GEVP estimator Eq. (147) with $t_0 = 5$ and $t_d = 10$ (green). Block Lanczos solves for the true overlap factors exactly at $m = N_t/4 = 8$ steps. The sign convention $Z_{k0} > 0$ has been enforced by hand for the GEVP estimate of Z_{01} .

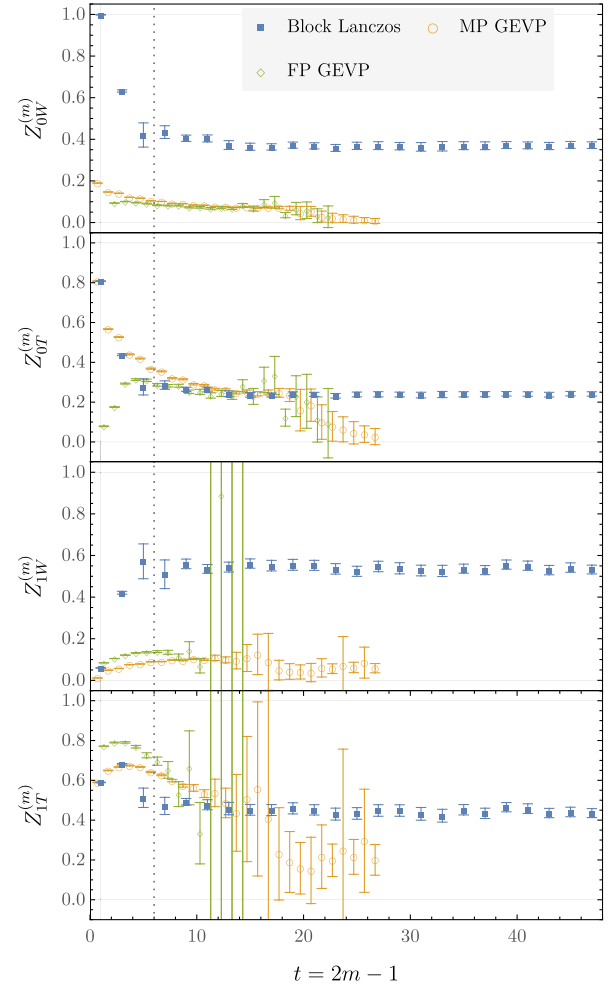


FIG. 30. For the nucleon correlator matrix example, overlap factors as computed by block Lanczos (blue) and by GEVP with a moving pivot at $t_0 = \lfloor t/2 \rfloor$ as in Ref. [4] (orange), and GEVP with a fixed pivot with $t_0 = 11$ and $t_d = 18$ chosen using the algorithm of Ref. [66] (green). Bootstrap-median estimators are used for block Lanczos. In all cases, uncertainties are computed using idiomatic bootstrap confidence intervals.

any convincing plateau before signal is lost; the FP GEVP scheme for Z_{1a} presents a brief but potentially deceptive pseudoplateau. In contrast, block Lanczos provides stable results with no SNR decay and none of the pathological behaviors observed for the GEVP estimators.

4. Matrix elements

For the noiseless mock-data example of Sec. III, Fig. 31 compares convergence of block Lanczos estimates of the diagonal matrix elements J_{00} and J_{11} with the moving-pivot GEVP estimator Eq. (148), as well as different estimators constructed from the fixed-pivot GEVP three-point function Eq. (149). As expected, the moving-pivot GEVP estimator Eq. (148) with $m = \lfloor t/4 \rfloor$ coincides with the block Lanczos one at $t = 1$. Block Lanczos provides a

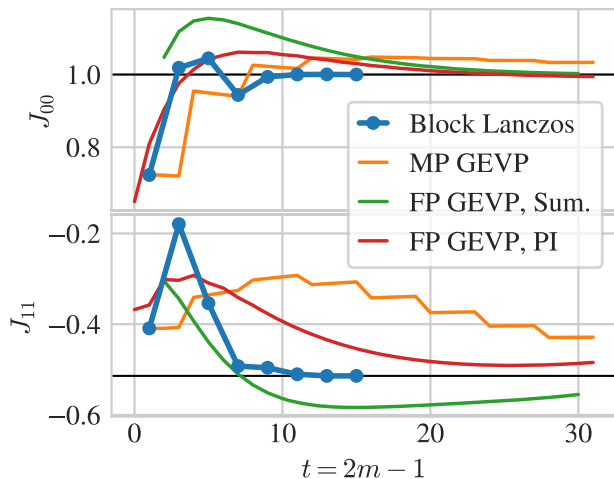


FIG. 31. For the off-diagonal noiseless example Eq. (75), estimates of diagonal matrix elements for the ground state J_{00} (top) and first excited state J_{11} as computed by block Lanczos (blue), by GEVP with a moving pivot at $t_0 = \lfloor t/2 \rfloor$ per Eq. (148), and using the summation [Eq. (78), green] and power iteration [Eq. (77), red] estimators applied to the diagonals of a GEVP three-point constructed per Eq. (149) for a fixed pivot at $t_0 = 5$ and $t_d = 10$. Block Lanczos solves for the true J_{00} exactly after $m = N_t/4 = 8$ steps.

qualitatively improved extraction of each. For either state, the initial Lanczos estimate changes nonsmoothly—a feature advertising nonconvergence—then converges quickly to the true value. In contrast, the GEVP estimates vary smoothly²⁴ and converge poorly. While the FP estimators for J_{00} provide a good estimate by the final step, the MP estimator shows potentially deceptive behavior: the estimate changes only slowly while still visibly offset from the true value. For the excited-state J_{11} , the MP estimator clearly does not converge before the end of the correlator. The two FP estimators drift near the true J_{11} but will not provide a reliable estimate; in the FP scheme each will eventually converge to J_{00} , which behavior is suggested by the upwards turn in each by the end of the correlator.

VII. CONCLUSIONS

We present a new method to extract energies, overlaps, and operator matrix elements from correlator matrix data based on an oblique block Lanczos algorithm and a new physically motivated test for spurious state filtering. The method extends and improves upon GEVP methods—the present standard and state of the art—in control of excited-state effects, systematic uncertainty quantification, and signal-to-noise ratio.

²⁴Up to expected even-odd effects in the MP estimator due to $t_0 = \lfloor t/2 \rfloor$, and in the power-iteration FP estimator due to the piecewise definition in Eq. (77).

Reformulating the Cullum-Willoughby test in terms of a cut on overlap factors provides a fully physical picture of spurious state filtering. Noise results in an apparent expansion of Krylov space outside the part of Hilbert space containing the states of physical interest, and the spurious states are a consequence of this expansion. Following this motivation, the ZCW test identifies spurious states as those with “spuriously small” overlap factors, with a physically motivated threshold for “small.” After filtering, the remaining nonspurious Ritz values and vectors provide reliable estimates for some eigenstates of the physical transfer matrix. The resulting estimators converge to the true spectrum faster than either GEVP or scalar Lanczos estimators, have directly calculable two-sided systematic uncertainty bounds, and have constant signal-to-noise ratios at large iteration count.

For spectroscopy applications, block Lanczos analysis is immediately applicable to existing datasets and can be adopted immediately. As in the scalar case, the data requirements to use block Lanczos to extract operator matrix elements require evenly spaced three-point data at small sink times. This is in tension with the common strategy for sequential-source calculations where data are generated for only a few intermediate sink times where excited-state effects are reduced. Data generation strategies should be adjusted. This is especially true where correlator matrices are employed, as they are typically used in matrix element analyses only when excited-state effects (or properties) are a concern; the badly contaminated but well-resolved measurements at early sink times that are not typically computed are exactly those that Lanczos uses to control excited-state effects.

As demonstrated, filtering states, sorting on Ritz values, and using an idiomatically bootstrapped median estimator to define values and uncertainties provides a simple and effective analysis. However, state identification (i.e., associating measurements between bootstraps to define the set to be averaged over) still poses some outstanding challenges. In Sec. V, we observed that simply sorting measurements on the Ritz values led to some obvious misidentifications which complicate the analysis of relatively high-energy excited states. Although ground-state energies and overlaps appear to be more robust, and there are indications that this effect is less severe at high statistics, it is nevertheless critical to check for such misidentification before blindly trusting the output of a Lanczos analysis. Better schemes for state identification which can provide more reliable excited-state extractions at finite statistics are an important topic for further study.

The key improvements offered by Lanczos over previous methods are spurious state filtering and the ability to construct rigorous, computable two-sided bounds, rather than just the estimators derived in Sec. II absent this additional machinery. As discussed above and in Refs. [16,17,75,76], the unfiltered Lanczos estimators of

Sec. II provide numerically identical outputs to other methods in certain cases and with appropriately aligned definitions. This includes not only standard effective estimators and GEVP in one-step limits, but also Prony's method²⁵ and its block generalization as introduced in Ref. [5]. Further connections between these conceptually distinct but numerically coincident methods—and opportunities for improvements suggested by these connections—will be explored in future work [30].

Besides correlator matrix analyses, the block Lanczos spurious eigenvalue filtering algorithms defined in this work may also be useful for the more traditional application of computing eigenvalues and eigenvectors of numerical matrices. To the best of our knowledge, this work represents the first construction of the ZCW test and its block generalization, as well as the block CW test obtained by consideration of the eigenvalue-eigenvector identity. (Z)CW tests provide alternatives to selective reorthogonalization—the present standard approach to using Lanczos methods with finite-precision arithmetic—which do not require storage and use of converged Ritz vectors. These may now be used with block Lanczos, which could provide practical advantages in some applications.

The Chroma [77], QUDAA [78–80], QDPhromaJIT [81], and Chromaform [82] software libraries were used to generate the data in this work. Numerical analysis was performed using NumPy [83], SciPy [84], PANDAS [85,86], LSQFIT [87], GVAR [88], MPMATH [31], and *Mathematica* [89]. Figures were produced using MATPLOTLIB [90], SEABORN [91], and *Mathematica* [89].

ACKNOWLEDGMENTS

We thank Anthony Grebe, Ryan Abbott, Will Detmold, George Fleming, Christoph Lehner, Johann Ostmeyer, Robert Perry, Dimitra Pefkou, and Fernando Romero-López for stimulating discussions and helpful comments. We thank Robert Edwards, Rajan Gupta, Balint Joó, Kostas Orginos, and the NPLQCD Collaboration for generating the gauge-field ensembles used in this study. Quark propagators and nucleon correlator matrices were computed by the NPLQCD Collaboration. The calculations were performed using an allocation from the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program using the resources of the Oak Ridge Leadership Computing Facility located in the Oak Ridge National Laboratory, which is supported by the Office of

Science of the Department of Energy under Contract No. DE-AC05-00OR22725. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. We acknowledge USQCD computing allocations and PRACE for awarding access to Marconi100 at CINECA, Italy. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. This research used facilities of the USQCD Collaboration, which are funded by the Office of Science of the U.S. Department of Energy. This manuscript has been authored by FermiForward Discovery Group, LLC under Contract No. 89243024CSC000002 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

DATA AVAILABILITY

The data supporting this study's findings are available within the article.

APPENDIX A: LANCZOS VECTOR BIORTHOGONALITY

The orthogonality of Lanczos vectors within each block j follows directly from the definition $\Delta_j = \beta_j \gamma_j$ [Eq. (8)] of β_j and γ_j in terms of the residual norm Δ_j ,

$$\begin{aligned} \langle v_{ja}^L | v_{jb}^R \rangle &= \sum_{cd} \beta_{jac}^{-1} \langle r_{jc}^L | r_{jd}^R \rangle \gamma_{jdb}^{-1} \\ &= [\beta_j^{-1} \Delta_j \gamma_j^{-1}]_{ab} \\ &= [\beta_j^{-1} \beta_j \gamma_j \gamma_j^{-1}]_{ab} \\ &= \delta_{ab} \end{aligned} \quad (\text{A1})$$

using Eq. (9) in the first equality and the definition $\langle r_{jc}^L | r_{jd}^R \rangle = \Delta_{jcd}$ [Eq. (7)] in the second equality. Orthogonality between Lanczos vectors from different blocks follows from the definitions of α_j , β_j , and γ_j in Eq. (5) as described next.

Starting with the orthogonality relation $\langle v_{ja}^L | v_{(j+1)b}^R \rangle = 0$, take the induction hypothesis $\langle v_{ia}^L | v_{i'b}^R \rangle = \delta_{ii'} \delta_{ab}$ for all $i, i' \leq j$. The base case $j = 1$ is established above. By Eqs. (6) and (9),

$$\begin{aligned} \langle v_{ja}^L | v_{(j+1)b}^R \rangle &= \sum_{cd} \left[\delta_{cd} \langle v_{ja}^L | T | v_{jd}^R \rangle - \langle v_{ja}^L | v_{jc}^R \rangle \alpha_{jcd} \right. \\ &\quad \left. - \langle v_{ja}^L | v_{(j-1)c}^R \rangle \beta_{jcd} \right] \gamma_{(j+1)db}^{-1}. \end{aligned} \quad (\text{A2})$$

²⁵Equivalence between unfiltered Ritz values and Prony solutions, and the Prony generalized eigenvalue method (PGEVM) [6] with particular parameters, was first noted in Ref. [16] and proven analytically in Refs. [75,76]. In contrast to the claims of Refs. [75], the generalization of PGEVM to correlator matrices discussed in Ref. [6] is not equivalent to block Lanczos; see Appendix G.

The third term vanishes and second term simplifies by the induction hypothesis, giving

$$\langle v_{ja}^L | v_{(j+1)b}^R \rangle = \sum_d \left[\langle v_{ja}^L | T | v_{jd}^R \rangle - \alpha_{jad} \right] \gamma_{(j+1)db}^{-1}. \quad (\text{A3})$$

The orthogonality relation $\langle v_{ja}^L | v_{(j+1)b}^R \rangle = 0$ is therefore achieved by defining α_{jab} as in Eq. (5). The same definition leads to $\langle v_{(j+1)a}^L | v_{jb}^R \rangle = 0$.

The corresponding orthogonality relation involving block $j - 1$ is the vanishing of

$$\begin{aligned} \langle v_{(j-1)a}^L | v_{(j+1)b}^R \rangle &= \sum_{cd} \left[\delta_{cd} \langle v_{(j-1)a}^L | T | v_{jd}^R \rangle \right. \\ &\quad - \langle v_{(j-1)a}^L | v_{jc}^R \rangle \alpha_{jcd} \\ &\quad \left. - \langle v_{(j-1)a}^L | v_{(j-1)c}^R \rangle \beta_{jcd} \right] \gamma_{(j+1)db}^{-1}. \end{aligned} \quad (\text{A4})$$

The second term vanishes by the first piece of the induction step, while the third term simplifies using Eq. (A1), which leads to

$$\langle v_{(j-1)a}^L | v_{(j+1)b}^R \rangle = \sum_d \left[\langle v_{(j-1)a}^L | T | v_{jd}^R \rangle - \beta_{jad} \right] \gamma_{(j+1)db}^{-1}. \quad (\text{A5})$$

The orthogonality relation $\langle v_{(j-1)a}^L | v_{(j+1)b}^R \rangle = 0$ is therefore achieved by defining β_{jab} as in Eq. (5). An analogous induction for $\langle v_{(j+1)a}^L | v_{(j-1)b}^R \rangle$ shows this orthogonality relation is also achieved by Eq. (5). Orthogonality with blocks $1, \dots, j - 2$ follows automatically from the fact that the recurrences for $T | v_{(j+1)a}^R \rangle$ and $\langle v_{(j+1)a}^L | T$ only involve Lanczos vectors from blocks $j + 1$, j , and $j - 1$. These definitions therefore ensure

$$\langle v_{ia}^L | v_{jb}^R \rangle = \delta_{ij} \delta_{ab}. \quad (\text{A6})$$

Similarly applying the recursion relation to the normalization condition

$$\begin{aligned} \langle v_{ja}^L | v_{jb}^R \rangle &= \sum_{cd} \left[\delta_{cd} \langle v_{ja}^L | T | v_{(j-1)d}^R \rangle - \langle v_{ja}^L | v_{(j-1)c}^R \rangle \alpha_{jcd} \right. \\ &\quad \left. - \langle v_{ja}^L | v_{(j-2)c}^R \rangle \beta_{jcd} \right] \gamma_{jdb}^{-1}, \end{aligned} \quad (\text{A7})$$

and noting that the second and third lines vanish by the orthogonality relations above gives

$$\langle v_{ja}^L | v_{jb}^R \rangle = \sum_d \langle v_{ja}^L | T | v_{(j-1)d}^R \rangle \gamma_{jdb}^{-1}. \quad (\text{A8})$$

This is equal to δ_{ab} iff

$$\gamma_{jab} = \langle v_{ja}^L | T | v_{(j-1)b}^R \rangle. \quad (\text{A9})$$

An analogous derivation shows that

$$\beta_{jab} = \langle v_{(j-1)a}^L | T | v_{jb}^R \rangle. \quad (\text{A10})$$

These relations, along with Eqs. (6) and (9), can be used to prove the oblique block Lanczos three-term recurrences, Eq. (4).

APPENDIX B: KPS BOUND

In this section, we reproduce the block KPS bound of Ref. [21], suitably translated to the notation used in this work, and derive the form provided in the main text. The derivation of Ref. [21] applies for nonoblique block Lanczos only, so no L/R distinctions are made and the result only applies in the infinite-statistics limit where the underlying Hermiticity of the transfer matrix is manifest in the data. We thus necessarily assume an $r \times r$ Hermitian correlator matrix

$$\langle \psi_a | T^t | \psi_b \rangle = C_{ab}(t) = C_{ba}^*(t) \quad (\text{B1})$$

as in the rest of this work. In the nonoblique block Lanczos algorithm there is no freedom of convention, so symmetric definitions apply, i.e.,

$$C(0) = \beta_1 \gamma_1 = \beta_1^2 \quad (\text{B2})$$

and

$$|v_j^R\rangle = |v_j^L\rangle \equiv |v_{1a}\rangle = \sum_b |\psi_b\rangle \beta_{1ba}^{-1}. \quad (\text{B3})$$

Consider the spectra of true eigenvalues λ_k and Ritz values after m steps $\lambda_k^{(m)}$, ordered such that²⁶

$$\begin{aligned} \lambda_0 &> \lambda_1 > \dots > \lambda_\infty, \quad \text{and} \\ \lambda_0^{(m)} &> \lambda_1^{(m)} > \dots > \lambda_{rm-1}^{(m)}. \end{aligned} \quad (\text{B4})$$

Equation 3.10 of Ref. [21] bounds the relative deviation of $\lambda_k^{(m)}$ from λ_k as

$$0 \leq \frac{\lambda_k - \lambda_k^{(m)}}{\lambda_k - \lambda_\infty} \leq \left[\frac{K_k^{(m)}}{T_{m-k-1}(\Gamma_k)} \right]^2 \| |k\rangle - |\hat{x}_k\rangle \|^2 \quad (\text{B5})$$

where $K_k^{(m)}$ and Γ_k are as defined in Eq. (80)(81) above, T_n is an n th order Chebyshev polynomial of the first kind,

²⁶Note that our convention is that λ_k are zero-indexed as $k = 0, 1, \dots$, which differs from the one-indexed convention $k = 1, 2, \dots$ of Ref. [21]. All expressions here have been suitably translated to the zero-indexed convention.

and $|k\rangle$ is a true eigenstate satisfying $T|k\rangle = \lambda_k|k\rangle$. In what follows, we define and compute the final term $\| |k\rangle - |\hat{x}_k\rangle \|^2$.

As defined in Eqs. 3.1 and 3.2 of Ref. [21], the vector $|\hat{x}_k\rangle$ is the unique vector in the r -dimensional Krylov space $\mathcal{K}^{(1)}$ spanned by the initial states $|v_{1a}\rangle$ that satisfies

$$\langle \hat{x}_k | l \rangle = \delta_{kl} \quad (\text{B6})$$

for $l \in \{k, k+1, \dots, k+r-1\}$, i.e.,

$$\begin{aligned} \langle \hat{x}_k | k \rangle &= 1, \\ \langle \hat{x}_k | l \rangle &= 0, \quad (k < l < k+r). \end{aligned} \quad (\text{B7})$$

This vector may be written as²⁷

$$|\hat{x}_k\rangle = \sum_{l=k}^{k+r-1} t_l^{(k)} \pi_1 |l\rangle \quad (\text{B8})$$

in terms of r coefficients $t_l^{(k)}$ and a projection operator π_1 onto $\mathcal{K}^{(1)}$; a convenient choice is

$$\pi_1 = \sum_a |v_{1a}\rangle \langle v_{1a}| \quad (\text{B9})$$

noting the lack of L/R distinction. The coefficients $t_l^{(k)}$ may be obtained as the solution to the $r \times r$ linear system

$$\sum_{l=k}^{k+r-1} \langle j | \pi_1 | l \rangle t_l^{(k)} = \delta_{jk} \quad (\text{B10})$$

where the free index j is considered only over the range $[k, k+r-1]$. The right-hand side should not be thought of as a matrix but rather a length- r vector with components $[1, 0, \dots, 0]$. The matrix defining the system may be rewritten as

$$\begin{aligned} \langle j | \pi_1 | l \rangle &= \sum_a \langle j | v_{1a} \rangle \langle v_{1a} | l \rangle \\ &= \sum_{abc} \langle j | \psi_b \rangle \frac{\Omega_{1ba}}{\sqrt{\Lambda_{1a}}} \frac{\Omega_{1ca}^*}{\sqrt{\Lambda_{1a}}} \langle \psi_c | l \rangle \\ &= \sum_{bc} Z_{jb} [C^{-1}(0)]_{bc} Z_{lc}^* \\ &\equiv X_{jl} \end{aligned} \quad (\text{B11})$$

recalling the definition $Z_{ja} = \langle j | \psi_a \rangle$ of the overlap factors. Although X_{jl} can be considered over the full basis of states,

Eq. (B10) only involves $j, l \in [k, k+r-1]$, so the relevant part is the $r \times r$ block on the diagonal beginning at $(j, l) = (k, k)$. Denoting this block as $X_{(k,r)}$, the coefficients $t_l^{(k)}$ can be obtained by inverting

$$t_l^{(k)} = \sum_{j=k}^{k+r-1} [X_{(k,r)}^{-1}]_{lj} \delta_{jk} = [X_{(k,r)}^{-1}]_{lk} \quad (\text{B12})$$

i.e., $t_l^{(k)}$ are obtained as the first column of the inverted block $X_{(k,r)}^{-1}$, recalling that j, l in the above expression are indexed starting at k .

Finally, we expand

$$\begin{aligned} \| |k\rangle - |\hat{x}_k\rangle \|^2 &= \langle k | k \rangle - \langle k | \hat{x}_k \rangle - \langle \hat{x}_k | k \rangle + \langle \hat{x}_k | \hat{x}_k \rangle \\ &= \langle \hat{x}_k | \hat{x}_k \rangle - 1 \end{aligned} \quad (\text{B13})$$

using $\langle k | \hat{x}_k \rangle = 1$ by Eq. (B6). The remaining nontrivial term can be simplified as

$$\begin{aligned} \langle \hat{x}_k | \hat{x}_k \rangle &= \sum_{j,l=k}^{k+r-1} t_j^{(k)*} \langle j | \pi_1^2 | l \rangle t_l^{(k)} \\ &= \sum_{jl} t_j^{(k)*} \langle j | \pi_1 | l \rangle t_l^{(k)} \\ &= \sum_j t_j^{(k)*} \delta_{jk} = t_k^{(k)*} \end{aligned} \quad (\text{B14})$$

using projector idempotency $\pi_1^2 = \pi_1$ in the second equality and Eq. (B10) in the third. We thus arrive at the result used in the main text, that

$$t_k^{(k)*} = [X_{(k,r)}^{-1}]_{kk}, \quad (\text{B15})$$

i.e., it is the upper-leftmost element of the inverted block $X_{(k,r)}^{-1}$ recalling the k -indexing, and that

$$\| |k\rangle - |\hat{x}_k\rangle \|^2 = [X_{(k,r)}^{-1}]_{kk} - 1. \quad (\text{B16})$$

The block KPS bound reduces to the scalar version in the limit $r = 1$. As presented in Refs. [17,92], the scalar KPS bound reads

$$0 \leq \frac{\lambda_k - \lambda_k^{(m)}}{\lambda_k - \lambda_\infty} \leq \left[\frac{K_k^{(m)} \tan \arccos z_k}{T_{m-k-1}(\Gamma_k)} \right]^2 \quad (\text{B17})$$

where $z_k = \langle k | v_1 \rangle = Z_k / \sqrt{C(0)}$ is the normalized overlap. The symbols $K_k^{(m)}$ and Γ_k are identical with their scalar-case definitions when $r = 1$, so it only remains to show that

$$[X_{(k,r)}^{-1}]_{kk} - 1 \xrightarrow{r=1} [\tan \arccos z_k]^2. \quad (\text{B18})$$

²⁷Note that there is a typographical error in the corresponding expression of Ref. [21].

Note that $[X_{(k,1)}^{-1}]_{kk}$ reduces to simply the reciprocal of the k th diagonal element of the matrix

$$\sum_{bc} Z_{jb} [C^{-1}(0)]_{bc} Z_{lc}^* \xrightarrow{r=1} \frac{Z_j Z_l^*}{C(0)}, \quad (\text{B19})$$

trivializing block indices, and so

$$[X_{(k,r)}^{-1}]_{kk} \xrightarrow{r=1} \frac{C(0)}{|Z_k|^2} = \frac{1}{z_k^2}. \quad (\text{B20})$$

The desired reduction Eq. (B18) finally follows from the identity $(\tan \arccos z)^2 = 1/z^2 - 1$.

APPENDIX C: PSEUDOPATEAUS AND RESIDUAL BOUNDS FOR NOISELESS SCALAR LANCZOS

Section III A compares block and scalar Lanczos extractions of the spectrum on a noiseless example defined as Eq. (72). The results therein for scalar Lanczos—reproduced in the markers of Fig. 32—raise the question of whether Lanczos can exhibit “pseudoplateau” behavior. This appendix addresses this concern. To be concrete, consider two different definitions of a pseudoplateau:

- (a) An apparent but false asymptote to a value unrelated to any energy in the true spectrum.
- (b) After some finite number of iterations, an energy is converging to that of a different state than expected.

In the example of Sec. III, the two different interpolators have suppressed overlaps with different (even and odd) subsets of states. This leads to scalar Lanczos initially yielding Ritz values which asymptote toward states with

unsuppressed overlaps, then only later providing estimates for suppressed states. For scalar Lanczos applied to C_{11} , this results in the largest Ritz value initially converging toward the first excited state, then later dropping to the ground state. This satisfies definition (b) above, if the largest Ritz value is expected to converge immediately to the ground state. The general nonmonotonicity of the scalar extractions results in various other examples satisfying (b), given appropriate choices of expectations. As always, cautious interpretation of results is required.

However, examination of the residual bounds shown in Fig. 32 provides evidence that the scalar results of Sec. III A do not satisfy definition (a). For C_{11} , the residual bounds for the lowest energy tighten quickly; for all $m \geq 3$, the bounds indicate that the Ritz value near E_1 is consistent with *only* the true E_1 and not with any other state. Similar pictures apply for various other examples across both scalar Lanczos extractions of the spectrum. Taken together, this suggests that the story told, e.g., by simply examining the sequence of lowest-lying energy estimates in the C_{11} analysis—that the ground-state energy estimator initially converges to the wrong value—is not the best interpretation of these dynamics. A better one is simply that scalar Lanczos provides no estimate at all for the ground-state energy until $m = 6$, at which point a new, separate state which converges to the ground state appears. Although state identification will always be an ambiguity, the value of the estimators in this example reliably correspond to physical values. This is in contrast with the potentially more dangerous behavior indicated by definition (a), wherein near cancellations result in false asymptotes and analyses which yield values with no physical meaning.

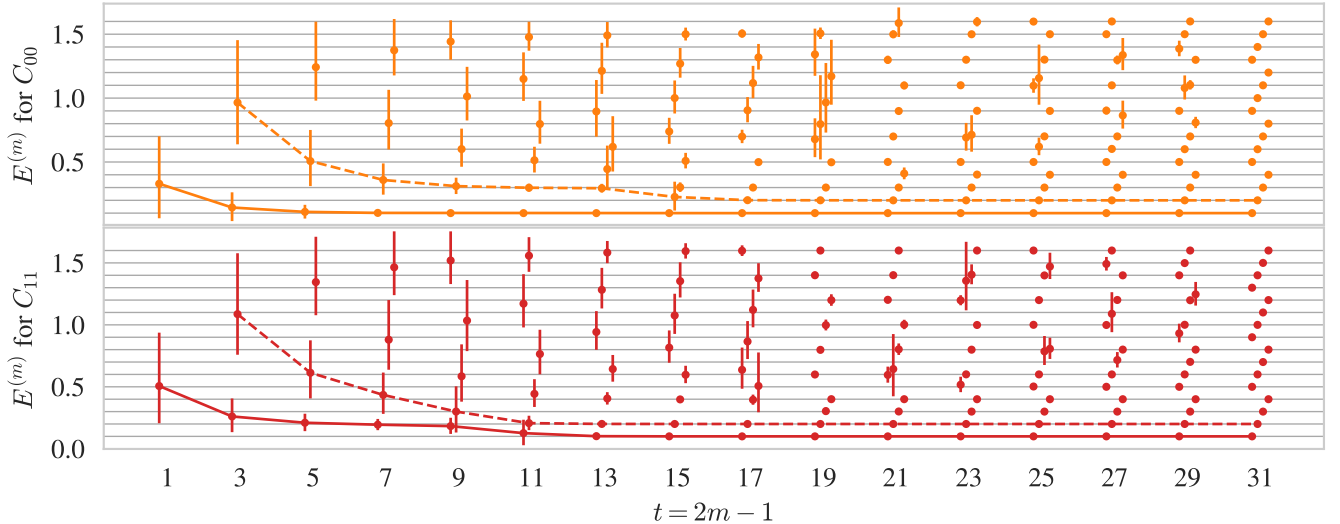


FIG. 32. Demonstration of scalar Lanczos residual bounds for the noiseless example Eq. (72), similar to Fig. 7. Error bars are not statistical but rather represent the extent of values allowed by the residual bound Eq. (92). The window of allowed values is computed for Ritz values and mapped through the logarithm, and thus asymmetric for energies. Lines connecting the lowest two energies at each m are drawn to guide the eye.

APPENDIX D: CORRELATOR MATRIX REALITY

1. C , P , and $R_{2\pi}$ symmetry

Charge conjugation (C) transformations are defined for LQCD gauge fields as

$$U(C)U_\mu(x)U(C)^\dagger = U_\mu(x)^*, \quad (\text{D1})$$

and for quark fields as

$$\begin{aligned} U(C)q(x)U(C)^\dagger &= C\bar{q}^T, \\ U(C)\bar{q}(x)U(C)^\dagger &= q^T C^\dagger, \end{aligned} \quad (\text{D2})$$

where $C = \gamma_4\gamma_2$ satisfies $C^T = C^\dagger = C^{-1} = -C$. The Dirac-Pauli basis in which $\gamma_4 = \text{diag}(1, 1, -1, -1)$ is used throughout this section; see Ref. [93] for details. Most LQCD actions possess a C symmetry in the sense that the action is invariant under this transformation, and the existence of such a symmetry will be assumed throughout this work. For more on C transformation properties in LQCD see Refs. [74,94,95]. The quark propagator therefore transforms as

$$\begin{aligned} U(C)S(x, y)U(C)^\dagger &= CS(y, x)^T C^\dagger \\ &= C\gamma_5 S(x, y)^* \gamma_5 C^\dagger, \end{aligned} \quad (\text{D3})$$

where γ_5 -Hermiticity $S(y, x) = \gamma_5 S(x, y)^\dagger \gamma_5$ has been assumed in going to the second line. This can be further reexpressed as

$$\begin{aligned} U(C)S(x, y)U(C)^\dagger &= (iC\gamma_5)S(x, y)^*(iC\gamma_5)^\dagger \\ &= (i\gamma_1\gamma_3)S(x, y)^*(i\gamma_1\gamma_3)^\dagger, \end{aligned} \quad (\text{D4})$$

where $\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4$ has been used in the last line.

In this form, the action of C on quark propagators can be recognized as complex conjugation in conjugation with the action of a 2π rotation as follows. The spinor representation $D(\vec{\omega})$ of a rotation—an element of the double cover of the cubic group $R(\vec{\omega}) \in O_h^D$ —whose magnitude and direction are given by the vector $\vec{\omega}$ is

$$D(\vec{\omega}) = \exp\left(-\frac{1}{8} \sum_k \omega_k \epsilon_{ijk} [\gamma_i, \gamma_j]\right). \quad (\text{D5})$$

For more on cubic symmetry in LQCD see Refs. [65,93,96–103]. For the vector $\vec{\omega} = (0, \omega_2, 0)$, this becomes

$$\begin{aligned} D(\omega_2 \hat{e}_2) &= \exp\left(-\frac{\omega_2}{4} \gamma_1 \gamma_3\right) \\ &= \cos\left(-\frac{\omega_2}{4}\right) - i\gamma_1 \gamma_3 \sin\left(-\frac{\omega_2}{4}\right). \end{aligned} \quad (\text{D6})$$

Choosing $\vec{\omega}_c = (0, -2\pi, 0)$ then gives

$$D(\vec{\omega}_c) = -i\gamma_1 \gamma_3. \quad (\text{D7})$$

Therefore, define $R_{2\pi} \in O_h^D$ as the corresponding $-2\pi \hat{e}_2$ rotation operator and $U(R)$ as the Hilbert space operator whose action is given by $U(R_{2\pi})S(x, y)U(R_{2\pi})^\dagger = D(\vec{\omega}_c)S(x, y)D(\vec{\omega}_c)^\dagger$. We assume that $R_{2\pi}$ is a symmetry of the action. The combined symmetry $CR_{2\pi}$ acts on quark propagators as

$$\begin{aligned} U(CR_{2\pi})S(x, y)U(CR_{2\pi})^\dagger &= (i\gamma_1 \gamma_3)D(\vec{\omega}_c)S(x, y)^*D(\vec{\omega}_c)^\dagger(i\gamma_1 \gamma_3)^\dagger \\ &= S(x, y)^*. \end{aligned} \quad (\text{D8})$$

Since R acts trivially on the gauge field, $U(CR_{2\pi})U_\mu(x)U(CR_{2\pi})^\dagger = U_\mu(x)^*$.

The parity transformation $P \in O_h^D$ has the spinor representation $D(P) = \gamma_4$ and transforms propagators by

$$U(P)S(x, y)U(P)^\dagger = \gamma_4 S((- \vec{x}, x_4), (- \vec{y}, y_4)) \gamma_4. \quad (\text{D9})$$

The action of parity on the gauge field is

$$U(P)U_\mu(x)U(P)^\dagger = U_{-\mu}((- \vec{x}, x_4)), \quad (\text{D10})$$

where $U_{-\mu}(x) = U_\mu(x - \hat{e}_\mu)^\dagger$. It will be convenient below to denote the set of gauge fields at time t by

$$U_{(t)} = \{U_\mu(x) | x_4 = t\}, \quad (\text{D11})$$

and their parity conjugates by

$$U_{(t)}^P = \{U(P)U_\mu(x)U(P)^\dagger | x_4 = t\}. \quad (\text{D12})$$

Further denote the set of quark propagators starting and ending at time t by

$$S_{(t)} = \{S(x, y) | x_4 = y_4 = t\}, \quad (\text{D13})$$

and their parity conjugates by

$$S_{(t)}^P = \{U(P)S(x, y)U(P)^\dagger | x_4 = y_4 = t\}. \quad (\text{D14})$$

2. Correlator matrix definitions

Consider a generic correlator matrix

$$C_{ab}(t) = \langle \chi_a(t) \bar{\psi}_b(0) \rangle, \quad (\text{D15})$$

with source interpolating operator

$$\begin{aligned} \bar{\psi}_b(0) = & \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\ & \times \bar{q}(\vec{y}_1, 0) \cdots \bar{q}(\vec{y}_K, 0) q(\vec{v}_1, 0)^T \cdots q(\vec{v}_L, 0)^T \\ & \times \psi_b(U_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L), \end{aligned} \quad (\text{D16})$$

where $\psi_b(U_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L)$ is a rank $K + L$ spin-color-flavor tensor that can depend on the gauge field at the source timeslice, $U_{(0)}$. The corresponding sink interpolating operator is

$$\begin{aligned} \chi_a(t) = & \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} \\ & \times \chi_a(U_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^\dagger \\ & \times q(\vec{x}_1, 0)^T \cdots q(\vec{x}_{K'}, 0)^T \bar{q}(\vec{u}_1, 0) \cdots \bar{q}(\vec{u}_{L'}, 0), \end{aligned} \quad (\text{D17})$$

where $\chi_a(U_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^\dagger$ is a rank $K' + L'$ spin-color-flavor tensor that can depend on the gauge field at the sink timeslice, $U_{(t)}$. Note that time translation invariance can be used to trivially generalize this result to nonzero source times. The total momentum carried by quark and antiquark fields in the initial and final state is therefore given by

$$\begin{aligned} \vec{P}_q &= \sum_{i=1}^K \vec{p}_i + \sum_{j=1}^L \vec{k}_j, \\ \vec{P}'_q &= \sum_{i=1}^{K'} \vec{p}'_i + \sum_{j=1}^{L'} \vec{k}'_j. \end{aligned} \quad (\text{D18})$$

The spatial translation properties of ψ_b and χ_a encode the momenta $\vec{P}_g$ and $\vec{P}'_g$ carried by gluons in the initial and final states, respectively, and should be chosen to ensure total momentum conservation $\vec{P}_q + \vec{P}_g = \vec{P}'_q + \vec{P}'_g$.

The total number of quark fields, $F \equiv K + K'$, must equal the total number of antiquarks, $L + L'$, for all nonvanishing correlators in an $R_{2\pi}$ -symmetric theory. Wick's theorem allows the F quark/antiquark fields in $C_{ab}(t)$ to be replaced by linear combinations of terms with F quark propagators with coefficients given by products of permutation signatures. The F quark propagators can be partitioned into sets of N quark propagators connecting

source/sink sites, M quark propagators connecting source/source sites, and M' propagators connecting sink/sink sites, where $N + M + M' = F$. Different such partitions

$$\sigma \equiv \{(N, M, M') \in \mathbb{Z}_3 | N + M + M' = F\} \quad (\text{D19})$$

of F correspond to different linear combinations of Wick contractions. This allows $C_{ab}(t)$ to be expressed as

$$C_{ab}(t) = \sum_{\sigma} \text{sig}(\sigma) \sum_{p \in S_N} \text{sig}(p) C_{ab}^{\sigma p}(t), \quad (\text{D20})$$

where $\text{sig}(\sigma)$ is the signature of the permutation required to reorder the quark fields into the appropriate normal ordering for N source/sink, M source/source, and M' sink/sink quark propagators, $p \in S_N$ is an element of the permutation group of N objects, $\text{sig}(p)$ is its signature, and

$$\begin{aligned} C_{ab}^{\sigma p}(t) = & \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \\ & \times e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\ & \times \langle \Xi_a(U_{(t)}, S_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^\dagger \\ & \times \bigotimes_{i,j=1}^N S((\vec{x}_i, t), (\vec{y}_{p(j)}, 0)) \\ & \times \Psi_b(U_{(0)}, S_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L) \rangle, \end{aligned} \quad (\text{D21})$$

where the spin-color-flavor tensor indices of the quark propagator product corresponding to the $\bar{q}(\vec{y}_i)$ fields are permuted into the same order as the source spatial index labels. Here, the source wave function including source/source propagator factors is defined by

$$\begin{aligned} \Psi_b(U_{(0)}, S_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L) \\ \equiv \sum_{p \in S_M} \text{sig}(p) \bigotimes_{i=1}^M S((\vec{v}_i, 0), (\vec{y}_{N+p(i)}, 0)) \\ \times \psi_b(U_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L), \end{aligned} \quad (\text{D22})$$

and the corresponding sink wave function is defined by

$$\begin{aligned} \Xi_a(U_{(t)}, S_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^\dagger \\ \equiv \chi_a(U_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^\dagger \\ \times \sum_{p \in S_{M'}} \text{sig}(p) \bigotimes_{i=1}^{M'} S((\vec{x}_{N+i}, 0), (\vec{u}_{p(i)}, 0)). \end{aligned} \quad (\text{D23})$$

Since the coefficients appearing Eq. (D20) are ± 1 permutation signatures, it is sufficient to prove that $C_{ab}^{\sigma p}(t) \in \mathbb{R}$ for arbitrary σ, p in order to show that $C_{ab}(t) \in \mathbb{R}$.

3. Correlator matrix transformations

Applying a $CR_{2\pi}$ transformation to $C_{ab}^{\sigma p}(t)$ gives

$$\begin{aligned} U(CR_{2\pi})C_{ab}^{\sigma p}(t)U(CR_{2\pi})^\dagger &= \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\ &\times \left\langle \Xi_a(U_{(t)}^*, S_{(t)}^*, \vec{x}_1, \dots, \vec{x}_N, \vec{u}_1, \dots, \vec{u}_{M'})^\dagger \bigotimes_{i,j=1}^N S((\vec{x}_i, t), (\vec{y}_{p(j)}, 0))^* \right. \\ &\times \left. \Psi_b(U_{(0)}^*, S_{(0)}^*, \vec{y}_1, \dots, \vec{y}_N, \vec{v}_1, \dots, \vec{v}_M) \right\rangle. \end{aligned} \quad (D24)$$

Further including a P transformation gives

$$\begin{aligned} U(CPR_{2\pi})C_{ab}^{\sigma p}(t)U(CPR_{2\pi})^\dagger &= \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\ &\times \left\langle \Xi_a(U_{(t)}^{P*}, S_{(t)}^{P*}, \vec{x}_1, \dots, \vec{x}_N, \vec{u}_1, \dots, \vec{u}_{M'})^\dagger \gamma_4 \bigotimes_{i,j=1}^N S((-\vec{x}_i, t), (-\vec{y}_{p(j)}, 0))^* \right. \\ &\times \left. \gamma_4 \Psi_b(U_{(0)}^{P*}, S_{(0)}^{P*}, \vec{y}_1, \dots, \vec{y}_N, \vec{v}_1, \dots, \vec{v}_M) \right\rangle. \end{aligned} \quad (D25)$$

Changing summation variables to reverse the orientations of all momenta, e.g., $\vec{x}_i \rightarrow -\vec{x}_i$, allows this to be expressed as

$$\begin{aligned} U(CPR_{2\pi})C_{ab}^{\sigma p}(t)U(CPR_{2\pi})^\dagger &= \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \left(e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \right)^* \\ &\times \left\langle \Xi_a(U_{(t)}^*, \gamma_4 S_{(t)}^* \gamma_4, -\vec{x}_1, \dots, -\vec{x}_N, -\vec{u}_1, \dots, -\vec{u}_{M'})^\dagger \bigotimes_{i,j=1}^N \gamma_4 S((\vec{x}_i, t), (\vec{y}_{p(j)}, 0))^* \gamma_4 \right. \\ &\times \left. \Psi_b(U_{(0)}^*, \gamma_4 S_{(0)}^* \gamma_4, -\vec{y}_1, \dots, -\vec{y}_N, -\vec{v}_1, \dots, -\vec{v}_M) \right\rangle, \end{aligned} \quad (D26)$$

where $\gamma_4 S_{(t)}^* \gamma_4 \equiv \{\gamma_4 S(x, y)^* \gamma_4 | x_4 = y_4 = t\}$.

To complete the proof of correlator matrix reality, it is convenient to decompose quark propagators into positive- and negative-parity components using the parity projectors $(1 \pm \gamma_4)/2$. Propagators can be decomposed as

$$S(x, y) = \sum_{\rho, \rho' = \pm 1} S_{\rho' \rho}(x, y), \quad (D27)$$

where

$$S_{\rho' \rho}(x, y) \equiv \left(\frac{1 + \rho' \gamma_4}{2} \right) S(x, y) \left(\frac{1 + \rho \gamma_4}{2} \right). \quad (D28)$$

Wave functions can similarly be decomposed as

$$\Xi_a^T = \sum_{\vec{\rho} \in Z_2^N} \Xi_{a\vec{\rho}}^T, \quad \Psi_b = \sum_{\vec{\rho} \in Z_2^N} \Psi_{b\vec{\rho}}, \quad (D29)$$

where

$$\begin{aligned} \Xi_{a\vec{\rho}}^T &\equiv \bigotimes_{i=1}^N \left(\frac{1 + \rho_i \gamma_4}{2} \right) \Xi_a^T, \\ \Psi_{b\vec{\rho}} &\equiv \bigotimes_{i=1}^N \left(\frac{1 + \rho_i \gamma_4}{2} \right) \Psi_b. \end{aligned} \quad (D30)$$

Since $(\frac{1+\gamma_4}{2})(\frac{1-\gamma_4}{2}) = 0$, this decomposition allows the original correlation function to be expressed as

$$\begin{aligned} C_{ab}^{\sigma p}(t) &= \sum_{\vec{\rho}, \vec{\rho}' \in Z_2^N} \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \\ &\times e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\ &\times \left\langle \Xi_{a\vec{\rho}}(U_{(t)}, S_{(t)}, \vec{x}_1, \dots, \vec{x}_N, \vec{u}_1, \dots, \vec{u}_{2M'})^\dagger \right. \\ &\times \bigotimes_{i,j=1}^N S_{\rho'_i \rho_j}((\vec{x}_i, t), (\vec{y}_{p(j)}, 0)) \\ &\times \left. \Psi_{b\vec{\rho}}(U_{(0)}, S_{(0)}, \vec{y}_1, \dots, \vec{y}_N, \vec{v}_1, \dots, \vec{v}_{2M}) \right\rangle. \end{aligned} \quad (D31)$$

Applying the same decomposition to the transformed correlation function gives

$$\begin{aligned}
 U(CPR_{2\pi})C_{ab}^{\sigma P}(t)U(CPR_{2\pi})^\dagger &= \sum_{\vec{\rho}, \vec{\rho}' \in Z_2^N} \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \left(e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \right)^* \\
 &\times \left\langle \Xi_{a\vec{\rho}'}(U_{(t)}^*, \gamma_4 S_{(t)}^* \gamma_4, -\vec{x}_1, \dots, -\vec{x}_N, -\vec{u}_1, \dots, -\vec{u}_{M'})^\dagger \gamma_4 \bigotimes_{i,j=1}^N S_{\rho'_i \rho_j}((\vec{x}_i, t), (\vec{y}_{p(j)}, 0))^* \right. \\
 &\times \left. \gamma_4 \Psi_{b\vec{\rho}}(U_{(0)}^*, \gamma_4 S_{(0)}^* \gamma_4, -\vec{y}_1, \dots, -\vec{y}_N, -\vec{v}_1, \dots, -\vec{v}_M) \right\rangle. \tag{D32}
 \end{aligned}$$

The action of γ_4 on $\chi_{a\vec{\rho}'}$ and $\psi_{b\vec{\rho}}$ then simply results in sign factors,

$$\begin{aligned}
 U(CPR_{2\pi})C_{ab}^{\sigma P}(t)U(CPR_{2\pi})^\dagger &= \sum_{\vec{\rho}, \vec{\rho}' \in Z_2^N} \det(\vec{\rho}' \vec{\rho}^T) \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \left(e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \right)^* \\
 &\times \left\langle \Xi_{a\vec{\rho}'}(U_{(t)}^*, \gamma_4 S_{(t)}^* \gamma_4, -\vec{x}_1, \dots, -\vec{x}_N, -\vec{u}_1, \dots, -\vec{u}_{M'})^\dagger \bigotimes_{i,j=1}^N S_{\rho'_i \rho_j}((\vec{x}_i, t), (\vec{y}_{p(j)}, 0))^* \right. \\
 &\times \left. \Psi_{b\vec{\rho}}(U_{(0)}^*, \gamma_4 S_{(0)}^* \gamma_4, -\vec{y}_1, \dots, -\vec{y}_N, -\vec{v}_1, \dots, -\vec{v}_M) \right\rangle. \tag{D33}
 \end{aligned}$$

The same decomposition can be applied to the propagators appearing in disconnected quark loops. Defining

$$\begin{aligned}
 S_{(t)}^{\vec{\xi} \vec{\xi}} &\equiv \{S_{\xi'_i \xi_i}(x, y) | x_4 = y_4 = t, i = 1, M'\}, \\
 S_{(0)}^{\vec{\xi} \vec{\xi}} &\equiv \{S_{\xi'_i \xi_i}(x, y) | x_4 = y_4 = t, i = 1, M\}, \tag{D34}
 \end{aligned}$$

the original correlation function can be expressed as

$$\begin{aligned}
 C_{ab}^{\sigma P}(t) &= \sum_{\vec{\rho}, \vec{\rho}' \in Z_2^N} \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \sum_{\vec{x}_1, \dots, \vec{x}_N} \sum_{\vec{u}_1, \dots, \vec{u}_{M'}} \sum_{\vec{y}_1, \dots, \vec{y}_N} \sum_{\vec{v}_1, \dots, \vec{v}_M} e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \\
 &\times \left\langle \Xi_{a\vec{\rho}'}(U_{(t)}, S_{(t)}^{\vec{\xi} \vec{\xi}}, \vec{x}_1, \dots, \vec{x}_N, \vec{u}_1, \dots, \vec{u}_{2M'})^\dagger \bigotimes_{i,j=1}^N S_{\rho'_i \rho_j}((\vec{x}_i, t), (\vec{y}_{p(j)}, 0)) \Psi_{b\vec{\rho}}(U_{(0)}, S_{(0)}^{\vec{\xi} \vec{\xi}}, \vec{y}_1, \dots, \vec{y}_N, \vec{v}_1, \dots, \vec{v}_{2M}) \right\rangle. \tag{D35}
 \end{aligned}$$

Applying the same decomposition to the transformed correlation function and simplifying the action of γ_4 ,

$$\begin{aligned}
 U(CPR_{2\pi})C_{ab}^{\sigma P}(t)U(CPR_{2\pi})^\dagger &= \sum_{\vec{\rho}, \vec{\rho}' \in Z_2^N} \sum_{\xi_1, \dots, \xi_M} \sum_{\tilde{\xi}_1, \dots, \tilde{\xi}_M} \sum_{\xi'_1, \dots, \xi'_M} \sum_{\tilde{\xi}'_1, \dots, \tilde{\xi}'_M} \Pi_S \Pi'_S \sum_{\vec{x}_1, \dots, \vec{x}_{K'}} \sum_{\vec{u}_1, \dots, \vec{u}_{L'}} \sum_{\vec{y}_1, \dots, \vec{y}_K} \sum_{\vec{v}_1, \dots, \vec{v}_L} \\
 &\times \left(e^{i[\sum_i \vec{p}'_i \cdot \vec{x}_i + \sum_j \vec{k}'_j \cdot \vec{u}_j]} e^{-i[\sum_i \vec{p}_i \cdot \vec{y}_i + \sum_j \vec{k}_j \cdot \vec{v}_j]} \right)^* \\
 &\times \left\langle \Xi_{a\vec{\rho}'}(U_{(t)}^*, [S_{(t)}^{\vec{\xi} \vec{\xi}}]^*, -\vec{x}_1, \dots, -\vec{x}_N, -\vec{u}_1, \dots, -\vec{u}_{M'})^\dagger \bigotimes_{i,j=1}^N S_{\rho'_i \rho_j}((\vec{x}_i, t), (\vec{y}_{p(j)}, 0))^* \right. \\
 &\times \left. \Psi_{b\vec{\rho}}(U_{(0)}^*, [S_{(0)}^{\vec{\xi} \vec{\xi}}]^*, -\vec{y}_1, \dots, -\vec{y}_N, -\vec{v}_1, \dots, -\vec{v}_M) \right\rangle, \tag{D36}
 \end{aligned}$$

where the total parities of all source/sink quark propagator factors are given by

$$\Pi_S = \prod_{i=1}^N \rho_i \prod_{j=1}^M \xi_j \tilde{\xi}_j, \quad \Pi'_S = \prod_{i=1}^N \rho'_i \prod_{j=1}^{M'} \xi'_j \tilde{\xi}'_j. \tag{D37}$$

It is therefore sufficient for the source and sink wave function to satisfy

$$\begin{aligned} \Psi_{b\bar{p}}\left(U_{(0)}^*, \left[S_{(0)}^{\vec{\xi}\vec{\xi}}\right]^*, -\vec{y}_1, \dots, -\vec{y}_K, -\vec{v}_1, \dots, -\vec{v}_L\right) \\ = \Pi_S \Psi_{b\bar{p}}\left(U_{(0)}, S_{(0)}^{\vec{\xi}\vec{\xi}}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L\right)^*, \end{aligned} \quad (\text{D38})$$

and

$$\begin{aligned} \Xi_{a\bar{p}'}\left(U_{(t)}^*, \left[S_{(t)}^{\vec{\xi}\vec{\xi}}\right]^*, -\vec{x}_1, \dots, -\vec{x}_{K'}, -\vec{u}_1, \dots, -\vec{u}_{L'}\right)^\dagger \\ = \Pi'_S \Xi_{a\bar{p}'}\left(U_{(t)}, S_{(t)}^{\vec{\xi}\vec{\xi}}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'}\right)^T, \end{aligned} \quad (\text{D39})$$

in order to guarantee that

$$U(CPR_{2\pi})C_{ab}^{\sigma p}(t)U(CPR_{2\pi})^\dagger = C_{ab}^{\sigma p}(t)^*. \quad (\text{D40})$$

Because there are path integral changes of variables that induce $CPR_{2\pi}$ transformations and they have equal weight by $CPR_{2\pi}$ symmetry of the action, averaging over all quark-and-gluon field configurations leads to the averaging of $C_{ab}^{\sigma p}(t)$ with $U(CPR_{2\pi})C_{ab}^{\sigma p}(t)U(CPR_{2\pi})^\dagger = C_{ab}^{\sigma p}(t)^*$. The imaginary part of $C_{ab}^{\sigma p}(t)$ therefore vanishes, or in other words Eqs. (D38) and (D39) are sufficient to ensure that $C_{ab}^{\sigma p}(t) \in \mathbb{R}$ and therefore that $C_{ab}(t) \in \mathbb{R}$.

4. Correlator matrix reality conditions

It is straightforward to obtain corresponding reality conditions for the quark-field coefficient functions appearing in the interpolating operator definitions Eqs. (D16) and (D17). Define a quark-field-parity decomposition of these wave functions by

$$\chi_a^T = \sum_{\vec{\tau} \in Z_2^{K'+L'}} \chi_{a\vec{\tau}}^T, \quad \psi_b = \sum_{\vec{\tau} \in Z_2^{K+L}} \psi_{b\vec{\tau}}, \quad (\text{D41})$$

where

$$\begin{aligned} \chi_{a\vec{\tau}}^T &\equiv \bigotimes_{i=1}^{K'+L'} \left(\frac{1 + \tau_i \gamma_4}{2} \right) \chi_a^T, \\ \psi_{b\vec{\tau}} &\equiv \bigotimes_{i=1}^{K+L} \left(\frac{1 + \tau_i \gamma_4}{2} \right) \psi_b. \end{aligned} \quad (\text{D42})$$

The fact that the total fermion-field parity

$$\begin{aligned} \Pi &\equiv \Pi_S \Pi_{S'} = \prod_{i=1}^N \rho_i \rho'_i \prod_{j=1}^M \xi_j \tilde{\xi}_j \prod_{k=1}^M \xi'_k \tilde{\xi}'_k \\ &= \prod_{i=1}^{K+L} \tau_i \prod_{j=1}^{K'+L'} \tau'_j \end{aligned} \quad (\text{D43})$$

is independent of how the quark fields are partitioned into source/sink, source/source, and sink/sink propagators

implies that the pair of conditions Eq. (D38) and (D39) is equivalent to

$$\begin{aligned} \psi_{b\vec{\tau}}(U_{(0)}^*, -\vec{y}_1, \dots, -\vec{y}_K, -\vec{v}_1, \dots, -\vec{v}_L) \\ = \eta \psi_{b\vec{\tau}}(U_{(0)}, \vec{y}_1, \dots, \vec{y}_K, \vec{v}_1, \dots, \vec{v}_L)^*, \end{aligned} \quad (\text{D44})$$

and

$$\begin{aligned} \chi_{a\vec{\tau}'}(U_{(t)}^*, -\vec{x}_1, \dots, -\vec{x}_{K'}, -\vec{u}_1, \dots, -\vec{u}_{L'}) \\ = \eta' \chi_{a\vec{\tau}'}(U_{(t)}, \vec{x}_1, \dots, \vec{x}_{K'}, \vec{u}_1, \dots, \vec{u}_{L'})^*, \end{aligned} \quad (\text{D45})$$

where $\eta, \eta' \in \mathbb{C}$ are any normalization factors satisfying

$$\eta \eta' = \Pi. \quad (\text{D46})$$

The conditions Eqs. (D44)–(D46) are sufficient to ensure that $C_{ab}(t) = C_{ab}(t)^*$ is real, and will be assumed for all correlator matrices studied in this work. This means that Hermitian correlator matrices for which $\chi = \psi$ are real, symmetric matrices.

The use of standard source and sink interpolating operators leads to wave functions satisfying Eqs. (D44)–(D46). In particular, Gaussian (momentum-)smeared [55–57] sources and sinks are complex conjugated by replacing $U_\mu(x)$ with $U_\mu(x)^*$ and simultaneously replacing $\vec{x}$ with $-\vec{x}$ in momentum-smearing phases. To determine whether the spin structure of a set of interpolating operators is compatible with Eqs. (D44)–(D46), it is helpful to note that a sufficient (although not strictly necessary) set of conditions for satisfying Eqs. (D44)–(D46) is as follows:

- (i) The spin-color-flavor weights multiplying all quark and antiquark fields in the source/sink interpolating operators are real in the Dirac-Pauli basis where $\gamma_4 = \text{diag}(1, 1, -1, -1)$.
- (ii) All spatial wave functions (including iterative smearing kernels) are complex conjugated by replacing $U_\mu(x)$ with $U_\mu(x)^*$ and simultaneously replacing all spatial coordinates $\vec{x}$ with $-\vec{x}$.
- (iii) The product of source and sink interpolating operators includes an even number of odd-parity quark-field components.

The last condition will be true, for instance, for any correlation function formed from products of nucleon fields with $(q^T C \gamma_5 q) q$ spin structures times an even number of total pion fields (note those involving an odd number of pion fields vanish in an isospin- and G -parity symmetric theory).

It is also possible to obtain real correlation functions when there are an odd number of odd-parity quark-field components if the spatial wave function is replaced by minus its complex conjugate when the spatial coordinates are inverted, $\vec{x}$ with $-\vec{x}$. If either but not both of the net quark-field parity or the spatial inversion parity are odd, then a correlation function will be purely imaginary.

APPENDIX E: LEFT GEVP

As presented in Sec. II, block Lanczos applied to noisy correlator data involves distinct L/R quantities with potentially different values. The identification of GEVP as one step of block Lanczos in Sec. VI A makes clear that the same L/R distinction must apply for GEVP quantities as well, even when the correlator matrices are Hermitian or symmetric and real. The moving-pivot estimators Eq. (146) (148) correspond to one-step block Lanczos R estimators, assuming G^R is real. Inequivalent L quantities may be defined in terms of the left eigenvectors G^L that arise in the left GEVP

$$\sum_b G_{kb}^L C_{ba}(t_d) = \lambda_k^L(t_d, t_0) \sum_b G_{kb}^L C_{ba}^{-1}(t_0). \quad (\text{E1})$$

For Hermitian $C(t)$, in the absence of statistical noise, all $\lambda_k^L = \lambda_k$ are real, and $G_{ka}^L = (G_{ak}^R)^*$ such that all R and L GEVP estimators are identical. While this is known to hold asymptotically at large t_0, t_d , this is nontrivial to prove algebraically from the GEVP perspective [3]; it is obvious from the Lanczos perspective.

Statistical noise at large t may generically result in complex eigenvalues and break the L/R coincidence, per the discussion of Hermitian subspaces in Sec. V B. It is clear from the Lanczos perspective that in general there is only a single set of (complex) eigenvalues $\lambda_k^L = \lambda_k = \bar{\lambda}_k^{(1)}$, the right GEVP eigenvectors are given by Eq. (139), and the left GEVP eigenvectors are given analogously by

$$G_{ak}^L(t_d, t_0) = \sum_b (\bar{\omega}^{-1})_{k1b}^{(1)} \bar{\beta}_{1ba}^{-1}. \quad (\text{E2})$$

For complex eigenvalues, $G^L \neq G^R$ in general, but they may also differ for real eigenvalues; in either case, this leads to inequivalent L/R estimators for overlaps and matrix elements. From the Lanczos perspective, this is a clear indication of a spurious state outside the Hermitian subspace that should be discarded.

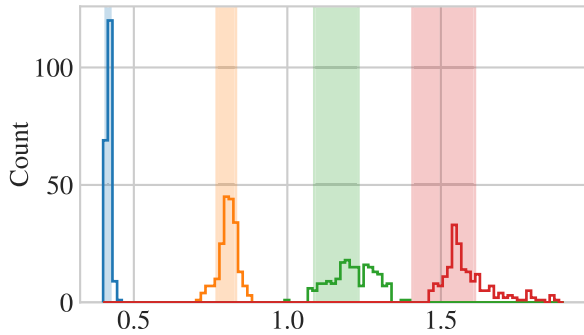


FIG. 33. Histogram of the union of the bootstrap distributions of $E_n^{(m_{\max})}$ for all nonspurious states computed using bootstrap-median estimators. Shaded regions show Gaussian estimates for the 1σ confidence interval, $E_n^{(m_{\max})} \pm \delta E_n^{(m_{\max})}$.

TABLE VII. Results of the Kolmogorov-Smirnov and Shapiro-Wilk tests for Gaussianity for the bootstrap distributions of bootstrap median block Lanczos energy estimators in Fig. 33 (bottom) and their sample mean counterparts (top).

Gaussianity test	Sample mean			
	E_0	E_1	E_2	E_3
Kolmogorov-Smirnov	0	0.001	10^{-5}	0
Shapiro-Wilk	10^{-26}	10^{-5}	10^{-17}	10^{-15}

Gaussianity test	Bootstrap median			
	E_0	E_1	E_2	E_3
Kolmogorov-Smirnov	0.350	0.191	0.501	10^{-4}
Shapiro-Wilk	0.005	0.191	0.487	10^{-7}

APPENDIX F: CONFIDENCE INTERVAL ESTIMATION

Histograms of Lanczos energy estimators in Fig. 33 clearly show qualitative signs of non-Gaussianity that are much larger for sample-mean estimators but still present for bootstrap-median estimators. These deviations from Gaussianity can be quantified using the Kolmogorov-Smirnov and Shapiro-Wilk tests. As shown in Table VII, the sample-mean estimators are highly non-Gaussian, while the bootstrap-median estimators are much closer to Gaussian but still display clear signs of non-Gaussianity.

Empirical bootstrap confidence intervals provide a standard estimator for computing bootstrap uncertainties that is more outlier robust than the bootstrap variance and applicable to non-Gaussian distributions [61,63]. The empirical bootstrap formula providing $s\sigma$ confidence intervals $\delta_s X$ for any quantity X with bootstrap samples $X^{(b)}$ is²⁸

$$\delta_s X = \frac{1}{2} \left[\text{Quantile} \left(X^{(b)}, \frac{1 + \text{erf}(s/\sqrt{2})}{2} \right) - \text{Quantile} \left(X^{(b)}, \frac{1 - \text{erf}(s/\sqrt{2})}{2} \right) \right]. \quad (\text{F1})$$

Results for $1 - 3\sigma$ confidence intervals can be computed using this definition; with $N_{\text{boot}} = 200$ the 3σ interval may be significantly impacted by finite N_{boot} effects and larger confidence intervals are likely to be unreliable.

These empirical bootstrap confidence interval results for $\delta_s X$ can be compared with the expectation of Gaussian uncertainties: $\delta_s X \approx s\delta X$ where δX is defined from the bootstrap variance in Eq. (124). As shown in Fig. 34 for $E_n^{(m_{\max})}$, the approximate equality of these confidence

²⁸Empirical and percentile bootstrap constructions give identical expressions for the width of the symmetric confidence interval [63].

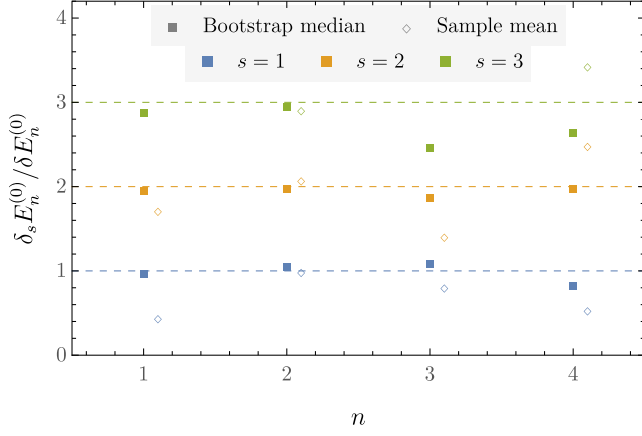


FIG. 34. Comparisons of $s\sigma$ empirical bootstrap confidence intervals $\delta_s E_n^{(m)}$ with $s \in \{1, 2, 3\}$ with the Gaussian approximations $s\delta E_n^{(m)}$ involving the bootstrap variance. Open (closed) markers show results obtained using idiomatic bootstrap uncertainties for sample-mean (bootstrap-median) estimators.

interval estimates holds within 6% for $n \in \{0, 1\}$ states for $s\sigma$ confidence intervals with $s \in \{1, 2, 3\}$. These deviations may be due in part to statistical fluctuations of the bootstrap estimators, which we do not quantify here but are proportional to $1/\sqrt{N_{\text{boot}}}$. For $n \in \{2, 3\}$ the agreement is less good and deviations of up to 20% are seen. Identification of $s\sigma$ confidence intervals with $s\delta E_n^{(m_{\text{max}})}$ therefore includes systematic uncertainties of these sizes.

APPENDIX G: INEQUIVALENCE OF BLOCK LANCZOS AND GEVM/PGEVM

Reference [75] asserts that block Lanczos is equivalent to “the block PGEVM approach introduced in Ref. [6].” That work discusses correlator matrix analysis via GEVP solutions, which they refer to as applications of the generalized eigenvalue method (GEVM). The only discussion of generalizing PGEVM to correlator matrices present in Ref. [6] (in which the word “block” does not appear) is “There is one straightforward way to combine GEVM and PGEVM: we noted already above that the principal correlators of the GEVM are again a sum of exponentials, and, hence, the PGEVM can be applied to them. This means a sequential application of first the GEVM with a correlator matrix of size n_0 to determine principal correlators λ_k and then of the PGEVM with size n_1 and the λ_k ’s as input, which we denote as GEVM/PGEVM.”

It was first pointed out in Ref. [16] that “the Ritz values are numerically identical to the solutions of the Prony generalized eigenvalue method (PGEVM) with $n = m$, $\tau_0 = 0$, and $\Delta = \delta t = 1$, which by construction coincide with the roots from Prony’s method.” Therefore, the GEVM/PGEVM method is equivalent to applying scalar Lanczos (without spurious state filtering) to the principal

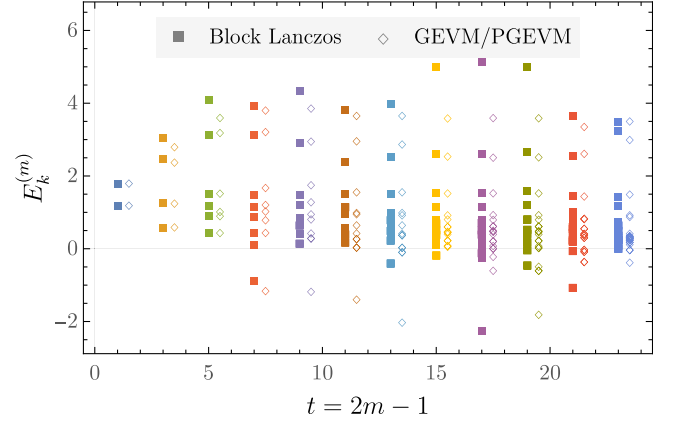


FIG. 35. Comparison of block Lanczos with the GEVM/PGEVM method [6] applied to the 2×2 proton correlator matrix discussed in the main text. Both methods coincide with GEVP for $m = 1$. Sample-mean estimators are shown; no spurious state filtering is performed.

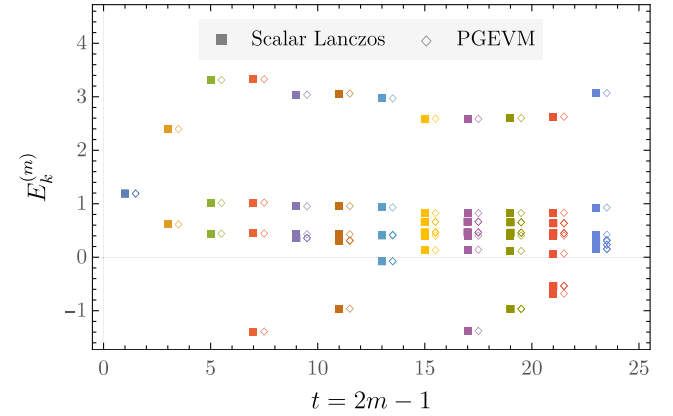


FIG. 36. Comparison of scalar Lanczos with the PGEVM method [6] applied to the WW diagonal element of the proton correlator matrix discussed in the main text. Sample-mean estimators are shown; no spurious state filtering is performed.

correlators λ_k in Ref. [6], which according to the GEVP definition in Eq. (6) of that work are simply the GEVP eigenvalues that are also denoted $\lambda_k(t, t_0)$ here.

Applying the GEVM/PGEVM method for m steps to a rank r correlator matrix is therefore equivalent to separate Lanczos analysis of r Krylov spaces, each of dimension m , generated by individually analyzing each λ_k . For $m = 1$, both methods involve analyzing the same r -dimensional Krylov space as GEVP, and both are therefore equivalent to GEVP and hence one another for $m = 1$. For $m > 1$, the rm dimensional Krylov space analysis provided by block Lanczos is not equivalent to (and is strictly more optimal than) the Lanczos analyses of r individual m -dimensional Krylov spaces provided by GEVM/PGEVM. This proves that GEVM/PGEVM cannot be equivalent to block Lanczos.

To verify this result numerically [35], we show results for block Lanczos and GEVM/PGEVM applied to the central values of the 2×2 proton correlator matrix discussed in Secs. V and VI in Fig. 35. The results are

not identical for $m > 1$. This can be contrasted with the identical results for scalar Lanczos and PGEVM applied to the central values of a single diagonal correlator matrix entry shown in Fig. 36.

-
- [1] G. Fox, R. Gupta, O. Martin, and S. Otto, Monte Carlo estimates of the mass gap of the O(2) and O(3) spin models in $(1+1)$ -dimensions, *Nucl. Phys.* **B205**, 188 (1982).
 - [2] C. Michael and I. Teasdale, Extracting glueball masses from lattice QCD, *Nucl. Phys.* **B215**, 433 (1983).
 - [3] M. Lüscher and U. Wolff, How to calculate the elastic scattering matrix in two-dimensional quantum field theories by numerical simulation, *Nucl. Phys.* **B339**, 222 (1990).
 - [4] 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, *J. High Energy Phys.* **04** (2009) 094.
 - [5] G. T. Fleming, Beyond generalized eigenvalues in lattice quantum field theory, [arXiv:2309.05111](https://arxiv.org/abs/2309.05111).
 - [6] M. Fischer, B. Kostrzewa, J. Ostmeyer, K. Ottnad, M. Ueding, and C. Urbach, On the generalised eigenvalue method and its relation to Prony and generalised pencil of function methods, *Eur. Phys. J. A* **56**, 206 (2020).
 - [7] G. C. F. M. Prony, Essai Experimental et Analytique, *J. l'école Polytech. Paris* **1**, 24 (1795).
 - [8] G. T. Fleming, What can lattice QCD theorists learn from NMR spectroscopists?, [arXiv:hep-lat/0403023](https://arxiv.org/abs/hep-lat/0403023).
 - [9] H.-W. Lin and S. D. Cohen, Lattice QCD beyond ground states, [arXiv:0709.1902](https://arxiv.org/abs/0709.1902).
 - [10] G. T. Fleming, S. D. Cohen, H.-W. Lin, and V. Pereyra, Excited-state effective masses in lattice QCD, *Phys. Rev. D* **80**, 074506 (2009).
 - [11] C. Aubin and K. Orginos, A new approach for Delta form factors, *AIP Conf. Proc.* **1374**, 621 (2011).
 - [12] C. Aubin and K. Orginos, An improved method for extracting matrix elements from lattice three-point functions, *Proc. Sci. LATTICE2011* (2011) 148.
 - [13] J. R. Green, J. W. Negele, A. V. Pochinsky, S. N. Syritsyn, M. Engelhardt, and S. Krieg, Nucleon electromagnetic form factors from lattice QCD using a nearly physical pion mass, *Phys. Rev. D* **90**, 074507 (2014).
 - [14] K. Ottnad, T. Harris, H. Meyer, G. von Hippel, J. Wilhelm, and H. Wittig, Nucleon average quark momentum fraction with $N_f = 2 + 1$ Wilson fermions, *EPJ Web Conf.* **175**, 06026 (2018).
 - [15] K. Ottnad, Excited states in nucleon structure calculations, *Eur. Phys. J. A* **57**, 50 (2021).
 - [16] M. L. Wagman, Lanczos, the transfer matrix, and the signal-to-noise problem, *Phys. Rev. Lett.* **134**, 241901 (2025).
 - [17] D. C. Hackett and M. L. Wagman, Lanczos for lattice QCD matrix elements, [arXiv:2407.21777](https://arxiv.org/abs/2407.21777).
 - [18] G. H. Golub, Some uses of the Lanczos algorithm in numerical algebra, in *Topics in numerical analysis: Proceedings of the Royal Irish Academy Conference on Numerical Analysis, Dublin, 1972*, edited by J. J. H. Miller (Academic Press, New York, 1973), pp. 173–184.
 - [19] J. Cullum and W. E. Donath, A block lanczos algorithm for computing the q algebraically largest eigenvalues and a corresponding eigenspace of large, sparse, real symmetric matrices, in *1974 IEEE Conference on Decision and Control including the 13th Symposium on Adaptive Processes* (IEEE Press, New York, 1974), pp. 505–509.
 - [20] G. Golub and R. Underwood, The block lanczos method for computing eigenvalues, in *Mathematical Software*, edited by J. R. Rice (Academic Press, New York, 1977), pp. 361–377.
 - [21] Y. Saad, On the rates of convergence of the Lanczos and the block-Lanczos methods, *SIAM J. Numer. Anal.* **17**, 687 (1980).
 - [22] B. Parlett, *The Symmetric Eigenvalue Problem*, Classics in Applied Mathematics (Society for Industrial and Applied Mathematics, Philadelphia, 1980).
 - [23] H. M. Kim and R. R. Craig Jr., Structural dynamics analysis using an unsymmetric block lanczos algorithm, *Int. J. Numer. Methods Eng.* **26**, 2305 (1988).
 - [24] H. M. Kim and R. R. Craig Jr., Computational enhancement of an unsymmetric block Lanczos algorithm, *Int. J. Numer. Methods Eng.* **30**, 1083 (1990).
 - [25] Z. Bai, D. Day, and Q. Ye, ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems, *SIAM J. Matrix Anal. Appl.* **20**, 1060 (1999).
 - [26] A. M. Abdel-Rehim, A. Stathopoulos, and K. Orginos, Extending the eigCG algorithm to nonsymmetric Lanczos for linear systems with multiple right-hand sides, *Num. Linear Algebra Appl.* **21**, 473 (2014).
 - [27] M. Lüscher and P. Weisz, Definition and general properties of the transfer matrix in continuum limit improved lattice gauge theories, *Nucl. Phys.* **B240**, 349 (1984).
 - [28] J. B. Kogut and L. Susskind, Hamiltonian formulation of Wilson's lattice gauge theories, *Phys. Rev. D* **11**, 395 (1975).
 - [29] Y. Saad, The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems, *SIAM J. Numer. Anal.* **19**, 485 (1982).
 - [30] R. Abbott, G. T. Fleming, D. C. Hackett, D. A. Pefkou, and M. L. Wagman, Filtered Rayleigh-Ritz is all you need, [arXiv:2503.17357](https://arxiv.org/abs/2503.17357).
 - [31] T. mpmath development team, MPMATH: A Python Library for Arbitrary-Precision Floating-Point Arithmetic (version 1.3.0) (2023), <http://mpmath.org/>.
 - [32] B. Parlett, Misconvergence in the Lanczos algorithm, in *Reliable Numerical Computation* (Oxford University Press, New York, 1990).

- [33] A. B. J. Kuijlaars, Which eigenvalues are found by the Lanczos method?, *SIAM J. Matrix Anal. Appl.* **22**, 306 (2000).
- [34] J. Garza-Vargas and A. Kulkarni, The Lanczos algorithm under few iterations: Concentration and location of the output, *SIAM J. Matrix Anal. Appl.* **41**, 1312 (2020).
- [35] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/fp74-q35q> for data used to generate the noiseless matrix element example.
- [36] L. Maiani, G. Martinelli, M. L. Paciello, and B. Taglienti, Scalar densities and baryon mass differences in lattice QCD with Wilson fermions, *Nucl. Phys.* **B293**, 420 (1987).
- [37] S. J. Dong, K. F. Liu, and A. G. Williams, Lattice calculation of the strangeness magnetic moment of the nucleon, *Phys. Rev. D* **58**, 074504 (1998).
- [38] S. Capitani, M. Della Morte, G. von Hippel, B. Jager, A. Jüttner, B. Knippschild, H. B. Meyer, and H. Wittig, The nucleon axial charge from lattice QCD with controlled errors, *Phys. Rev. D* **86**, 074502 (2012).
- [39] D. C. Hackett, P. R. Oare, D. A. Pefkou, and P. E. Shanahan, Gravitational form factors of the pion from lattice QCD, *Phys. Rev. D* **108**, 114504 (2023).
- [40] R. A. Briceño, M. T. Hansen, and A. Walker-Loud, Multi-channel $1 \rightarrow 2$ transition amplitudes in a finite volume, *Phys. Rev. D* **91**, 034501 (2015).
- [41] L. Leskovec, Electroweak transitions involving resonances, *Proc. Sci. LATTICE2023* (2024) 119 [arXiv:2401.02495].
- [42] S. Kaniel, Estimates for some computational techniques in linear algebra, *Math. Comput.* **20**, 369 (1966).
- [43] C. C. Paige, The computation of eigenvalues and eigenvectors of very large sparse matrices, Ph.D. thesis, London University, London, 1971.
- [44] B. N. Parlett and D. S. Scott, The Lanczos algorithm with selective orthogonalization, *Math. Comput.* **33**, 217 (1979).
- [45] B. N. Parlett, Do we fully understand the symmetric Lanczos algorithm yet (1995), <https://apps.dtic.mil/sti/citations/tr/ADA289614>.
- [46] J. Cullum and R. A. Willoughby, Computing eigenvalues of very large symmetric matrices—an implementation of a Lanczos algorithm with no reorthogonalization, *J. Comput. Phys.* **44**, 329 (1981).
- [47] J. K. Cullum and R. A. Willoughby, Lanczos procedures, in *Lanczos Algorithms for Large Symmetric Eigenvalue Computations Vol. I Theory* (Birkhäuser Boston, Boston, MA, 1985), pp. 92–163.
- [48] S. Park, R. Gupta, B. Yoon, S. Mondal, T. Bhattacharya, Y.-C. Jang, B. Joó, and F. Winter (Nucleon Matrix Elements (NME) Collaboration), Precision nucleon charges and form factors using $(2+1)$ -flavor lattice QCD, *Phys. Rev. D* **105**, 054505 (2022).
- [49] B. Yoon *et al.*, Isovector charges of the nucleon from $2+1$ -flavor QCD with clover fermions, *Phys. Rev. D* **95**, 074508 (2017).
- [50] S. Mondal, R. Gupta, S. Park, B. Yoon, T. Bhattacharya, B. Joó, and F. Winter (Nucleon Matrix Elements (NME) Collaboration), Nucleon momentum fraction, helicity and transversity from $2+1$ -flavor lattice QCD, *J. High Energy Phys.* **04** (2020) 004.
- [51] R. Edwards, R. Gupta, N. Joó, K. Orginos, D. Richards, F. Winter, and B. Yoon, U.S. $2+1$ flavor clover lattice generation program (unpublished).
- [52] M. Lüscher and P. Weisz, On-shell improved lattice gauge theories, *Commun. Math. Phys.* **98**, 433 (1985); **98**, 433(E) (1985).
- [53] B. Sheikholeslami and R. Wohlert, Improved continuum limit lattice action for QCD with Wilson fermions, *Nucl. Phys.* **B259**, 572 (1985).
- [54] C. Morningstar and M. J. Peardon, Analytic smearing of SU(3) link variables in lattice QCD, *Phys. Rev. D* **69**, 054501 (2004).
- [55] S. Gusken, U. Low, K. H. Mutter, R. Sommer, A. Patel, and K. Schilling, Nonsinglet axial vector couplings of the baryon octet in lattice QCD, *Phys. Lett. B* **227**, 266 (1989).
- [56] S. Gusken, A study of smearing techniques for hadron correlation functions, *Nucl. Phys. B, Proc. Suppl.* **17**, 361 (1990).
- [57] G. S. Bali, B. Lang, B. U. Musch, and A. Schäfer, Novel quark smearing for hadrons with high momenta in lattice QCD, *Phys. Rev. D* **93**, 094515 (2016).
- [58] D. C. Hackett, D. A. Pefkou, and P. E. Shanahan, Gravitational form factors of the proton from lattice QCD, *Phys. Rev. Lett.* **132**, 251904 (2024).
- [59] P. B. Denton, S. J. Parke, T. Tao, and X. Zhang, Eigenvectors from Eigenvalues: A survey of a basic identity in linear algebra, *Bull. Am. Math. Soc.* **59**, 31 (2022).
- [60] C. Jacobi, De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas tranformandis quae solis quadratis variabilium constant; una cum variis theorematis de tranformatione et determinatione integralium multiplicium., *J. Reine Angew. Math.* **1834**, 1 (1834).
- [61] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*, Regional Conference Series in Applied Mathematics Vol. 38 (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1982).
- [62] T. J. DiCiccio and B. Efron, Bootstrap confidence intervals, *Stat. Sci.* **11**, 189 (1996).
- [63] A. C. Davison and D. V. Hinkley, The basic bootstraps, in *Bootstrap Methods and their Application*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, England, 1997), pp. 11–69.
- [64] P. Young, Everything you wanted to know about data analysis and fitting but were afraid to ask, [arXiv:1210.3781](https://arxiv.org/abs/1210.3781).
- [65] S. Amarasinghe, R. Baghdadi, Z. Davoudi, W. Detmold, M. Illa, A. Parreño, A. V. Pochinsky, P. E. Shanahan, and M. L. Wagman, Variational study of two-nucleon systems with lattice QCD, *Phys. Rev. D* **107**, 094508 (2023).
- [66] W. Detmold, M. Illa, W. I. Jay, A. Parreño, R. J. Perry, P. E. Shanahan, and M. L. Wagman, Constraints on the finite volume two-nucleon spectrum at $m_\pi \approx 806$ MeV, *Phys. Rev. D* **111**, 114501 (2025).
- [67] J. Bulava, R. G. Edwards, E. Engelson, B. Joo, H.-W. Lin, C. Morningstar, D. G. Richards, and S. J. Wallace,

- Nucleon, Δ and Ω excited states in $N_f = 2 + 1$ lattice QCD, *Phys. Rev. D* **82**, 014507 (2010).
- [68] J. Bulava, B. Fahy, B. Hörz, K. J. Juge, C. Morningstar, and C. H. Wong, $I = 1$ and $I = 2$ $\pi - \pi$ scattering phase shifts from $N_f = 2 + 1$ lattice QCD, *Nucl. Phys.* **B910**, 842 (2016).
- [69] J. Dragos, R. Horsley, W. Kamleh, D. B. Leinweber, Y. Nakamura, P. E. L. Rakow, G. Schierholz, R. D. Young, and J. M. Zanotti, Nucleon matrix elements using the variational method in lattice QCD, *Phys. Rev. D* **94**, 074505 (2016).
- [70] S. R. Beane *et al.* (NPLQCD, QCDSF Collaborations), Charged multihadron systems in lattice QCD + QED, *Phys. Rev. D* **103**, 054504 (2021).
- [71] C. Stein, Inadmissibility of the usual estimator for the mean of a multivariate normal distribution, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics* (University of California Press, Berkeley, CA, 1956), pp. 197–206.
- [72] O. Ledoit and M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivariate Anal.* **88**, 365 (2004).
- [73] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control* **19**, 716 (1974).
- [74] E. Rinaldi, S. Syritsyn, M. L. Wagman, M. I. Buchoff, C. Schroeder, and J. Wasem, Lattice QCD determination of neutron-antineutron matrix elements with physical quark masses, *Phys. Rev. D* **99**, 074510 (2019).
- [75] J. Ostmeyer, A. Sen, and C. Urbach, On the equivalence of Prony and Lanczos methods for Euclidean correlation functions, *Eur. Phys. J. A* **61**, 26 (2025).
- [76] D. Chakraborty, D. Sood, A. Radhakrishnan, and N. Mathur, Estimating energy levels from lattice QCD correlation functions using a transfer matrix formalism, *arXiv*: 2412.01900.
- [77] R. G. Edwards and B. Joó (SciDAC, LHPC, UKQCD Collaborations), The Chroma software system for lattice QCD, *Nucl. Phys. B, Proc. Suppl.* **140**, 832 (2005).
- [78] M. A. Clark, R. Babich, K. Barros, R. C. Brower, and C. Rebbi (QUDA Collaboration), Solving lattice QCD systems of equations using mixed precision solvers on GPUs, *Comput. Phys. Commun.* **181**, 1517 (2010).
- [79] R. Babich, M. A. Clark, B. Joó, G. Shi, R. C. Brower, and S. Gottlieb (QUDA Collaboration), Scaling lattice QCD beyond 100 GPUs, *arXiv*:1109.2935.
- [80] M. A. Clark, B. Joó, A. Strelchenko, M. Cheng, A. Gambhir, and R. C. Brower (QUDA Collaboration), Accelerating lattice QCD multigrid on GPUs using fine-grained parallelization, *arXiv*:1612.07873.
- [81] F. Winter, M. Clark, R. Edwards, and B. Joó, A framework for lattice QCD calculations on GPUs, in *2014 IEEE 28th International Parallel and Distributed Processing Symposium* (2014), pp. 1073–1082.
- [82] E. Romero and C. Kallidonis, Chromaform, <https://github.com/eromero-vc/chromaform>.
- [83] C. R. Harris *et al.*, Array programming with NumPy, *Nature (London)* **585**, 357 (2020).
- [84] P. Virtanen *et al.* (SciPy 1.0 Contributors), SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nat. Methods* **17**, 261 (2020).
- [85] J. Reback *et al.*, pandas-dev/pandas: PANDAS 1.0.3 (2020).
- [86] Wes McKinney, Data structures for statistical computing in Python, in *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman (2010), pp. 56–61.
- [87] G. P. Lepage, LSQFIT v. 11.7, [10.5281/zenodo.4037174](https://doi.org/10.5281/zenodo.4037174) (2020).
- [88] G. P. Lepage, GVAR v. 11.9.1 [10.5281/zenodo.4290884](https://doi.org/10.5281/zenodo.4290884) (2020).
- [89] Wolfram Research Inc., Mathematica, Version 14.0, <https://www.wolfram.com/mathematica>.
- [90] J. D. Hunter, Matplotlib: A 2d graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [91] M. L. Waskom, SEABORN: Statistical data visualization, *J. Open Source Software* **6**, 3021 (2021).
- [92] M. L. Wagman and M. J. Savage, Statistics of baryon correlation functions in lattice QCD, *Phys. Rev. D* **96**, 114508 (2017).
- [93] W. Detmold, W. I. Jay, G. Kanwar, P. E. Shanahan, and M. L. Wagman, Multiparticle interpolating operators in quantum field theories with cubic symmetry, *Phys. Rev. D* **109**, 094516 (2024).
- [94] D. B. Leinweber, R. M. Woloshyn, and T. Draper, Electromagnetic structure of octet baryons, *Phys. Rev. D* **43**, 1659 (1991).
- [95] W. Melnitchouk, S. O. Bilson-Thompson, F. D. R. Bonnet, J. N. Hedditch, F. X. Lee, D. B. Leinweber, A. G. Williams, J. M. Zanotti, and J. B. Zhang, Excited baryons in lattice QCD, *Phys. Rev. D* **67**, 114506 (2003).
- [96] J. E. Mandula, G. Zweig, and J. Govaerts, Covariant lattice glueball fields, *Nucl. Phys.* **B228**, 109 (1983).
- [97] J. E. Mandula, G. Zweig, and J. Govaerts, Representations of the rotation reflection symmetry group of the four-dimensional cubic lattice, *Nucl. Phys.* **B228**, 91 (1983).
- [98] S. Basak, R. G. Edwards, G. T. Fleming, U. M. Heller, C. Morningstar, D. Richards, I. Sato, and S. Wallace, Group-theoretical construction of extended baryon operators in lattice QCD, *Phys. Rev. D* **72**, 094506 (2005).
- [99] S. Basak, R. Edwards, G. T. Fleming, U. M. Heller, C. Morningstar, D. Richards, I. Sato, and S. J. Wallace (Lattice Hadron Physics (LHPC) Collaboration), Clebsch-Gordan construction of lattice interpolating fields for excited baryons, *Phys. Rev. D* **72**, 074501 (2005).
- [100] T. Luu and M. J. Savage, Extracting scattering phase-shifts in higher partial-waves from lattice QCD calculations, *Phys. Rev. D* **83**, 114508 (2011).
- [101] C. E. Thomas, R. G. Edwards, and J. J. Dudek, Helicity operators for mesons in flight on the lattice, *Phys. Rev. D* **85**, 014507 (2012).
- [102] C. Morningstar, J. Bulava, B. Fahy, J. Foley, Y. C. Jhang, K. J. Juge, D. Lenkner, and C. H. Wong, Extended hadron and two-hadron operators of definite momentum for spectrum calculations in lattice QCD, *Phys. Rev. D* **88**, 014511 (2013).
- [103] S. Prelovsek, U. Skerbis, and C. B. Lang, Lattice operators for scattering of particles with spin, *J. High Energy Phys.* **01** (2017) 129.