Block Lanczos and signal-to-noise

MW, arXiv:2406.20009

Hackett, MW, arXiv:2407.21777

Hackett, MW, arXiv:2412.04444

Michael Wagman

NGT Algorithm Workshop — Lattice QCD at the large scale on exascale computing facilities

CERN

December 9, 2024

Transfer-matrix eigenvalues

Lattice theories do not have continuous time translation symmetry defining Hamiltonian

$$\mathcal{O}(t) = e^{-Ht}\mathcal{O}e^{Ht}$$

Discrete time translation symmetry enables definition of transfer matrix T

$$\mathcal{O}(ka) = T^k\mathcal{O}(T^{-1})^k$$

Energy spectrum = - ln (spectrum of eigenvalues of T)

$$T\big|n\big\rangle = \big|n\big\rangle\lambda_n \qquad E_n = -\ln\lambda_n$$

Correlation functions are matrix elements of powers of T

$$C(t) \equiv \langle \psi(t)\psi^\dagger(0)\rangle = \Big\langle \psi \Big| T^{t/a} \Big| \psi \Big\rangle + \ldots$$

Transfer-matrix eigenstates

Arbitrary LQCD states can be expressed in transfer matrix (energy) eigenstate basis:

$$|\psi\rangle = \sum_{n=0}^{\infty} |n\rangle\langle n|\psi\rangle \equiv \sum_{n=0}^{\infty} |n\rangle Z_n$$

Even for a single gauge-link (~collection of rigid rotors)

- The transfer matrix acts simply in this basis

$$T|\psi\rangle = \sum_n T|n\rangle Z_n = \sum_n \lambda_n |n\rangle Z_n = \sum_n e^{-aE_n} |n\rangle Z_n$$

Transfer-matrix eigenstates

Arbitrary LQCD states can be expressed in transfer matrix (energy) eigenstate basis:

$$|\psi\rangle = \sum_{n=0}^{\infty} |n\rangle\langle n|\psi\rangle \equiv \sum_{n=0}^{\infty} |n\rangle Z_n$$

Even for a single gauge-link (~collection of rigid rotors)

- The transfer matrix acts simply in this basis

$$T|\psi\rangle = \sum_n T|n\rangle Z_n = \sum_n \lambda_n |n\rangle Z_n = \sum_n e^{-aE_n}|n\rangle Z_n$$

- Repeatedly acting on any vector with a matrix filters out the component proportional to the eigenvector with the largest eigenvalue (= the ground state)

$$T^k|\psi\rangle = \sum_n T^k|n\rangle Z_n = \sum_n \lambda_n^k |n\rangle Z_n = \sum_n e^{-kaE_n}|n\rangle Z_n$$
$$= e^{-kaE_0}|0\rangle Z_0 + O\left(e^{-ka(E_1-E_0)}\right)$$

Backbone of the power-iteration algorithm for finding largest eigenvalue of a matrix:

von Mises and Pollaczek-Geiringer (1929)

The power-iteration algorithm

Start with an arbitrary normalized initial state: $|b_1\rangle = |\psi\rangle / |\psi|$

Iteration step: $|p_{k+1}\rangle = T|b_k\rangle$ $\quad |b_{k+1}\rangle = |p_{k+1}\rangle / |p_{k+1}|$

Convergence: $|b_k\rangle \propto T^{k-1}|\psi\rangle = e^{-(k-1)aE_0}|\psi\rangle Z_0 + O\left(e^{-k\delta}\right)$

The power-iteration algorithm

Start with an arbitrary normalized initial state: $\big|b_1\big\rangle = \big|\psi\big\rangle / |\psi|$

Iteration step: $\big|p_{k+1}\big\rangle = T\big|b_k\big\rangle$ $\big|b_{k+1}\big\rangle = \big|p_{k+1}\big\rangle / |p_{k+1}|$

Convergence: $\big|b_k\big\rangle \propto T^{k-1}\big|\psi\big\rangle = e^{-(k-1)aE_0}\big|\psi\big\rangle Z_0 + O\left(e^{-k\delta}\right)$

Energies from power-iteration eigenvalues:

$$-\ln\big\langle b_k\big|T\big|b_k\big\rangle = -\ln\left[\frac{\big\langle\psi\big|T^{2k-1}\big|\psi\big\rangle}{\big\langle\psi\big|T^{2k-2}\big|\psi\big\rangle}\right] = aE_0 + O\left(e^{-k\delta}\right)$$

$$= -\ln\left[\frac{C((2k-1)a)}{C((2k-2)a)}\right] = aE^{\mathrm{eff}}(t/a = 2k-1)$$

Standard effective mass = “apply power-iteration algorithm to the transfer matrix”

Lanczos = Krylov + Rayleigh-Ritz

Start with an arbitrary normalized initial state: $|v_1\rangle = |\psi\rangle / |\psi| = |\psi\rangle / \sqrt{C(0)}$

Iteration step: $|v_{j+1}\rangle \beta_{j+1} = (T - \alpha_j)|v_j\rangle - \beta_j |v_{j-1}\rangle$

Where $\alpha_j = \langle v_j | T | v_j \rangle \quad \beta_j = \langle v_{j-1} | T | v_j \rangle$

Lanczos (1950) See Parlett, "The Symmetric Eigenvalue Problem" (1980)

Novel features not present in power iteration

- Lanczos vectors form orthonormal basis for Krylov space

$$\mathcal{K}^{(m)} = \mathrm{span}\{|v_1\rangle, |v_2\rangle, \ldots, |v_m\rangle\}$$

$$\langle v_i | v_j \rangle = \delta_{ij}$$

- Krylov-space approximation to T directly computable

$$T_{ij}^{(m)} = \langle v_i | T | v_j \rangle = \delta_{ij}\alpha_j + \delta_{i(j-1)}\beta_j + \delta_{i(j+1)}\beta_{j+1}$$

Krylov space ~ span of data ~ computationally accessible part of Hilbert space

Optimal estimators given fixed data

Krylov-space approximation to T directly computed in Lanczos algorithm

- It's eigenvalues provide "best" Krylov-space approximations to T eigenvalues

$$T_{ij}^{(m)} = \langle v_i | T | v_j \rangle = \begin{pmatrix} \alpha_1 & \beta_2 & & & & 0 \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{m-1} & \\ & & & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & & & & \beta_m & \alpha_m \end{pmatrix}_{ij}$$

Diagonalize the Krylov-space transfer matrix:

$$T_{ij}^{(m)} = \sum_k \omega_{ik}^{(m)} \lambda_k^{(m)} (\omega^{-1})_{kj}^{(m)}$$

"Ritz vectors" = corresponding approximate eigenstates

$$|y_k^{(m)}\rangle = \sum_j |v_j\rangle \omega_{jk}^{(m)}$$

"Ritz values" = optimal Krylov-space approximation to T eigenvalues

$$\lambda_k^{(m)} = \langle y_k^{(m)} | T | y_k^{(m)} \rangle$$

Lanczos without Lanczos vectors

Problem: In LQCD, we don't have direct access to infinite-dimensional Hilbert space vectors

Lanczos without Lanczos vectors

Problem: In LQCD, we don't have direct access to infinite-dimensional Hilbert space vectors

Solution: Compute the matrix elements $T_{ij}^{(m)}$ directly from correlation functions via recursion relations:

MW, arXiv:2406.20009

$$\alpha_1 = \langle v_1 | T | v_1 \rangle = \frac{C(1a)}{C(0)} \qquad \beta_1 = 0$$

Recursive Lanczos iteration:

$$A_j^k = \langle v_j | T^k | v_j \rangle \quad B_j^k = \langle v_{j-1} | T^k | v_j \rangle$$

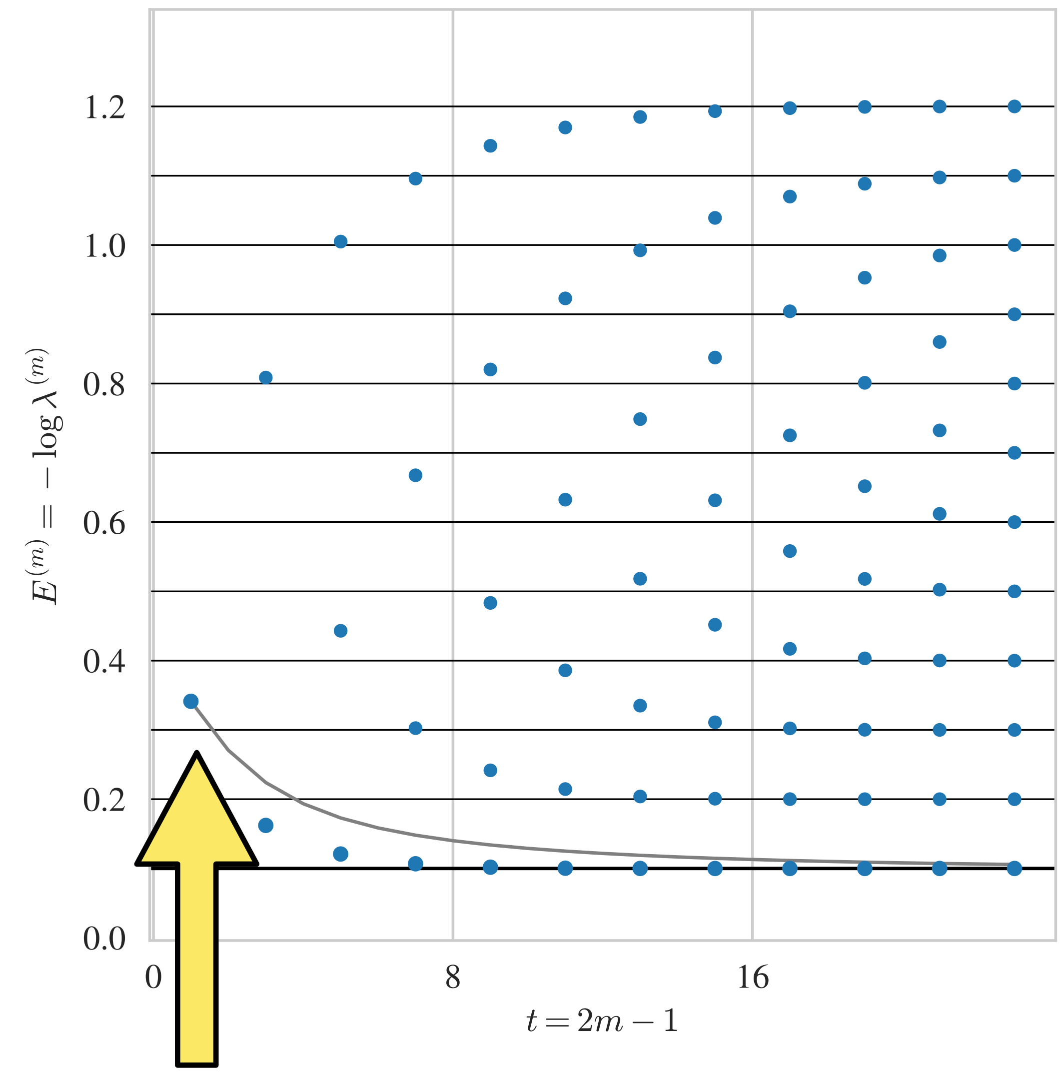
$$\beta_{j+1} = \sqrt{A_j^2 - \alpha_j^2 - \beta_j^2}$$

$$B_{j+1}^k = \frac{1}{\beta_{j+1}} [A_j^{k+1} - \alpha_j A_j^k - \beta_j B_j^k]$$

$\cdots$

Ritz values reproduce spectrum of 12-state toy model exactly after 12 steps:

$$C(t) = \sum_{n=1}^{12} \frac{1}{2(0.1n)} e^{-0.1nt}$$



Lanczos equals power iteration after $m = 1$ step, converges faster for $m > 1$

Correlator matrices

State-of-the-art LQCD studies often use multiple interpolators

$t \in \{0, \ldots, N_t - 1\}$

$N_t = $ length of time series

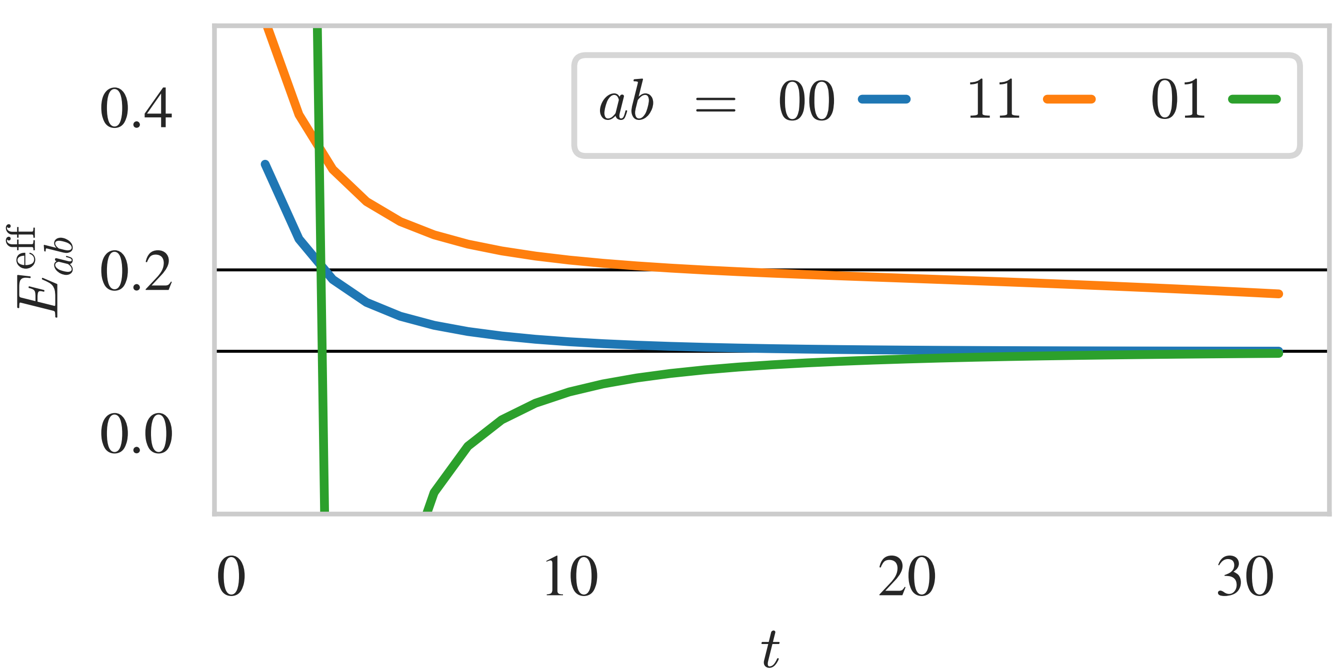
$$C_{ab}(t) = \langle \psi_a(t) \overline{\psi}_b(0) \rangle$$

$a, b \in \{0, \ldots, r - 1\}$

$r = $ matrix rank

Higher-dimensional dataset:

Correlator matrices — time series of matrix-valued correlator data



Useful when different interpolators have significant overlaps with multiple states

Standard analysis framework involves generalized eigenvalue problem (GEVP)

See e.g. Blossier et al, JHEP 04 094 (2009)

Block Lanczos

"Block Lanczos" algorithm developed for matrices with small spectral gaps

Golub, (1973) Cullum and Donath, (1974) Golub and Underwoord, (1977)

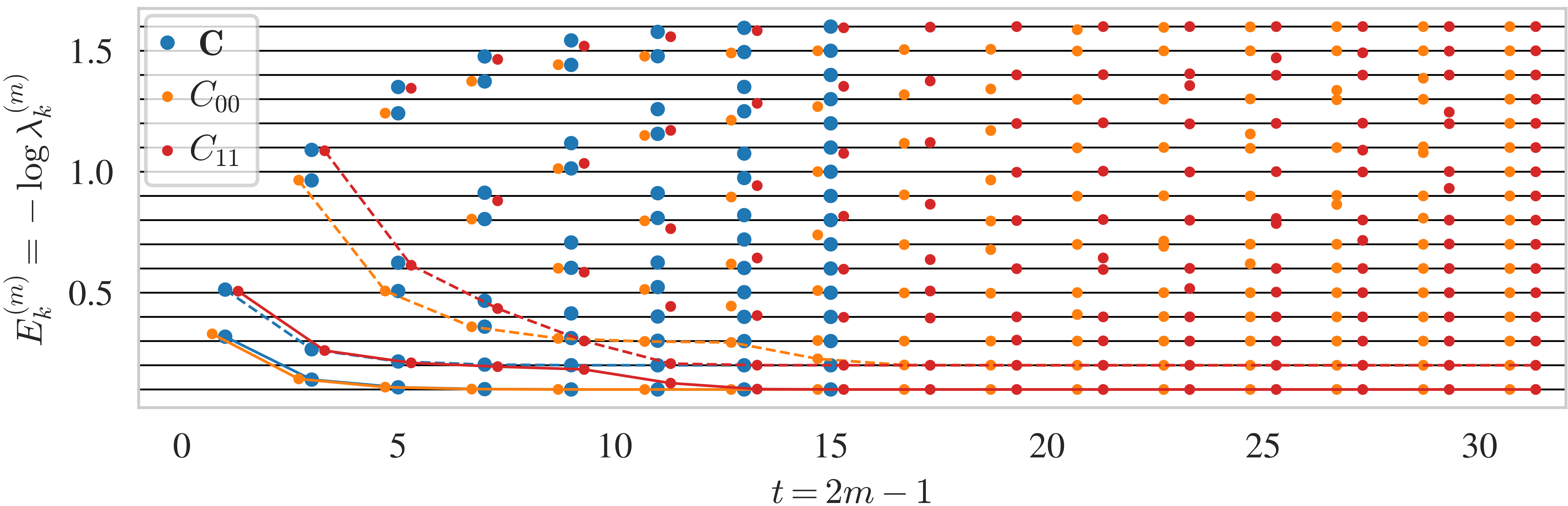
Generalized three-term recursion:

$$T\big|v_{ja}\big\rangle = \sum_b \big|v_{(j+1)b}\big\rangle\beta_{(j+1)ba} - \big|v_{jb}\big\rangle\alpha_{jba} + \big|v_{(j-1)b}\big\rangle\beta_{jba}$$

Ritz values: rm eigenvalues of block-diagonal $T^{(m)}_{iajb} = \langle v_{ia}|T|v_{jb}\rangle$

- rm-dimensional block Krylov space gives strictly faster convergence for $r > 1$

Hackett, MW, arXiv:2412.04444

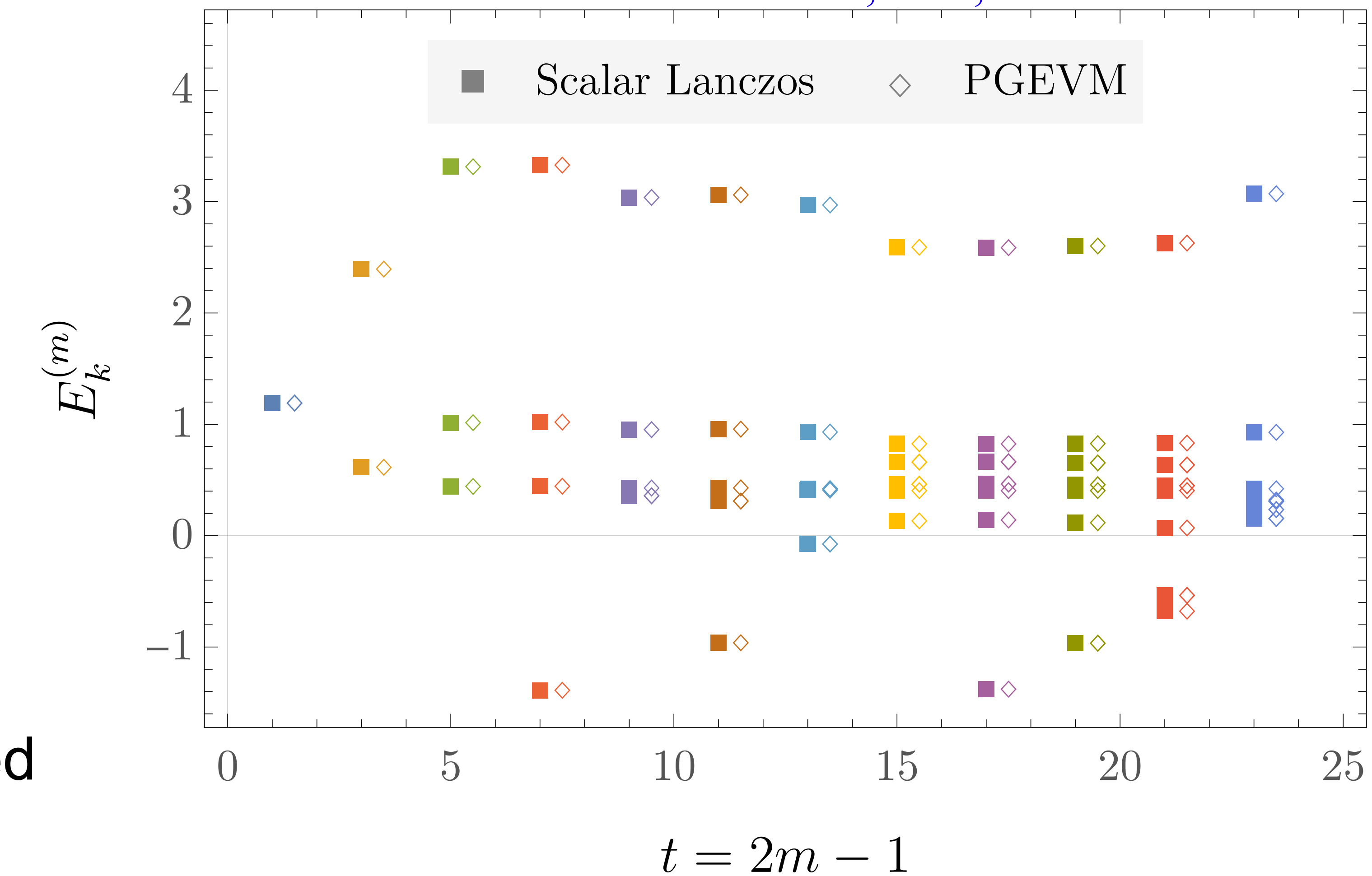


Lanczos = Krylov + RR = Prony

Algebraic methods for decomposing time series into sum of exponentials known since 1795

Prony and Gaspard (1795)

Hackett, MW, arXiv:2412.04444



Applications of Prony's method to LQCD first proposed by Fleming in 2004

Fleming arXiv:hep-lat/0403023 (2004)

Other equivalent implementations possible, e.g. Prony generalized eigenvalue method (PGEVM)

Fischer et al, Eur. Phys. J. A 56, 206 (2020)

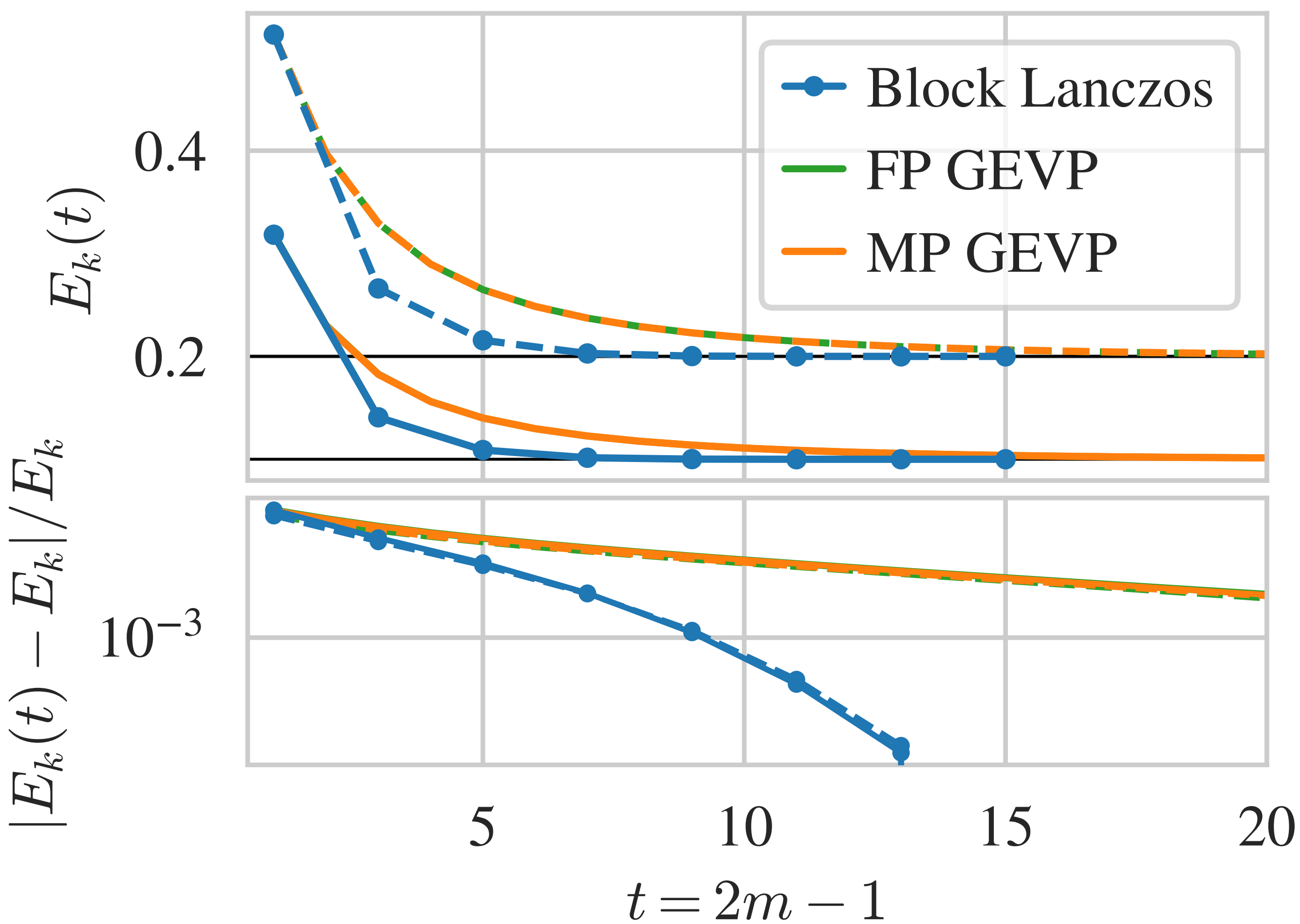
Lanczos and Prony produce identical energy estimators for noisy data

MW, arXiv:2406.20009v2 (August) Ostmeyer et al, arXiv:2411.14981 Chakraborty et al, arXiv:2412.01900

Block Lanczos, GEVP, block Prony, …

One step of block Lanczos = GEVP

Lüscher and Wolff, Nucl. Phys. B 339, 222 (1990)



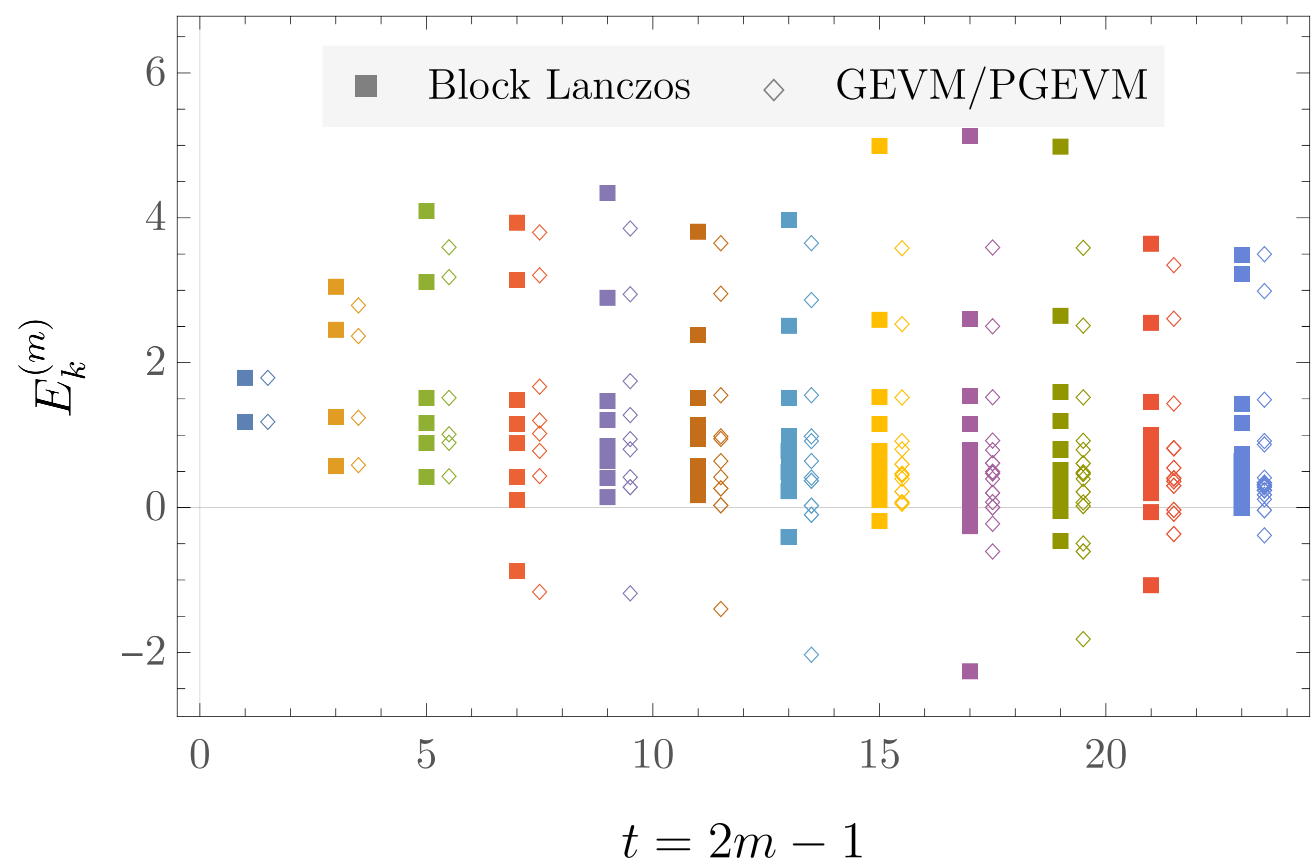
Hackett, MW, arXiv:2412.04444

Block Lanczos is a strict generalization of GEVP

In contrast to conjecture of Ostmeyer et al, arXiv:2411.14981

Block Lanczos ≠ GEVM/PGEVM

Fischer et al, Eur. Phys. J. A 56, 206 (2020)



Hackett, MW, arXiv:2412.04444

Block Lanczos = Block Prony

Fleming, LATTICE2023

What’s new?

What’s new?

1. Fast convergence, guaranteed

2. Residual bounds

3. Spurious state filtering

4. No fitting needed

5. Asymptotically constant SNR

6. Simple matrix element extractions

7. More excited states, cleanly

What's new?

1. **Fast convergence, guaranteed**

2. **Residual bounds**

Part 1: Signal

KPS convergence theory

Lanczos converges exponentially faster than power iteration for transfer matrices with small gaps (e.g. for small a)

Kaniel, Mathematics of Computation 20, 369 (1966)

Paige, PhD thesis 1971

Saad, SIAM 17 (1980)

$$\delta = a(E_1 - E_0)$$

$$\left|E_0 - E_0^{(m)}\right| \propto e^{-2t\sqrt{\delta}}$$

Lanczos

$$\left|E_0 - E_0^{\text{eff}}(t)\right| \propto e^{-t\delta}$$

Power iteration

- Convergence benefits largest near continuum limit where $1 \gg \sqrt{\delta} \gg \delta$
- Prony (= Lanczos) has identical convergence, but we didn't know the rate before

Block Lanczos converges exponentially faster than GEVP for transfer matrices with small gaps (e.g. for small a)

$$\delta_r = a(E_r - E_0)$$

Saad, SIAM 17 (1980)

$$\left|E_0 - E_0^{(m)}\right| \propto e^{-2t\sqrt{\delta_r}}$$

Block Lanczos

$$\left|E_0 - E_0^{\text{GEVP}}(t)\right| \propto e^{-t\delta_r}$$

GEVP

Residual bounds

- Lanczos approximation error after finite number of iterations directly computable:

$$\min_{\lambda \in \{\lambda_n\}} |\lambda_0^{(m)} - \lambda| \leq |\beta_{m+1} \omega_{m0}^{(m)}|$$

Eigenvectors of $T^{(m)}$

Matrix element $T^{(m)}_{m(m+1)}$

See Parlett, The Symmetric Eigenvalue Problem (1980)

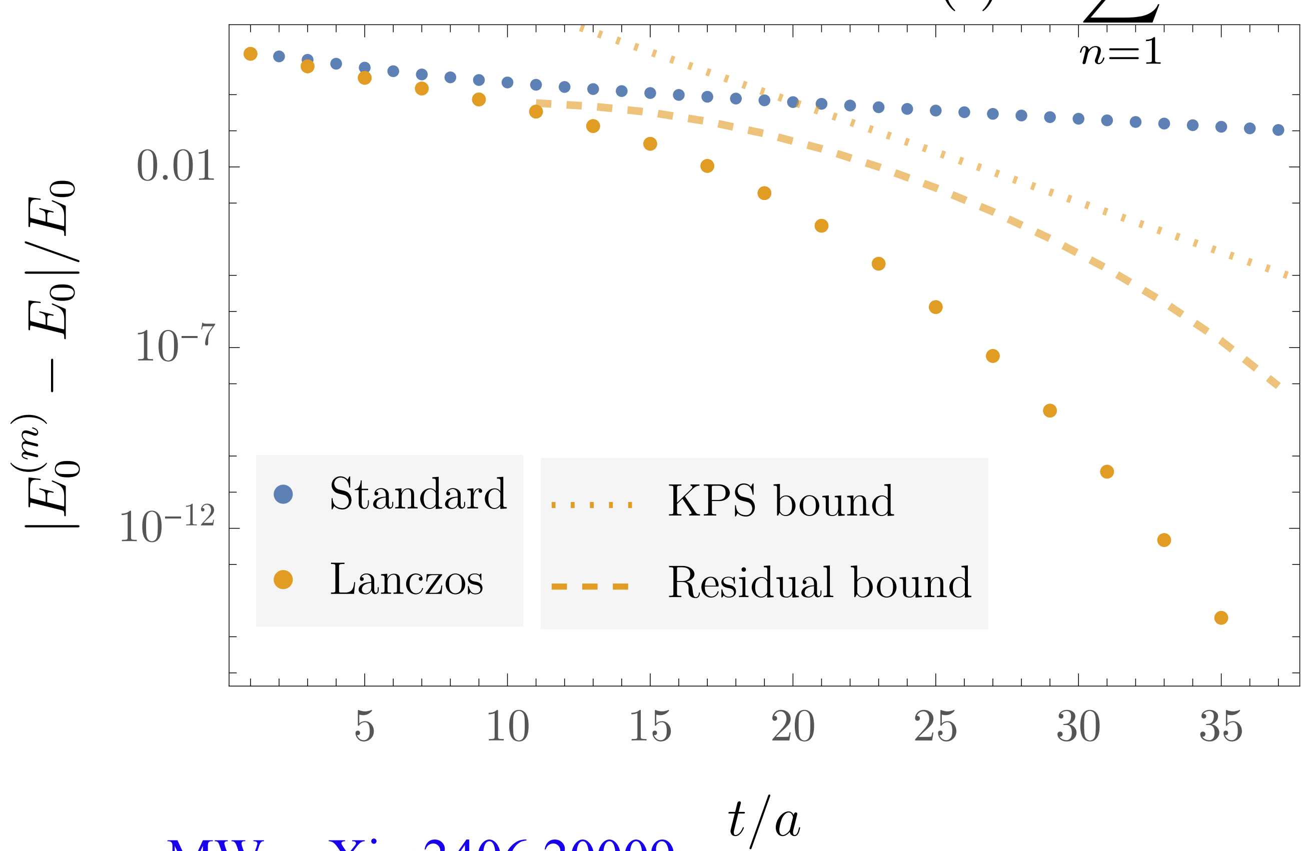
Rigorous quantification of excited-state effects!

Mock data tests demonstrate

- Lanczos converges exponentially faster than power iteration / effective mass
- Residual bound provides valid two-sided bound on errors from excited-state effects

Note: residual bound is on distance to closest eigenvalue, not e.g. "true ground state"

$$C(t) = \sum_{n=1}^{20} n^2 e^{-0.1nt}$$



MW, arXiv:2406.20009

What’s new?

1. Fast convergence, guaranteed ✓

2. Residual bounds ✓

3. Spurious state filtering

4. No fitting needed

5. Asymptotically constant SNR

6. Simple matrix element extractions

7. More excited states, cleanly

What's new?

Part 2: Noise

3. Spurious state filtering

4. No fitting needed

5. Asymptotically constant SNR

6. Simple matrix element extractions

7. More excited states, cleanly

Spurious eigenvalues

Decades of research on how roundoff affects Lanczos has led to an understanding of the “Lanczos phenomenon”

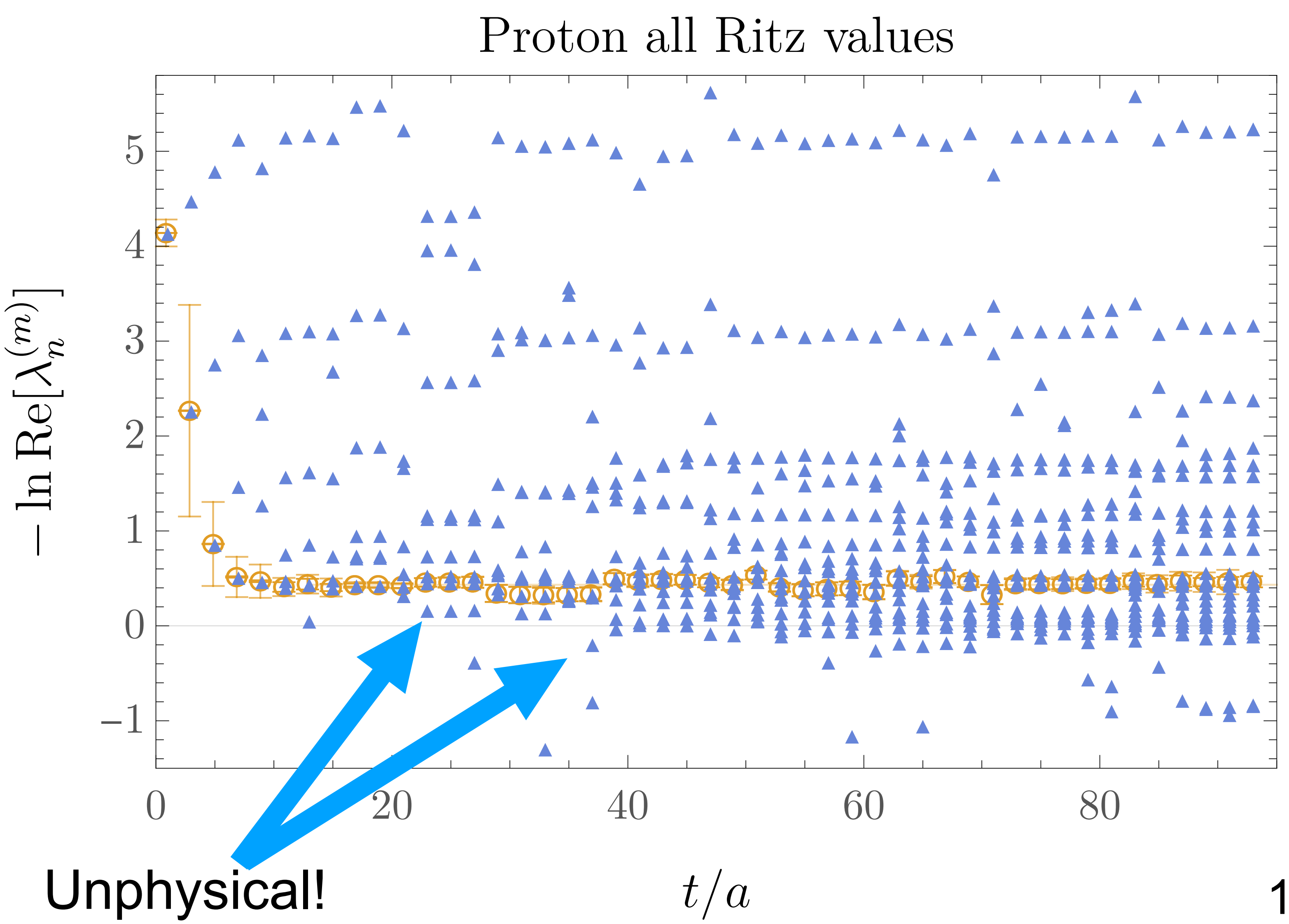
- Roundoff leads to O(1) errors in some “spurious” Ritz values that do not converge
- Remaining “non-spurious” Ritz values still accurate, converge to eigenvalues

Statistical noise leads to unphysical Ritz values:

MW, arXiv:2406.20009

- Most Ritz values complex even though transfer matrix eigenvalues real + positive
- Taking real parts at face value would give ground-state energy violating QCD inequality $M_N > m_\pi$

Does removing “spurious eigenvalues” fix this?



Cullum-Willoughby

- Jane Cullum and Ralph Willoughby developed a useful criterion for identifying spurious eigenvalues in 1981

Cullum and Willoughby, Journal of Computational Physics 44, 329 (1981)

DEFINITION 1. Spurious ≡ Outwardly similar or corresponding to something without having its genuine qualities.

$$T^{(m)} = \begin{pmatrix} \alpha_1 & \beta_2 & & & & 0 \\ \gamma_2 & \alpha_2 & \beta_3 & & & \\ & \gamma_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{m-1} & \\ & & & \gamma_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & & & & \gamma_m & \alpha_m \end{pmatrix}$$

$$T_2^{(m)} = \begin{pmatrix} \cancel{\alpha_1} & \cancel{\beta_2} & & & & 0 \\ \cancel{\gamma_2} & \alpha_2 & \beta_3 & & & \\ & \gamma_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{m-1} & \\ & & & \gamma_{m-1} & \alpha_{m-1} & \beta_m \\ 0 & & & & \gamma_m & \alpha_m \end{pmatrix}$$

DEFINITION 2. Any simple eigenvalue of T_m that is pathologically close to an eigenvalue of $\hat{T}_2$ will be called "spurious."

Think positive

Since transfer matrix is positive-definite by assumption, any eigenvalues with non-zero imaginary parts can be discarded as spurious on physical grounds

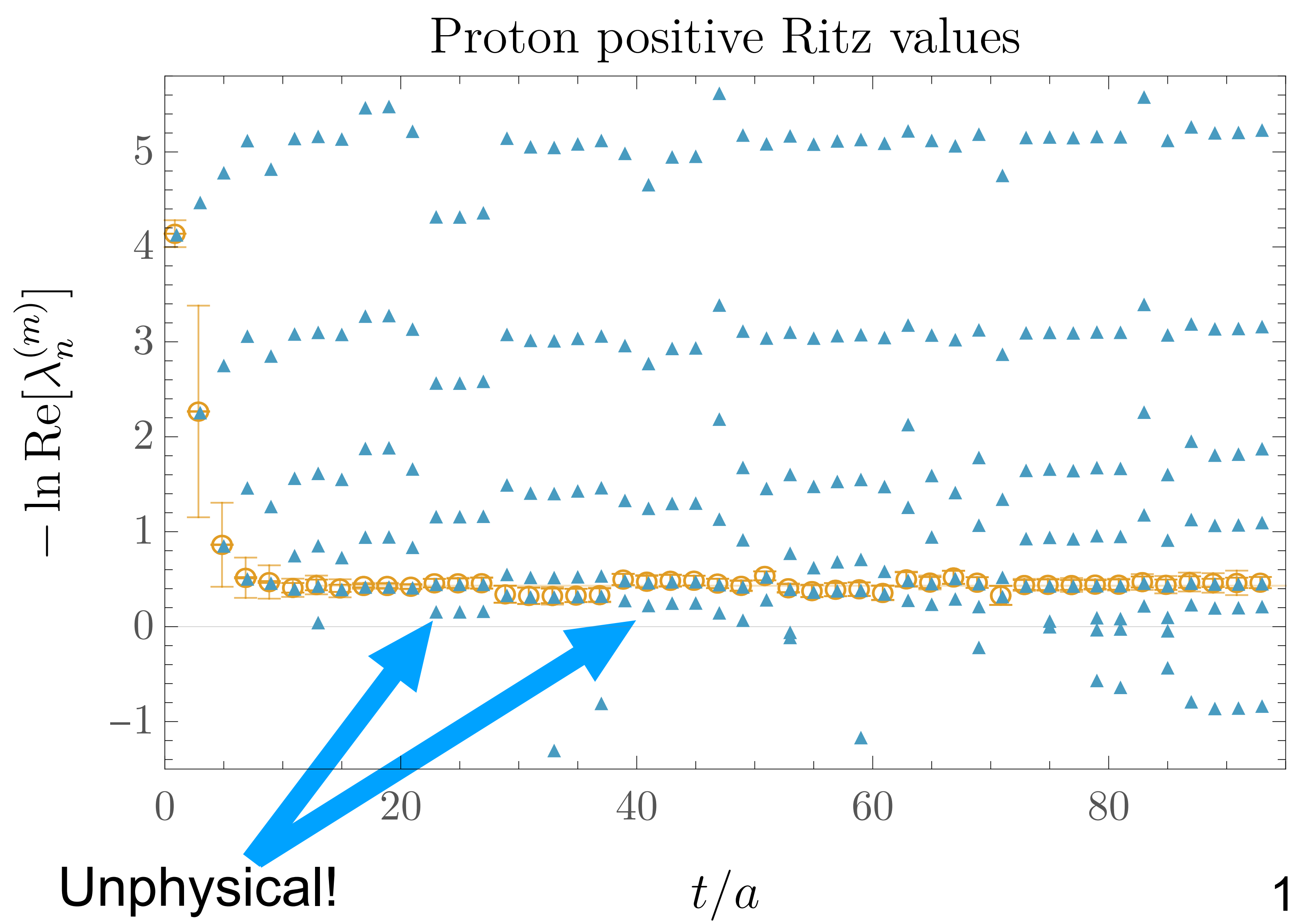
- Oblique Lanczos formalism provides rigorous framework for analyzing complex Ritz values

Saad, SIAM 19 (1982)

Not all unphysical Ritz values are complex, positive ones can still violate $M_N > m_\pi$

Non-positive eigenvalues have much smaller Cullum-Willoughby "knockout distances" than positive ones

- At least some unphysical eigenvalues are spurious by CW definition



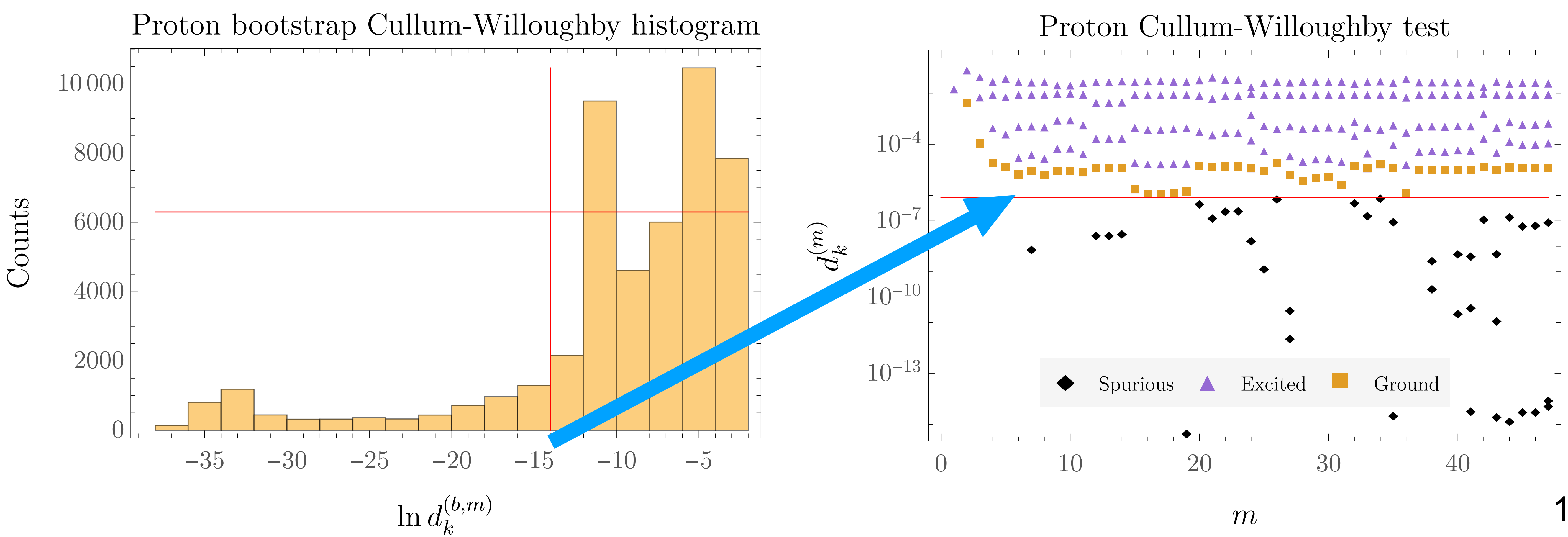
Bootstrapping Cullum-Willoughby

- Defining "pathologically close" is easy for finite matrices with floating-point roundoff error, harder for Monte Carlo simulations of infinite-dimensional matrices

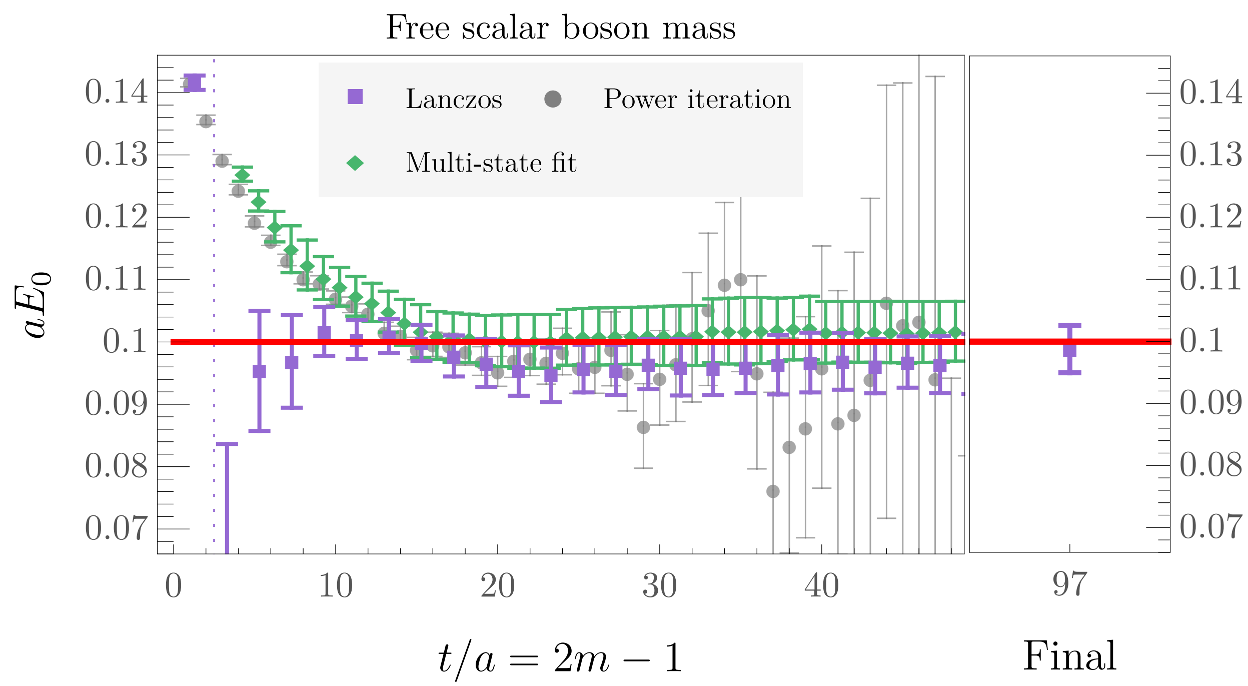
DEFINITION 1. Spurious ≡ Outwardly similar or corresponding to something without having its genuine qualities.

- Distances between $T^{(m)}$ and $T_2^{(m)}$ are consistently smaller for spurious than nonspurious eigenvalues — spurious ones also less stable vs iteration

- Use bootstrap histograms to define cutoff

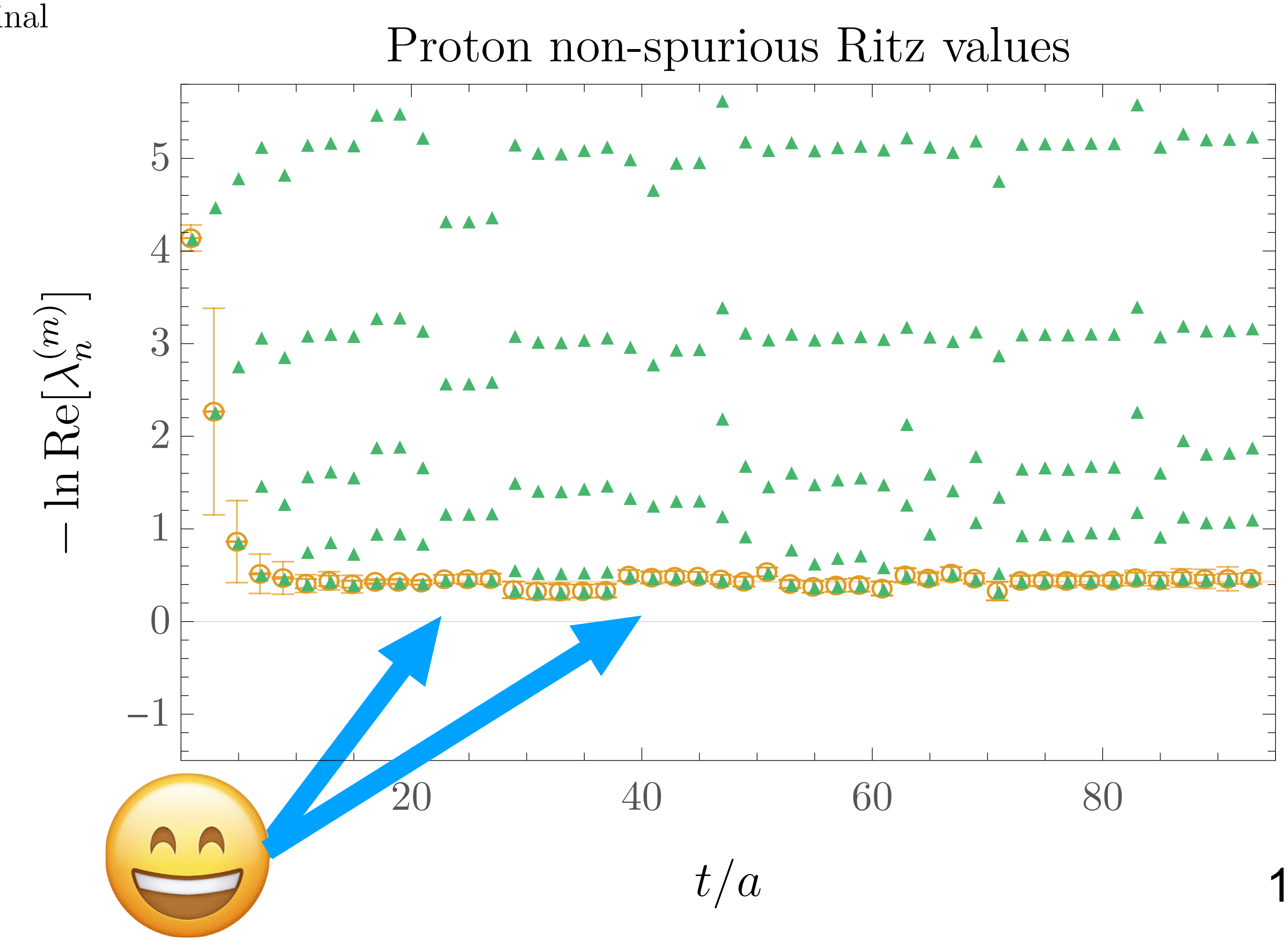


Non-spurious energies are accurate



Defining $\lambda_0^{(m)}$ as the largest “non-spurious” Ritz value leads to accurate ground-state energy determinations in solvable models (e.g. free scalar field)

All obviously unphysical proton eigenvalues removed by filtering on “spuriously small” CW knockout distance defined via bootstrap histogram



Who ordered that?

Unappealing features remain:

- Cullum-Willoughby test is essential for stabilizing SNR but not physically motivated, appears *ad hoc*

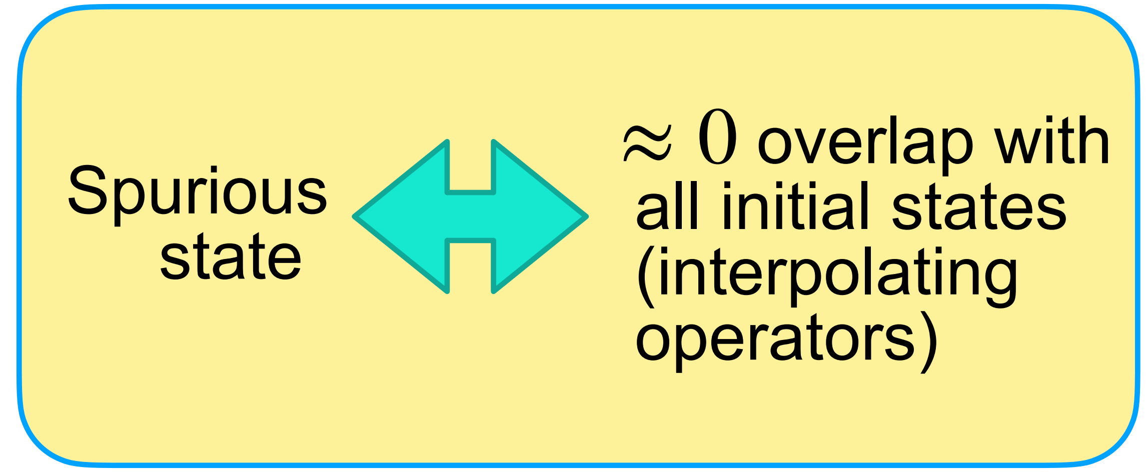
- CW test does not immediately generalize to block Lanczos

Who ordered that?

Unappealing features remain:

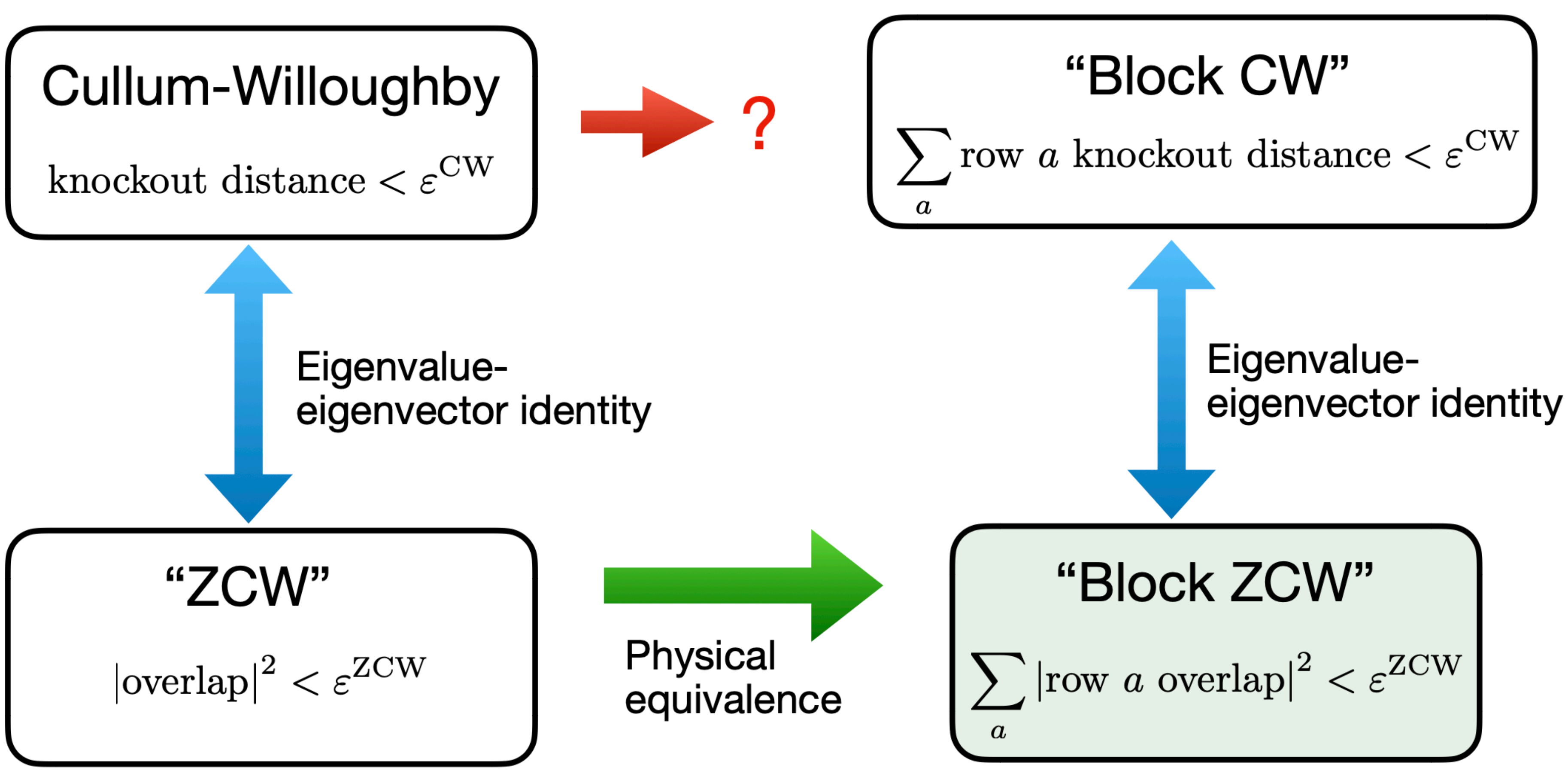
- Cullum-Willoughby test is essential for stabilizing SNR but not physically motivated, appears *ad hoc*
- CW test does not immediately generalize to block Lanczos

Introducing the ZCW test:



Hackett, MW, arXiv:2412.04444

Spurious eigenvalue tests



Physical motivation generalizes straightforwardly to block ZCW

The ZCW test

Roundoff (and noise) leads to errors in orthogonalization, artificially extend Krylov space in spurious directions Paige (1971) Parlett and Scott (1979)

- Motivation for CW test is that these spurious directions should only depend on numerical artifacts and be independent of initial vector

Physically: independence of initial vector ~ zero overlap with source ~ wrong quantum numbers

The ZCW test

Roundoff (and noise) leads to errors in orthogonalization, artificially extend Krylov space in spurious directions Paige (1971) Parlett and Scott (1979)

- Motivation for CW test is that these spurious directions should only depend on numerical artifacts and be independent of initial vector

Physically: independence of initial vector ~ zero overlap with source ~ wrong quantum numbers

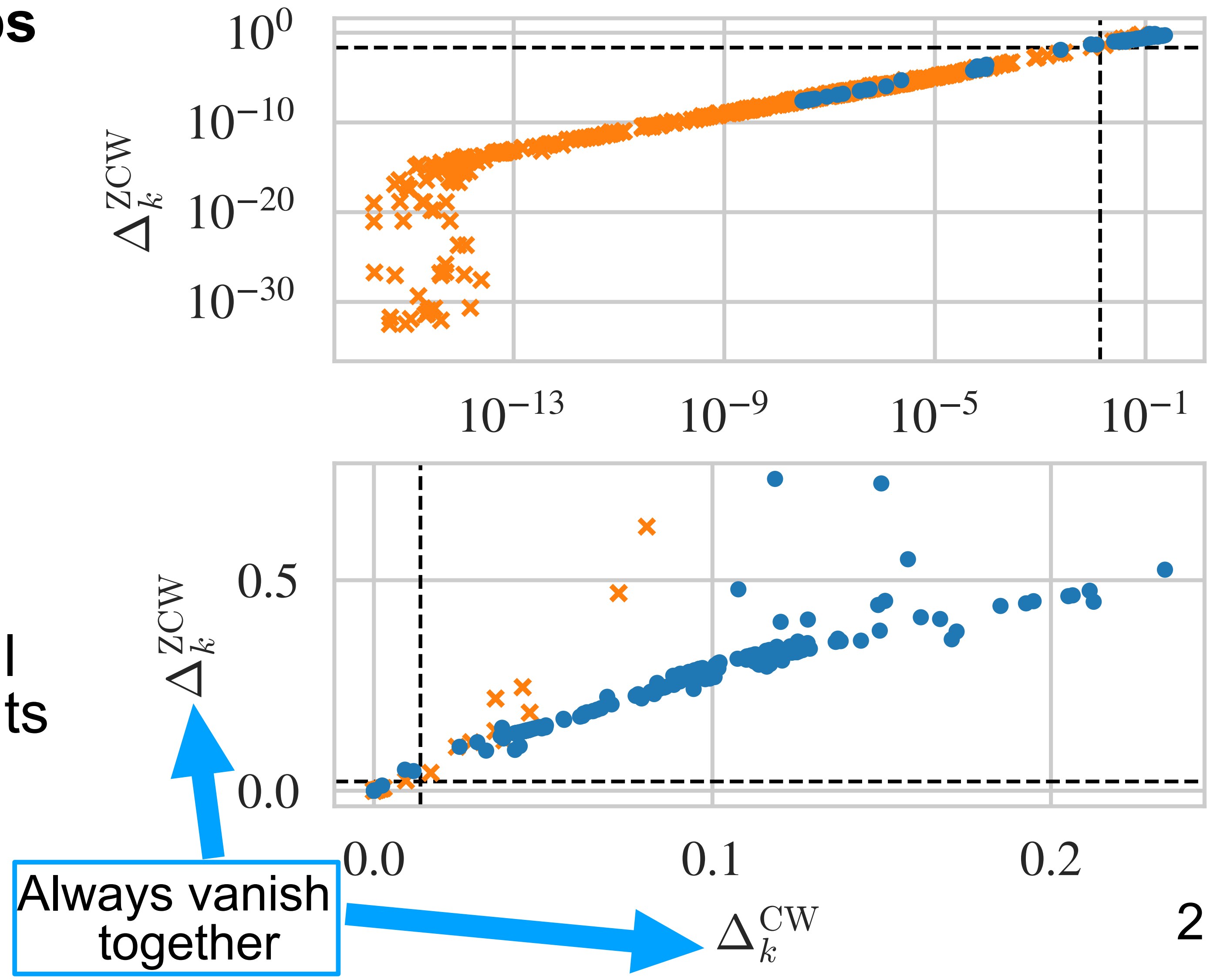
ZCW test for spuriously small overlaps

Hackett, MW, arXiv:2412.04444

$$\Delta_k^{\mathrm{ZCW}(m)} = \left| \frac{Z_k^{R(m)*} Z_k^{L(m)}}{C(0)} \right| < \varepsilon^{\mathrm{ZCW}}$$

Eigenvalue-eigenvector identity* can be used to prove equivalence of CW and ZCW tests for small $\varepsilon^{\mathrm{ZCW}} \sim \varepsilon^{\mathrm{CW}}$

Size of overlaps on last iteration where all Ritz values obey all physical constraints sets natural scale for $\varepsilon^{\mathrm{ZCW}}$



* See Denton, Parke, Tao, and Zhang, Bull. Am. Math Soc. 59, 31 (2022)

The physics of noise

Krylov space can be decomposed into sectors based off Ritz properties

$$T^{(m)} = \sum_{k\in\overline{\mathcal{S}}} \lambda_k^{(m)} \left|y_k^{(m)}\right\rangle\left\langle y_k^{(m)}\right| + \sum_{k\in\mathcal{S}} \lambda_k^{(m)} \left|y_k^{L(m)}\right\rangle\left\langle y_k^{R(m)}\right|$$

Non-spurious

Spurious

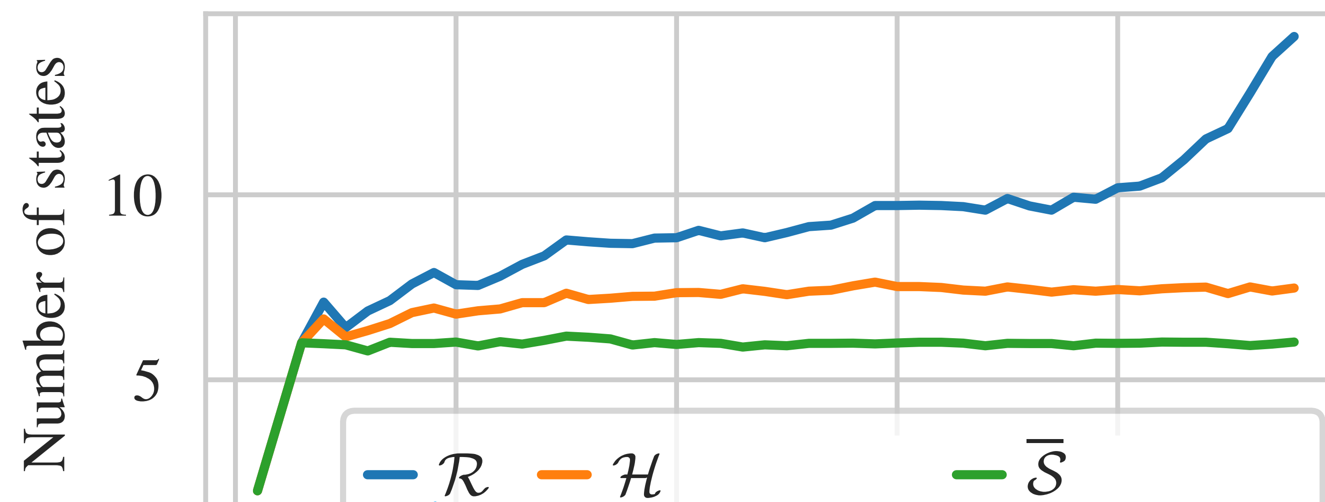
Further classification of spurious states possible:

- Non-spurious ⊂ Hermitian subspace ⊂ Real Ritz values

Unphysical states (e.g. with complex norms) needed to describe data that is non-convex in the presence of noise

States with non-zero initial-state overlap ("correct quantum numbers") are unaffected by spurious state filtering, can be interpreted physically

Hackett, MW, arXiv:2412.04444



Real Ritz values

Hermitian subspace

Non-spurious subspace

The block ZCW test

Physical picture provides immediate generalization to “block ZCW test”

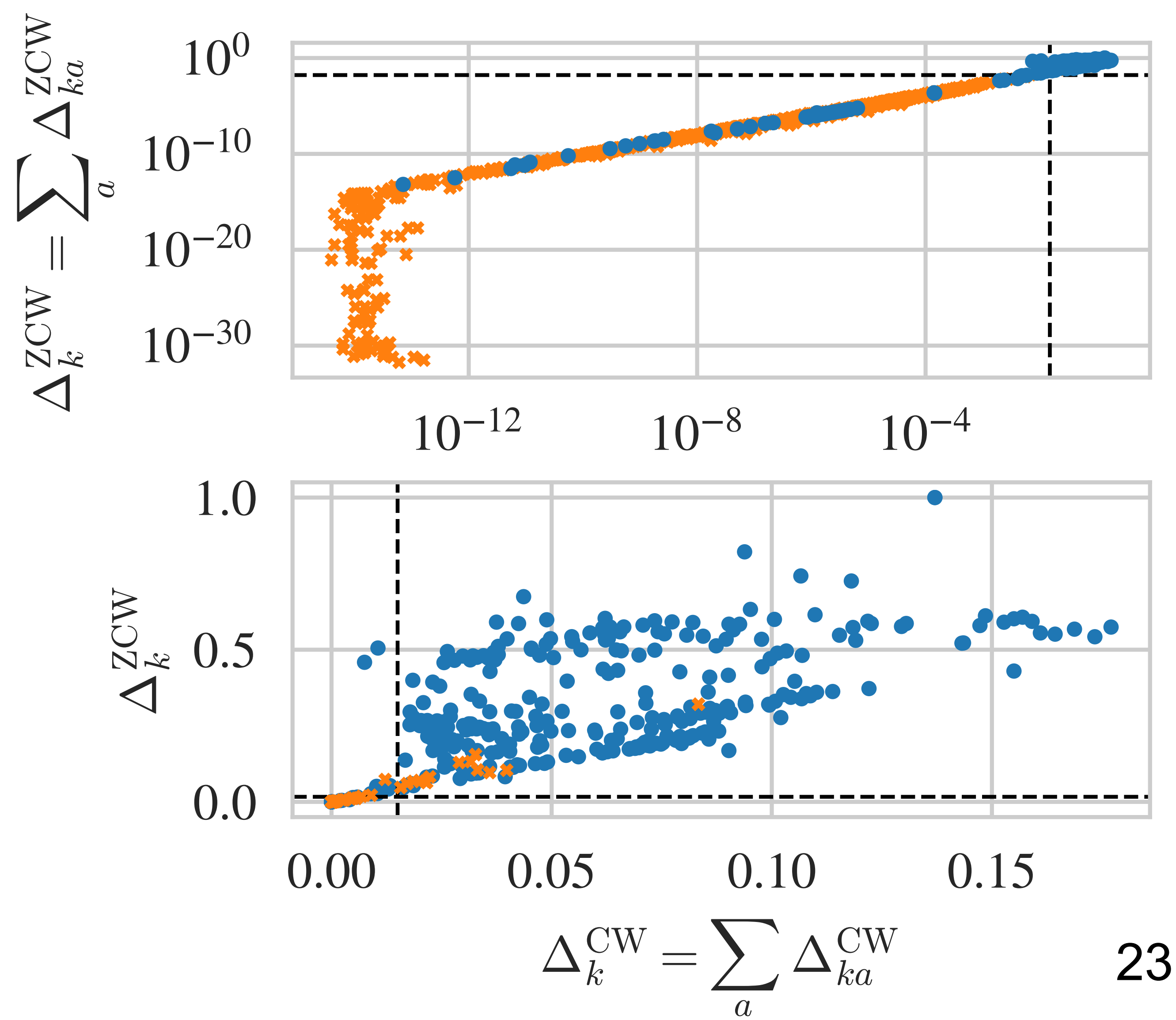
Hackett, MW, arXiv:2412.04444

$$\Delta_k^{\mathrm{ZCW}(m)} = \sum_{a,b} \left| Z_{ka}^{R(m)*} [\mathbf{C}(0)^{-1}]_{ab} Z_{kb}^{L(m)} \right| < \varepsilon^{\mathrm{ZCW}}$$

“Spurious" iff sum of overlaps with all interpolating operators ≈ 0

Eigenvalue-eigenvector identity can be used again to define an equivalent “block CW test” involving a sum over knockout distances

- Block ZCW is simpler to implement and (much) computationally cheaper, preferred in practice



What's new?

1. Fast convergence, guaranteed ✓

2. Residual bounds ✓

3. Spurious state filtering ✓

4. No fitting needed

5. Asymptotically constant SNR

6. Simple matrix element extractions

7. More excited states, cleanly

What's new?

1. Fast convergence, guaranteed ✓

2. Residual bounds ✓

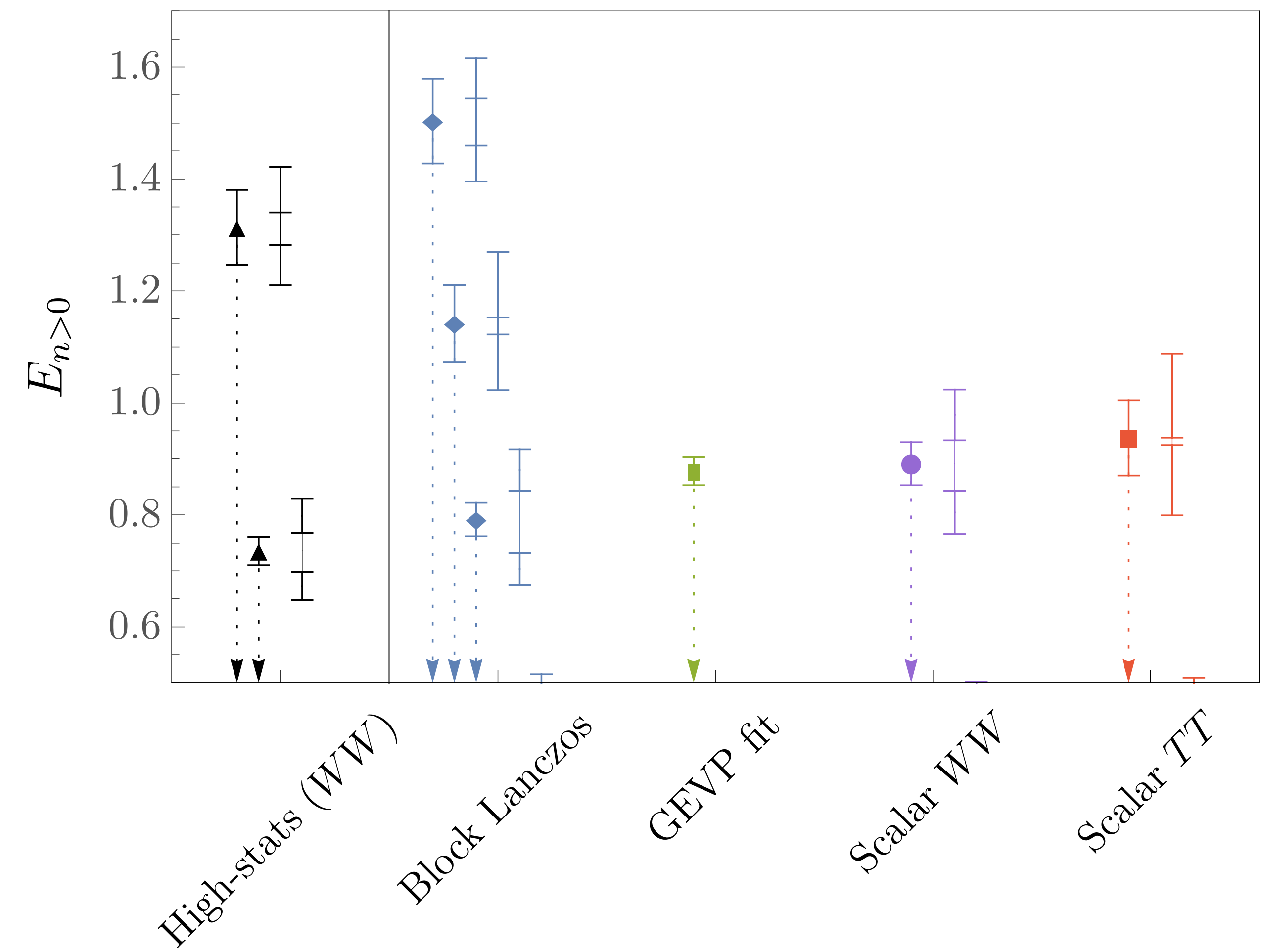
3. Spurious state filtering ✓

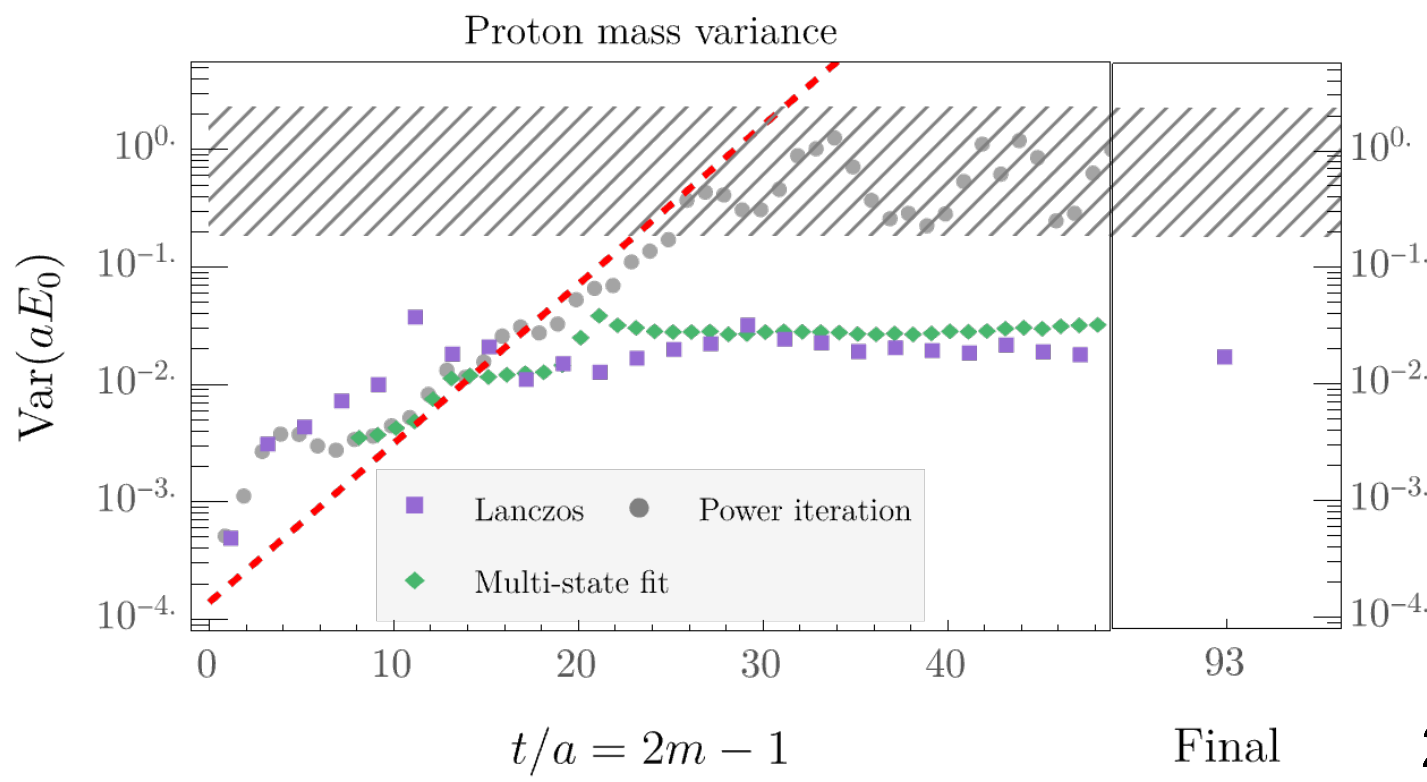
4. No fitting needed

5. Asymptotically constant SNR

6. Simple matrix element extractions

7. More excited states, cleanly

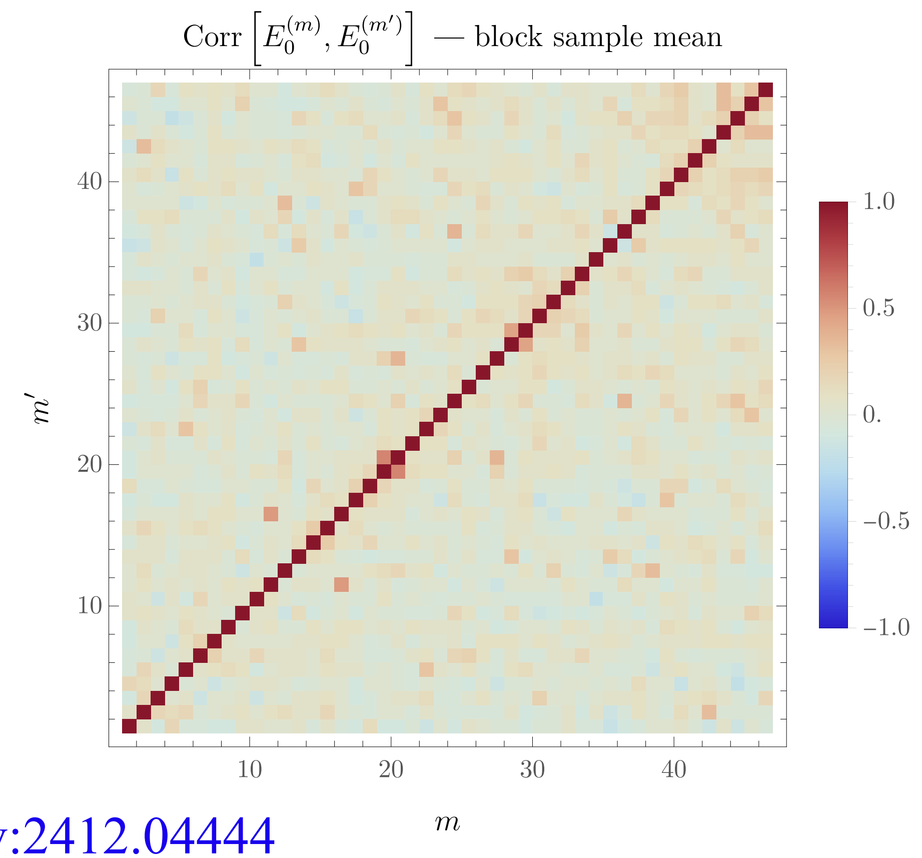




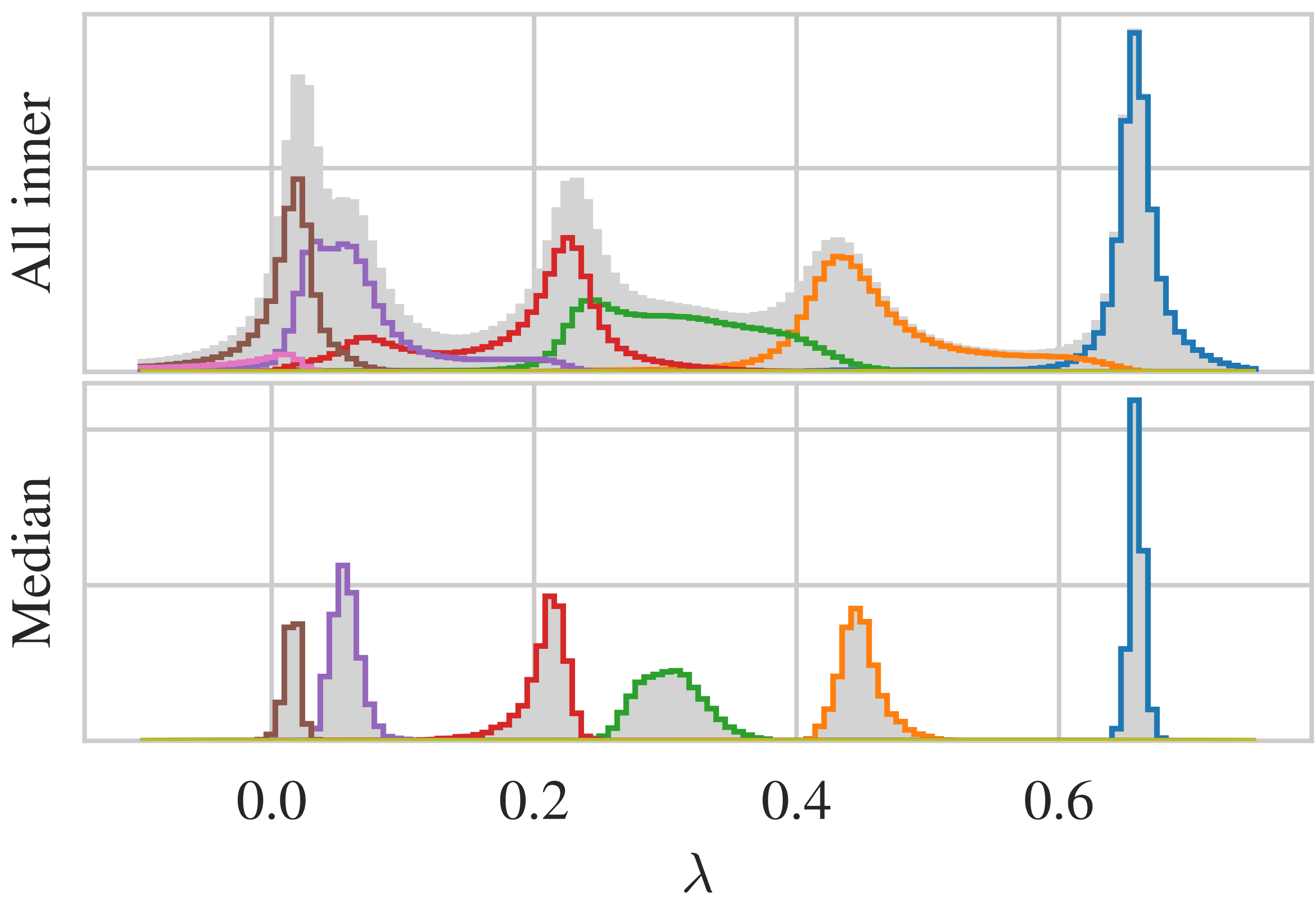
No fitting needed

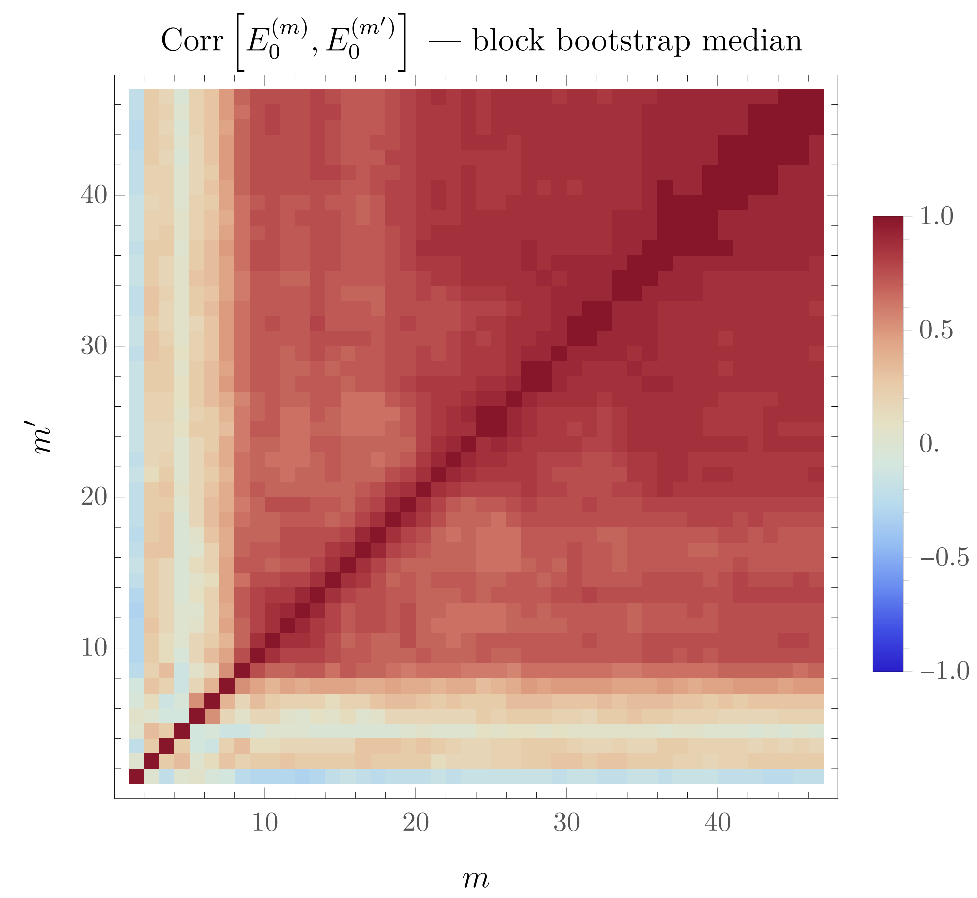
Spurious state filtering isn't perfect — outlier robust estimators can be both more precise + accurate

- Use bootstrap median as estimator, compute uncertainties with nested bootstrap
- Large correlations appear for large m with bootstrap median, washed out in sample mean
- Energy distributions closer to Gaussian for bootstrap median



Hackett, MW, arXiv:2412.04444

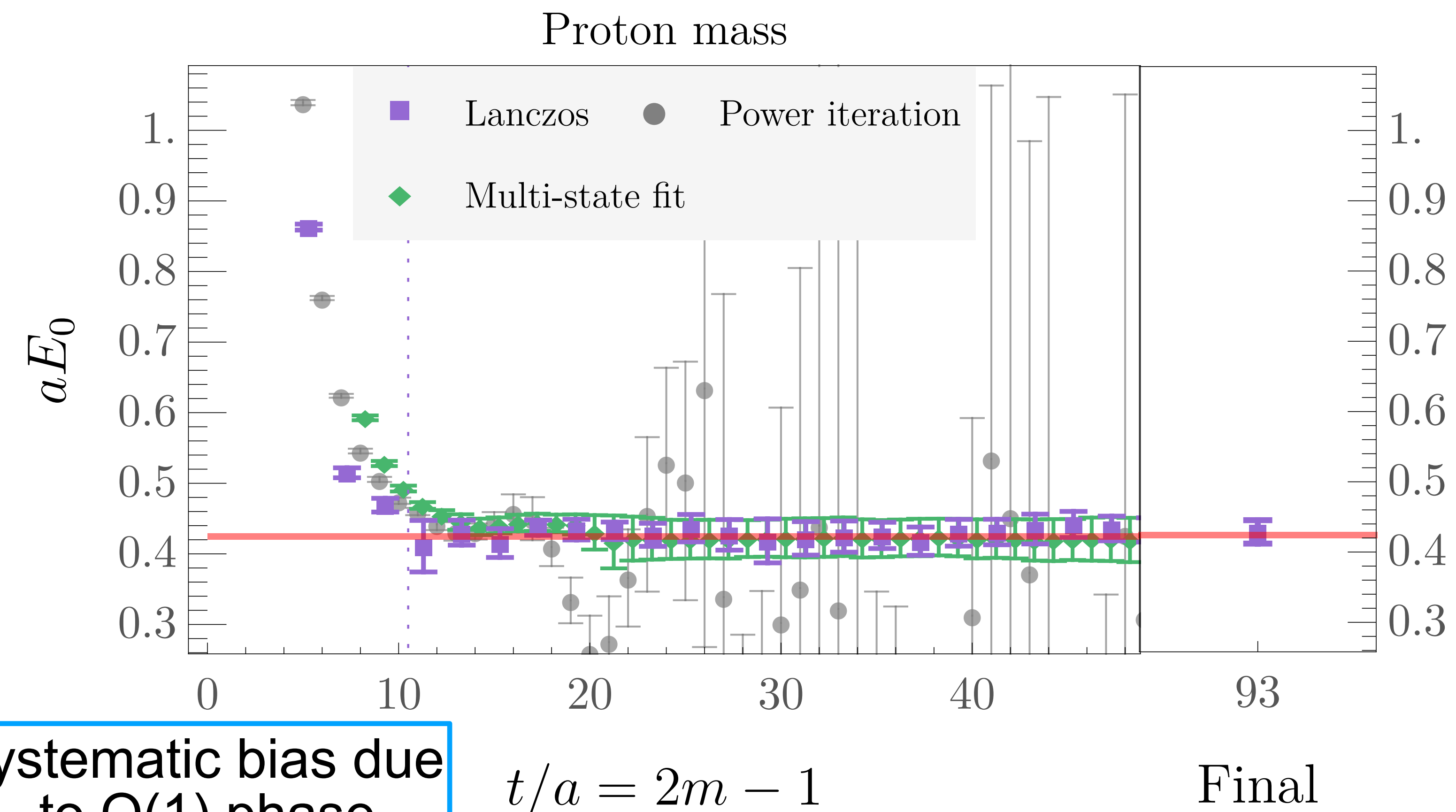




Asymptotically constant SNR

Bootstrap median estimators provide comparable uncertainties to multi-state fits with $t_{\max} = 2m - 1$

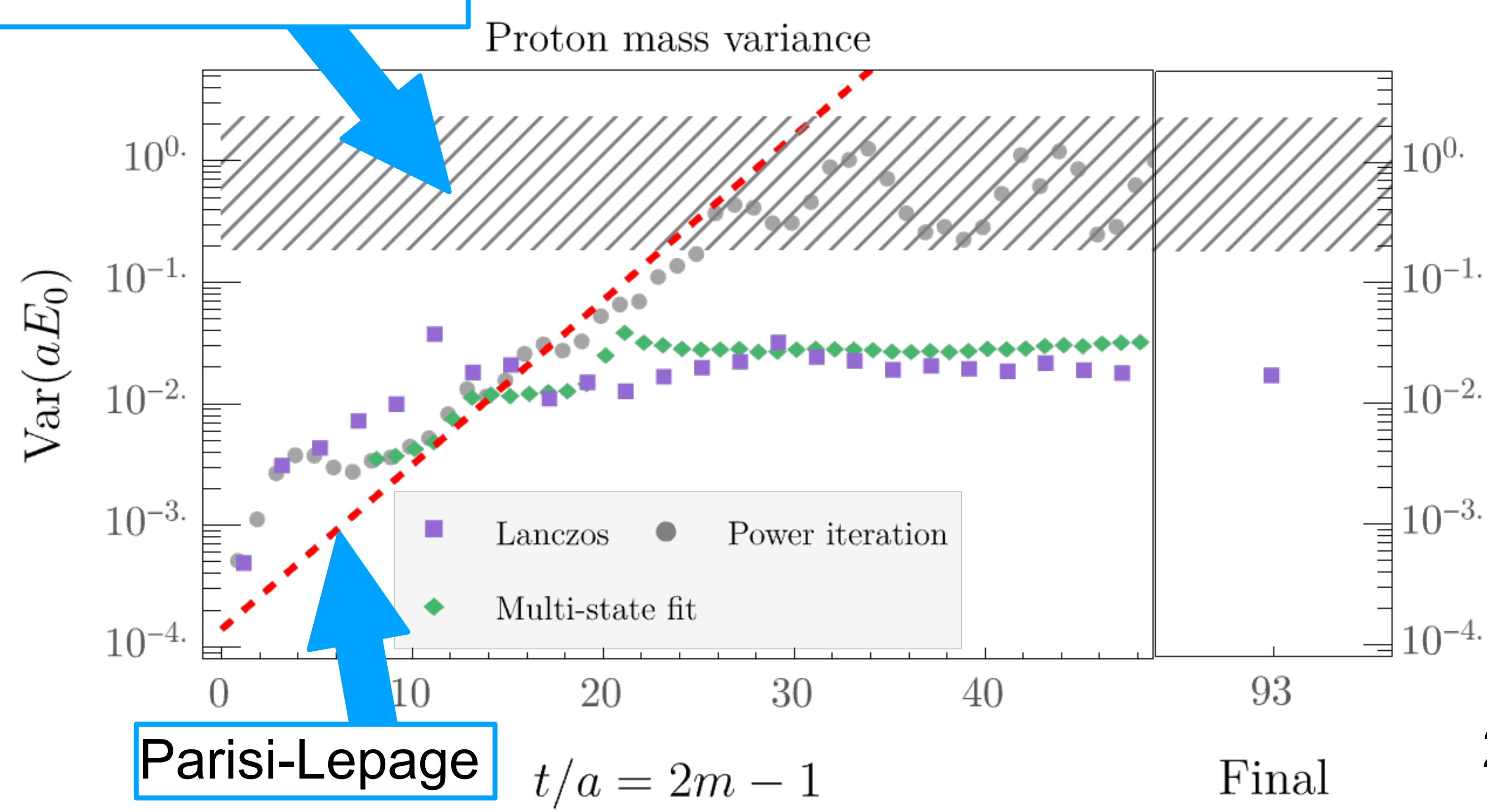
Given large correlations at large m, sufficient to define energy estimator from final iteration



Systematic bias due to O(1) phase fluctuations

Variance saturates to constant value for large m, comparable to saturation of multi-state fit results

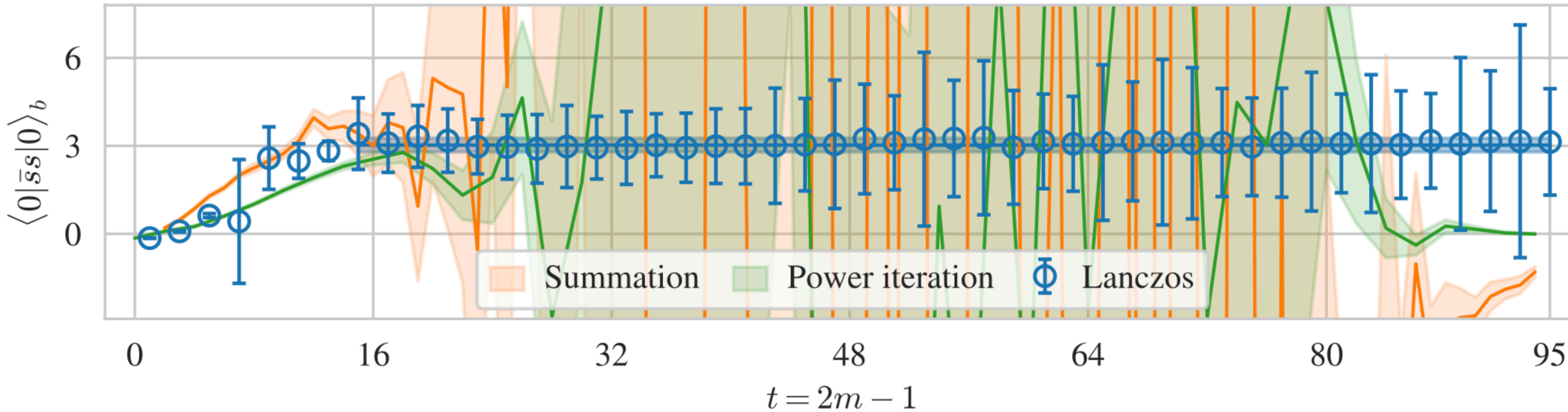
- Contrasts with power-iteration / effective mass, which exponentially approaches 0 SNR



Parisi-Lepage

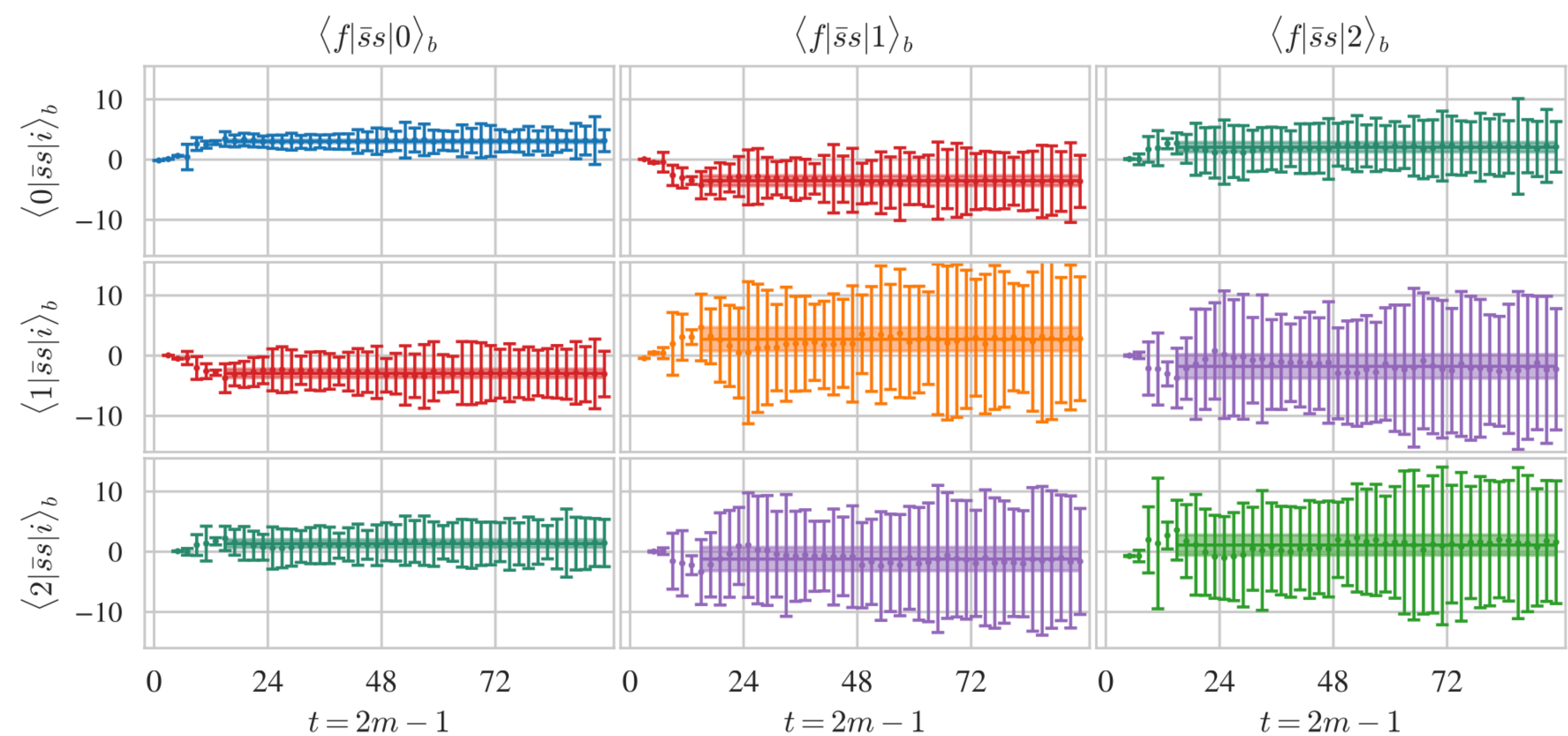
Operator matrix elements

Nucleon strange scalar (bare) matrix element, non-nested   Hackett, MW, arXiv:2407.21777



Lanczos eigenvectors provide change of basis allowing matrix elements to be extracted from 3pt functions with simple matrix multiplication

- Excited / transition matrix elements accessible



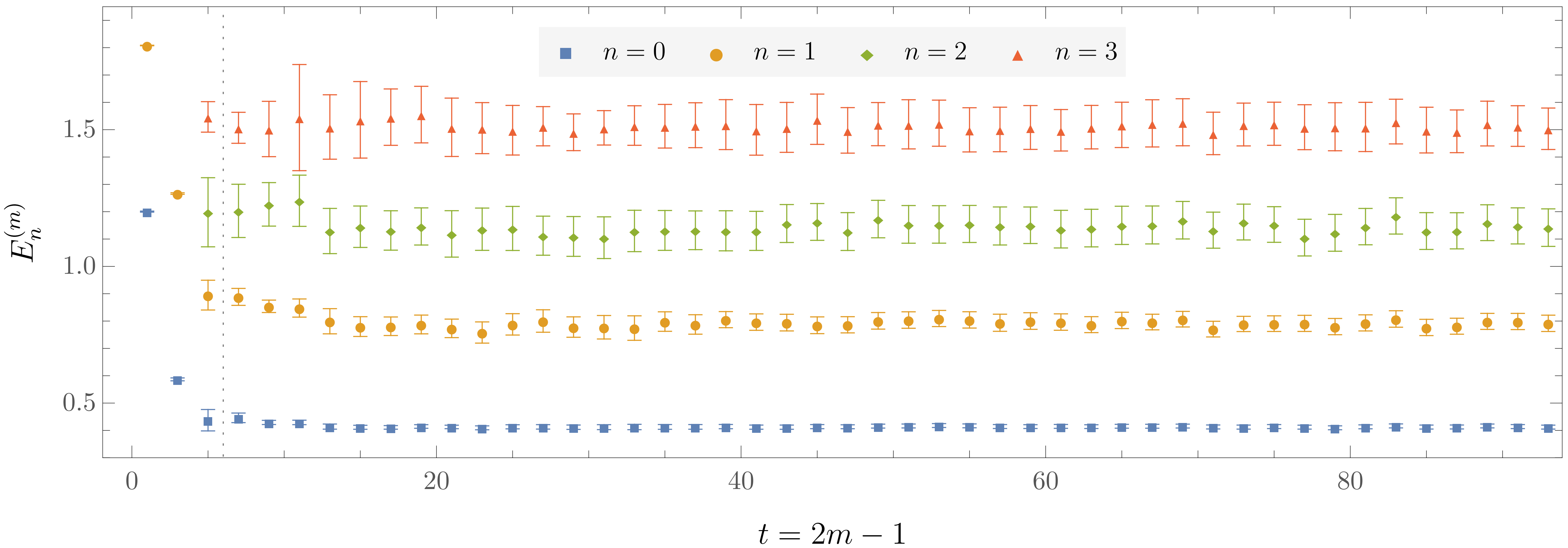
More excited states

Excited-states also show asymptotically constant SNR, rapid convergence

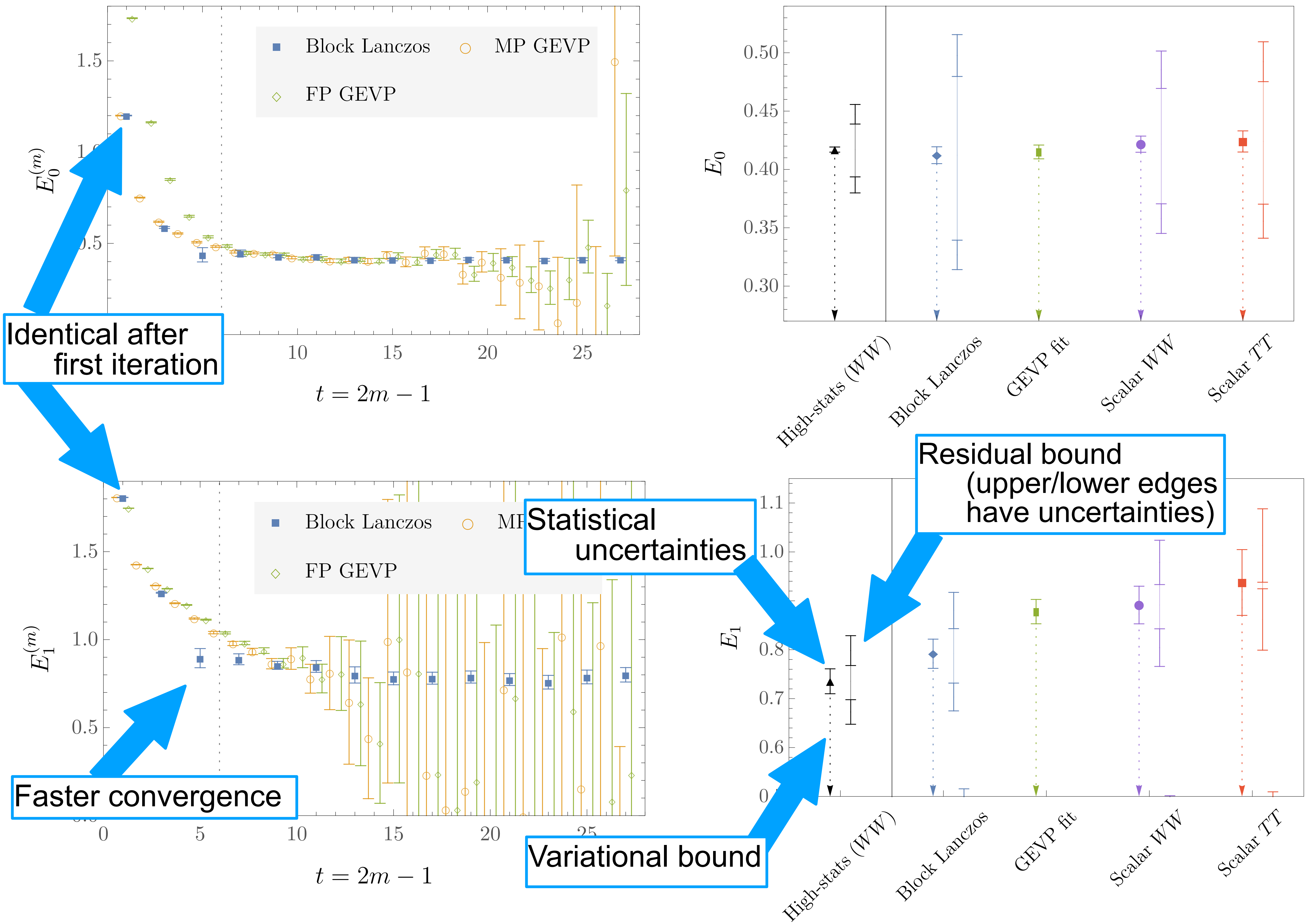
- Block Lanczos can cleanly extract e.g. 4 states from a noisy 2x2 matrix

Residual bounds provide rigorous two-sided systematics for excited states

Hackett, MW, arXiv:2412.04444



Block Lanczos and GEVP: Es

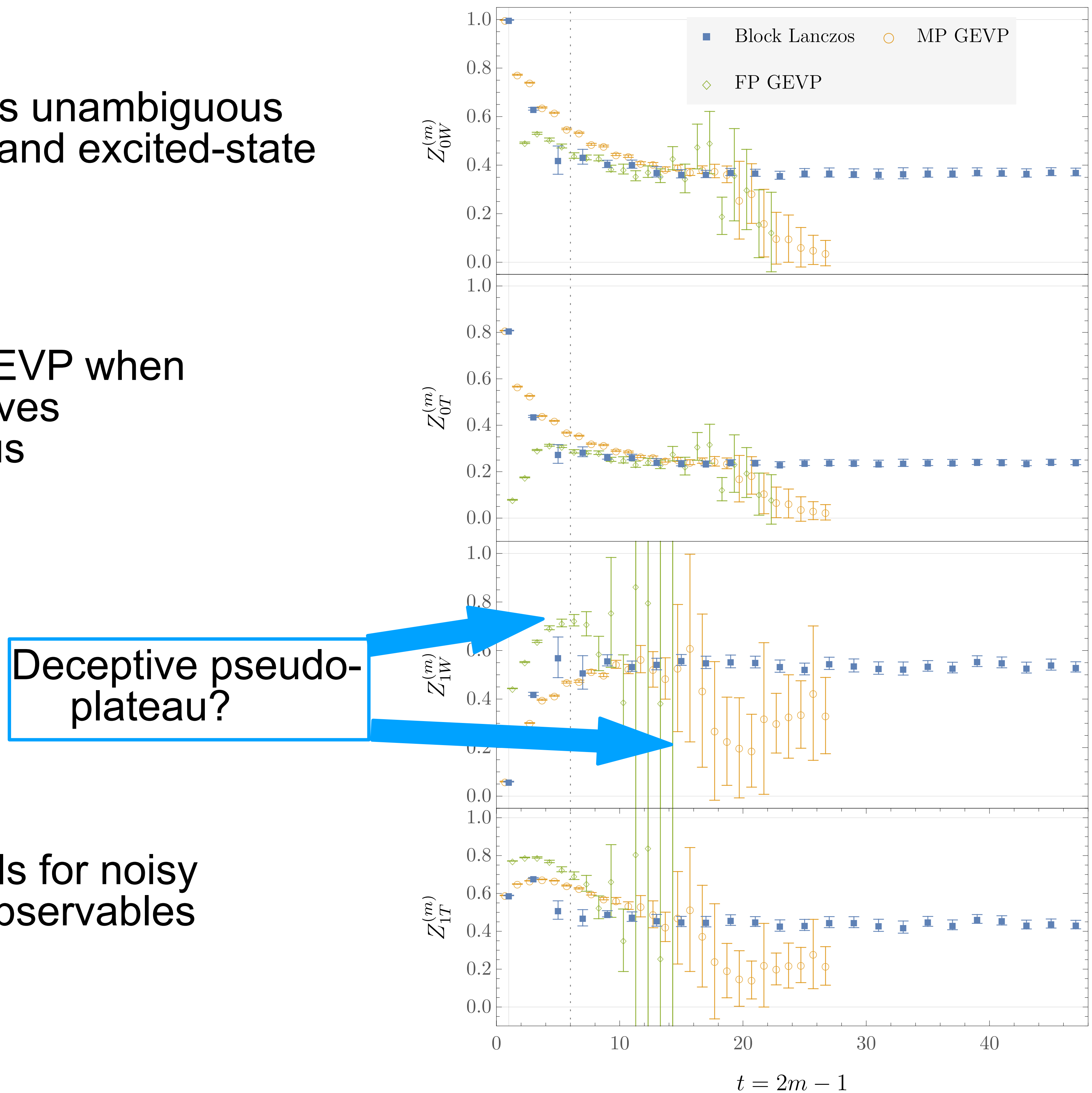


Block Lanczos and GEVP: Zs

Block Lanczos provides unambiguous signals for ground- and excited-state overlap factors

- Consistent with GEVP when the latter achieves reliable plateaus

- More robust signals for noisy excited-state observables

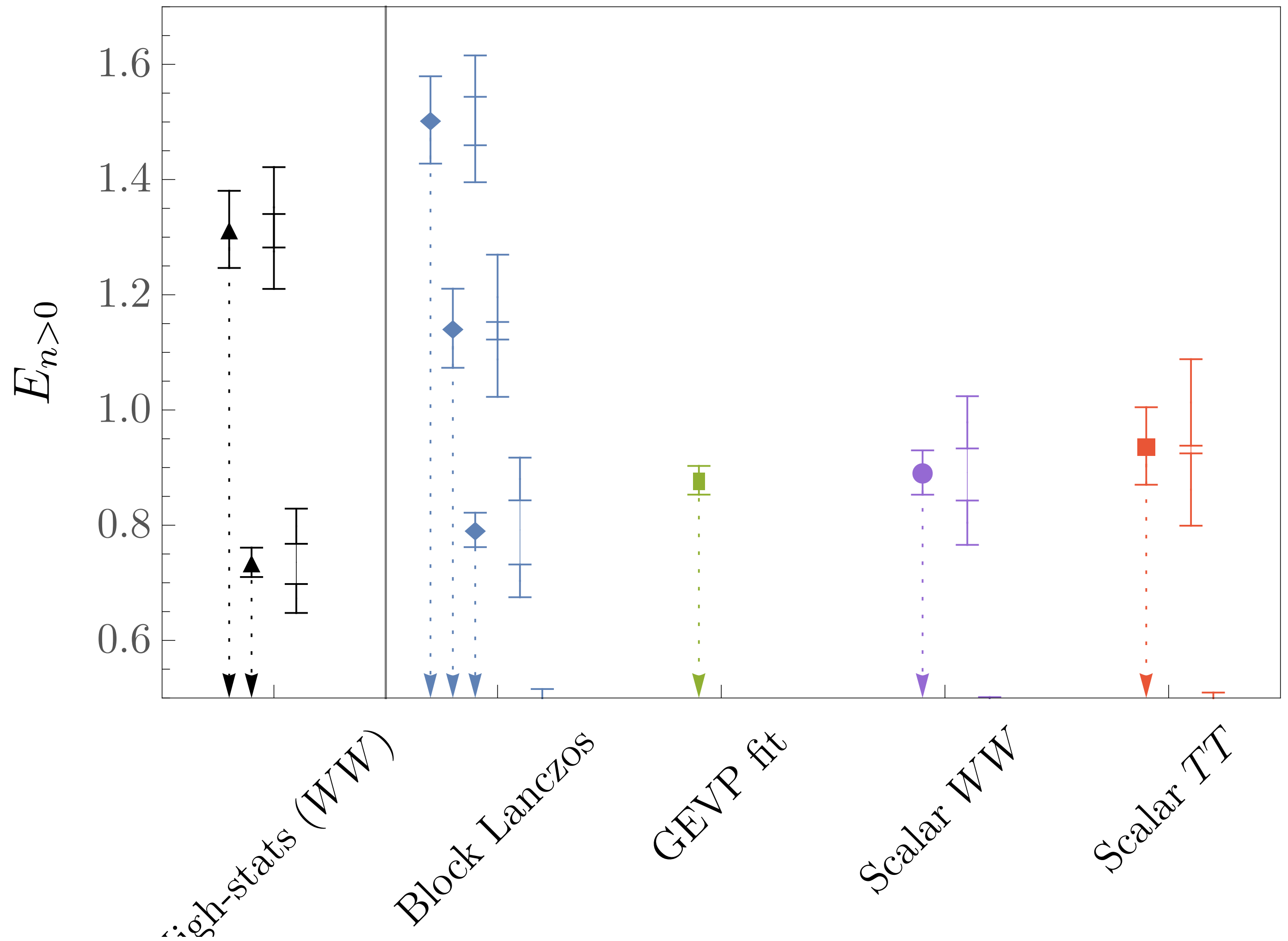


Towards higher excited states

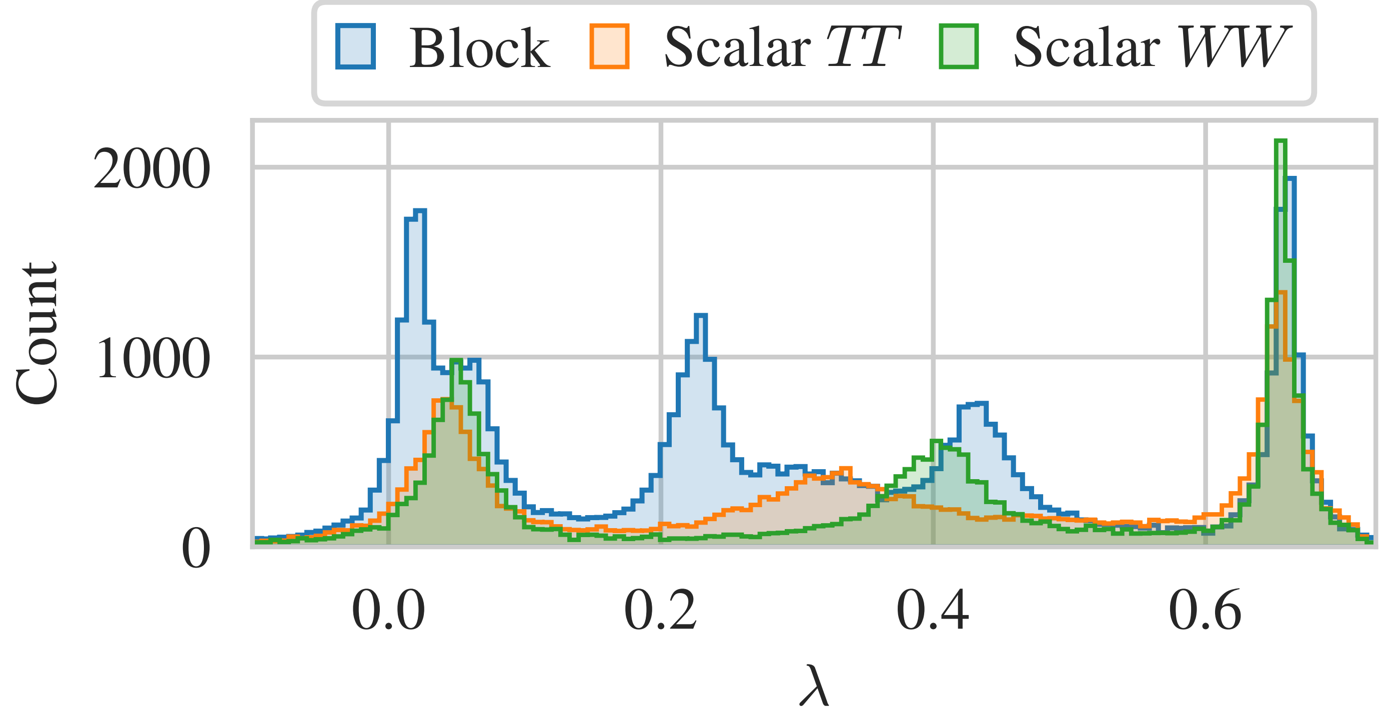
Eigenvalue multiplicity is non-trivial to determine with Lanczos

- Ritz values with overlapping residual bounds could be approximating the same eigenvalue

- Residual bounds key for ensuring an eigenvalue has been isolated, multiplicity ambiguous if they are statistically compatible

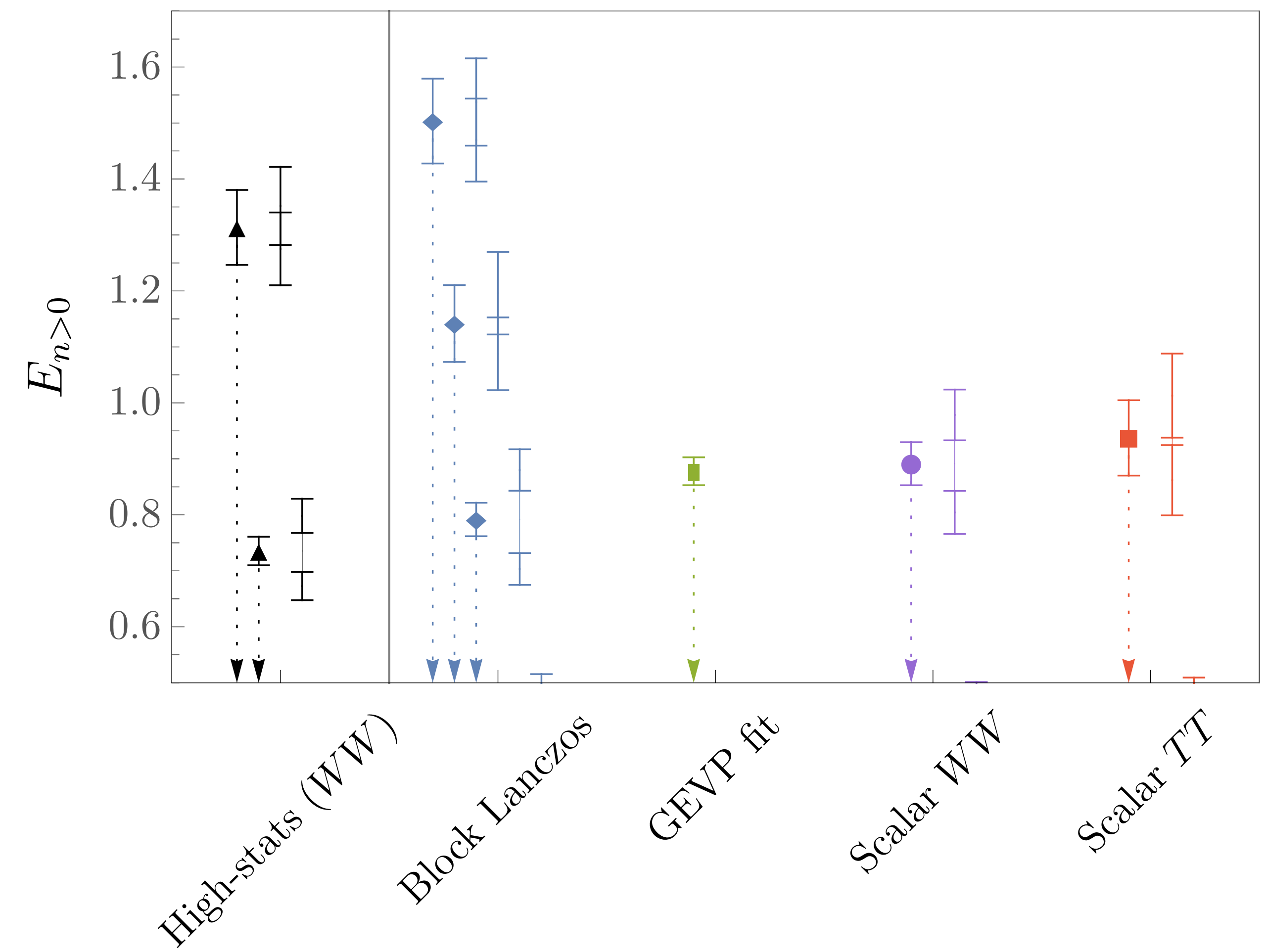


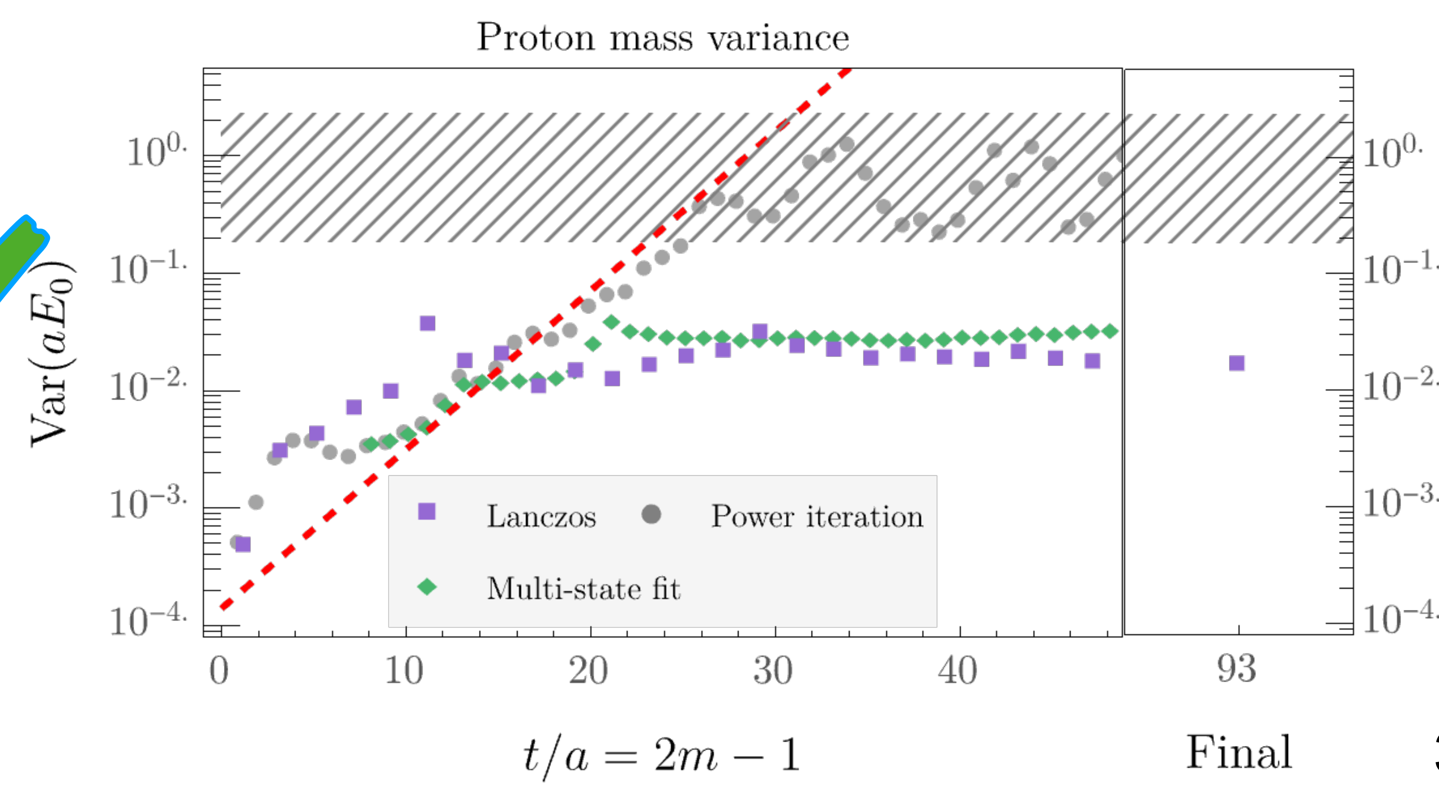
Physical states with sufficiently small interpolator overlap (e.g. $N\pi$ states here) are not resolved by Lanczos. Large sets of multi-hadron operators still needed to compute phase shifts with block Lanczos



What's new?

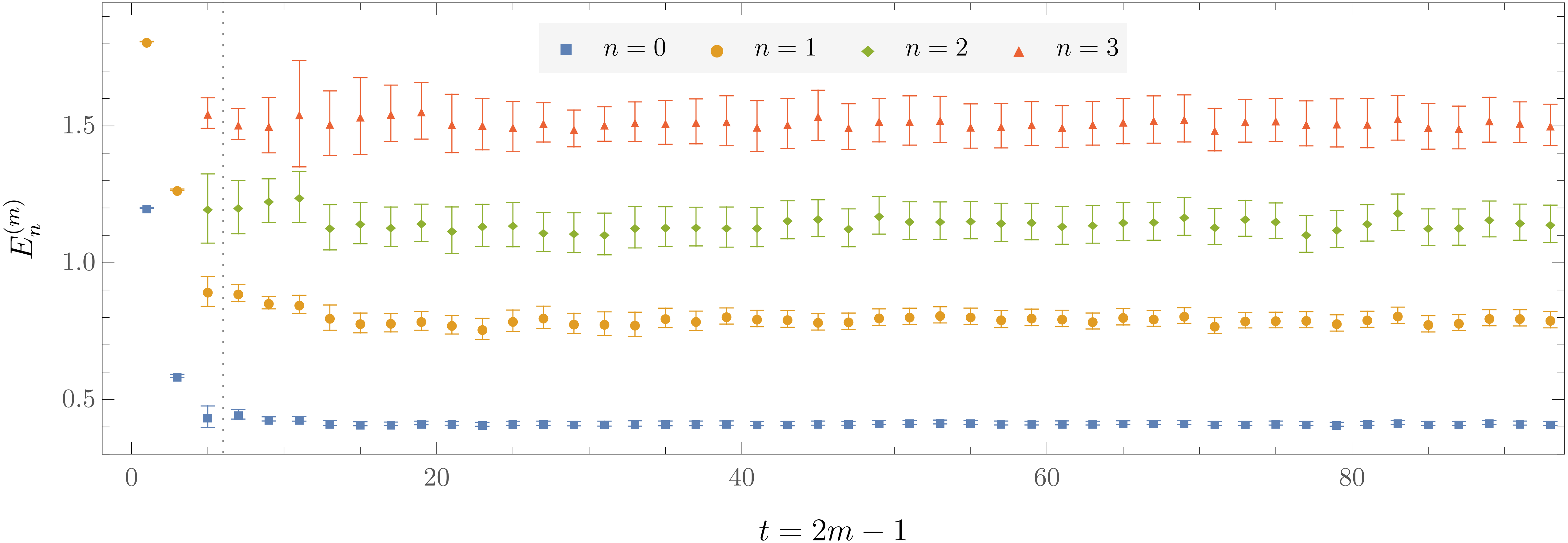
1. Fast convergence, guaranteed ✓
2. Residual bounds ✓
3. Spurious state filtering ✓
4. No fitting needed ✓
5. Asymptotically constant SNR ✓
6. Simple matrix element extractions ✓
7. More excited states, cleanly ✓





Thank you!

Daniel Hackett



Thank you!

Daniel Hackett

Cornelius Lanczos

Beresford Parlett

Jane Cullum

Backup - GEVP fits

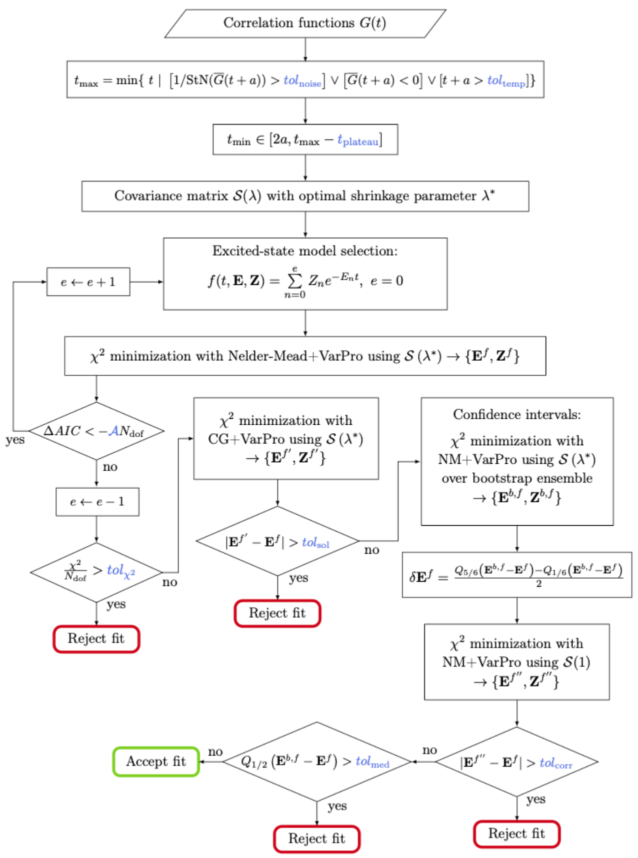
N_{states}	t_{min}	E_0	χ^2/N_{dof}	N_{dof}
3	2	0.403(12)	1.0	13
3	3	0.4221(50)	1.7	12
3	4	0.466(27)	29	11
3	5	0.4025(93)	1.3	10
3	6	0.4331(20)	4.1	9
3	7	0.4288(84)	4.0	8
3	8	0.4139(60)	1.1	7
3	9	0.4157(40)	1.4	6
3	10	0.4143(42)	1.6	5
3	11	0.4138(71)	1.9	4
3	12	0.416(10)	2.5	3
3	13	0.42(21)	3.7	2
3	14	0.422(12)	7.1	1
2	2	0.4293(30)	2.0	15
2	3	0.5213(51)	129	14
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

TABLE V. Multi-state fit results for the $n = 0$ FP GEVP correlator with $t_{\text{max}} = 19$; see the main text for details.

Fits performed for all t_{min}, AIC used to pick N_{states}, weighted average of all acceptable fit results:

Beane, MW et al, PRD 103 (2020)

Many other fitting strategies possible, should give statistically consistent results. Individual fits tabulated so you can compare with your favorite



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
3	4	1.09(13)	290	1
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

TABLE VI. Multi-state fit results for the $n = 1$ FP GEVP correlator with $t_{\text{max}} = 9$; see the main text for details.