Lanczos for matrix elements

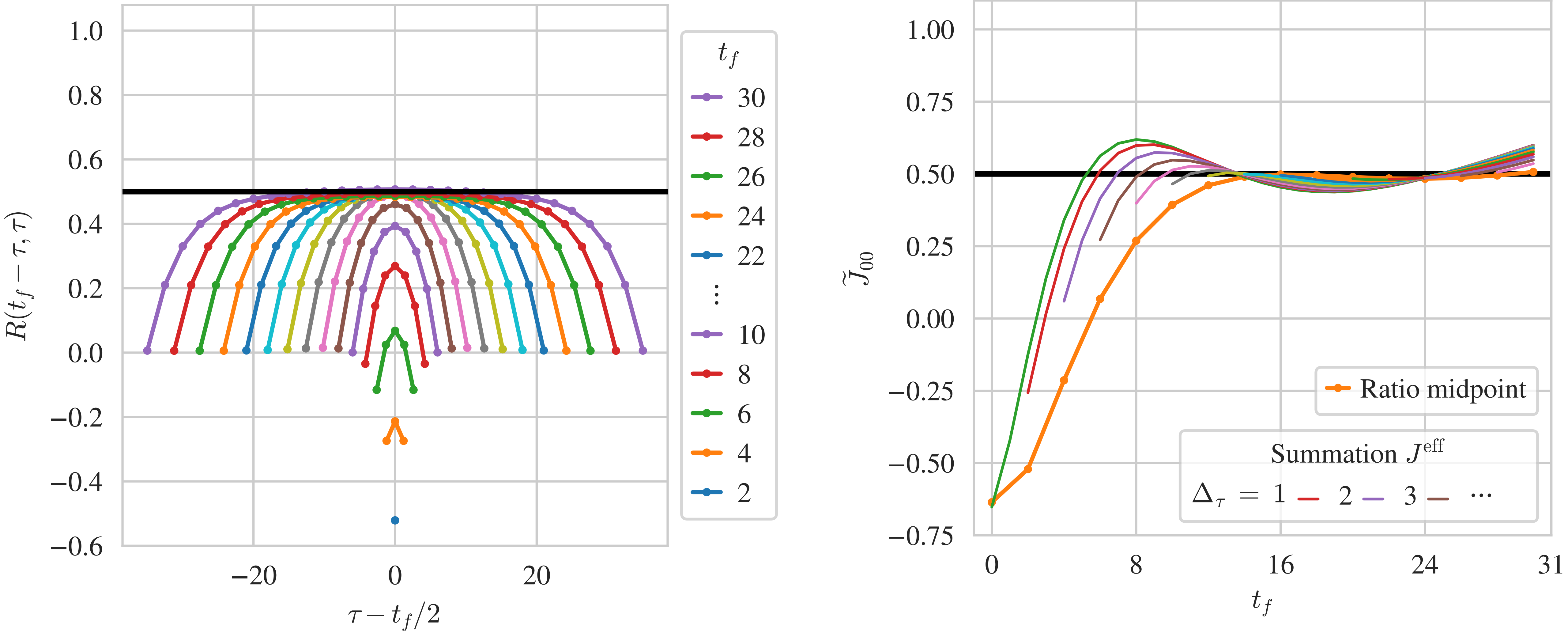
Lattice 2024
Aug 2, 2024
Liverpool, UK

Dan Hackett (FNAL)

Michael Wagman (FNAL)

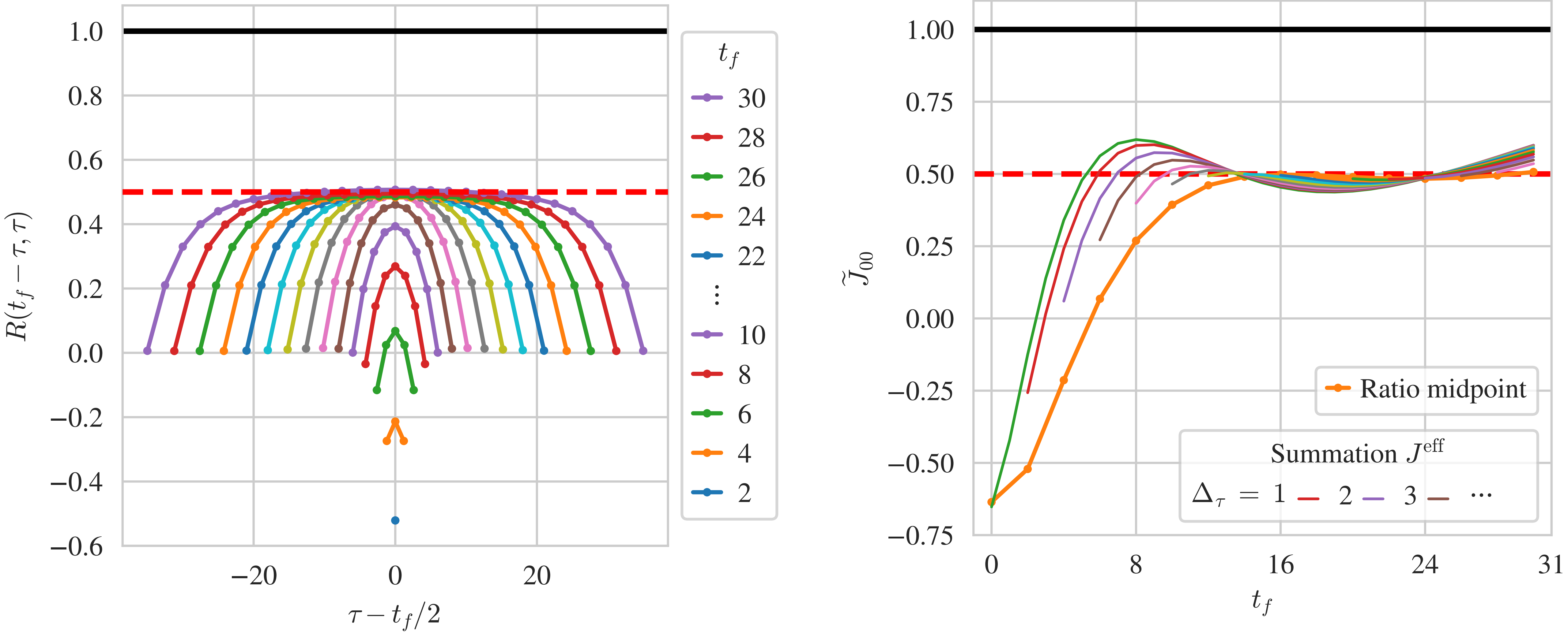
Matrix elements are hard

$$C(t) = \sum_k \frac{|Z_k|^2}{2E_k} e^{-E_k t} \qquad C^{3\mathrm{pt}}(\sigma,\tau) = \sum_{fi} \frac{Z_f Z_i}{4E_f E_i} \tilde{J}_{fi}\, e^{-E_f \tau - E_i \sigma} \qquad E_k = 0.1k \qquad Z_k = 1$$



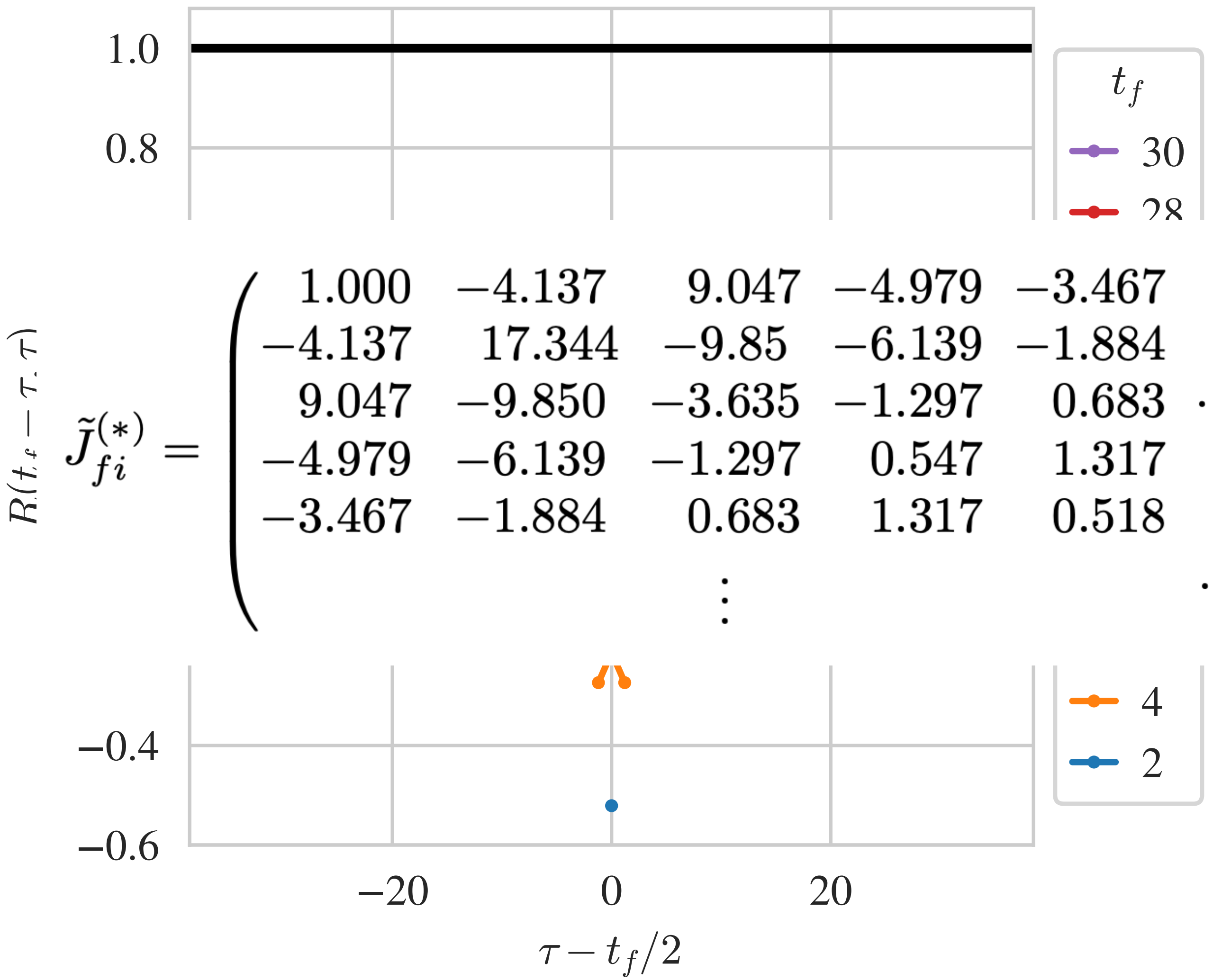
Matrix elements are hard

$$C(t) = \sum_k \frac{|Z_k|^2}{2E_k} e^{-E_k t} \qquad C^{3\mathrm{pt}}(\sigma,\tau) = \sum_{fi} \frac{Z_f Z_i}{4E_f E_i} \tilde{J}_{fi}\, e^{-E_f\tau - E_i\sigma} \qquad E_k = 0.1k \qquad Z_k = 1$$

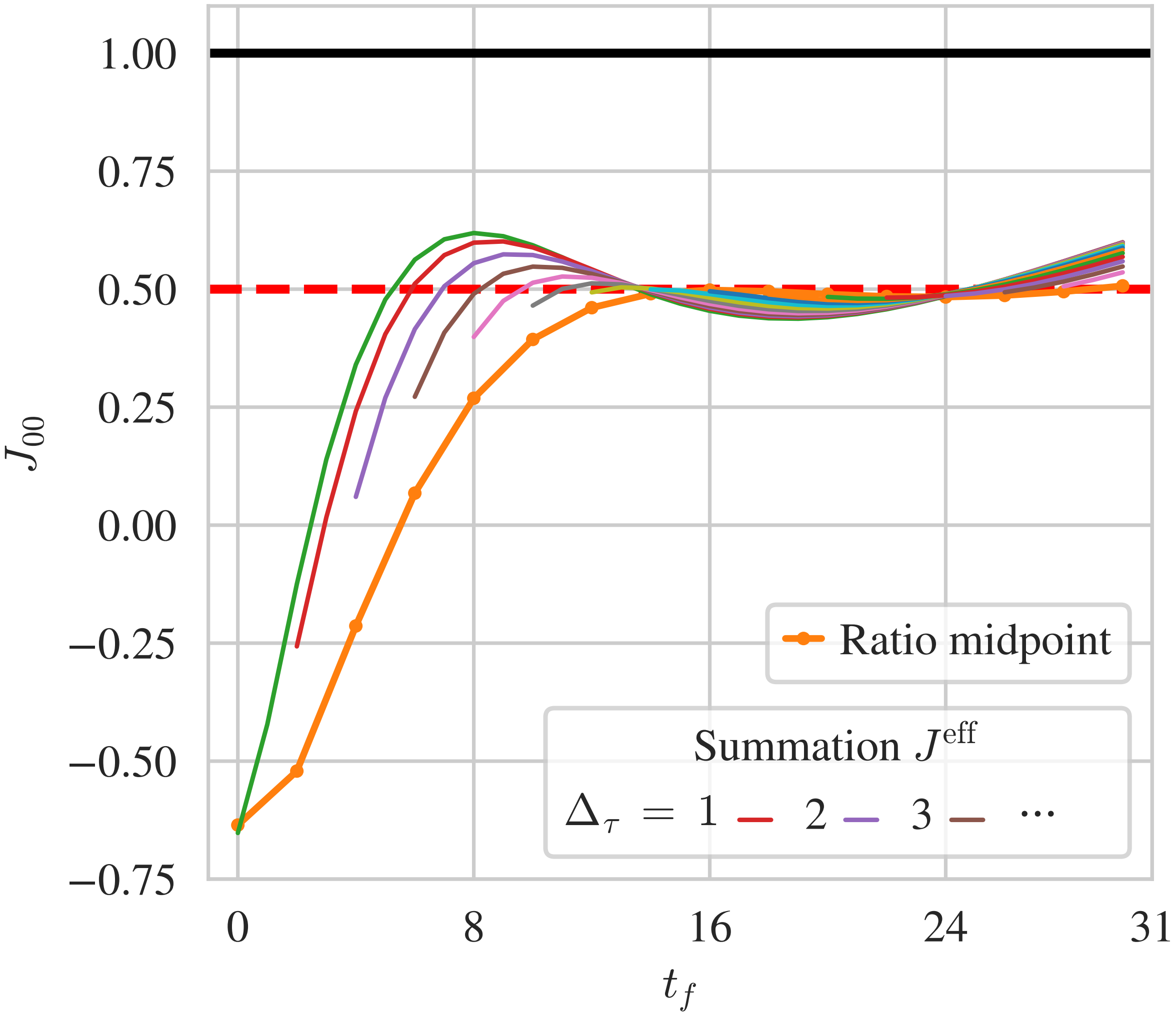


Matrix elements are hard

$C(t) = \sum_k \frac{|Z_k|^2}{2E_k} e^{-E_k t}$ $C^{3\mathrm{pt}}(\sigma,\tau) = \sum_{fi} \frac{Z_f Z_i}{4E_f E_i} \tilde{J}_{fi}\; e^{-E_f \tau - E_i \sigma}$ $E_k = 0.1k$ $Z_k = 1$

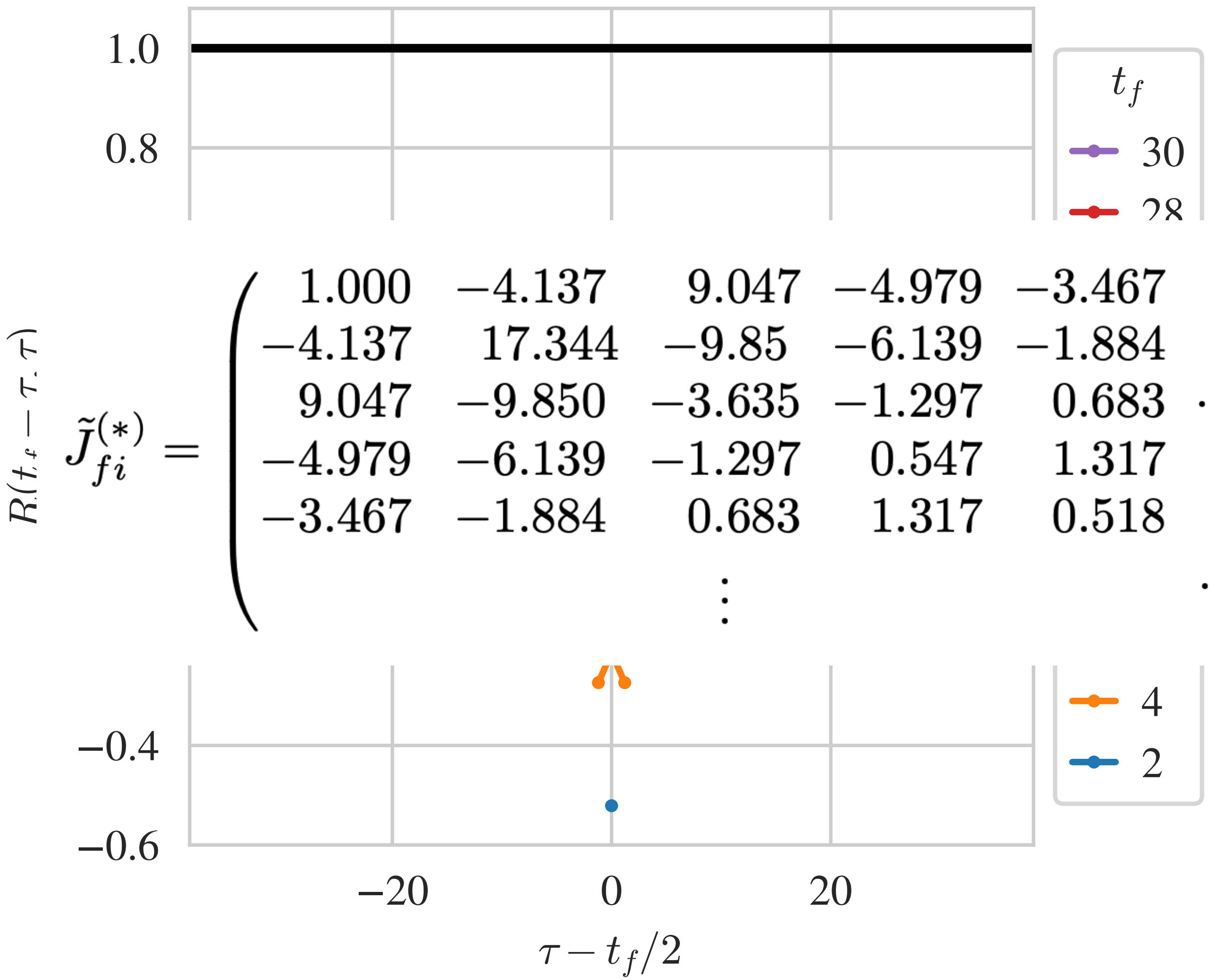


$$\tilde{J}^{(*)}_{fi} = \begin{pmatrix} 1.000 & -4.137 & 9.047 & -4.979 & -3.467 & \\ -4.137 & 17.344 & -9.85 & -6.139 & -1.884 & \\ 9.047 & -9.850 & -3.635 & -1.297 & 0.683 & \cdots \\ -4.979 & -6.139 & -1.297 & 0.547 & 1.317 & \\ -3.467 & -1.884 & 0.683 & 1.317 & 0.518 & \\ & & \vdots & & & \ddots \end{pmatrix}$$

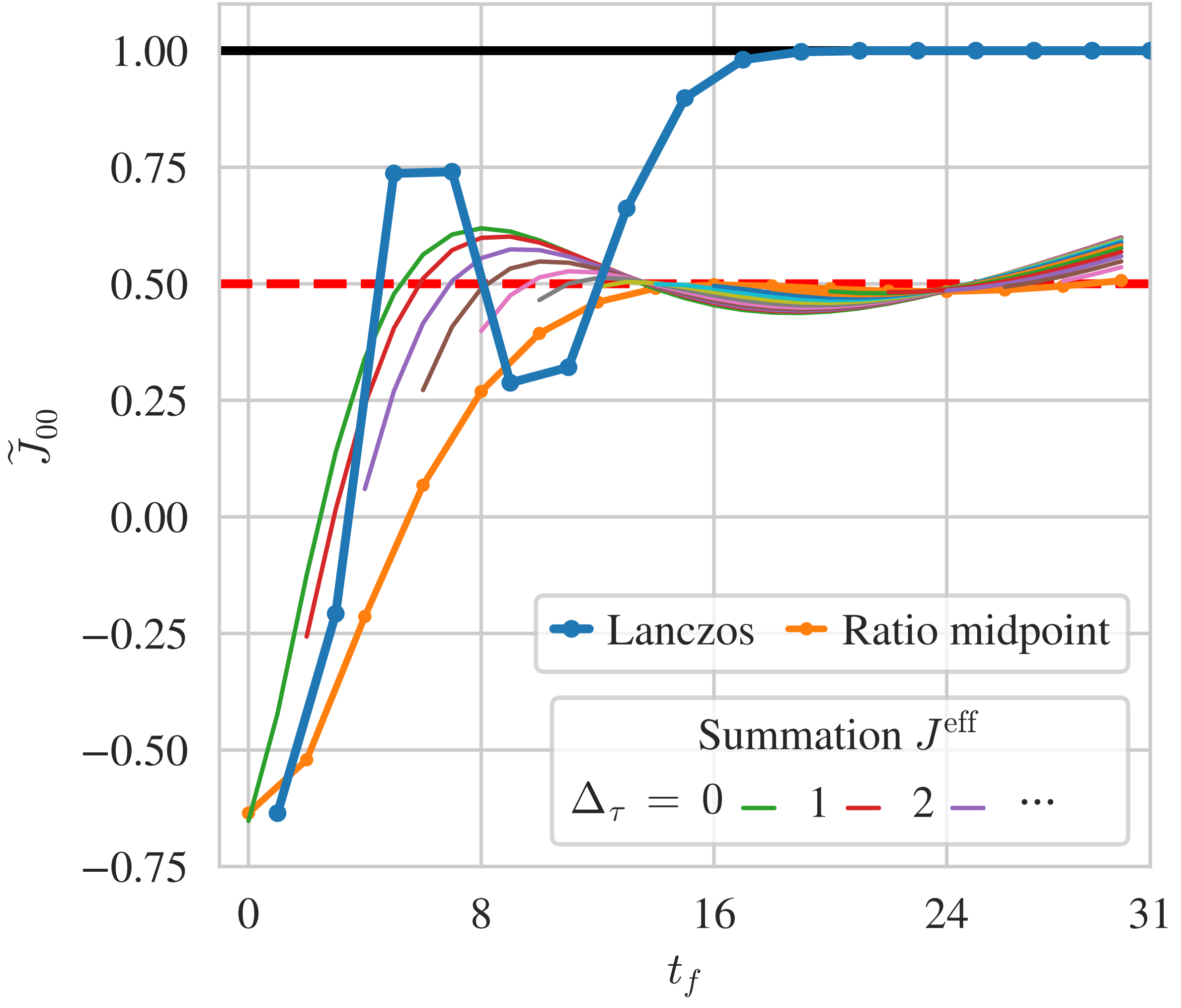


Matrix elements are hard

$$C(t) = \sum_k \frac{|Z_k|^2}{2E_k} e^{-E_k t} \qquad C^{3\mathrm{pt}}(\sigma,\tau) = \sum_{fi} \frac{Z_f Z_i}{4E_f E_i} \tilde{J}_{fi}\, e^{-E_f\tau - E_i\sigma} \qquad E_k = 0.1k \qquad Z_k = 1$$

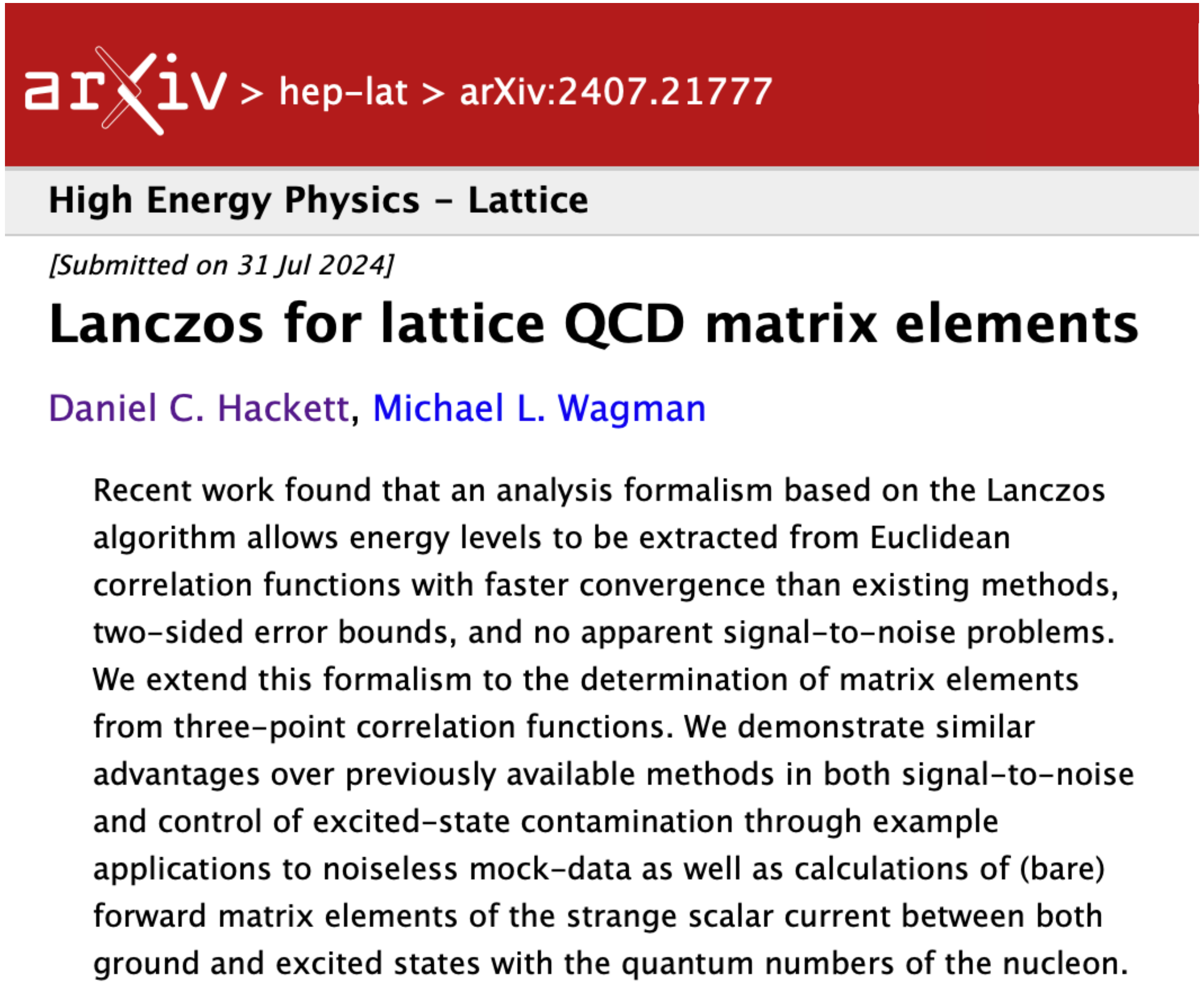


$$\tilde{J}_{fi}^{(*)} = \begin{pmatrix} 1.000 & -4.137 & 9.047 & -4.979 & -3.467 & \\ -4.137 & 17.344 & -9.85 & -6.139 & -1.884 & \\ 9.047 & -9.850 & -3.635 & -1.297 & 0.683 & \cdots \\ -4.979 & -6.139 & -1.297 & 0.547 & 1.317 & \\ -3.467 & -1.884 & 0.683 & 1.317 & 0.518 & \\ & & \vdots & & & \ddots \end{pmatrix}$$



Outline

- ~~Motivation~~
- Lanczos in Hilbert space
- Convergence in noiseless case
- Adjustments for noise
- Strange scalar current in the nucleon & excited states



arXiv > hep-lat > arXiv:2407.21777

High Energy Physics – Lattice

[Submitted on 31 Jul 2024]

Lanczos for lattice QCD matrix elements

Daniel C. Hackett, Michael L. Wagman

Recent work found that an analysis formalism based on the Lanczos algorithm allows energy levels to be extracted from Euclidean correlation functions with faster convergence than existing methods, two-sided error bounds, and no apparent signal-to-noise problems. We extend this formalism to the determination of matrix elements from three-point correlation functions. We demonstrate similar advantages over previously available methods in both signal-to-noise and control of excited-state contamination through example applications to noiseless mock-data as well as calculations of (bare) forward matrix elements of the strange scalar current between both ground and excited states with the quantum numbers of the nucleon.

Hilbert space & the Schrodinger picture

Interpolator excites the starting state

$$\bar{\psi}\,|\Omega\rangle = |\psi\rangle = \sum_k \langle k|\psi\rangle\,|k\rangle \equiv \sum_k Z_k |k\rangle$$

Hermitian transfer matrix

$$T = T^\dagger = \sum_k |k\rangle\,\lambda_k\,\langle k|$$

where $\lambda_k = e^{-E_k t}$

Euclidean time evolution

$$T^t|\psi\rangle = \sum_k Z_k\,e^{-E_k t}|k\rangle$$

$$\rightarrow Z_0 e^{-E_0 t}|0\rangle + (\mathrm{ESC})$$

Hilbert space & the Schrodinger picture

Interpolator excites the starting state

$$\bar{\psi}\,|\Omega\rangle = |\psi\rangle = \sum_k \langle k|\psi\rangle\,|k\rangle \equiv \sum_k Z_k|k\rangle$$

Hermitian transfer matrix

$$T = T^\dagger = \sum_k |k\rangle\,\lambda_k\,\langle k|$$

where $\lambda_k = e^{-E_k t}$

Euclidean time evolution

$$T^t|\psi\rangle = \sum_k Z_k\, e^{-E_k t}|k\rangle \rightarrow Z_0 e^{-E_0 t}|0\rangle + (\mathrm{ESC})$$

No access to states/operators, but can compute correlators:

$$C(t) = \langle\psi|T^t|\psi\rangle = \sum_k |Z_k|^2\,\lambda_k^t$$

$$C^{3\mathrm{pt}}(\sigma,\tau) = \langle\psi'|T^\sigma J T^\tau|\psi\rangle = \sum_{fi} Z_f'^*\,J_{fi}\,Z_i\,\lambda_f'^\sigma\,\lambda_i^\tau$$

Running the Lanczos algorithm in Hilbert space

Start:

$$|v_1\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} = \frac{|\psi\rangle}{\sqrt{C(0)}}$$

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

Running the Lanczos algorithm in Hilbert space

Start:

$$|v_1\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} = \frac{|\psi\rangle}{\sqrt{C(0)}}$$

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

Iterate:

1. Apply T and orthogonalize

$$|\tilde{v}_{j+1}\rangle = (T - \alpha_j)|v_j\rangle + \beta_j|v_{j-1}\rangle$$

2. Normalize & compute α

$$\beta_{j+1}^2 = \langle\tilde{v}_{j+1}|\tilde{v}_{j+1}\rangle$$

$$|v_{j+1}\rangle = \frac{1}{\beta_{j+1}}|\tilde{v}_{j=1}\rangle$$

$$\alpha_{j+1} = \langle v_{j+1}|T|v_{j+1}\rangle$$

Running the Lanczos algorithm in Hilbert space

Start:

$$|v_1\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} = \frac{|\psi\rangle}{\sqrt{C(0)}}$$

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

Iterate:

1. Apply T and orthogonalize
$$|\tilde{v}_{j+1}\rangle = (T - \alpha_j)|v_j\rangle + \beta_j|v_{j-1}\rangle$$
2. Normalize & compute α
$$\beta_{j+1}^2 = \langle\tilde{v}_{j+1}|\tilde{v}_{j+1}\rangle$$
$$|v_{j+1}\rangle = \frac{1}{\beta_{j+1}}|\tilde{v}_{j=1}\rangle$$
$$\alpha_{j+1} = \langle v_{j+1}|T|v_{j+1}\rangle$$

After m iterations:

Lanczos vectors $|v_j\rangle$ for $j = 1, \ldots, m$

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

T in Lanczos vector basis:

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

Running the Lanczos algorithm in Hilbert space

Start:

$$|v_1\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} = \frac{|\psi\rangle}{\sqrt{C(0)}}$$

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

Iterate:

1. Apply T and orthogonalize

$$|\tilde{v}_{j+1}\rangle = (T - \alpha_j)|v_j\rangle + \beta_j|v_{j-1}\rangle$$

2. Normalize & compute α

$$\beta_{j+1}^2 = \langle\tilde{v}_{j+1}|\tilde{v}_{j+1}\rangle$$

$$|v_{j+1}\rangle = \frac{1}{\beta_{j+1}}|\tilde{v}_{j=1}\rangle$$

$$\alpha_{j+1} = \langle v_{j+1}|T|v_{j+1}\rangle$$

After m iterations:

Lanczos vectors $|v_j\rangle$ for $j = 1, \dots, m$

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

T in Lanczos vector basis:

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

Can compute from

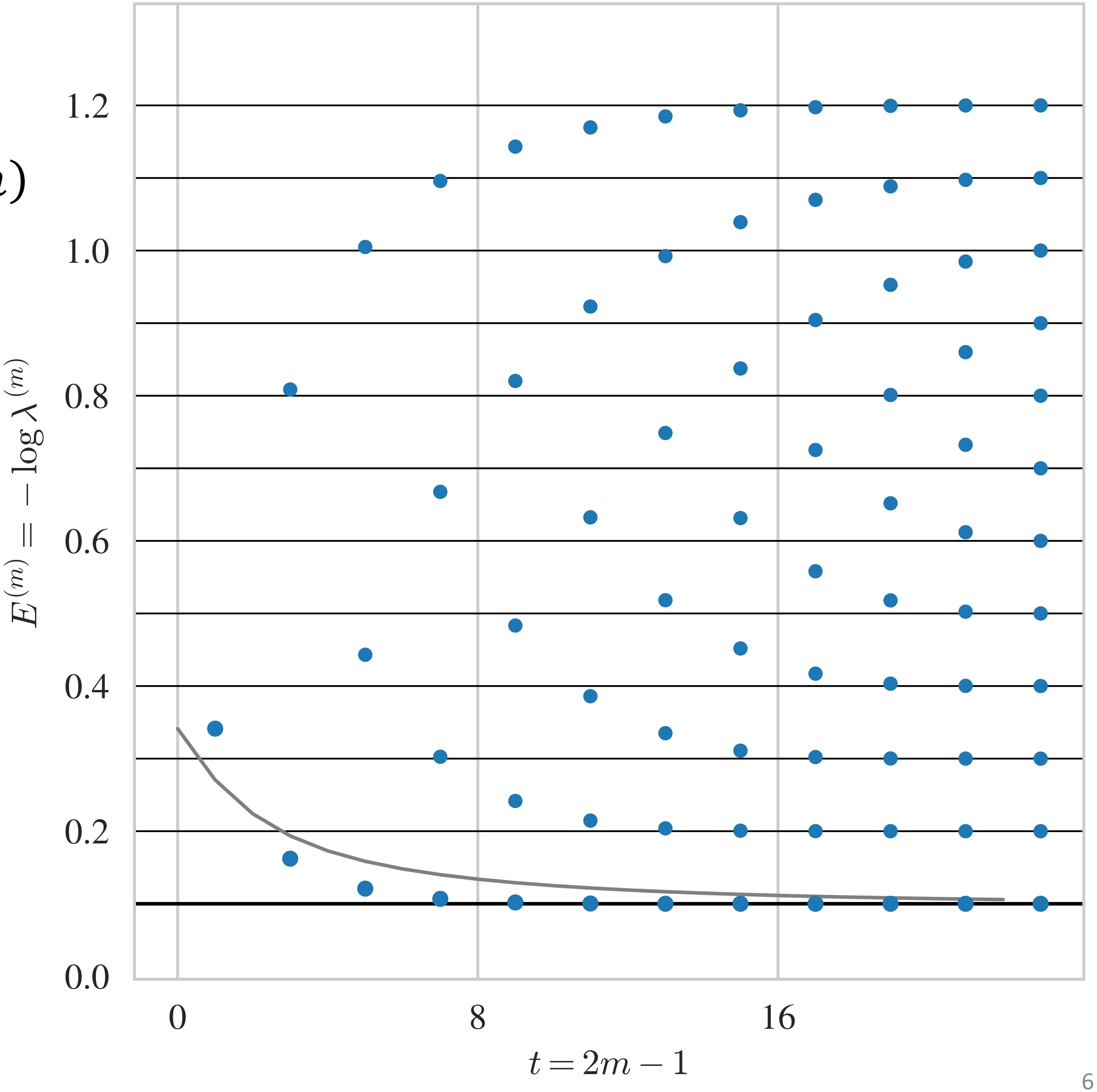
$C(t) = \langle\psi|T^t|\psi\rangle$!

(via a recursion relation)

Diagonalizing T

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

Ritz value $\approx e^{-E_k t}$



Diagonalizing T

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

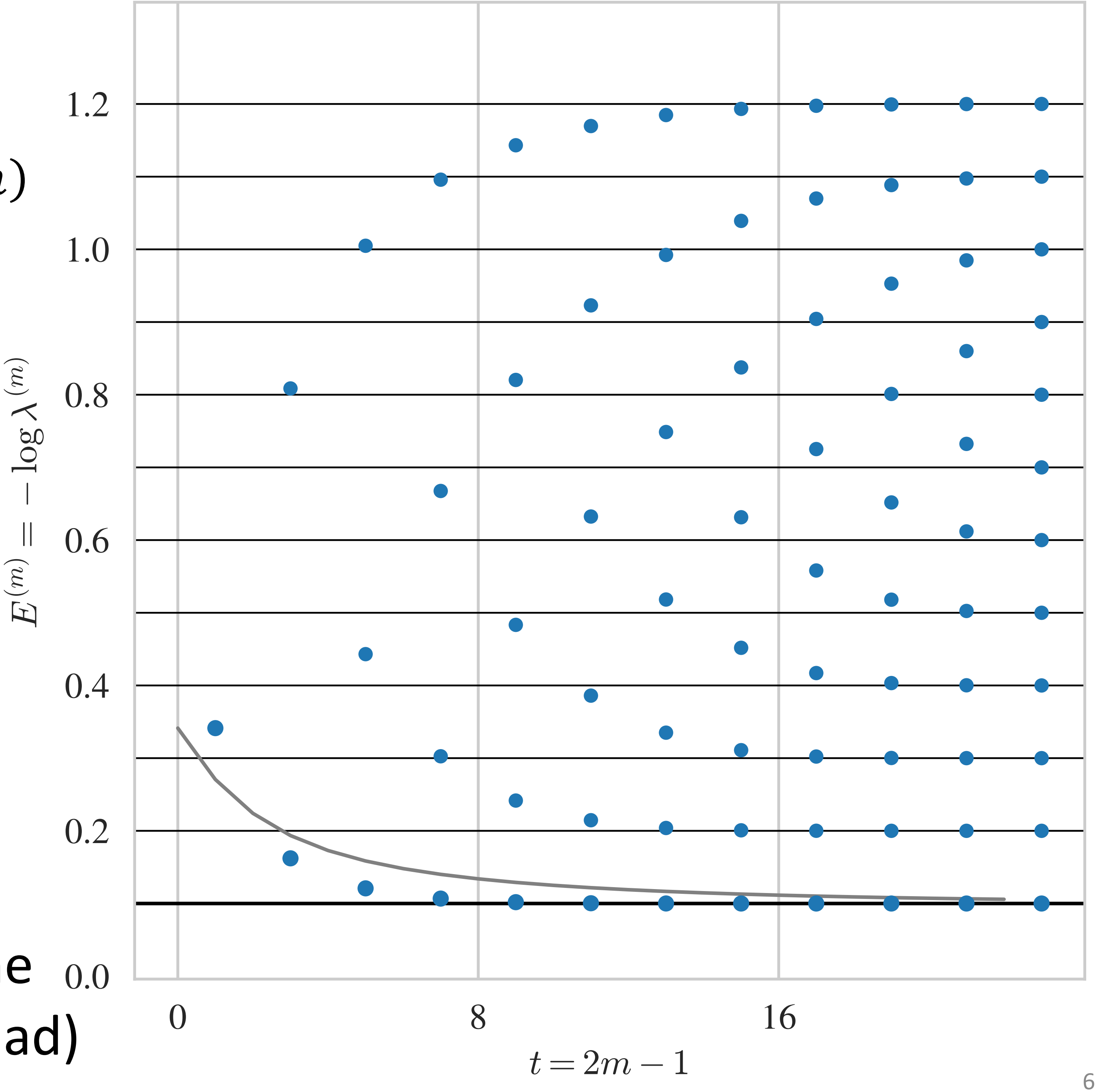
Ritz value $\approx e^{-E_k t}$

$$\min_{\lambda} \left|\lambda_k^{(m)} - \lambda\right|^2 \leq B_k^{(m)}$$

↑ Over all true eigenstates

↑ Computable from $C(t)$

Note: eigenvectors converge at same rate as eigenvalues! (Kaniel-Page-Saad)



Diagonalizing T

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

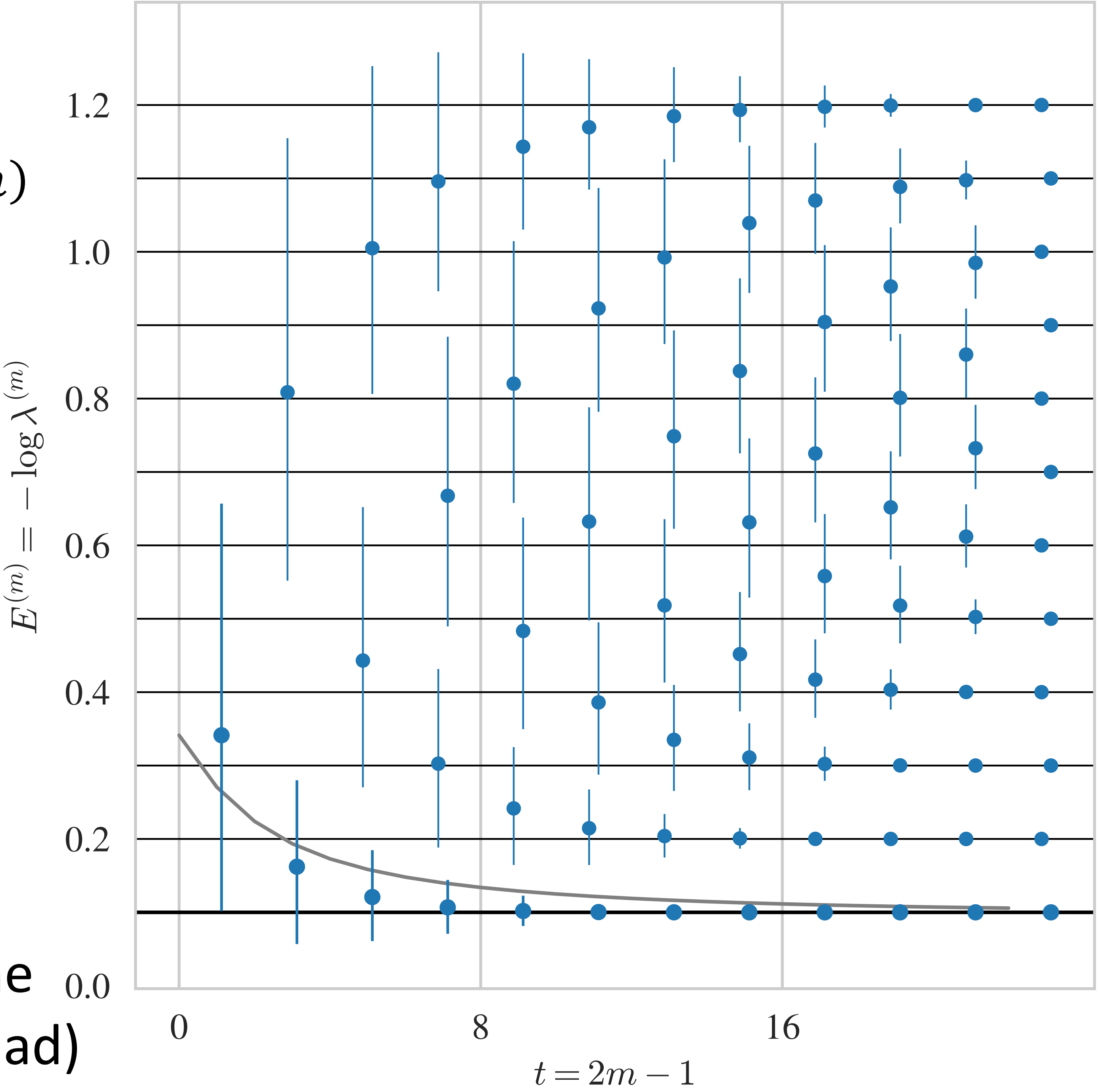
Ritz value $\approx e^{-E_k t}$

$$\min_{\lambda} \left| \lambda_k^{(m)} - \lambda \right|^2 \leq B_k^{(m)}$$

Over all true eigenstates

Computable from $C(t)$

Note: eigenvectors converge at same rate as eigenvalues! (Kaniel-Page-Saad)



Ritz vectors & changes of basis

Ritz vector $\approx |k\rangle$

Lanczos approximates

$$T \approx T^{(m)} = \sum_{ij} |v_i\rangle \, T_{ij}^{(m)} \langle v_j| = \sum_k \left|y_k^{(m)}\right\rangle \, \lambda_k^{(m)} \left\langle y_k^{(m)}\right|$$

Ritz vectors & changes of basis

Ritz vector $\approx |k\rangle$

Lanczos approximates

$$T \approx T^{(m)} = \sum_{ij} |v_i\rangle\, T_{ij}^{(m)} \langle v_j| = \sum_k \left|y_k^{(m)}\right\rangle \lambda_k^{(m)} \left\langle y_k^{(m)}\right|$$

Krylov → Lanczos: $\quad |v_j\rangle = \sum_t K_{jt}\, T^t \frac{|\psi\rangle}{\sqrt{C(0)}}$

Can compute K_{jt} from α_j, β_j via another recursion

Ritz vectors & changes of basis

Ritz vector ≈ $|k\rangle$

Lanczos approximates

$$T \approx T^{(m)} = \sum_{ij} |v_i\rangle\, T_{ij}^{(m)} \langle v_j| = \sum_k \left|y_k^{(m)}\right\rangle \lambda_k^{(m)} \left\langle y_k^{(m)}\right|$$

Krylov → Lanczos: $$|v_j\rangle = \sum_t K_{jt}\, T^t \frac{|\psi\rangle}{\sqrt{C(0)}}$$

Can compute K_{jt} from α_j, β_j via another recursion

Lanczos → Ritz: $$\left|y_k^{(m)}\right\rangle = \sum_j \omega_{jk}^{(m)}\, |v_j\rangle$$

Ritz vectors & changes of basis

Ritz vector $\approx |k\rangle$

Lanczos approximates

$$T \approx T^{(m)} = \sum_{ij} |v_i\rangle\, T^{(m)}_{ij} \langle v_j| = \sum_k \left|y_k^{(m)}\right\rangle \lambda_k^{(m)} \left\langle y_k^{(m)}\right|$$

Krylov → Lanczos: $\quad |v_j\rangle = \sum_t K_{jt}\, T^t \frac{|\psi\rangle}{\sqrt{C(0)}}$

Can compute K_{jt} from α_j, β_j via another recursion

Lanczos → Ritz: $\quad \left|y_k^{(m)}\right\rangle = \sum_j \omega_{jk}^{(m)}\, |v_j\rangle$

⇒ Krylov → Ritz

$$\left|y_k^{(m)}\right\rangle = \sum_{jt} \omega_{jk}^{(m)}\, K_{jt}\, T^t \frac{|\psi\rangle}{\sqrt{C(0)}} \equiv \sum_t P_{kt}^{(m)}\, T^t \frac{|\psi\rangle}{\sqrt{C(0)}}$$

Matrix elements with Lanczos

Construct initial/final state Ritz vectors

$$C(t) \to |y_i\rangle \qquad C'(t) \to |y_f'\rangle$$

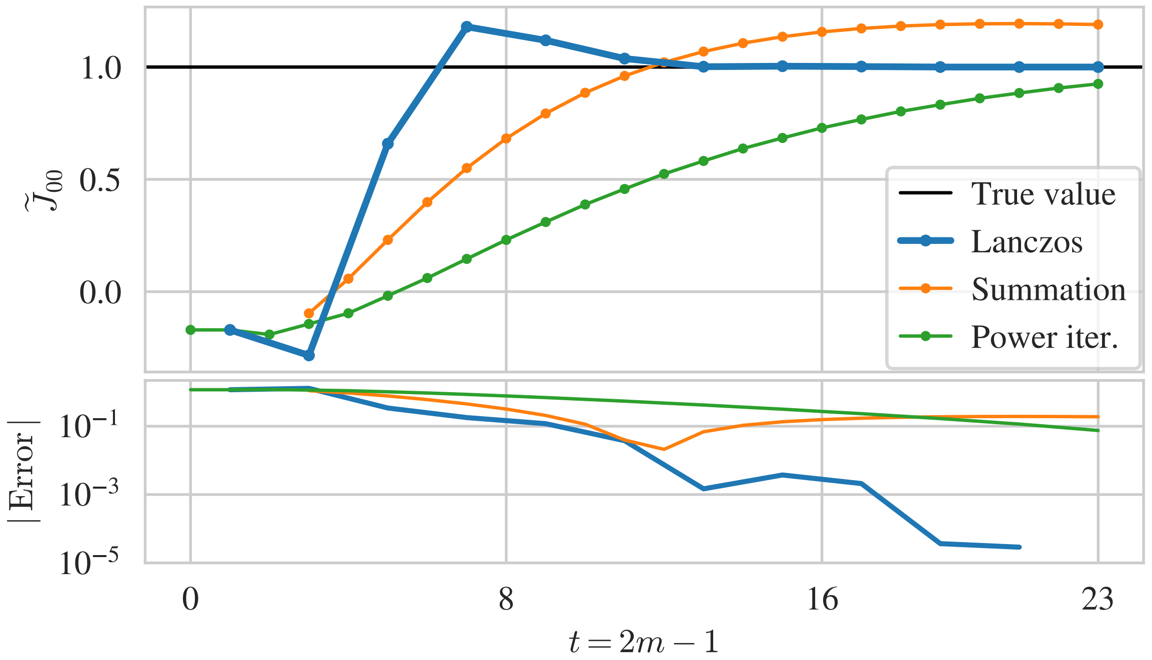
$$\langle y_f' |\, J \,| y_i \rangle = \sum_{\sigma\tau} P'^{\,*}_{f\sigma} \, \langle v_1 | T^{\sigma} J \, T^{\tau} | v_1 \rangle P_{i\tau}$$

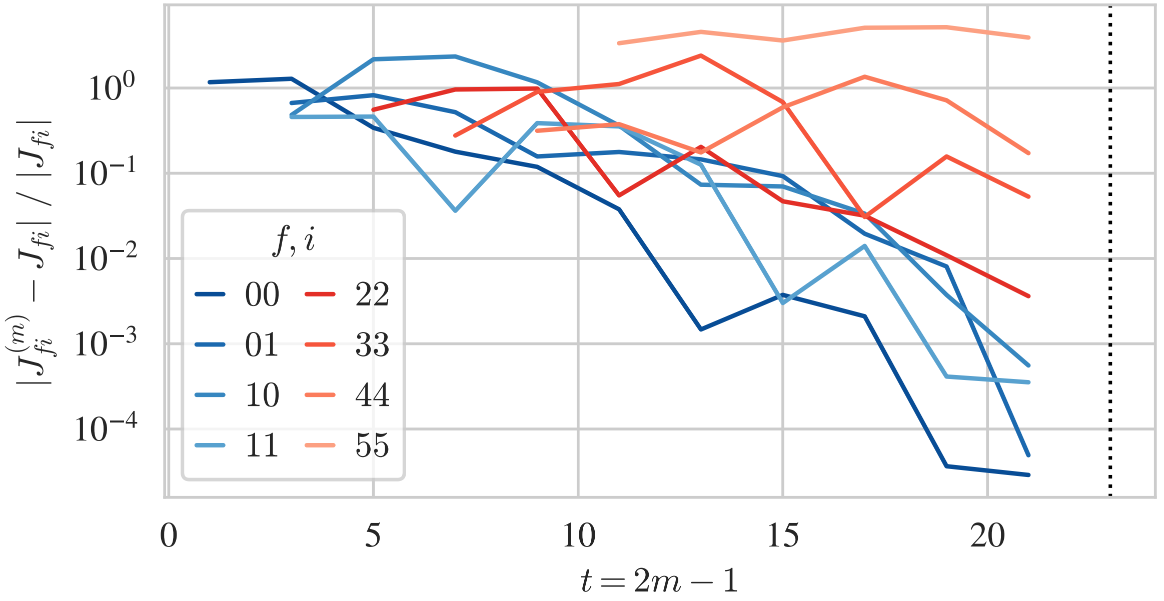
$$= \sum_{\sigma\tau} P'^{\,*}_{f\sigma} \, \frac{C^{3\mathrm{pt}}(\sigma,\tau)}{\sqrt{C'(0)C(0)}} \, P_{i\tau}$$

Just matrix multiplication!

Noiseless case: solves an $N_t/2$ state system after incorporating all N_t points

i.e. finds all $(N_t/2)^2$ matrix elements ***exactly***





Lanczos in the presence of noise

Noise → must use “oblique Lanczos” (non-Hermitian T / off-diag $C(t)$)

$\left|v_1^R\right\rangle = \left|v_1^L\right\rangle$ but $\left|v_j^R\right\rangle \neq \left|v_j^L\right\rangle$ and $\left|y_j^{R(m)}\right\rangle \neq \left|y_j^{L(m)}\right\rangle$

Must filter out ”spurious” states

Lanczos in the presence of noise

Noise → must use "oblique Lanczos" (non-Hermitian T / off-diag $C(t)$)

$\left|v_1^R\right\rangle = |v_1^L\rangle$ but $\left|v_j^R\right\rangle \neq |v_j^L\rangle$ and $\left|y_j^{R(m)}\right\rangle \neq \left|y_j^{L(m)}\right\rangle$

Must filter out "spurious" states

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

Lanczos in the presence of noise

Noise → must use "oblique Lanczos" (non-Hermitian T / off-diag $C(t)$)

$$\left|v_1^R\right\rangle = \left|v_1^L\right\rangle \quad \text{but} \quad \left|v_j^R\right\rangle \neq \left|v_j^L\right\rangle \quad \text{and} \quad \left|y_j^{R(m)}\right\rangle \neq \left|y_j^{L(m)}\right\rangle$$

Must filter out "spurious" states

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

Hermitian subspace

Real $\lambda_k^{(m)}$, $|\mathrm{y}^R\rangle = |y^L\rangle = |y\rangle$

Physically interpretable

Lanczos in the presence of noise

Noise → must use "oblique Lanczos" (non-Hermitian T / off-diag $C(t)$)

$$\left|v_1^R\right\rangle = \left|v_1^L\right\rangle \quad \text{but} \quad \left|v_j^R\right\rangle \neq \left|v_j^L\right\rangle \quad \text{and} \quad \left|y_j^{R(m)}\right\rangle \neq \left|y_j^{L(m)}\right\rangle$$

Must filter out "spurious" states

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

Noise artifact states

Complex $\lambda_k^{(m)}$, $|\mathrm{y}^R\rangle \neq |y^L\rangle$

$C(t) = \langle\psi|\left(T^{(m)}\right)^t|\psi\rangle$ for $t \leq 2m - 1$

(i.e. all data incorporated)

Intuition: need oscillating modes to replicate noisy correlator

Hermitian subspace

Real $\lambda_k^{(m)}$, $|\mathrm{y}^R\rangle = |y^L\rangle = |y\rangle$

Physically interpretable

Lanczos in the presence of noise

Noise → must use “oblique Lanczos” (non-Hermitian T / off-diag $C(t)$)

$|v_1^R\rangle = |v_1^L\rangle$ but $|v_j^R\rangle \neq |v_j^L\rangle$ and $\left|y_j^{R(m)}\right\rangle \neq \left|y_j^{L(m)}\right\rangle$

Must filter out ”spurious” states

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

Noise artifact states

Complex $\lambda_k^{(m)}$, $|y^R\rangle \neq |y^L\rangle$

$C(t) = \langle\psi|\left(T^{(m)}\right)^t|\psi\rangle$ for $t \leq 2m-1$

(i.e. all data incorporated)

Intuition: need oscillating modes to replicate noisy correlator

Hermitian subspace

Real $\lambda_k^{(m)}$, $|y^R\rangle = |y^L\rangle = |y\rangle$

Physically interpretable

+ Cullum-Willoughby test (see Mike Wagman’s talk)

Strange scalar matrix elements of the nucleon (& exciteds)

Action: Luscher-Weisz, 2+1 stout-smeared clover

$M_\pi \approx 170$ MeV $\quad a \approx 0.09$ fm

Nucleon $\quad \chi \sim (u\, C\, \gamma_5\, d)\, u$

Quarks smeared to $r = 4.5$

Strange scalar current $\quad J \sim \bar{s}s$

Forward ($\boldsymbol{p} = \boldsymbol{p}' = \boldsymbol{q} = \boldsymbol{0}$) matrix elements

1381 configs × 1024 sources ≈ 1.4×10^6 meas

Fully disconnected $C^{3\mathrm{pt}}$

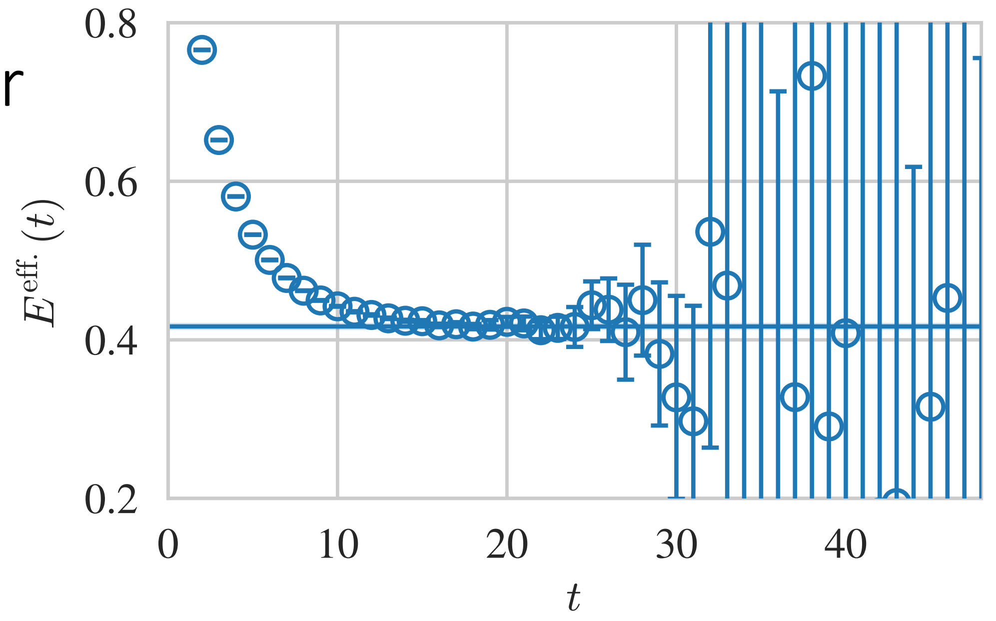
Hierarchical probing w/ 512 Hadamards

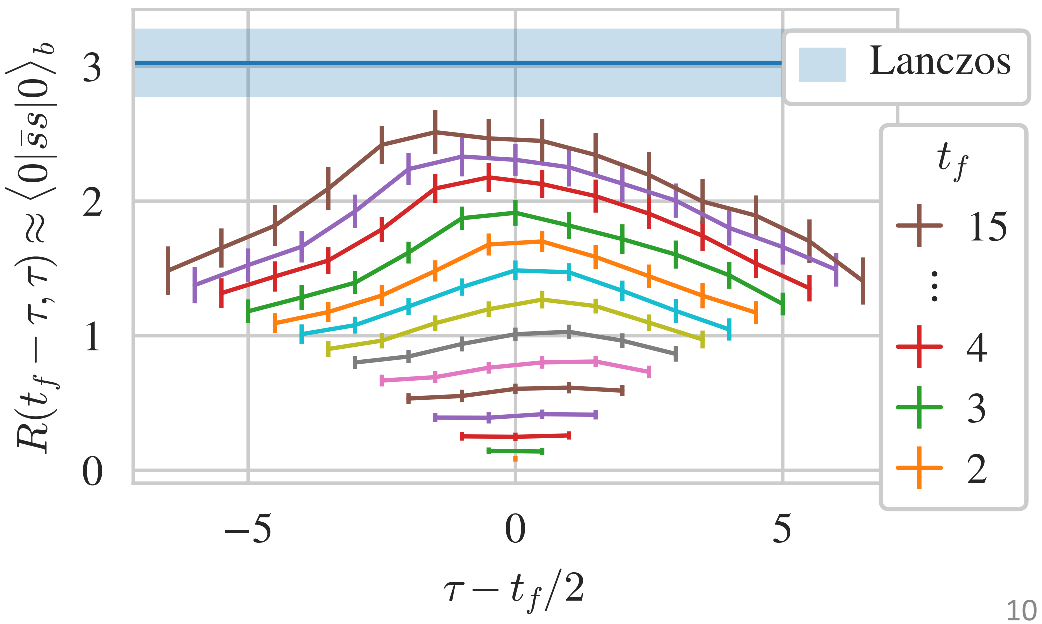
One shot Z_4 noise per config

No deflation / low-mode subtraction

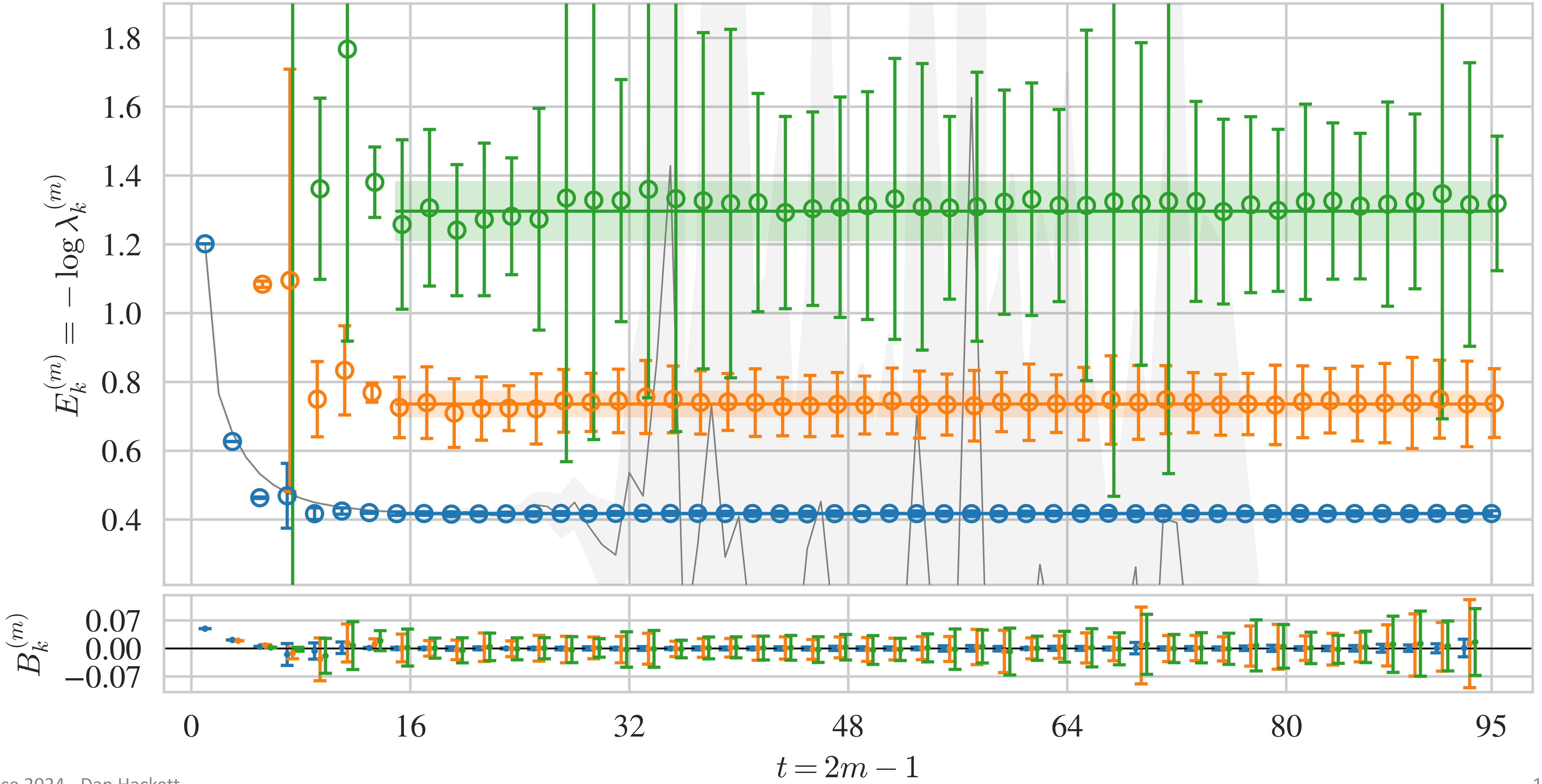
Not renormalized!

Mixes w/ light quarks

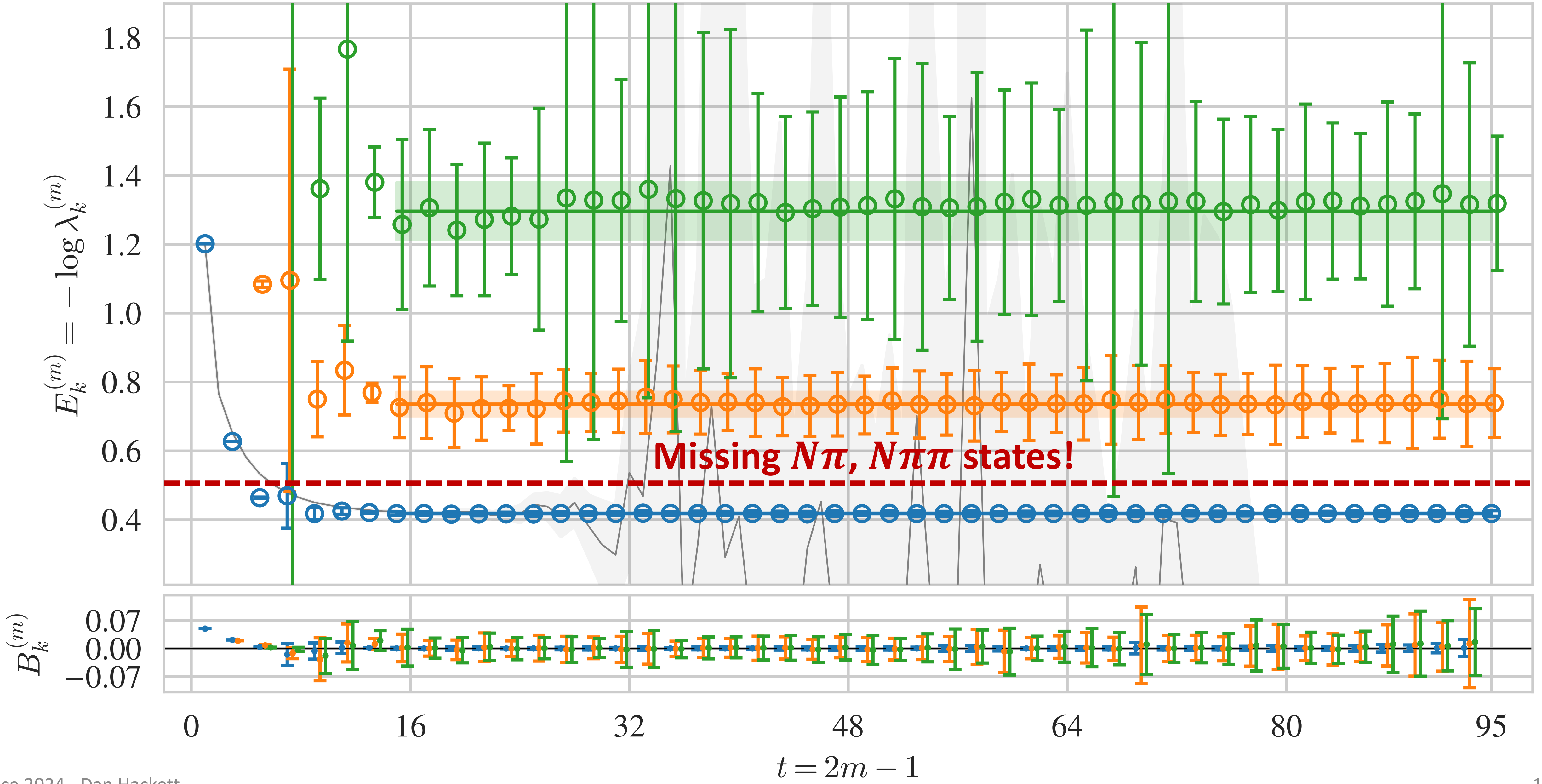




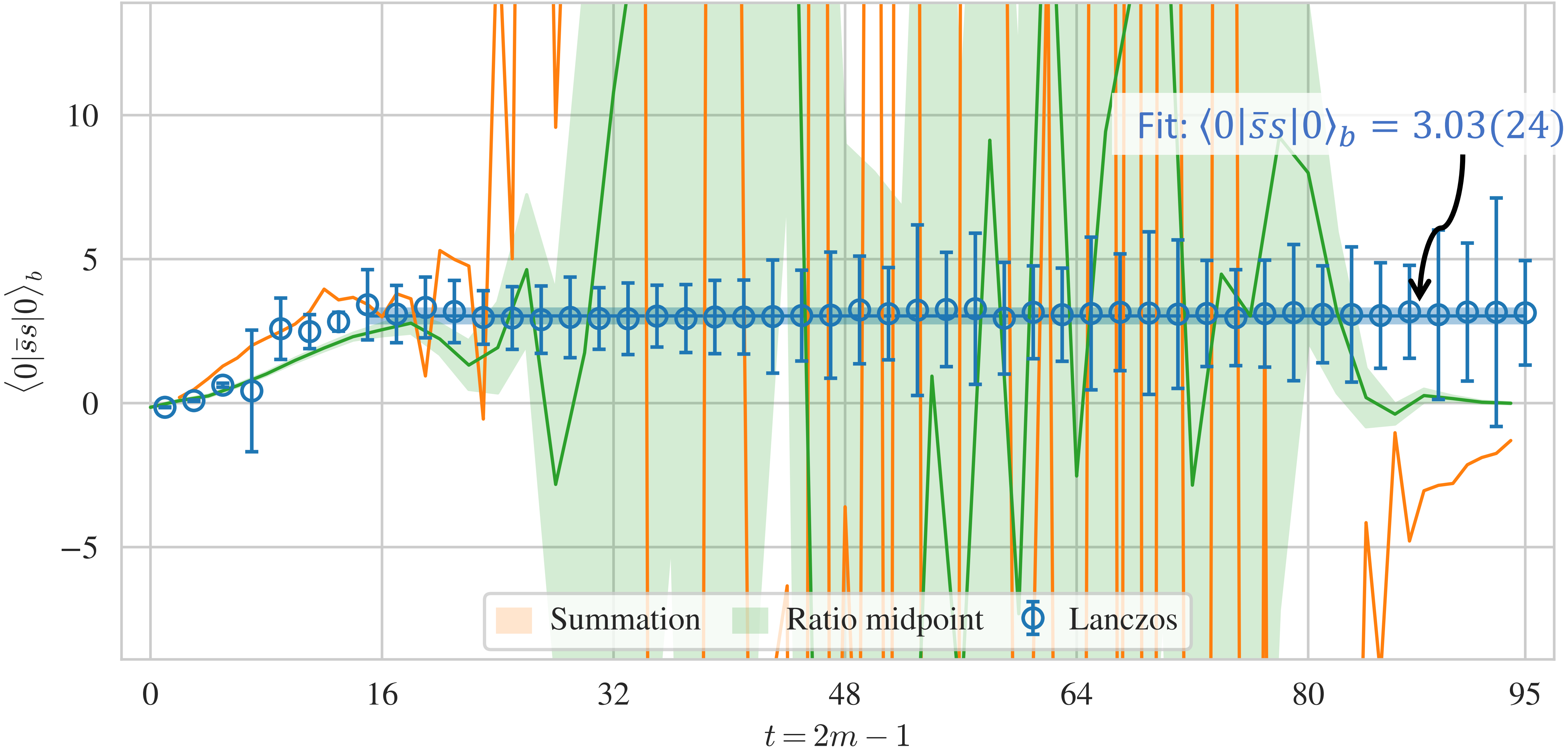
Results: nucleon spectrum



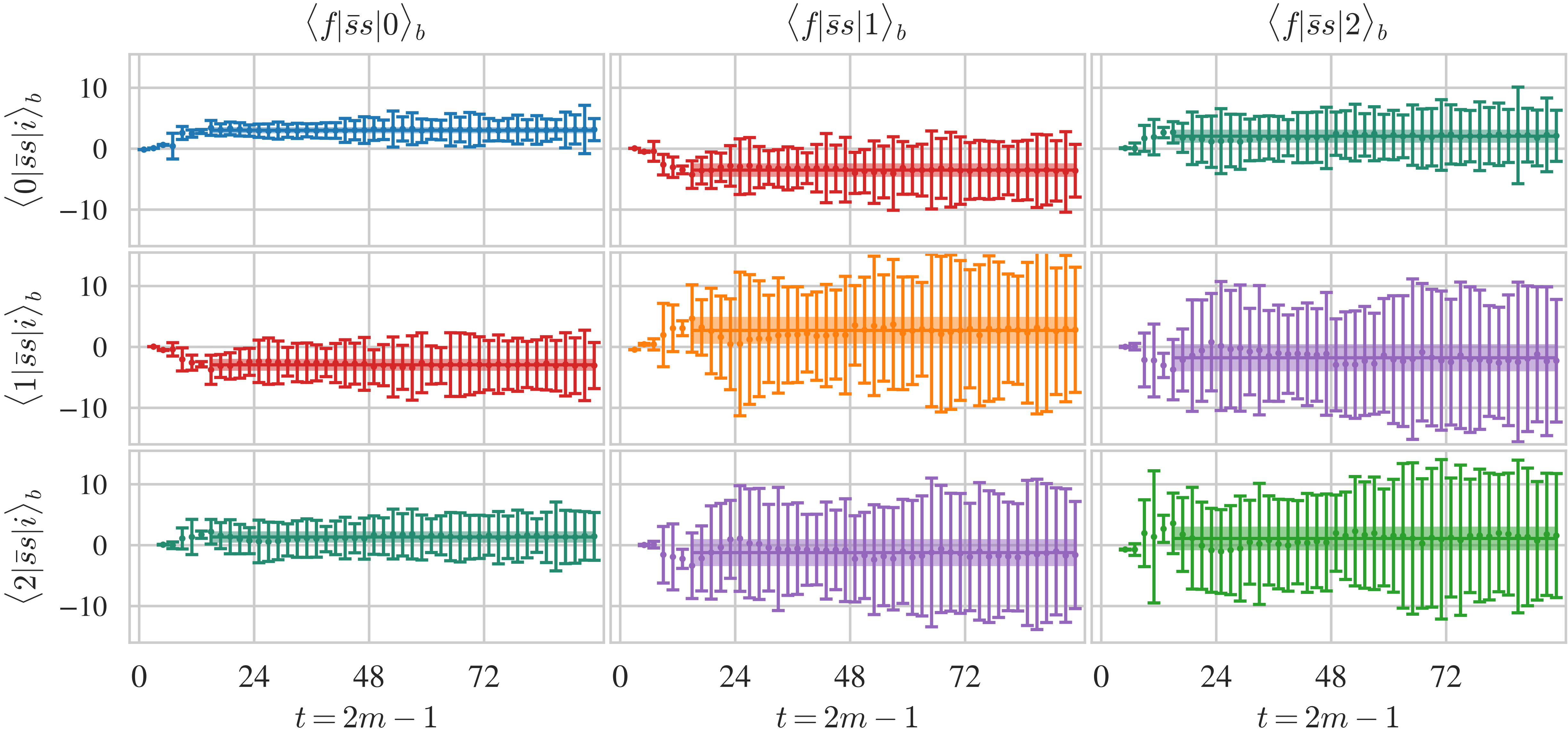
Results: nucleon spectrum



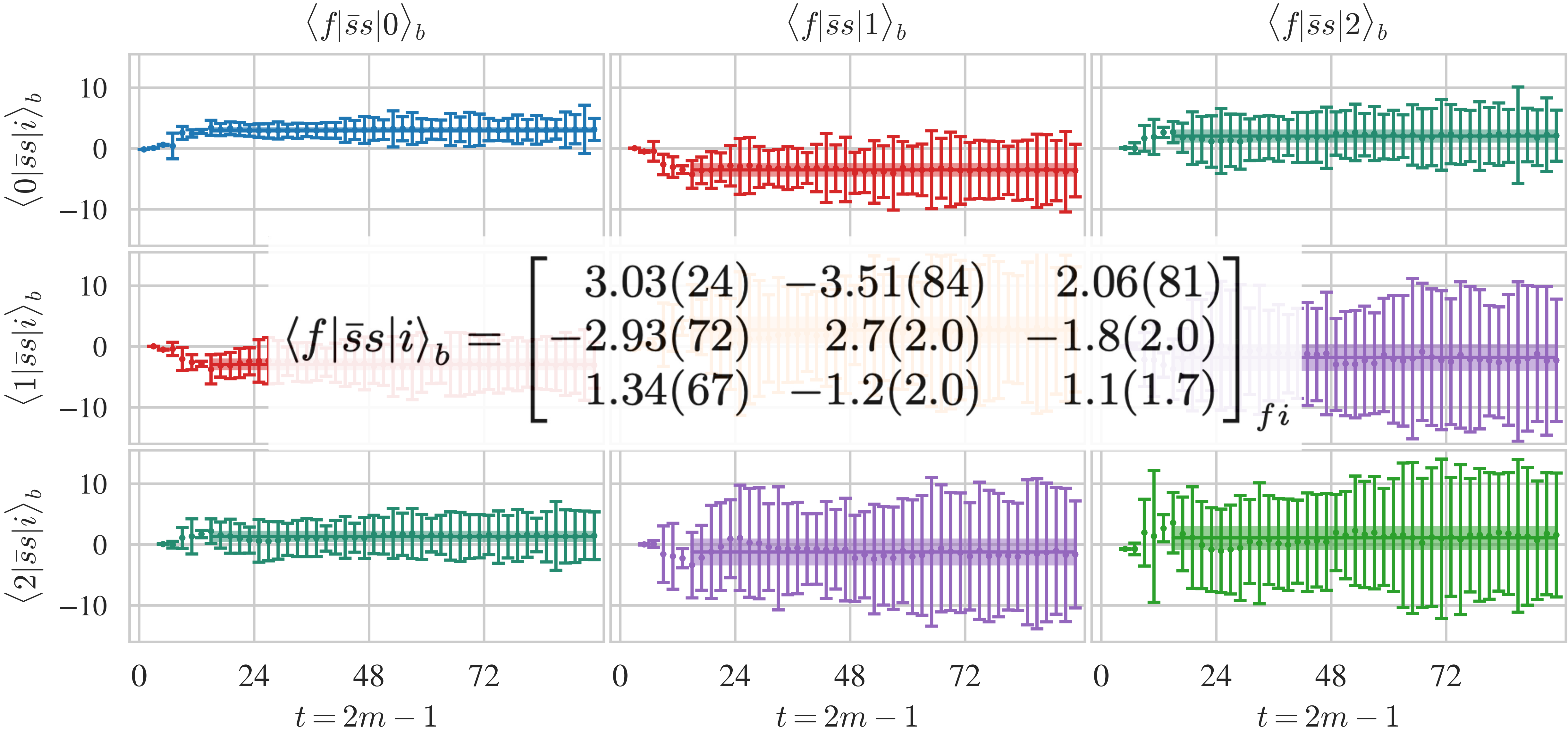
Ground-state matrix element



Excited & transition matrix elements



Excited & transition matrix elements



$$\langle f|\bar{s}s|i\rangle_b = \begin{bmatrix} 3.03(24) & -3.51(84) & 2.06(81) \\ -2.93(72) & 2.7(2.0) & -1.8(2.0) \\ 1.34(67) & -1.2(2.0) & 1.1(1.7) \end{bmatrix}_{fi}$$

Outlook

Versus previous methods:

- No SNR decay, no statistical modeling, no optimizers
- Better control of excited states
- Simple: only analysis hyperparameters are for state identification

Can also compute overlaps Z_k (decay constants!)

Can apply to >3-point functions: take $\mathcal{J} = J_1 T^{\delta} J_2$

~~Need dense evaluation of sink times~~

Only have data for $t_f \geq t_0$ → take $|\Psi\rangle = T^{t_0}|\psi\rangle$

Only have e.g. even t_f → take “T” $= T^2$

→ Can apply to existing ~~disconnected~~ data now

TODO: better statistics, state ID, convergence diagnostics; bounds?

What can we do now that we couldn't before?