Solving the signal-to-noise problem with Lanczos

Michael Wagman

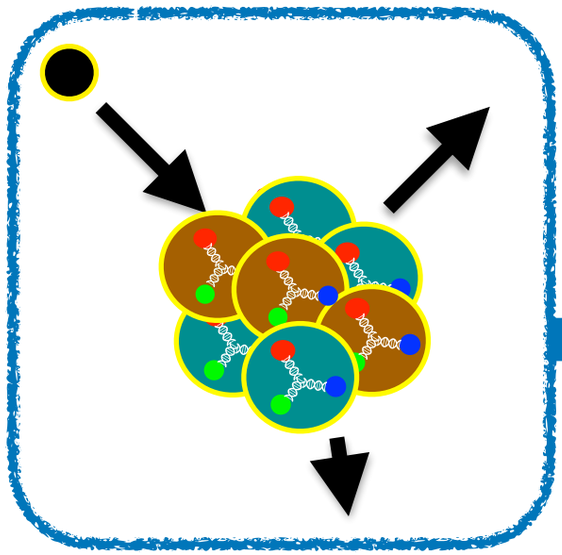
Fermilab Theory Seminar

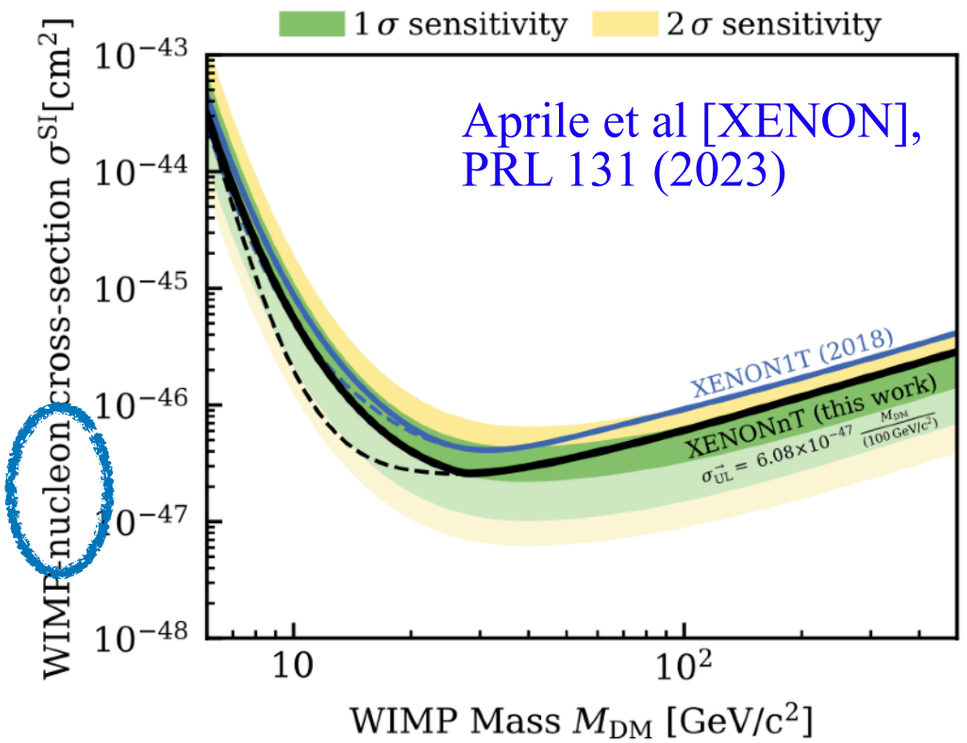
August 22, 2024

Nuclei and new physics

Nuclei are abundant and useful experimental targets

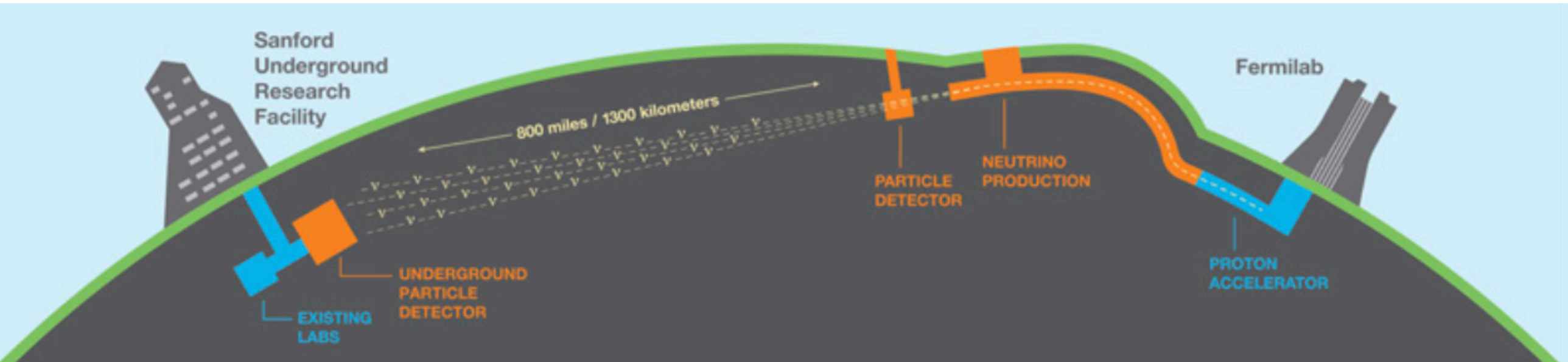
Relating new physics models to experimental data requires nuclear responses



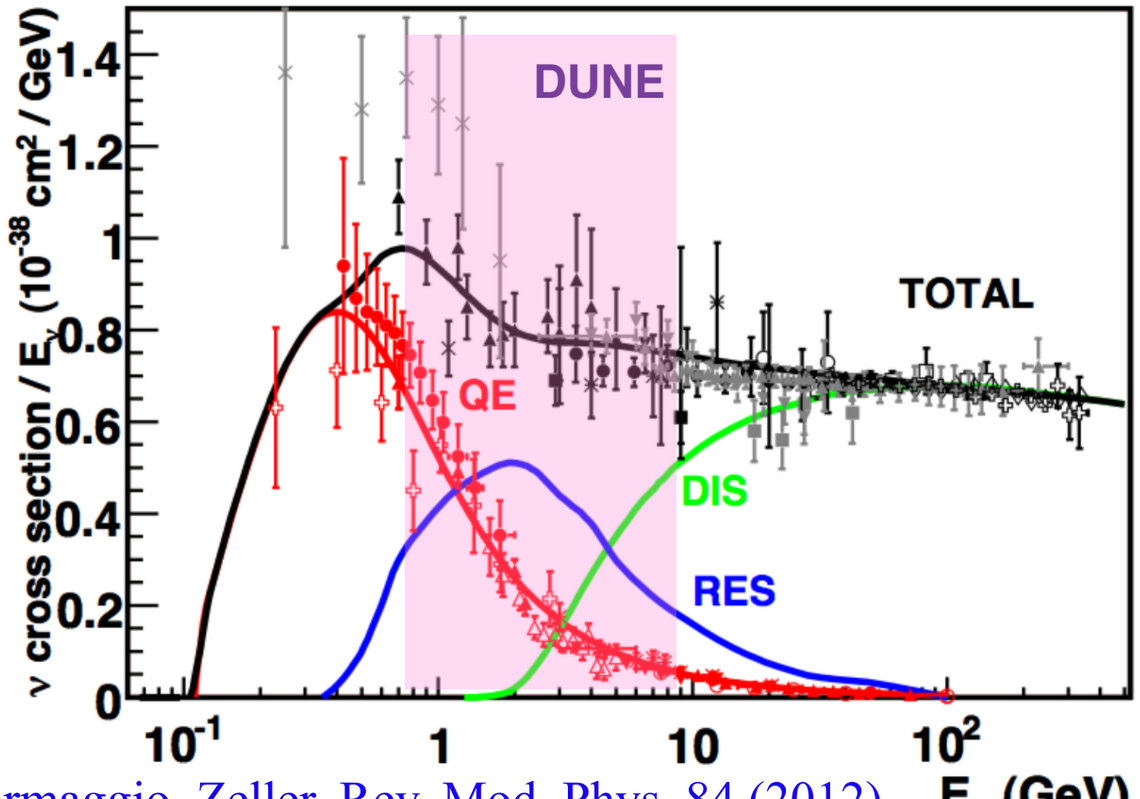


- Dark matter direct detection (low-energy scalar)
- Neutrinoless double-beta decay (low-energy axial)
- Neutrino-nucleus scattering (low- and high-energy axial)

DUNE



Neutrino-nucleus scattering



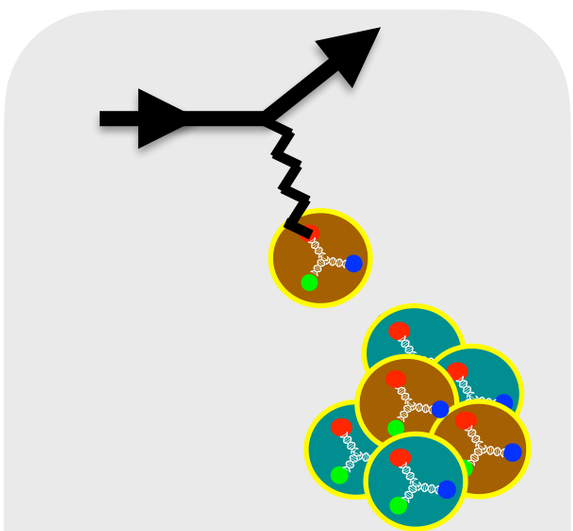
Formaggio, Zeller, Rev. Mod. Phys. 84 (2012)

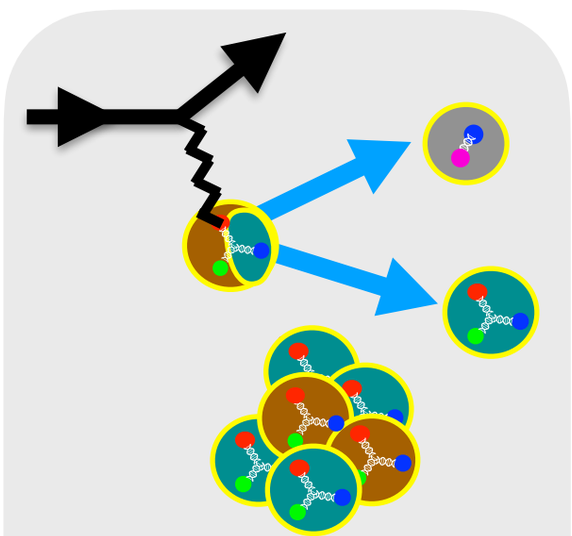
Accelerator neutrino fluxes cover a wide range of energies where different processes dominate cross-section:

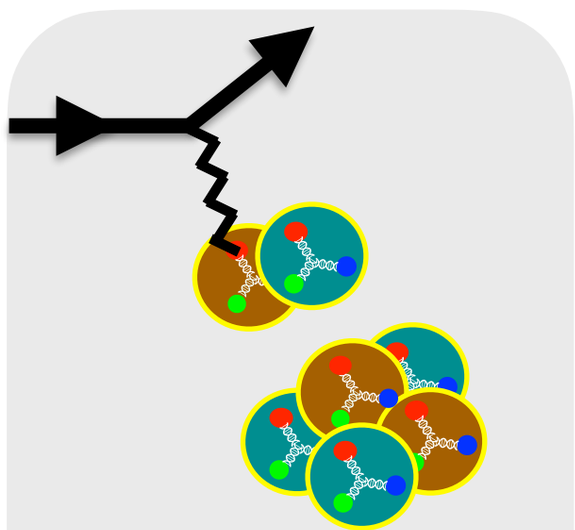
- Quasi-elastic nucleon scattering
- Resonance production
- Deep inelastic scattering

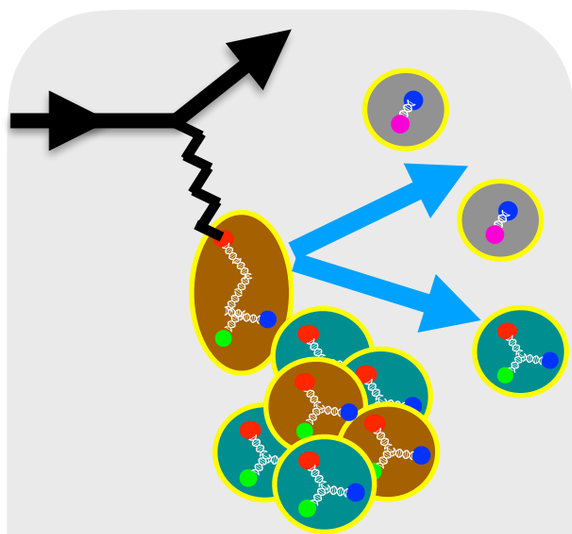
Theory input required to decompose cross section into such processes and therefore predict its energy dependence

Effective theories for different energies require different inputs





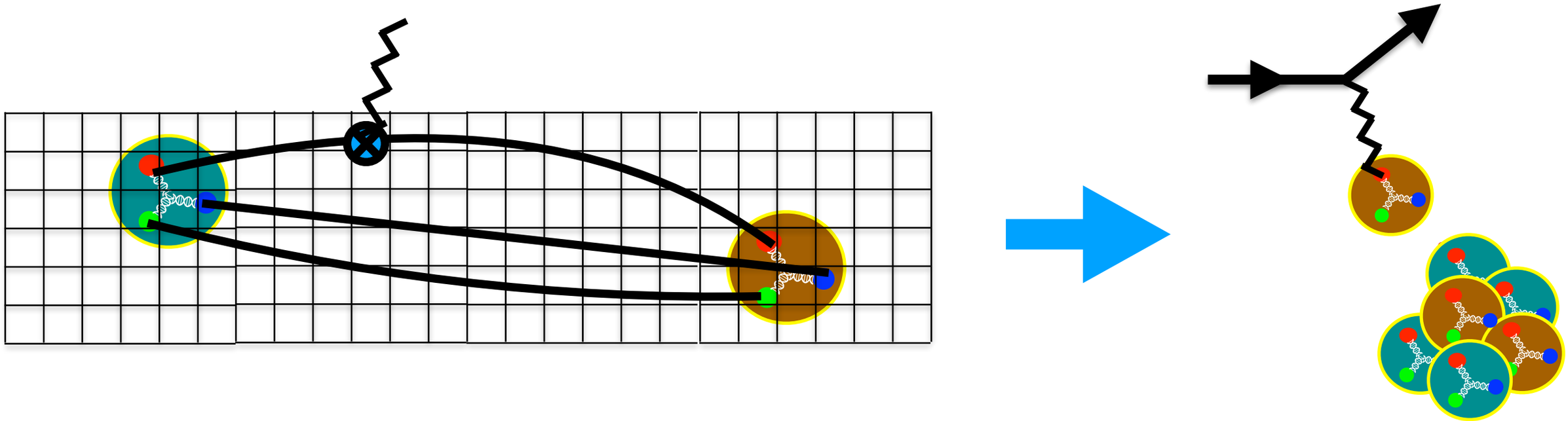




Lattice QCD and neutrino-nucleus

Lattice QCD provides reliable methods for numerically computing properties of QCD including nucleon form factors encoding responses to electroweak currents

Neutrino-nucleon scattering amplitudes can be computed straightforwardly once nucleon electroweak form factors known



Connecting **nucleon** form factors to neutrino-**nucleus** scattering is more complicated

- Lattice QCD can constrain inputs to nuclear EFTs and models
- Constraints from lattice QCD and experiment are often complementary

Monte Carlo path integrals

Lattice QCD uses a path integral version of quantum mechanics

— Quark propagators provide explicit solutions to the quark field path integral

— Gluon field path integrals are performed numerically using Monte Carlo:
random field values are drawn from a probability distribution similar to the integrand

Compromises:

Lattice spacing	Finite-volume	Imaginary time
$a \to 0$	$L \to \infty$	$\tau \to it$

Imaginary time turns complex quantum probability amplitudes

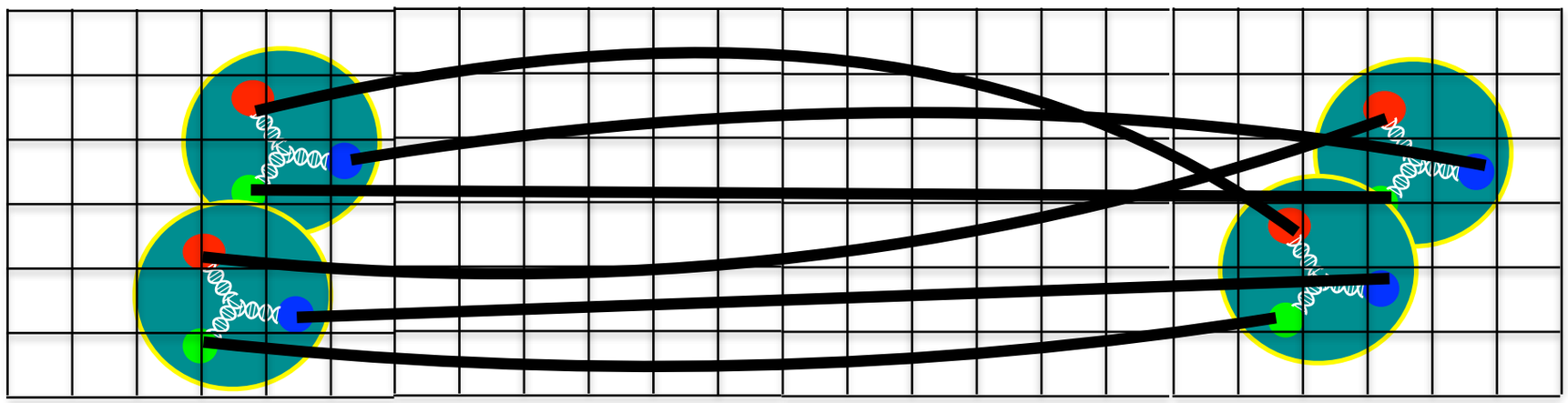
$$\text{probability amplitude} \sim e^{iS}$$

into positive-definite functions that can be interpreted as probabilities for random numbers in a Monte Carlo simulation

$$\text{probability amplitude} \sim e^{-S}$$

Observables in LQCD

LQCD energy spectrum determined from 2-point correlation functions



In imaginary time, correlation functions can be written as sums of exponentials

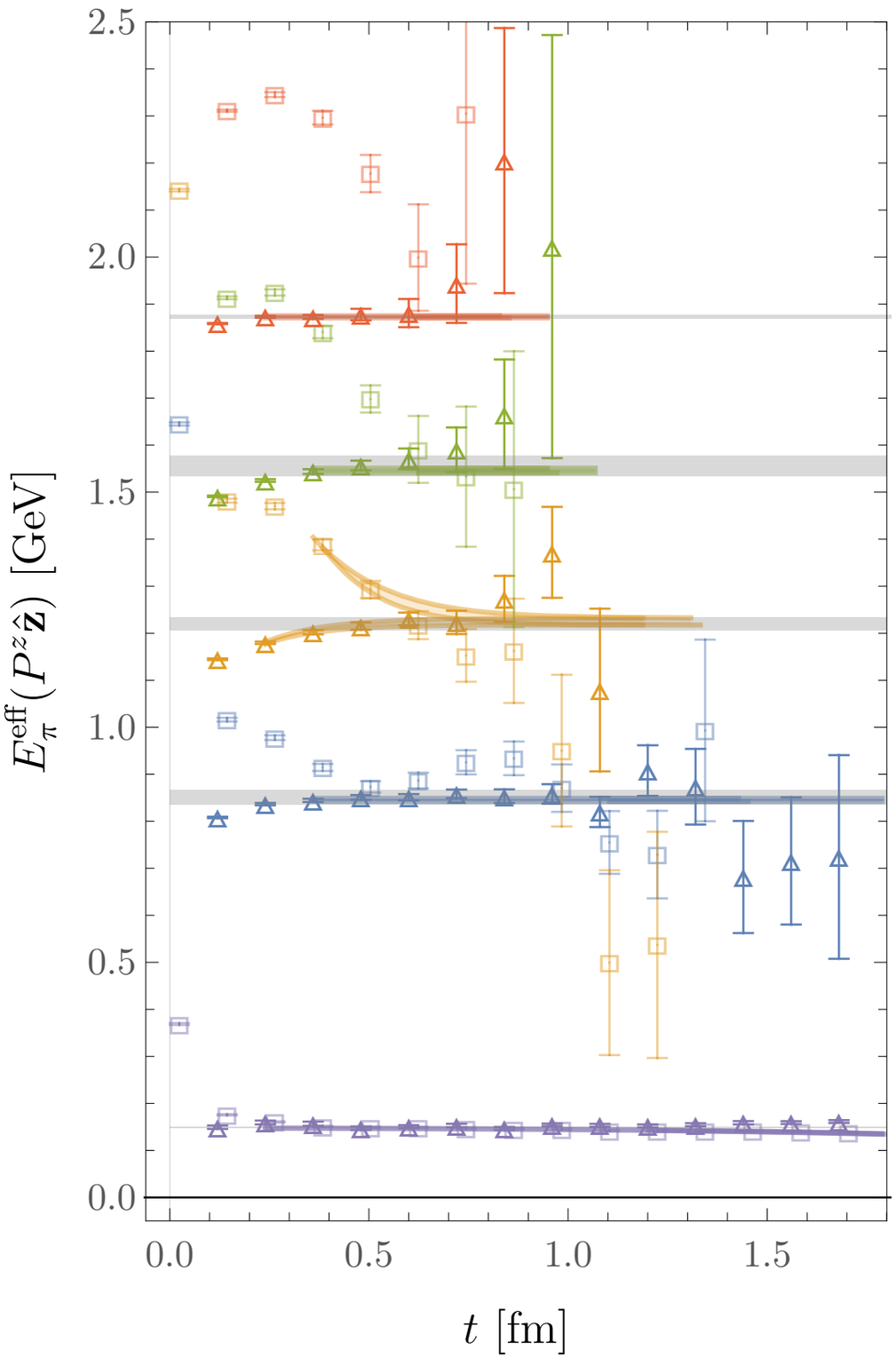
$$C_A(t) = \langle 0|A(t)A^\dagger(0)|0\rangle = \sum_n \langle 0|A(0)e^{-Ht}|n\rangle \langle n|A^\dagger(0)|0\rangle + \dots$$

$$= \sum_n |Z_n|^2 e^{-E_n t}$$

Imaginary time evolution $e^{-iHt_{\text{real}}} = e^{-H(it_{\text{real}})}$

Ground state dominates for large t: $C_A(t) \propto e^{-E_0 t} + O\left(e^{-(E_1 - E_0)t}\right)$

Effective masses

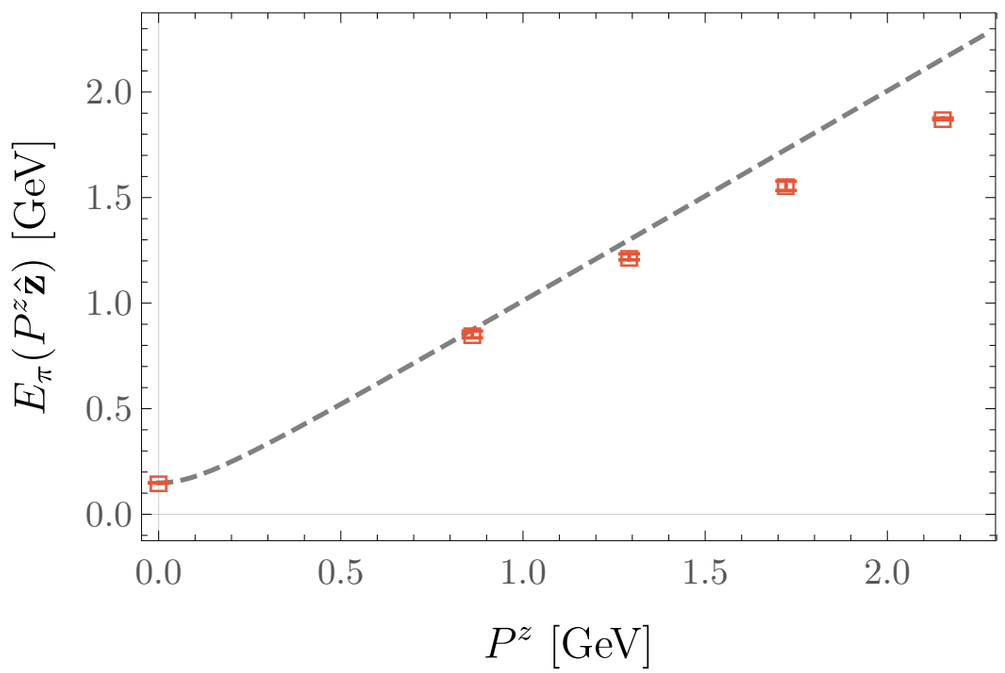


$$E^{\rm eff}(t) = \frac{1}{a}\ln\left[\frac{C_A(t+a)}{C_A(t)}\right] = E_0 + \mathcal{O}(e^{-(E_1-E_0)t})$$

Effective mass "plateau" signals ground state dominates correlation function at finite *t*

For simple states, e.g. low-momentum pion, simple interpolating operators and *t* ~ 1 fm appear sufficient

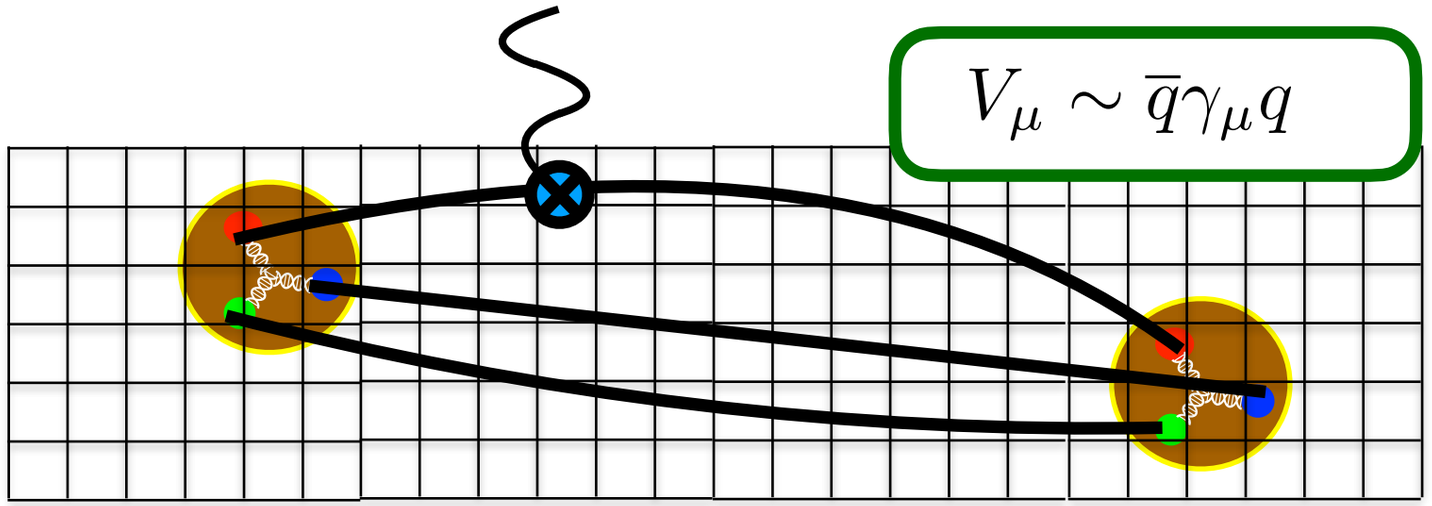
Fitted dispersion relations agree with continuum expectations + discretization effects



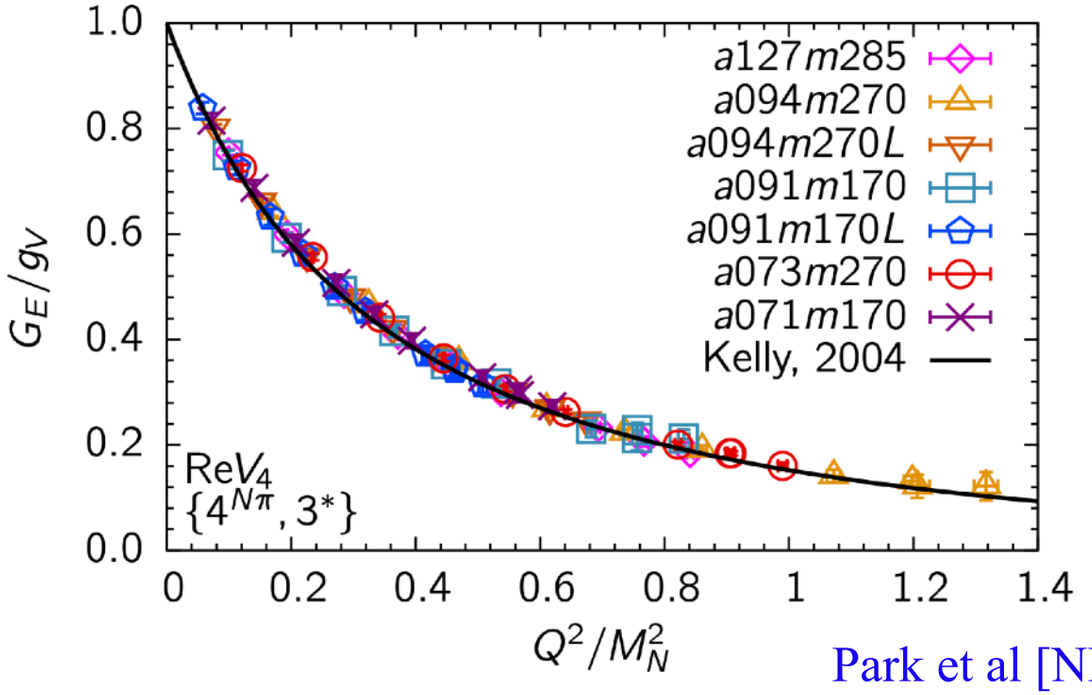
Avkhadiev, Shanahan, MW, Zhao, PRD 108 (2023)

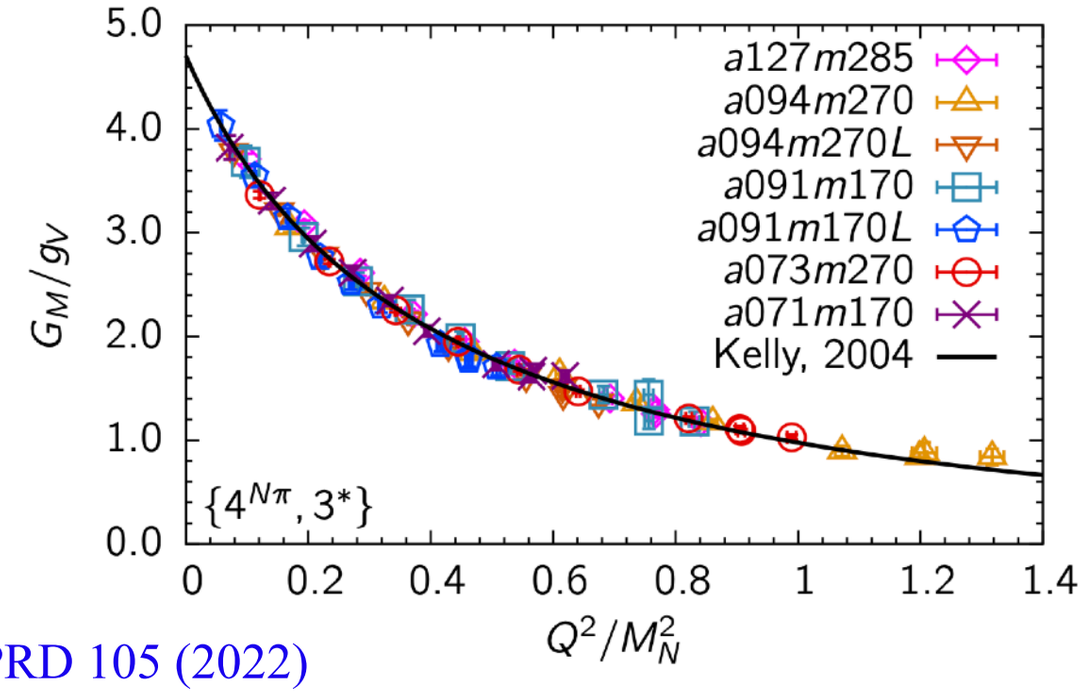
LQCD and nucleon form factors

Nucleon electric and magnetic form factors recently calculated using LQCD with approximately physical quark masses



$$\langle N(\mathbf{p}+\mathbf{q})|V^\mu|N(\mathbf{p})\rangle = \overline{u}(\mathbf{p}+\mathbf{q})\left[F_1(q^2)\gamma^\mu + i\sigma^{\mu\nu}q_\nu\frac{F_2(q^2)}{2M_N}\right]u(\mathbf{p})$$





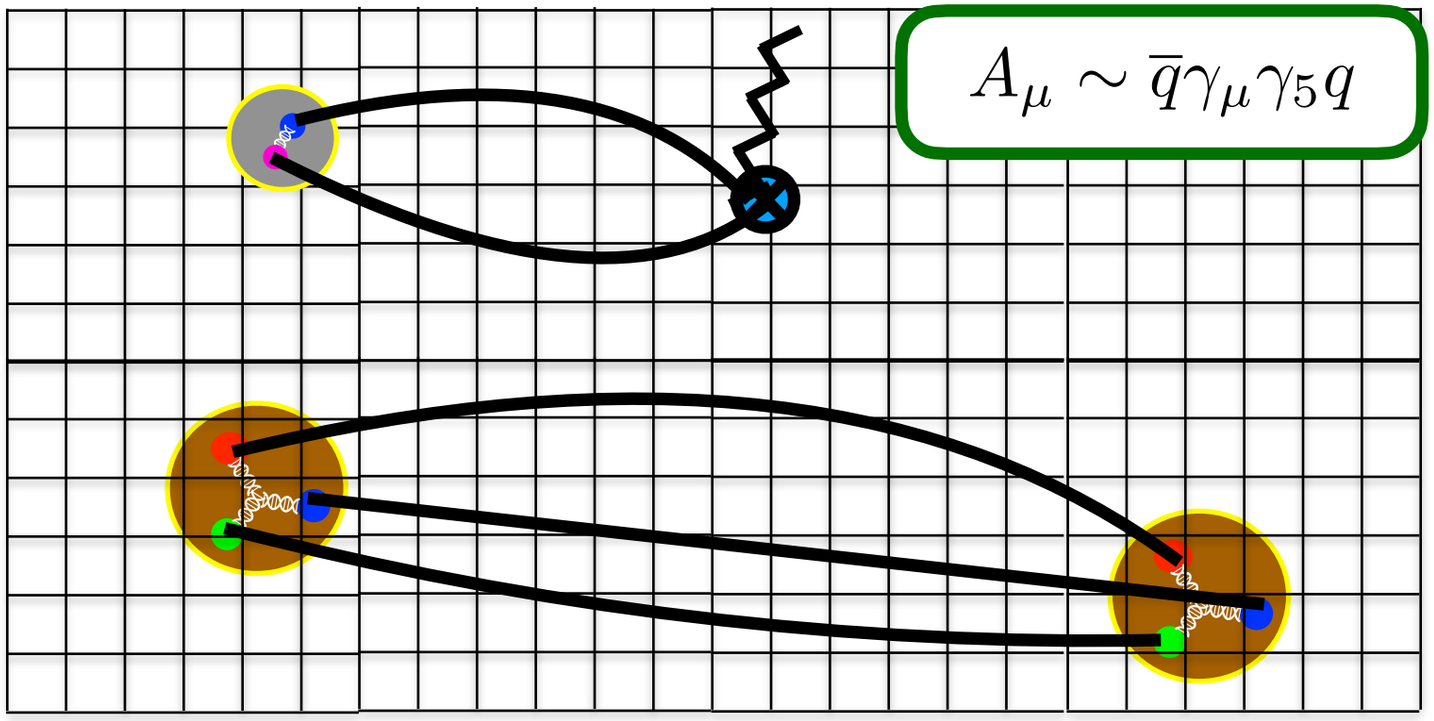
Park et al [NME], PRD 105 (2022)

LQCD results for nucleon electric and magnetic form factors (linear combinations of F_1 and F_2) show good consistency with phenomenological parameterizations

Excited states

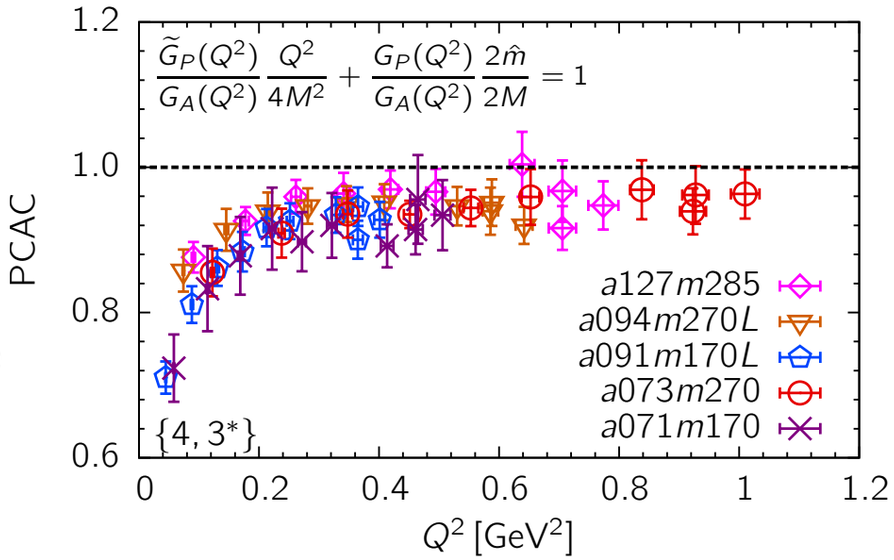
Axial form factor calculations have been performed using analogous methods

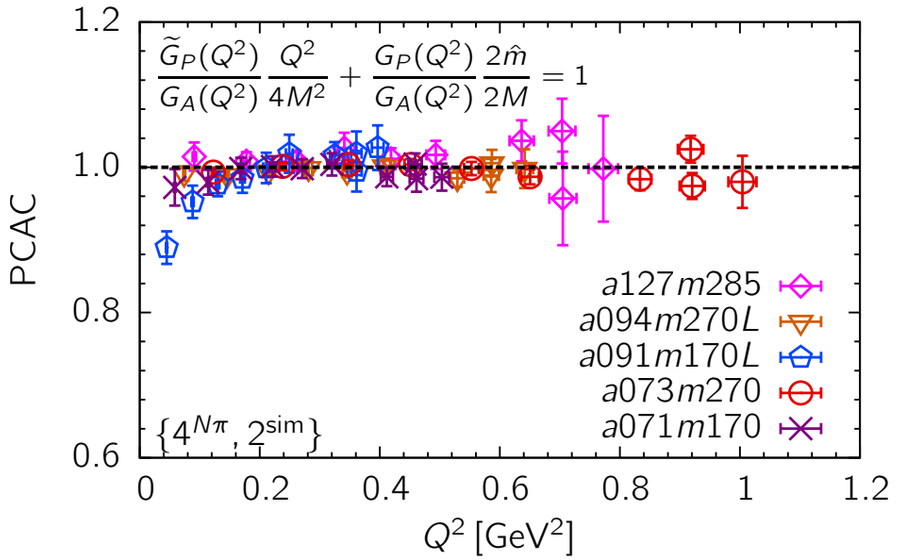
Additional excited-state effects arise from the fact that axial currents can act as pion sources



$$\langle N(\mathbf{p}+\mathbf{q})|A^\mu|N(\mathbf{p})\rangle = \overline{u}(\mathbf{p}+\mathbf{q})\left[G_A(q^2)\gamma^\mu\gamma_5 + q^\mu\gamma_5\frac{\widetilde{G}_P(q^2)}{2M_N}\right]u(\mathbf{p})$$

Careful treatment of $N\pi$ excited states required to reproduce known symmetry constraints assuming ground-state dominance of results





Park et al [NME], PRD 105 (2022)

Axial form factors

LQCD calculations of nucleon axial form factors with approximately physical quark masses and continuum extrapolations achieved by multiple groups

Bali et al [RQCD], JHEP 05 126 (2020)

Park et al [NME], PRD 105 (2022)

Djukanovic et al, PRD 103 (2021)

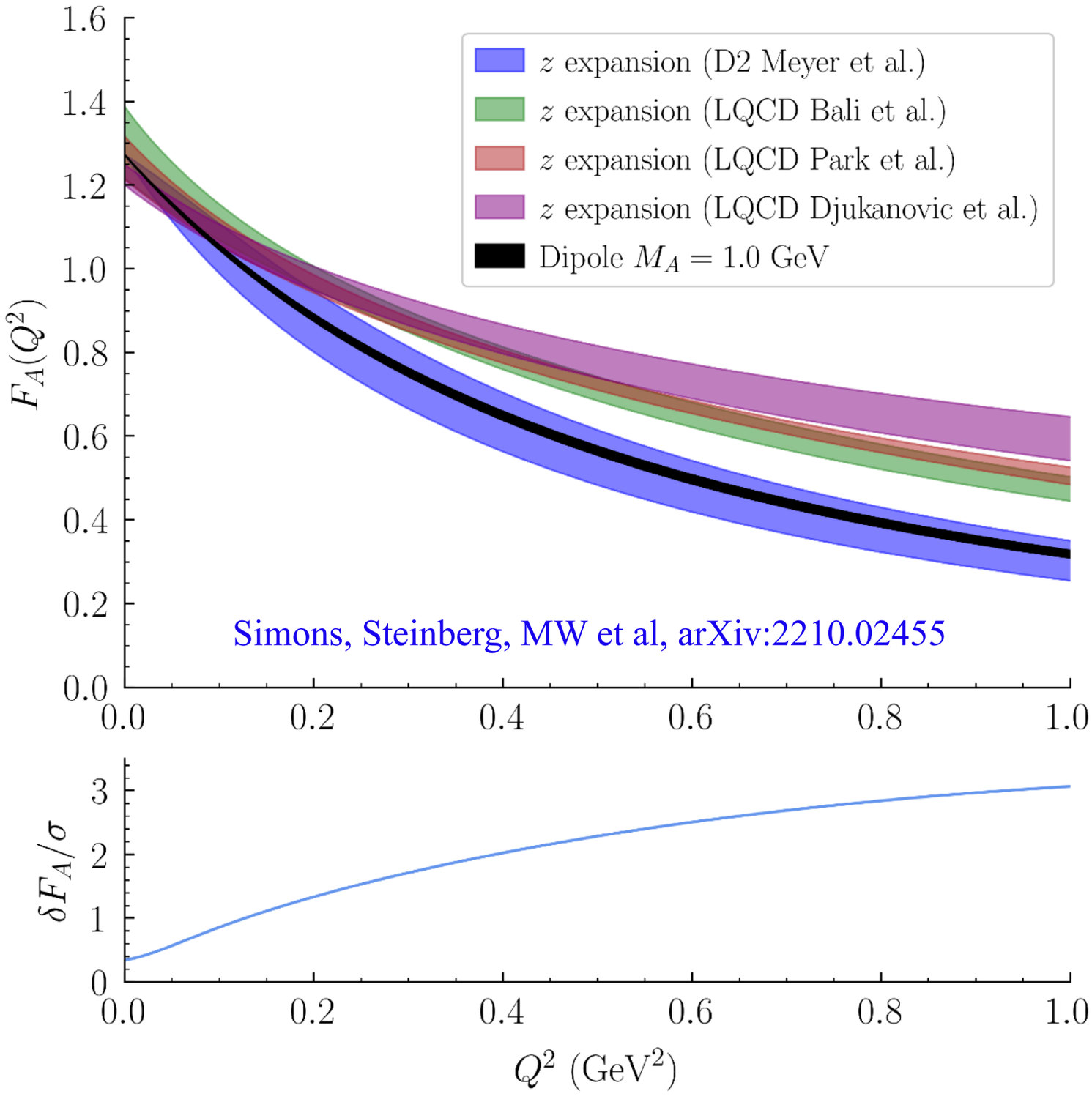
Alexandrou et al [ETMC], PRD 109 (2024)

Jang et al [NME], PRD 109 (2024)

Up to 3 sigma differences between LQCD and experimental axial form factor determinations, could arise from challenging LQCD systematic uncertainties

Differences could also arise from underestimated uncertainties in phenomenological form factor determinations using deuterium bubble chamber data

Meyer, Betancourt, Gran, and Hill, PRD 93 (2016)

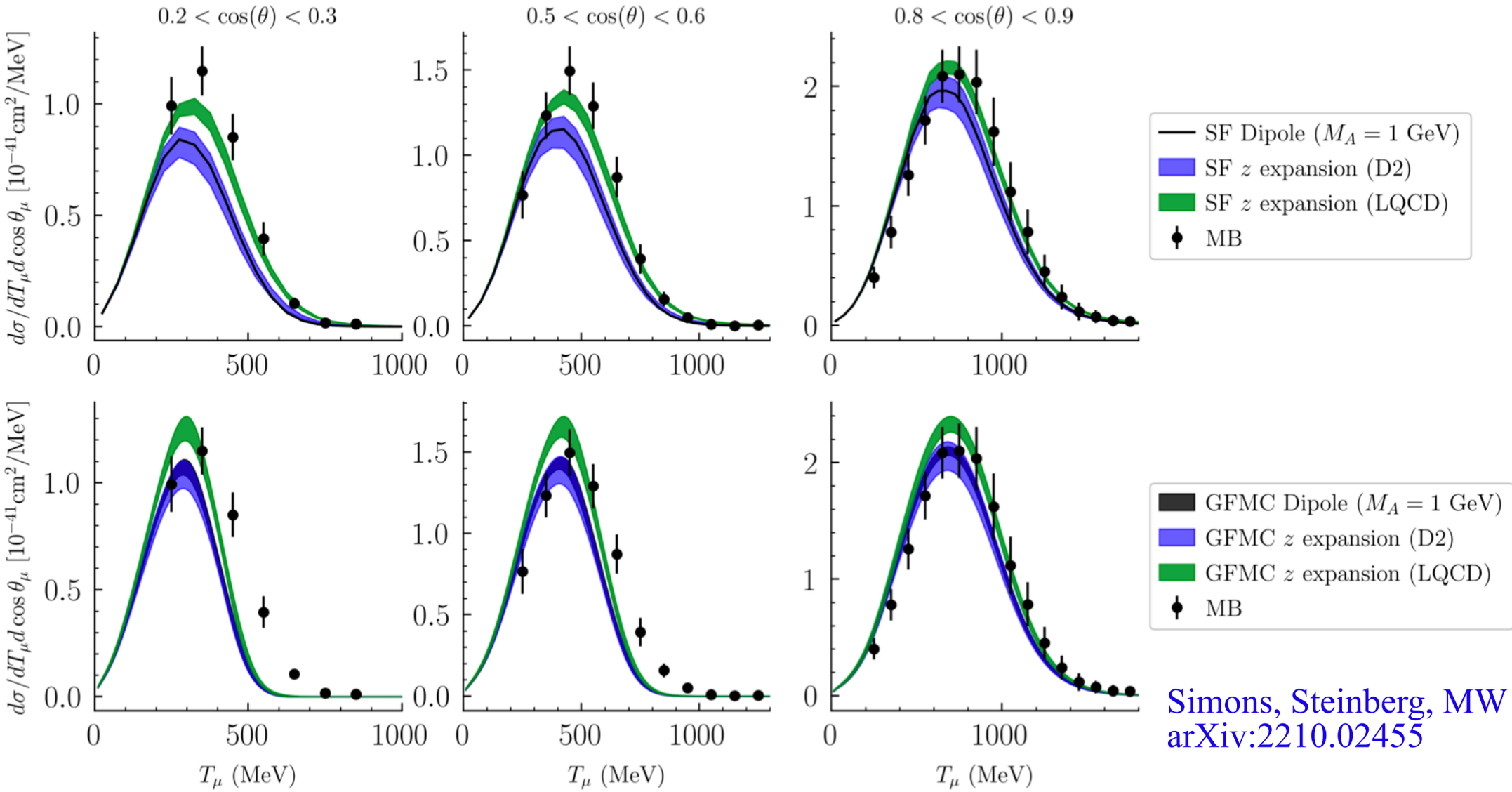


Simons, Steinberg, MW et al, arXiv:2210.02455

MiniBooNE Results

Impact of LQCD vs D2 form factors studied using GFMC and spectral function (SF) calculations of carbon relevant for MiniBooNE

- 10-20% differences found between LQCD and D2 form factors
- Disentangling effects of nucleon form-factors and nuclear interactions is non-trivial



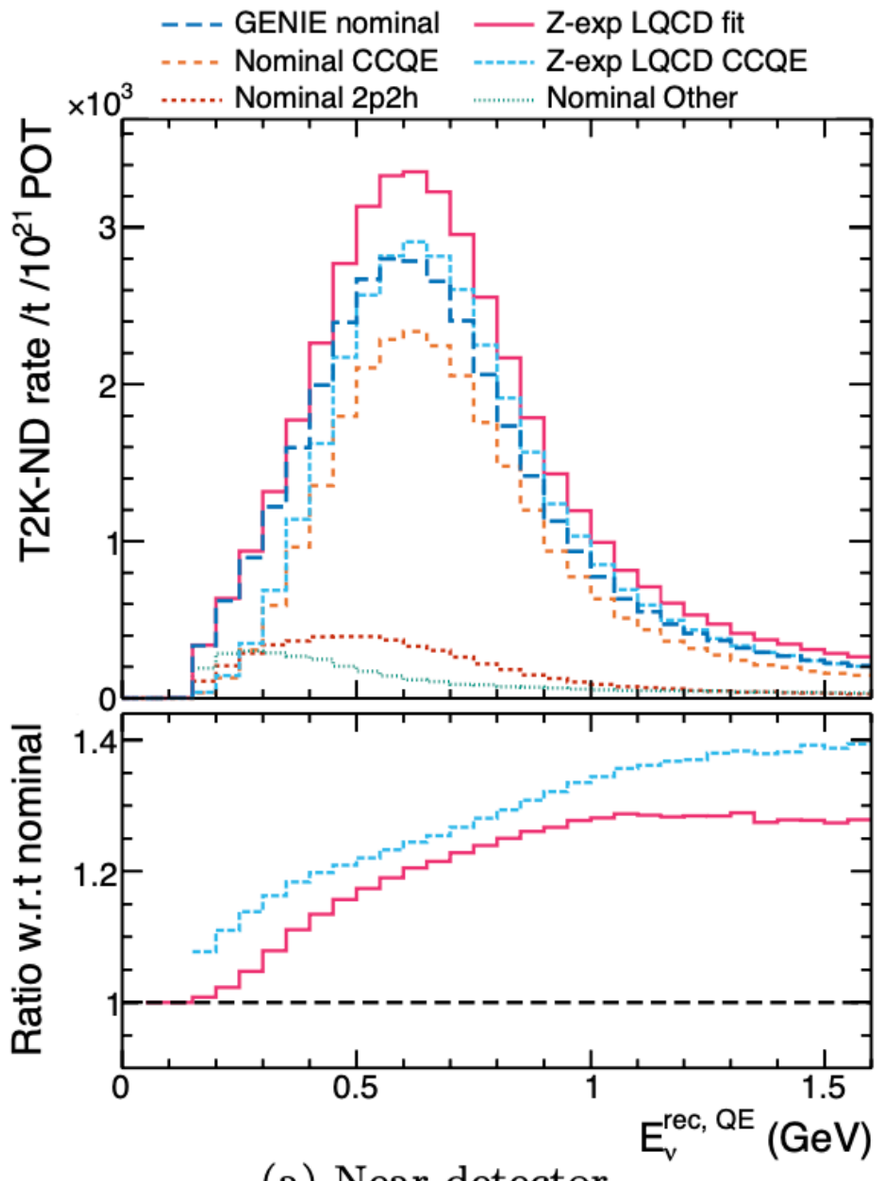
Simons, Steinberg, MW et al, arXiv:2210.02455

Axial form factor uncertainties

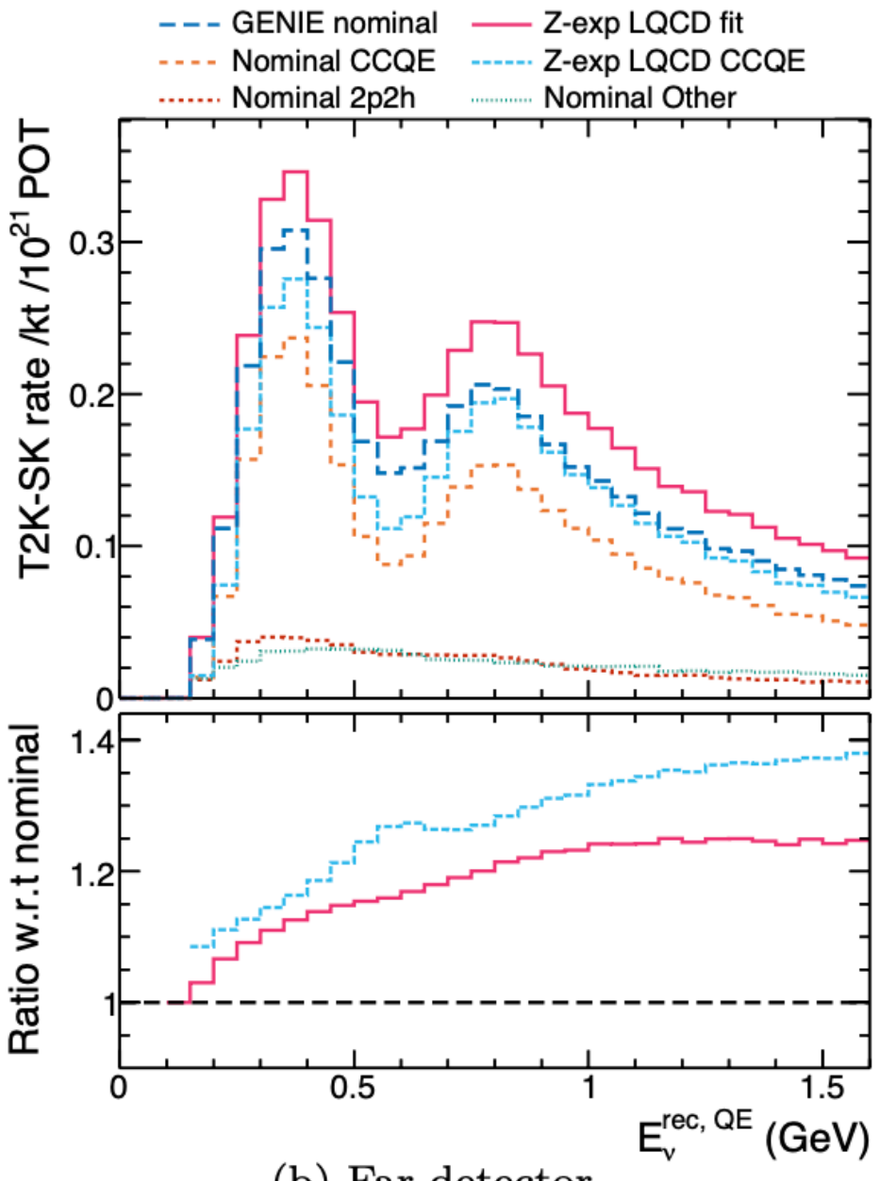
GENIE event generator predictions for T2K event rate using deuterium bubble chamber vs recent LQCD axial form factors differ by ~20%

- Effects on near and far detectors differ, understanding discrepancy essential for reliable neutrino oscillation analyses

Meyer, Walker-Loud, Wilkinson, Ann. Rev. Nucl. Part. Sci. 72 (2022)



(a) Near detector



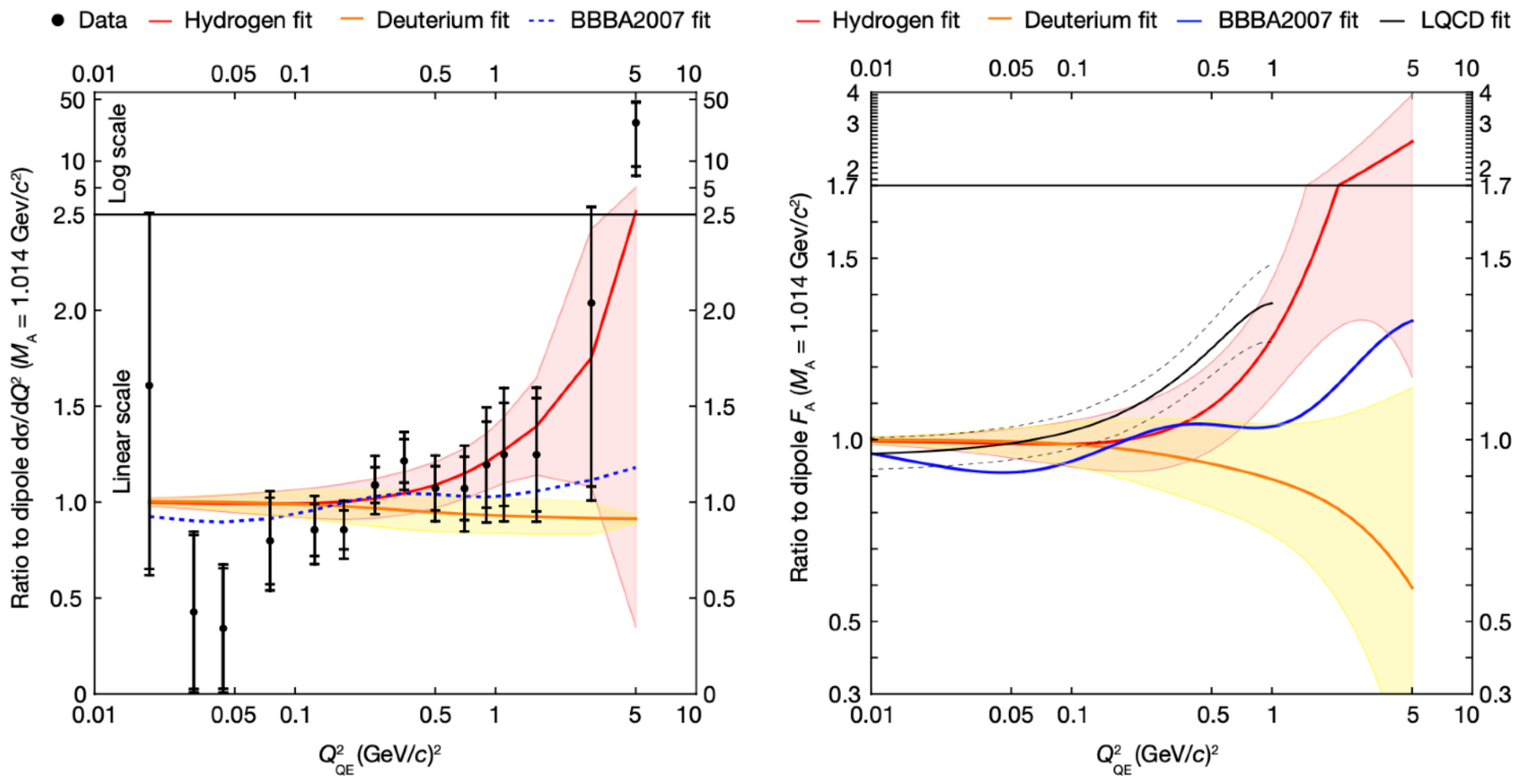
(b) Far detector

MINERvA Results

MINERvA has recently analyzed antineutrino scattering with a hydrocarbon scintillator target (8% H + 89% C + …) to extract nucleon axial form factor

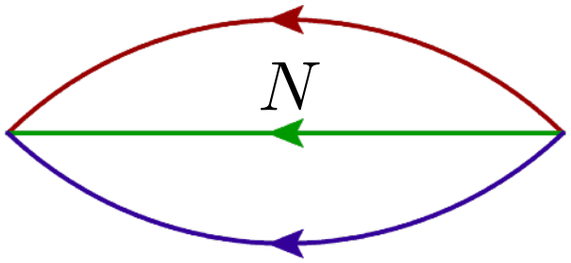
- Broad consistency is seen between MINERvA extraction and LQCD
- Separating nuclear / free nucleon events challenging

Cai et al. [MINERvA], Nature 614 (2023)



- Experimental determinations of nucleon axial form factor limited by nuclear theory systematics (+other systematics of old bubble chamber data)
- Competitive systematic uncertainties achieved with present-day lattice QCD…

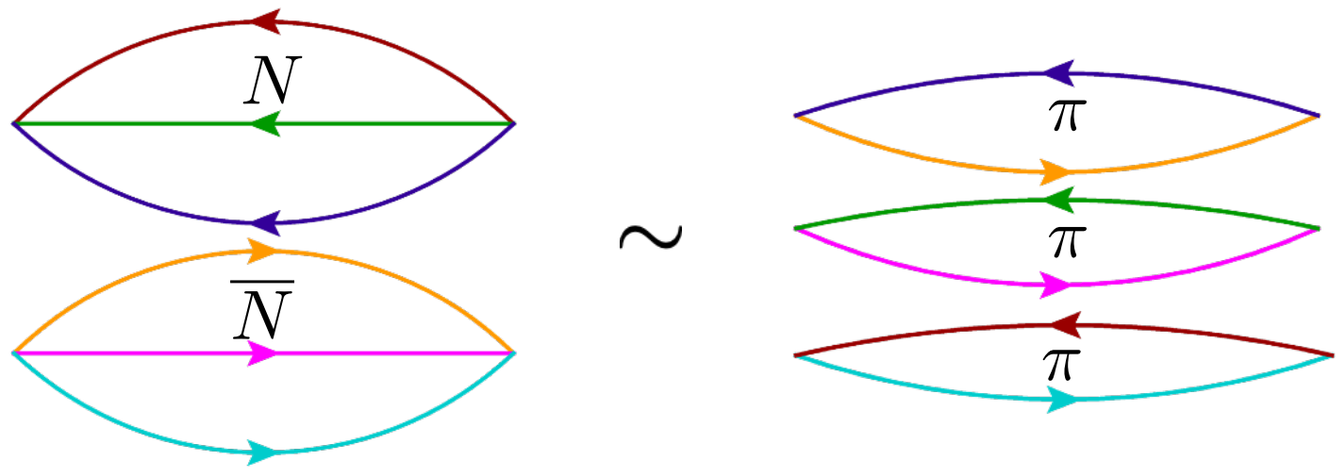
The signal-to-noise problem



Nucleon ground state dominates correlation function for large t

$$C_N(t) \sim e^{-M_N t}$$

Variance of nucleon correlation function is itself a correlation function with quantum numbers of $N\overline{N}$



The lightest allowed state is 3π

$$\mathrm{Var}[C_N(t)] \sim e^{-3m_\pi t}$$

Implies signal-to-noise ratios scale as

$$\mathrm{StN}[C_N(t)] = \frac{\langle C_N(t)\rangle}{\sqrt{\mathrm{Var}[C_N(t)]}} \sim e^{-(M_N - \frac{3}{2}m_\pi)t}$$

Parisi, Phys.Rept. 103 (1984)

Lepage, TASI (1989)

Same analysis for a system of *A* nucleons:

$$\mathrm{StN}[C_A(t)] = \frac{\langle C_A(t)\rangle}{\sqrt{\mathrm{Var}[C_A(t)]}} \sim e^{-A(M_N - \frac{3}{2}m_\pi)t}$$

The sign/al-to-noise problem

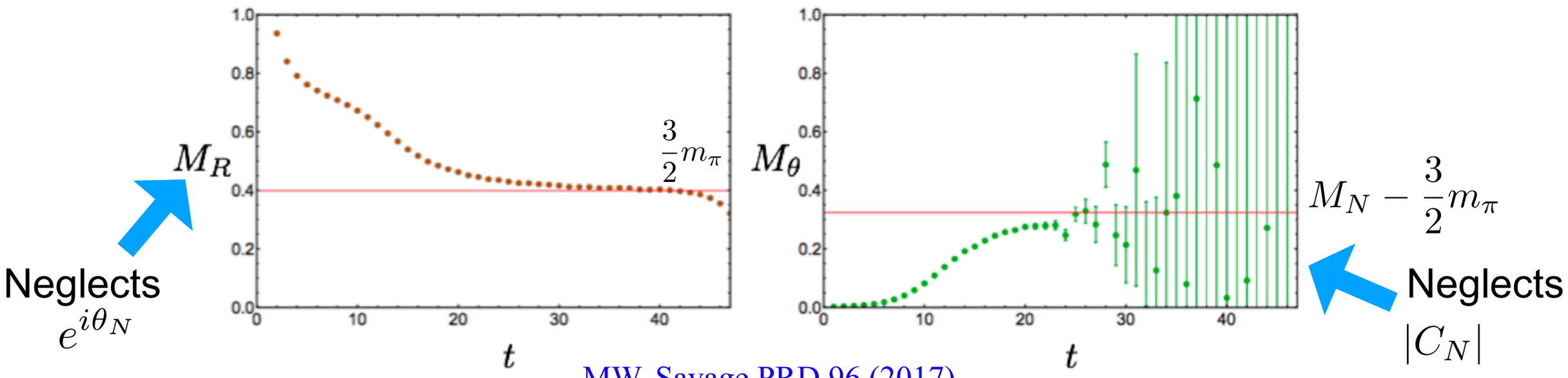
Nucleon correlation functions are complex in generic gauge field backgrounds

Complex phases of correlation functions give path integrals “sign problems”

$$C_N(t) = \frac{1}{Z}\int \mathcal{D}U \; e^{-S(U)} \; |C_N(t)|\, e^{i\theta_N(t)}$$

Integral can’t be interpreted as a probability

The same phase fluctuations are responsible for the full severity of the signal-to-noise problem for the nucleon and nuclei



MW, Savage PRD 96 (2017)

Noise reduction with AI/ML

Complex analysis guarantees that changing the integration manifold does not change integrals of analytic functions, but “complex phase” is nonanalytic

— changing contours can reduce phase fluctuations without changing the results for path integrals (i.e. the physics)

Witten AMS/IP Stud. Adv. Math. 50 (2011)

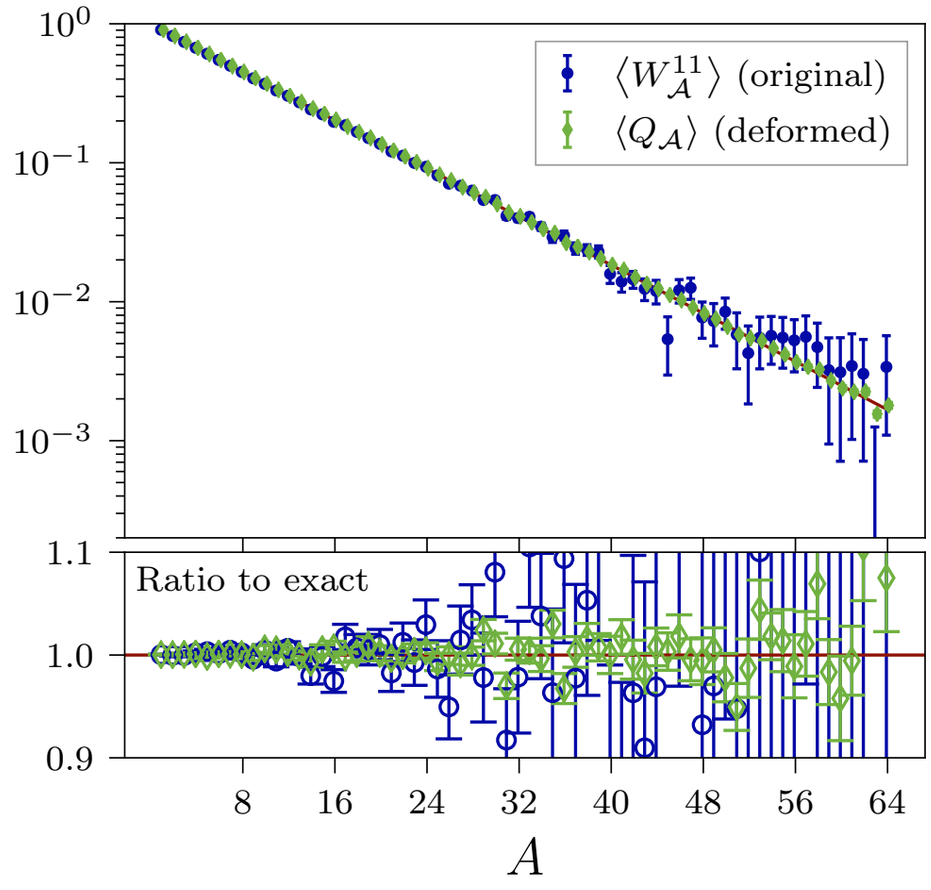
Review: Alexandru et al, Rev Mod Phys 94 (2022)

Path integral contour deformations can be applied to “observables with signal-to-noise problems” as well as “actions with sign problems”

$$\mathrm{Var}[\mathrm{Re}\,\mathcal{O}] = \frac{1}{2}\left\langle|\mathcal{O}|^2\right\rangle + \frac{1}{2}\left\langle\mathcal{O}^2\right\rangle - \left[\mathrm{Re}\,\left\langle\mathcal{O}\right\rangle\right]^2$$

Detmold, MW, et al, PRD 102 (2020)

Detmold, MW, et al, PRD 103 (2021)



Noise reduction with AI/ML

Complex analysis guarantees that changing the integration manifold does not change integrals of analytic functions, but “complex phase” is nonanalytic

— changing contours can reduce phase fluctuations without changing the results for path integrals (i.e. the physics)

Witten AMS/IP Stud. Adv. Math. 50 (2011)

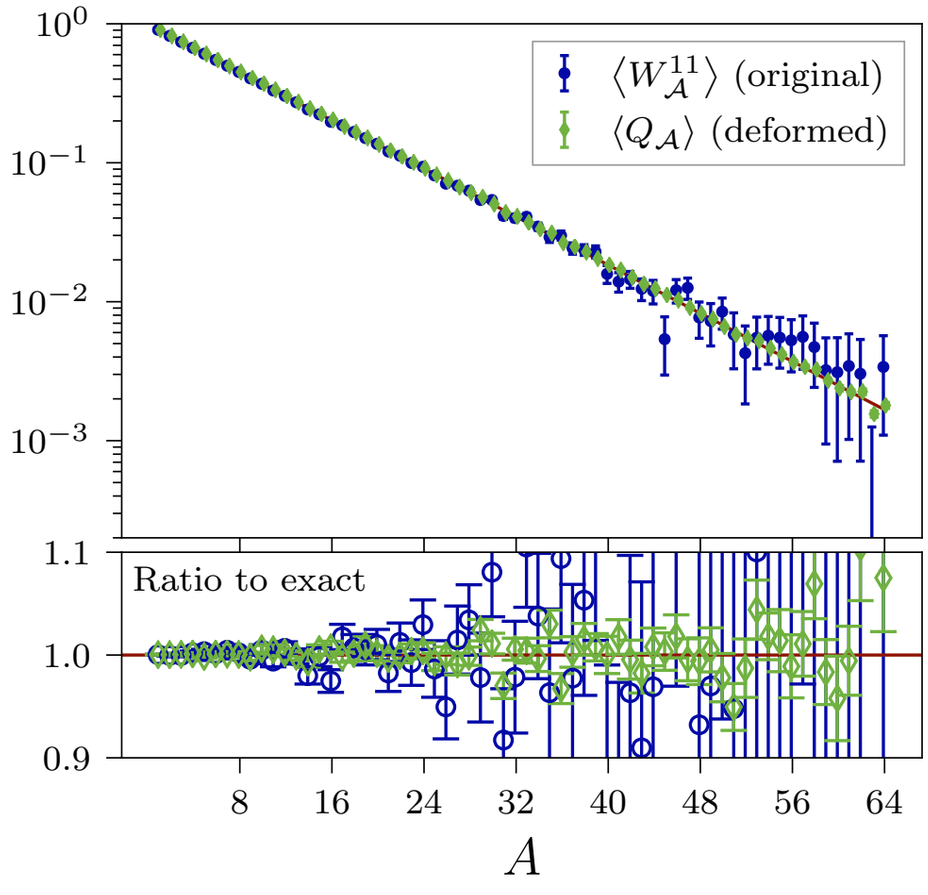
Review: Alexandru et al, Rev Mod Phys 94 (2022)

Path integral contour deformations can be applied to “observables with signal-to-noise problems” as well as “actions with sign problems”

$$\mathrm{Var}[\mathrm{Re}\,\mathcal{O}] = \frac{1}{2}\left\langle|\mathcal{O}|^2\right\rangle + \frac{1}{2}\left\langle\mathcal{O}^2\right\rangle - \left[\mathrm{Re}\,\left\langle\mathcal{O}\right\rangle\right]^2$$

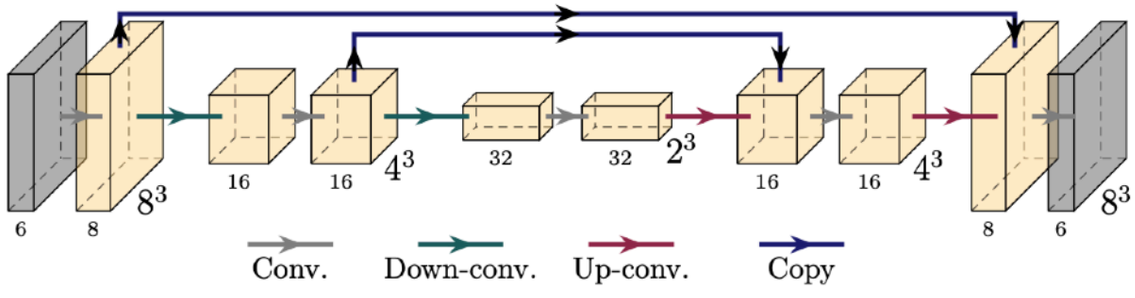
Detmold, MW, et al, PRD 102 (2020)

Detmold, MW, et al, PRD 103 (2021)



Extending from 2D - 4D requires more sophisticated AI/ML techniques, remains challenging

Lin, MW et al, NeurIPS 2023



Fitting is hard

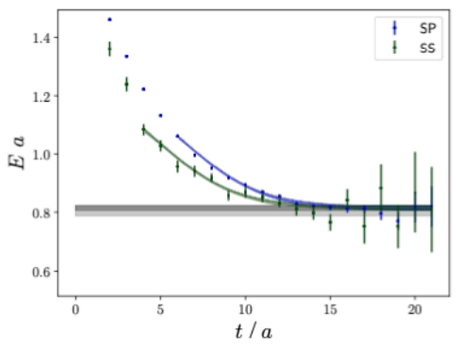
Accurate systematic uncertainty quantification requires sophisticated tools, heuristics, and human judgment

- t_{min} — *results sensitive, take weighted average*

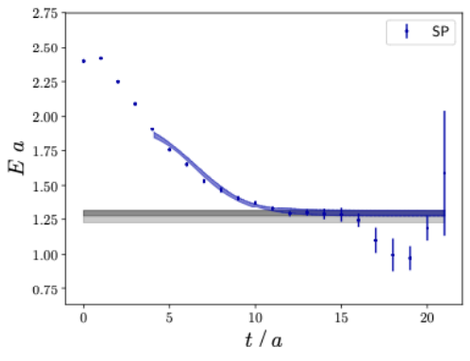
  Rinaldi, MW, et al, PRD 99 (2019)

  Jay and Neil, PRD 103 (2021)
- N_{states} — *results sensitive, take weighted average*
- t_{max} — *results insensitive (except when they're not), impose arbitrary cut on signal-to-noise*
- *Covariance matrix — SVD, shrinkage, …*

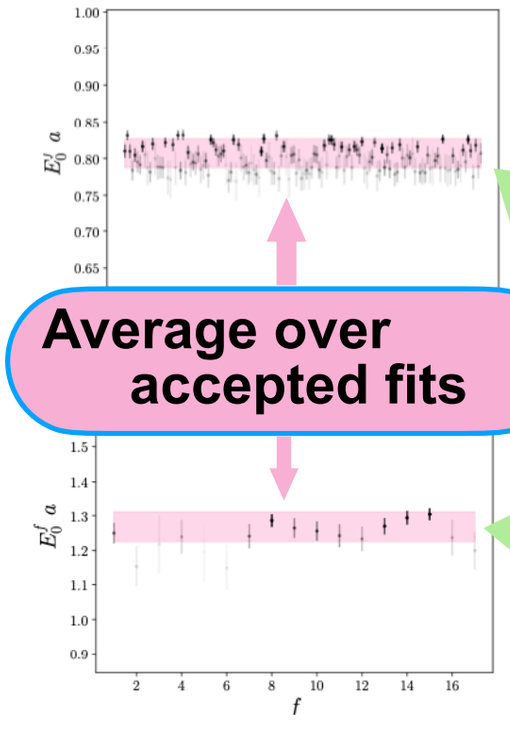
$pn(^1S_0)$
$L/a = 32$



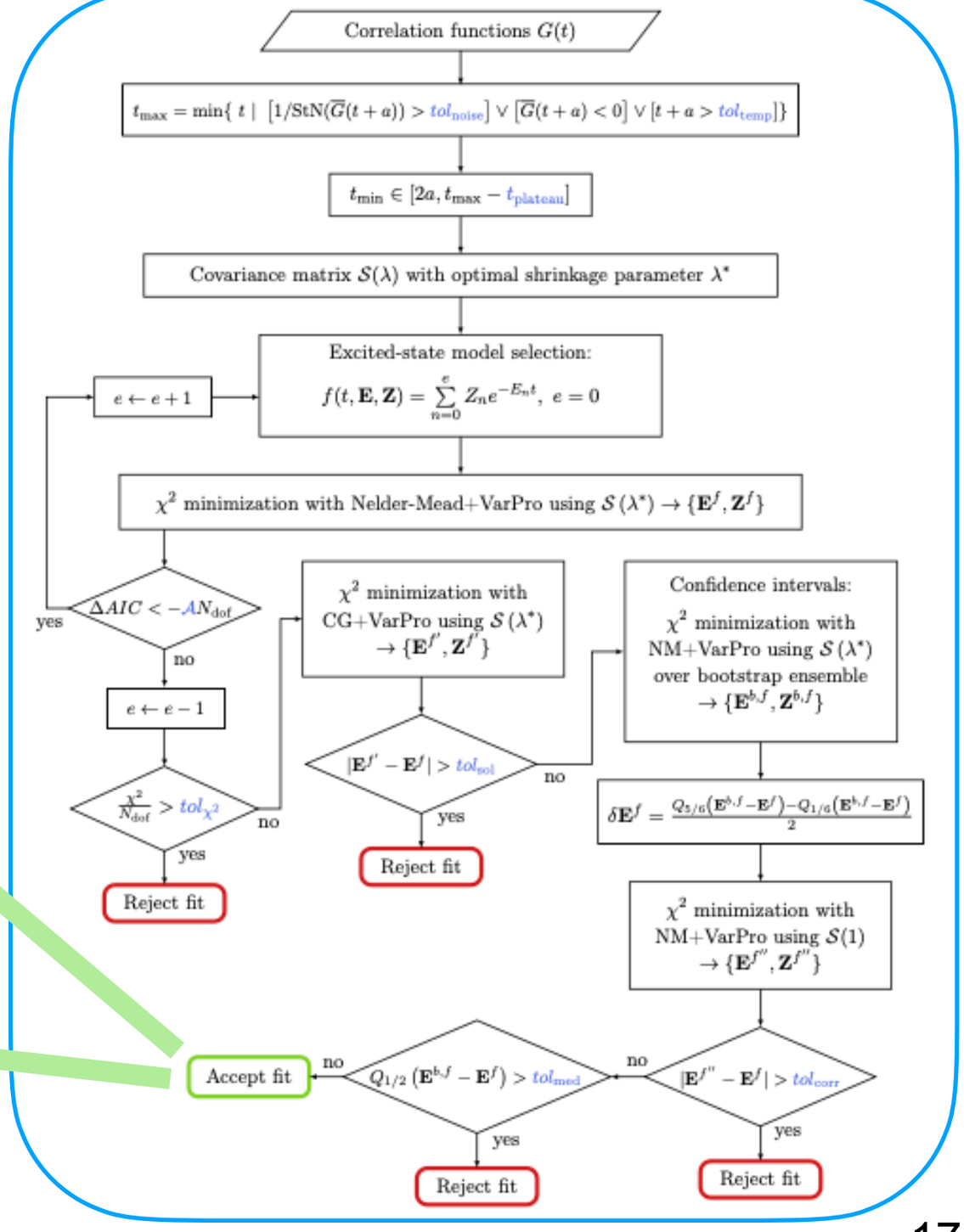
^{3}H
$L/a = 48$



Average over accepted fits



"Autofitter" flowchart



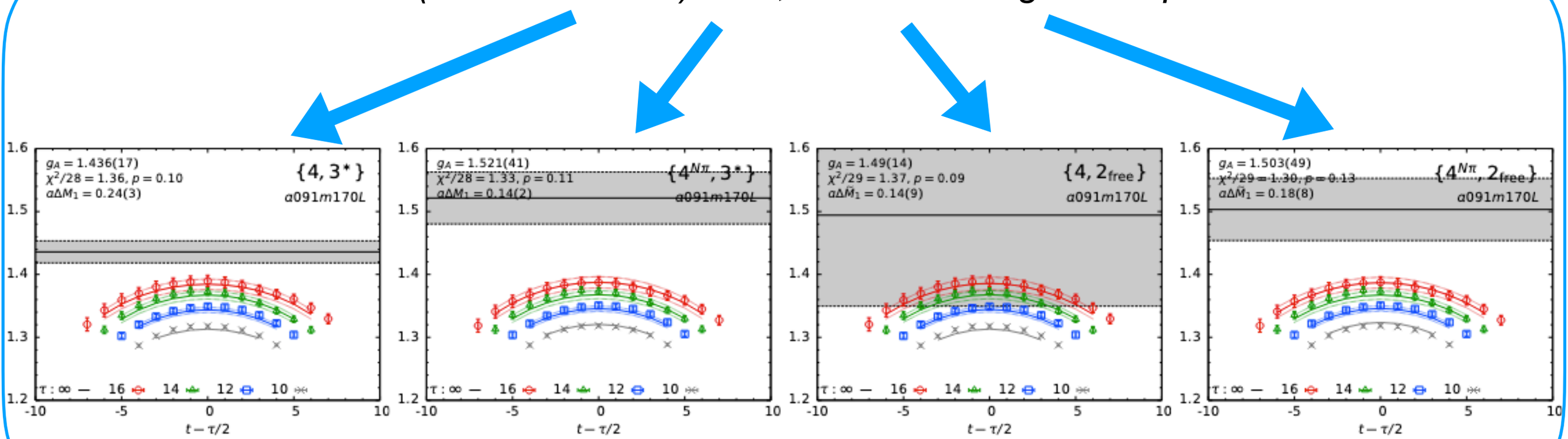
Beane, MW, et al [NPLCD+QCDSF+CSSM], PRD 103 (2021)

Matrix elements are really hard

$N\pi$ excited-state contamination in nucleon axial form factor manifests as

- Slow imaginary-time convergence
- Increased sensitivity to fit strategies

Same (state-of-the-art) data, different fitting assumptions

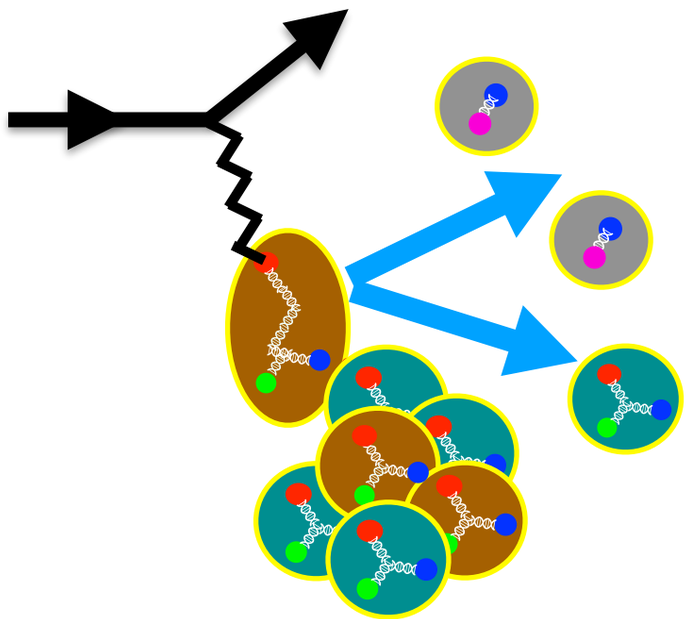


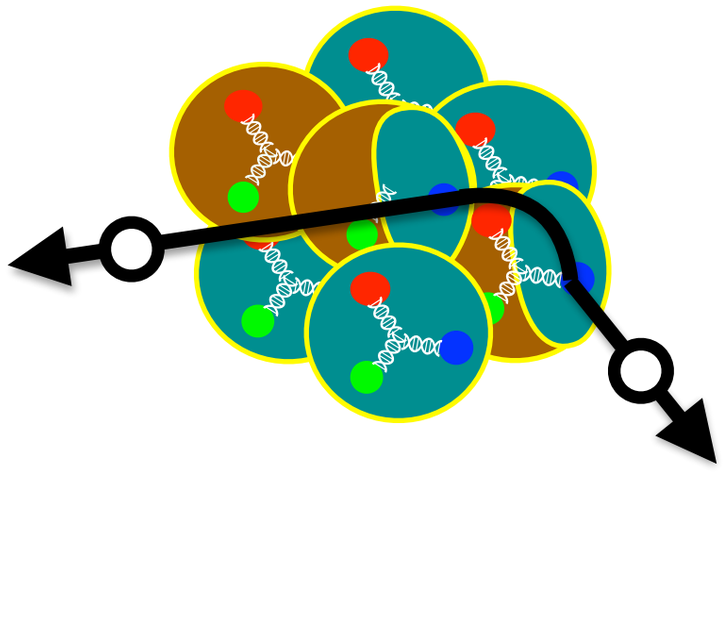
Park et al [NME], PRD 105 (2022)

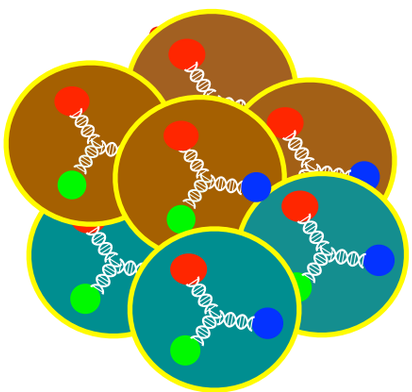
- Quantifying fitting systematics is challenging —- different "right" answers differ on central values by $\gtrsim 2\sigma$ and differ on error bar size by factors of 2 - 7

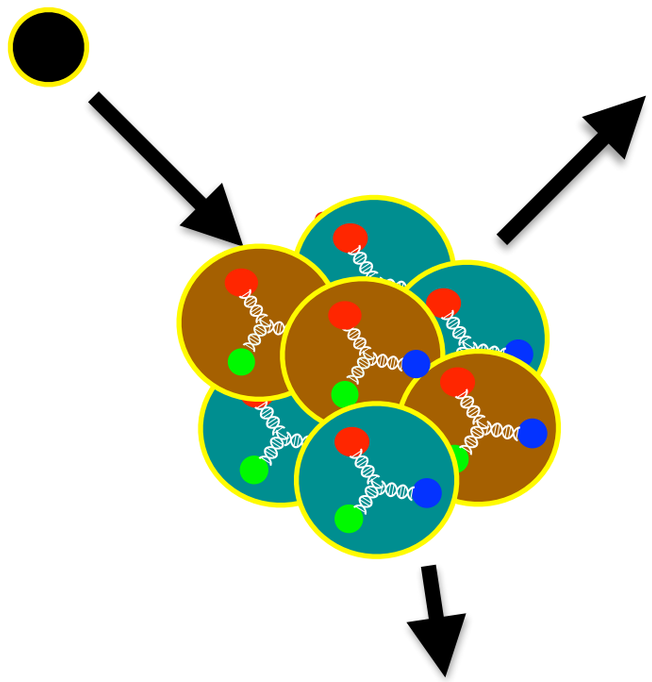
Why don't we just compute ν-argon?

Lattice QCD is a many-body method — just simulate a few 100 quarks





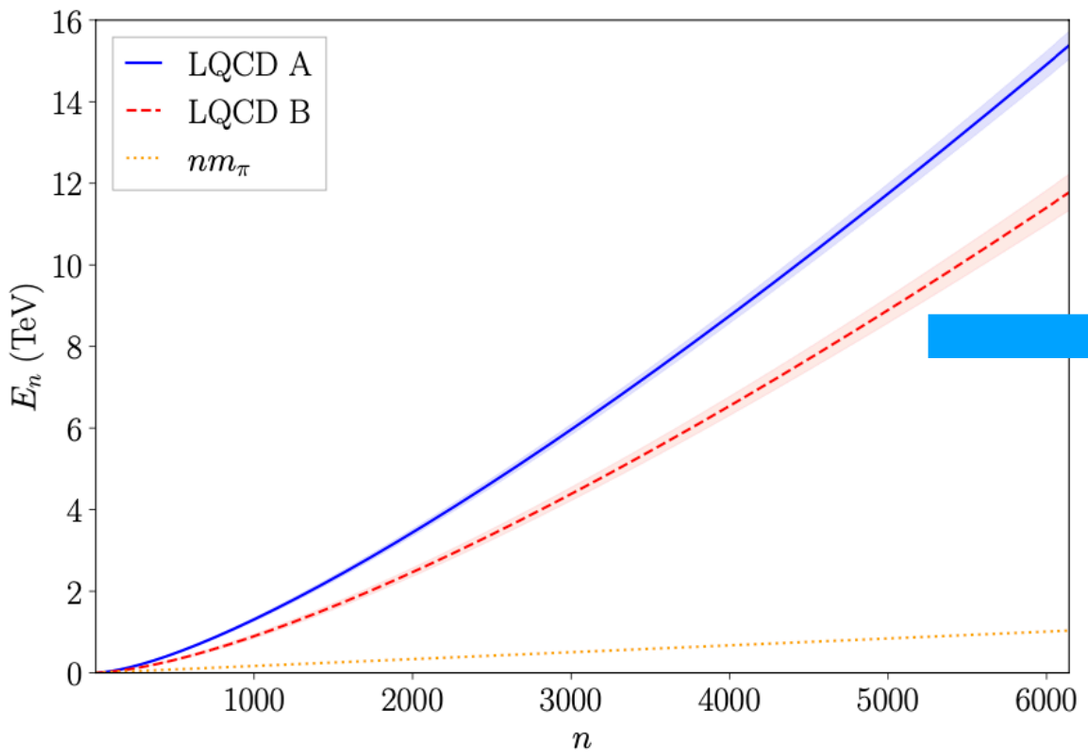




Why don't we just compute ν-argon?

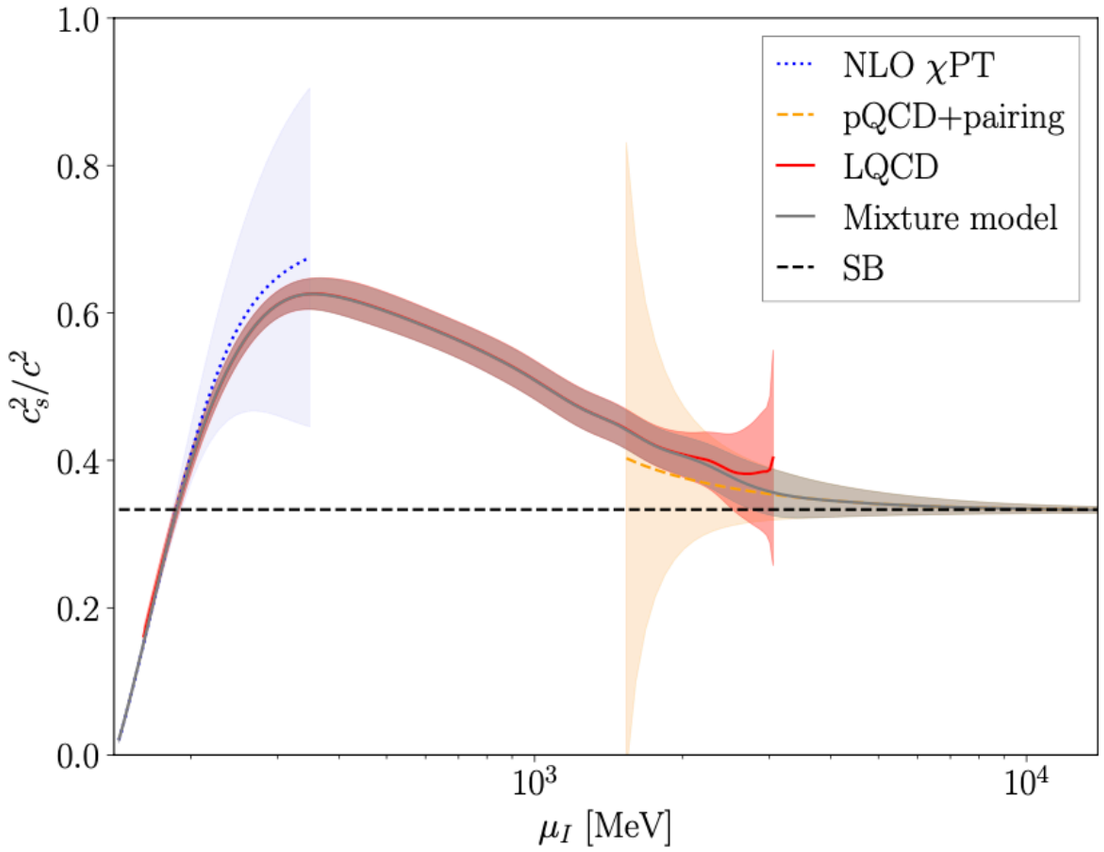
Lattice QCD is a many-body method — just simulate a few 100 quarks

Energy spectrum of up to 6000 pions in a box:



Abbot, Detmold, Romero-Lopez, MW et al [NPLQCD], PRD 108 (2023)

Speed of sound at large isospin density



Abbot, Detmold, Romero-Lopez, MW et al [NPLQCD], arXiv:2406.09273

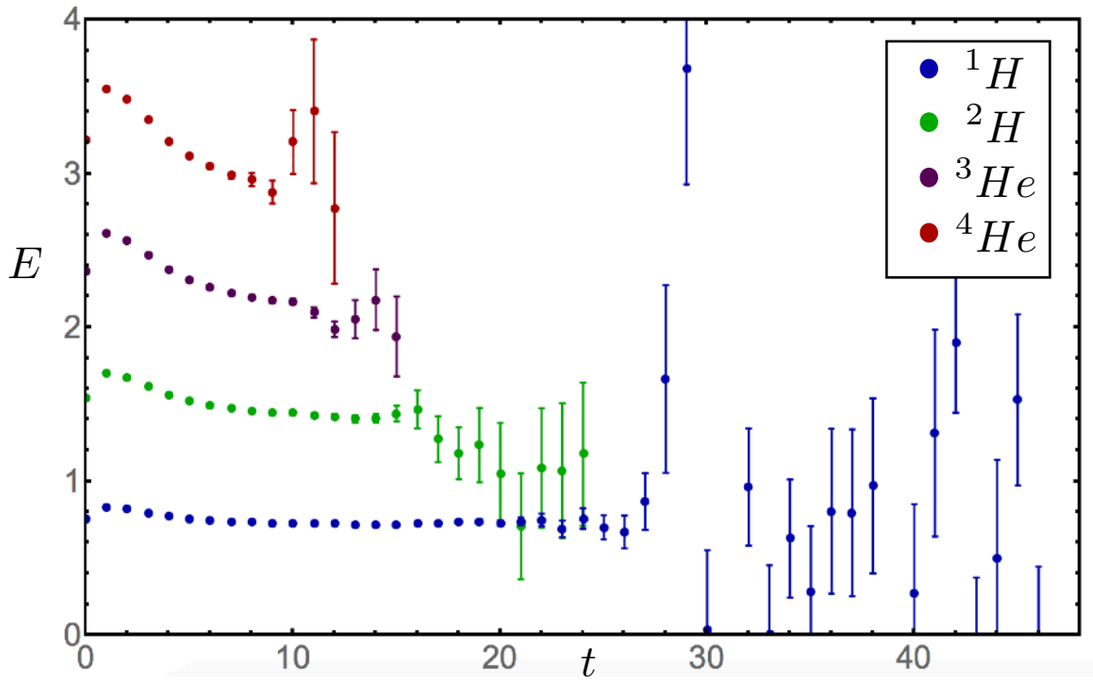
What's so hard about nuclei?

Lattice QCD is a many-body method — just simulate a few 100 quarks

1) Too many Wick contractions

2) Small energy gaps to excited states

3) Exponential signal-to-noise degradation



$$aE(t) = -\ln \frac{C(t+a)}{C(t)} = aE_0 + \dots$$

What's so hard about nuclei?

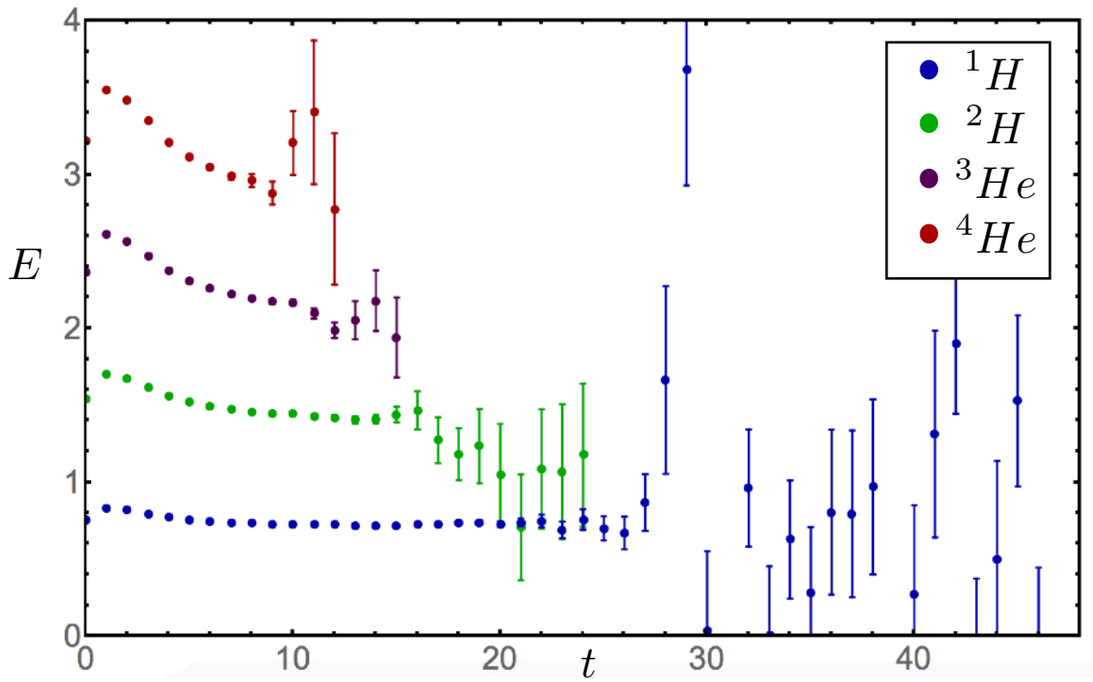
Lattice QCD is a many-body method — just simulate a few 100 quarks

~~**1) Too many Wick contractions**~~

Detmold and Orginos, PRD 87 (2013)

2) Small energy gaps to excited states

3) Exponential signal-to-noise degradation



$$aE(t) = -\ln \frac{C(t+a)}{C(t)} = aE_0 + \dots$$

What's so hard about nuclei?

Lattice QCD is a many-body method — just simulate a few 100 quarks

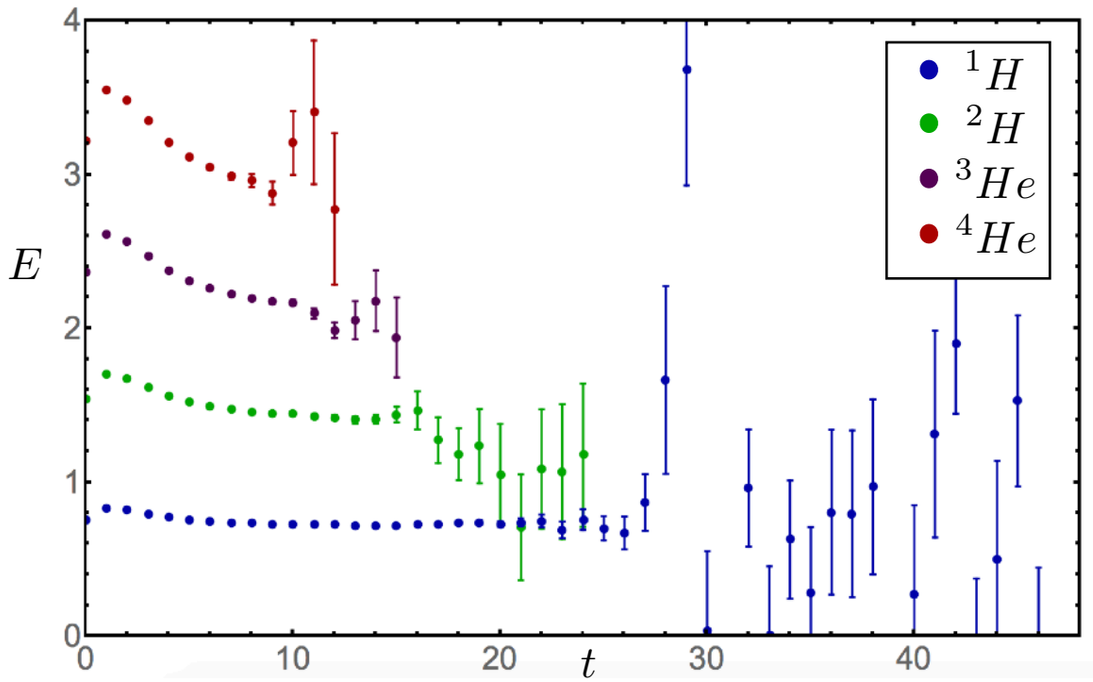
~~**1) Too many Wick contractions**~~

Detmold and Orginos, PRD 87 (2013)

2) Small energy gaps to excited states

$$\delta \approx 4\pi^2/(M_N L^2) \quad \text{or} \quad \delta \approx B_A$$

3) Exponential signal-to-noise degradation



$$aE(t) = -\ln \frac{C(t+a)}{C(t)} = aE_0 + \dots$$

What's so hard about nuclei?

Lattice QCD is a many-body method — just simulate a few 100 quarks

~~**1) Too many Wick contractions**~~

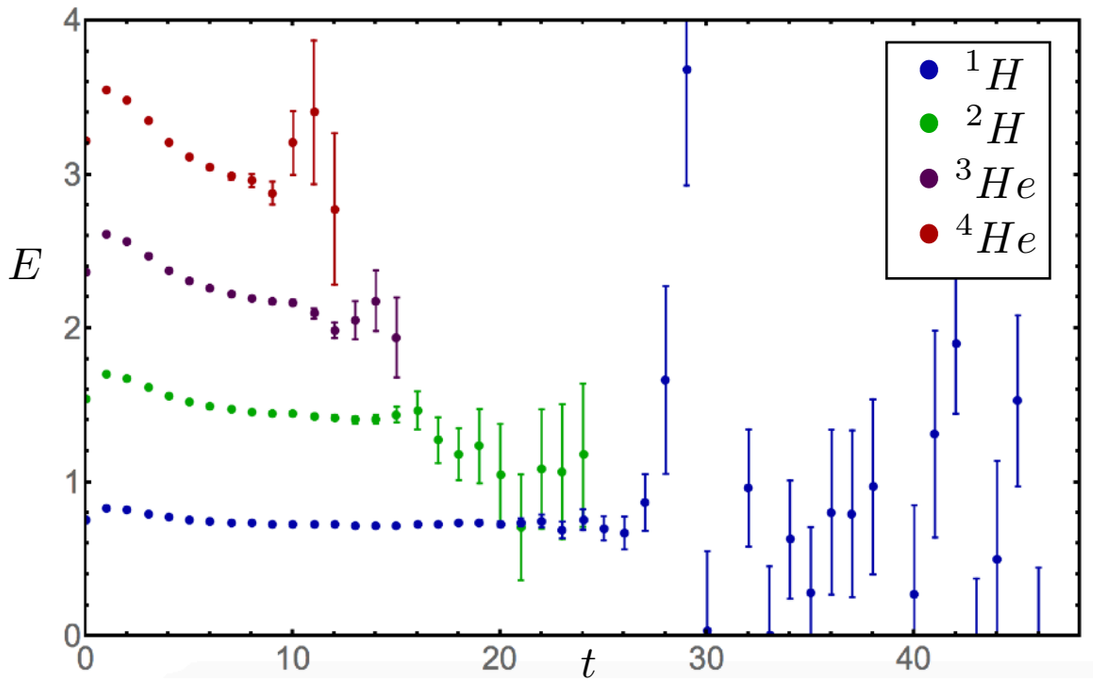
Detmold and Orginos, PRD 87 (2013)

2) Small energy gaps to excited states

$$\delta \approx 4\pi^2/(M_N L^2) \quad \text{or} \quad \delta \approx B_A$$

3) Exponential signal-to-noise degradation

$$\mathrm{StN} \sim e^{-A(M_N - \frac{3}{2}m_\pi)t}$$



$$aE(t) = -\ln \frac{C(t+a)}{C(t)} = aE_0 + \ldots$$

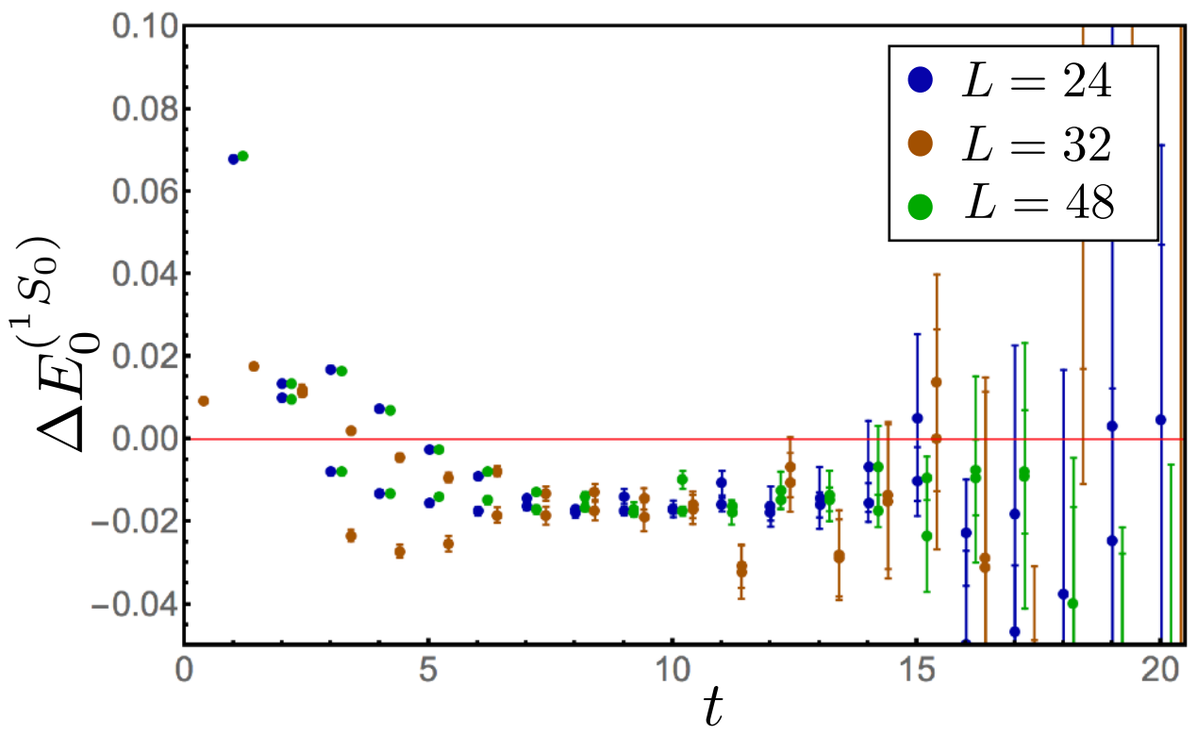
Getting large enough imaginary times to suppress excited-state effects can be challenging or impossible for multi-nucleon systems

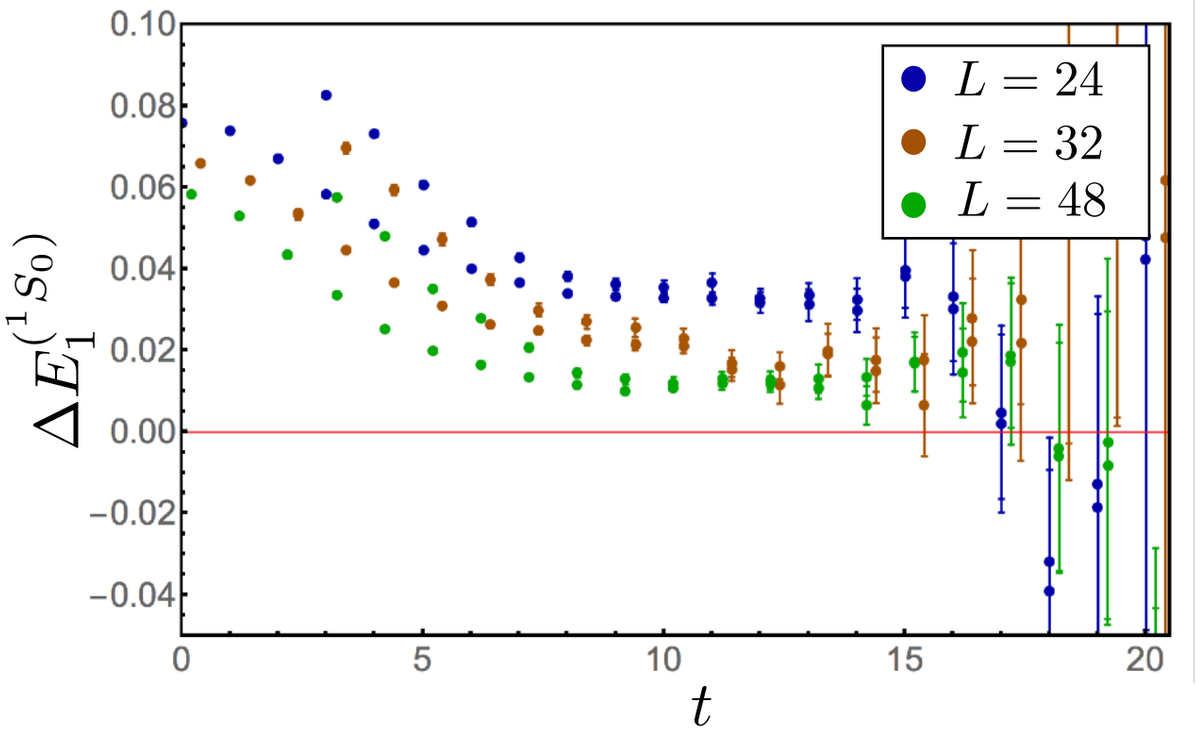
Nuclei from LQCD

First calculations of 2-5 baryon (asymmetric) correlation functions

Beane et al [NPLQCD], PRD 87 (2013)	$L = 2.9 \text{ fm} \to 5.8 \text{ fm}$	$a = 0.145 \text{ fm}$	$m_\pi \sim 806 \text{ MeV}$
Yamazaki et al, PRD 86 (2012)	$L = 3.5 \text{ fm} \to 7.0 \text{ fm}$	$a = 0.09 \text{ fm}$	$m_\pi \sim 510 \text{ MeV}$

- Ground state energy appears approximately volume independent
- First excited state shows volume dependence consistent with unbound
- Operators with two different smearings give consistent results



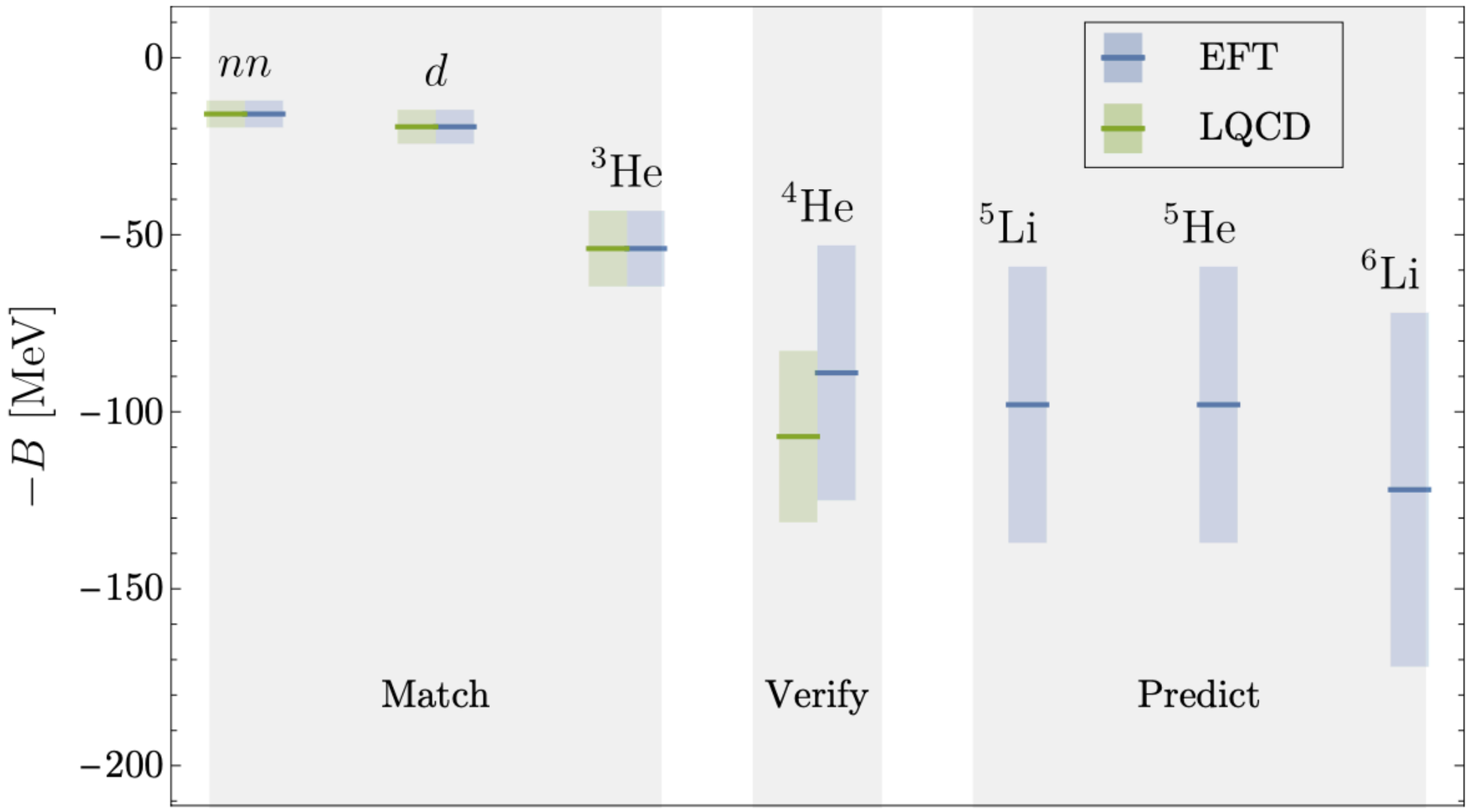


Data from Beane et al [NPLQCD], PRD 87 (2013)

Nuclei from LQCD + EFT

First calculations of 2-5 baryon (asymmetric) correlation functions

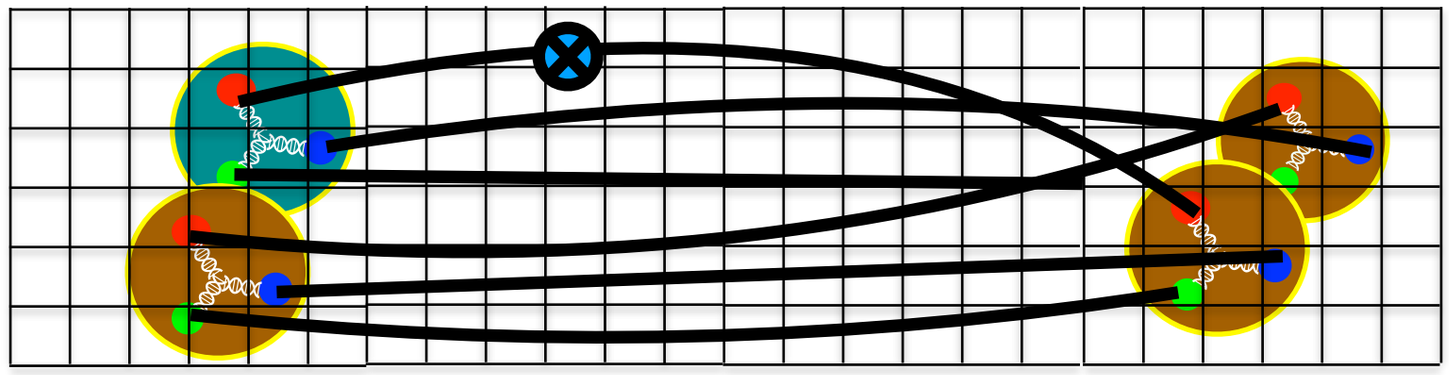
Beane et al [NPLQCD], PRD 87 (2013) $L = 2.9 \text{ fm} \to 5.8 \text{ fm}$ $a = 0.145 \text{ fm}$ $m_\pi \sim 806 \text{ MeV}$



EFT: Barnea et al, PRL 114 (2015)

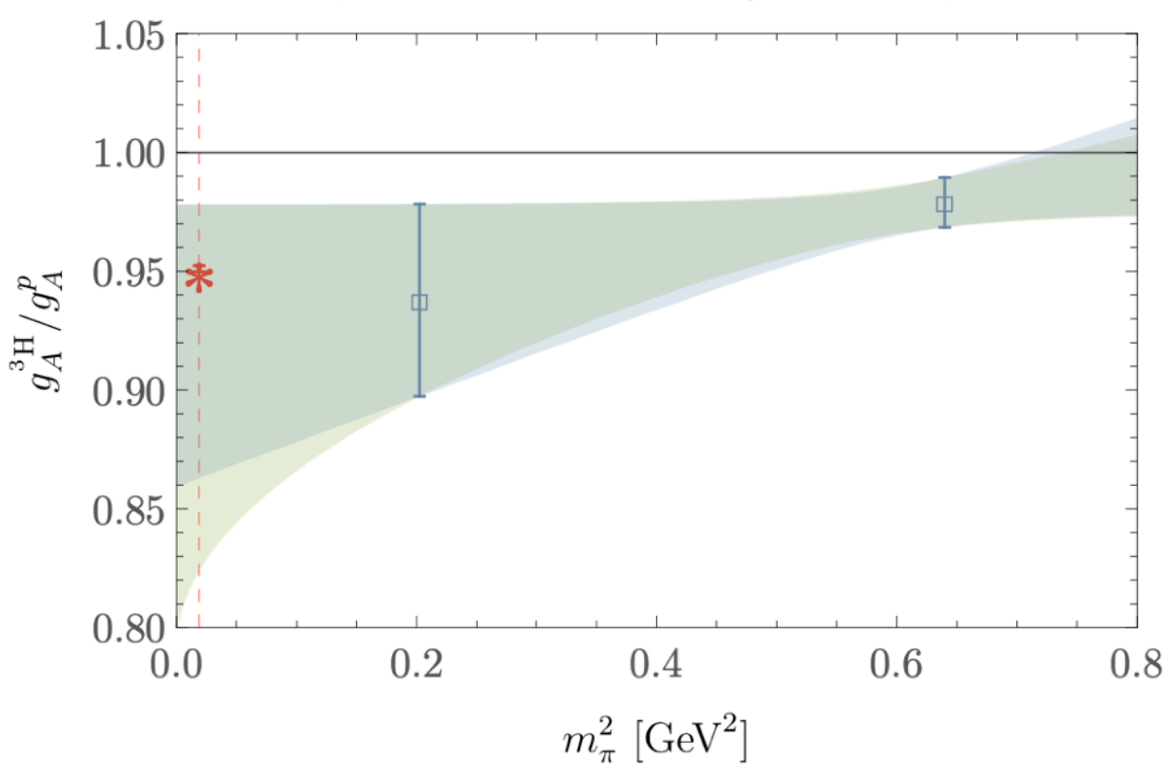
Two-body axial currents

Two-nucleon axial matrix elements relevant for proton-proton fusion computed, used to constrain two-body currents in pionless EFT for $m_\pi = 806\ \mathrm{MeV}$



Savage, MW et al [NPLQCD], PRL 119 (2017)

Axial current matrix element calculations with $m_\pi = 450\ \mathrm{MeV}$ permit preliminary extrapolations to physical quark masses



Parreño, MW et al [NPLQCD] PRD 103 (2021)

Matrix elements of two axial currents constrain $2\nu\beta\beta$ in pionless EFT

Shanahan, MW et al [NPLQCD], PRL 119 (2017)

Tiburzi, MW et al [NPLQCD], PRD 96 (2017)

More complicated two-body currents important for $0\nu\beta\beta$, first study:

Davoudi, Grebe, MW et al [NPLQCD], arXiv:2402.09362

Systematic uncertainties

Present-day LQCD studies of nuclei still have several systematic uncertainties that need to be studied in detail

- Heavier than physical quark masses only
- One lattice spacing
- Excited-state effects

Systematic uncertainties

Present-day LQCD studies of nuclei still have several systematic uncertainties that need to be studied in detail

- Heavier than physical quark masses only
- One lattice spacing
- Excited-state effects

Gap between ground and two-nucleon finite-volume “scattering” states becomes small for large volumes, ground-state dominance relies on overlap factors

$$Z_0 e^{-E_0 t}\left(1+\frac{Z_1}{Z_0}e^{-\delta t}+\ldots\right) \qquad \delta \sim \frac{4\pi^2}{ML^2}$$

For non-positive-definite correlation functions, cancellations between the ground and excited-state could in principle conspire to form a “false plateau”

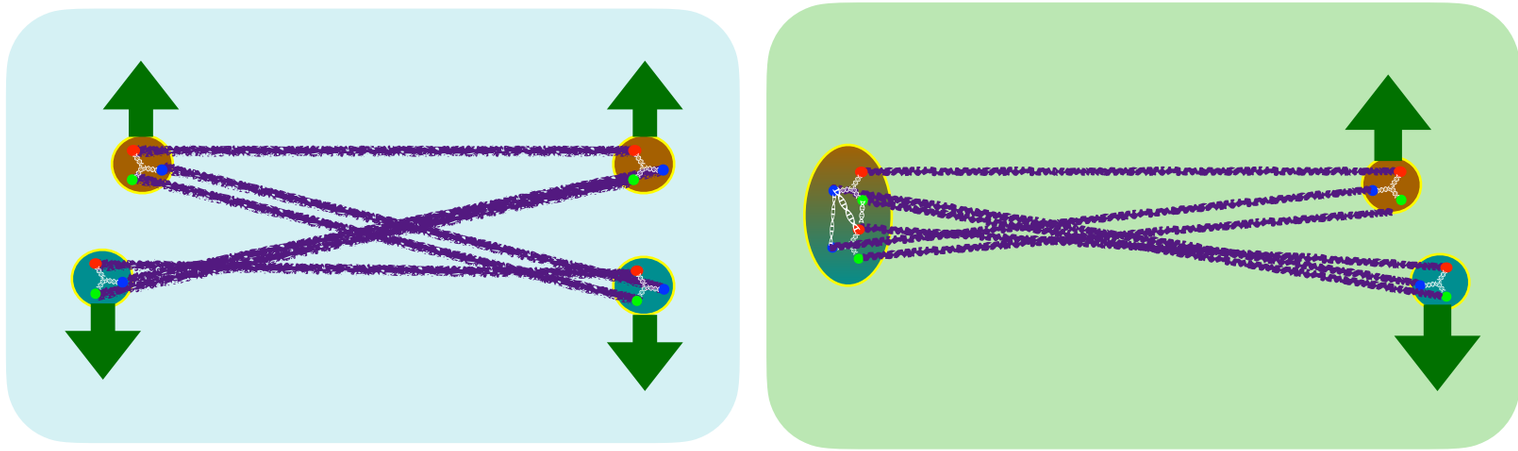
See e.g. Iritani et al, JHEP 10 (2016)

All Z factors in spectral representation guaranteed to be positive for symmetric correlation functions

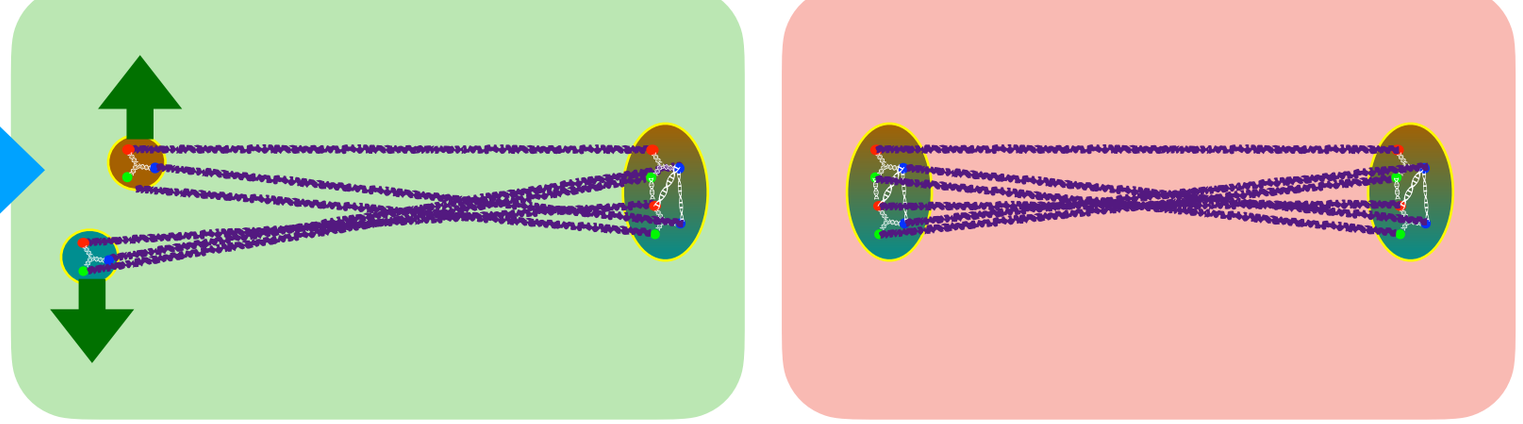
$$\langle \mathcal{O}\ \overline{\mathcal{O}}\rangle = \sum_n |Z_n|^2 e^{-E_n T}$$

Variational methods

Robust upper bounds on energy spectrum can be obtained by diagonalizing symmetric matrices of correlation functions



All LQCD nuclear matrix element calculations so far



Although application of variational methods to multi-nucleon systems has long been advocated, it has only recently become computationally feasible

Distillation:

Peardon et al PRD 80 (2009)

Morningstar et al PRD 83 (2011)

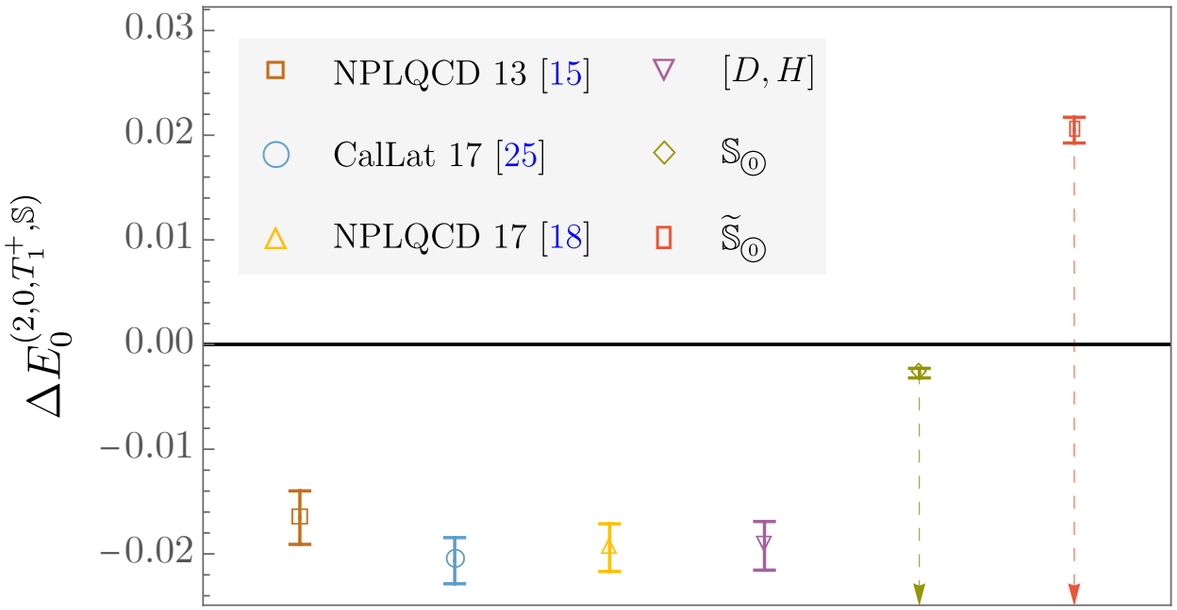
Sparsening:

Detmold, MW et al, PRD 104 (2021)

Li et al, PRD 103 (2021)

Two-nucleon variational bounds

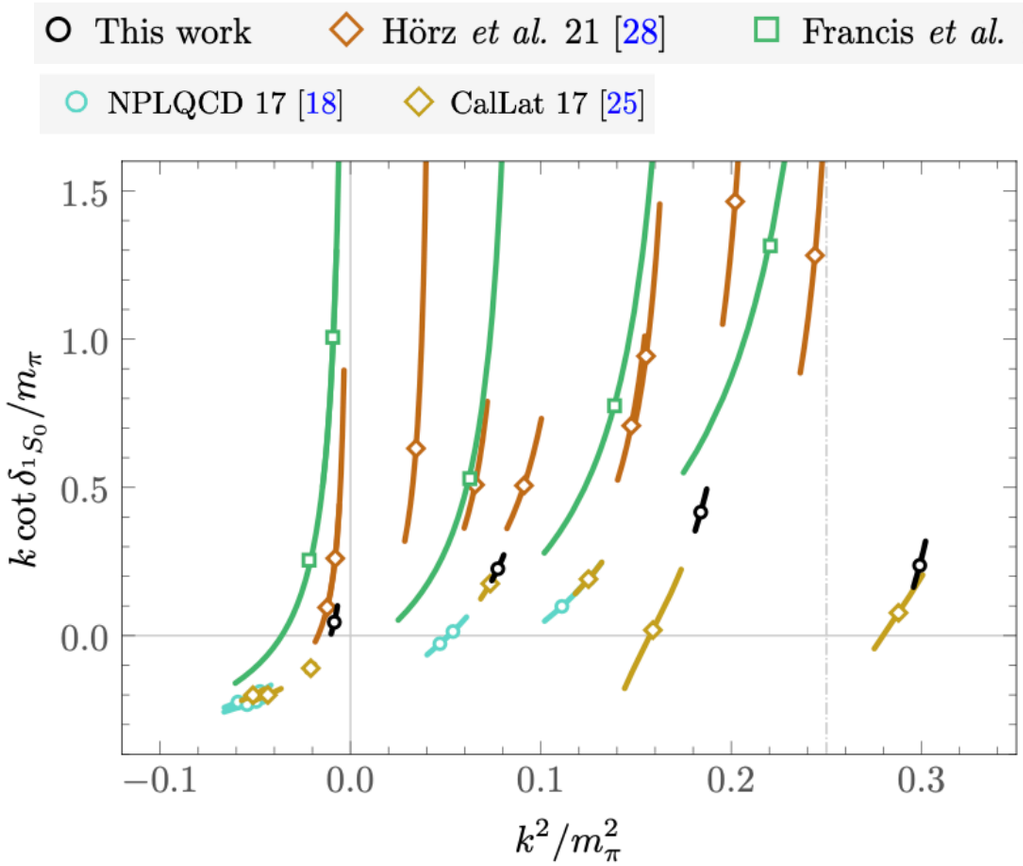
Variational upper bounds obtained using different interpolating operator sets are consistent



Ground-state energy **estimates** using different interpolating-operator sets show large discrepancies

Phase shifts obtained using asymmetric vs variational energy estimates suggest qualitatively different physics (bound vs unbound)

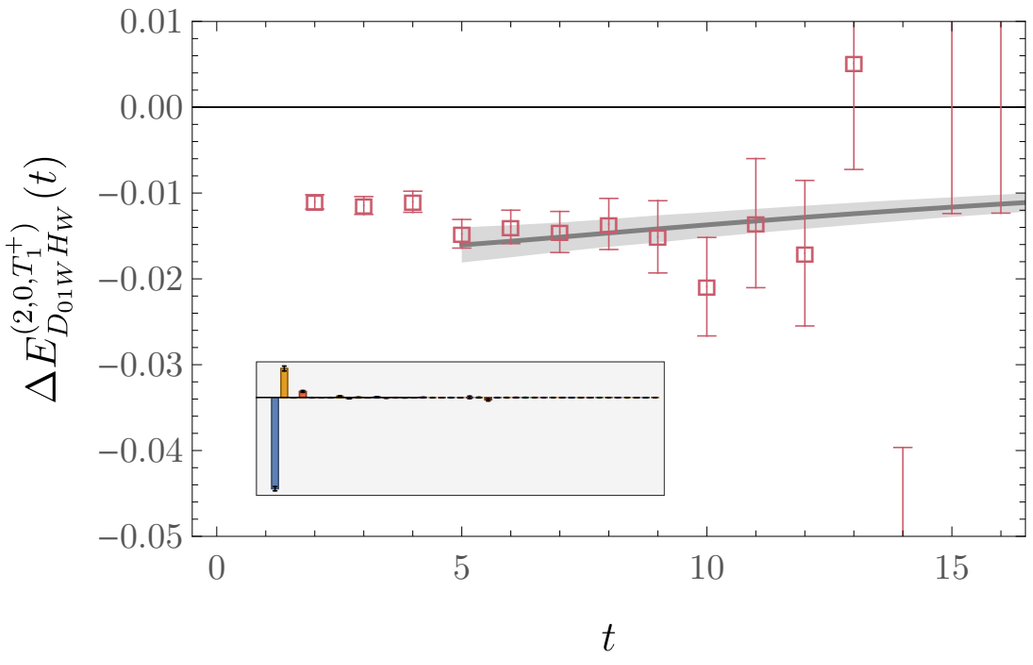
Amarasinghe, MW et al [NPLQCD], PRD 107 (2023)



Results by different groups using similar interpolating operators show good consistency

Excited states or overlap problem?

Apparent plateau of asymmetric correlators can be reproduced by spectral representation from variational results



Variational predicts asymmetric slowly approaches from below for

$$t \gg 40a \sim 6 \text{ fm}$$

Toy model: 2 operators, 3 states

$$Z_{\mathsf{n}}^{(A)} = (\epsilon, \sqrt{1-\epsilon^2}, 0)$$
$$Z_{\mathsf{n}}^{(B)} = (\epsilon, 0, \sqrt{1-\epsilon^2})$$

- Both operators have small overlap ϵ with ground state
- Operators are approximately orthogonal

GEVP eigenvalues controlled by first and second excited state (**not** ground state) for

$$\epsilon^2 \ll e^{t(E_1 - E_0)}$$

Off-diagonal correlator conversely has perfect ground-state overlap

Lanczos, the transfer matrix, and the signal-to-noise problem

MW, arXiv:2406.20009

Hackett, MW, arXiv:2407.21777

+ ongoing work with Grebe, Fleming, Rocco, Tame Narvaez, Van de Water, …

Cornelius Lanczos

Daniel Hackett

Spectroscopy = finding 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)

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

Correlation functions are matrix elements of powers of T

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

Life in Hilbert space

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

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

Hilbert space is big, even for a single gauge-field link

— standard in quantum mechanics, sometimes not in linear algebra

Life in Hilbert space

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

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

Hilbert space is big, even for a single gauge-field link

— standard in quantum mechanics, sometimes not in linear algebra

- The transfer matrix acts simply in this basis

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

Life in Hilbert space

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

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

Hilbert space is big, even for a single gauge-field link

— standard in quantum mechanics, sometimes not in linear algebra

- 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)

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

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

von Mises and Pollaczek-Geiringer, Zeitschrift Angewandte Mathematik und Mechanik 9, 58 (1929)

The power-iteration algorithm

Start with an arbitrary normalized initial state: $\left|b_0\right\rangle = \left|\psi\right\rangle / |\psi|$

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

Eigenvalue convergence: $\left|b_k\right\rangle \propto T^k \left|\psi\right\rangle = e^{-kaE_0}\left|\psi\right\rangle Z_0 + O\left(e^{-k\delta}\right)$

Energies from power-iteration eigenvalues:

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

The power-iteration algorithm

Start with an arbitrary normalized initial state: $|b_0\rangle = |\psi\rangle/|\psi|$

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

Eigenvalue convergence: $|b_k\rangle \propto T^k|\psi\rangle = e^{-kaE_0}|\psi\rangle Z_0 + O\left(e^{-k\delta}\right)$

Energies from power-iteration eigenvalues:

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

$$= -\ln\left[\frac{C((2k+1)a)}{C(2ka)}\right] = E_{\text{eff}}(t/a = 2k+1)$$

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

Lanczos and the transfer matrix

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

$$|b_k\rangle \propto T^k |\psi\rangle \quad \Rightarrow \quad -\ln\langle b_k|T|b_k\rangle = -\ln\left[\frac{C((2k+1)a)}{C(2ka)}\right] = E_{\text{eff}}((2k+1)a)$$

von Mises and Pollaczek-Geiringer, Zeitschrift Angewandte Mathematik und Mechanik 9, 58 (1929)

- Modern computational linear algebra uses more sophisticated methods, e.g.

Lanczos algorithm

Lanczos, J. Res. Natl. Bur. Stand. B 45, 255 (1950)

Applied to LQCD since at least Barbour et al (1984)

$$|v_j\rangle \propto [T - T^{(m)}]|v_{j-1}\rangle$$

$$T_{ij}^{(m)} = \langle v_i|T|v_j\rangle \quad \Rightarrow \quad E_k^{(m)} = -\ln\lambda_k^{(m)}$$

- Exponentially faster convergence for systems with small gaps $\delta = a(E_1 - E_0)$

Kaniel, Mathematics of Computation 20, 369 (1966)

Paige, PhD thesis 1971

Saad, SIAM 17 (1980)

$$\left|E_0 - E_0^{(m)}\right| \propto e^{-4m\sqrt{\delta}} \ll \left|E_0 - E^{\text{eff}}(2m)\right| \propto e^{-2m\delta}$$

Lanczos in Hilbert space

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

Novel features not present in power iteration

- Lanczos vector form an orthonormal basis for Krylov space

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

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

- 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 ~ finite-dimensional EFT where we know all the matrix elements

Living on the Ritz

Krylov space ~ finite-dimensional EFT where we know all the matrix elements

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

Diagonalize the Krylov-space transfer matrix:

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

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

"Ritz vectors" = corresponding approximation eigenstates

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

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

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:

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

Recursive Lanczos iteration:

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

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

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

• • •

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:

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

Recursive Lanczos iteration:

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

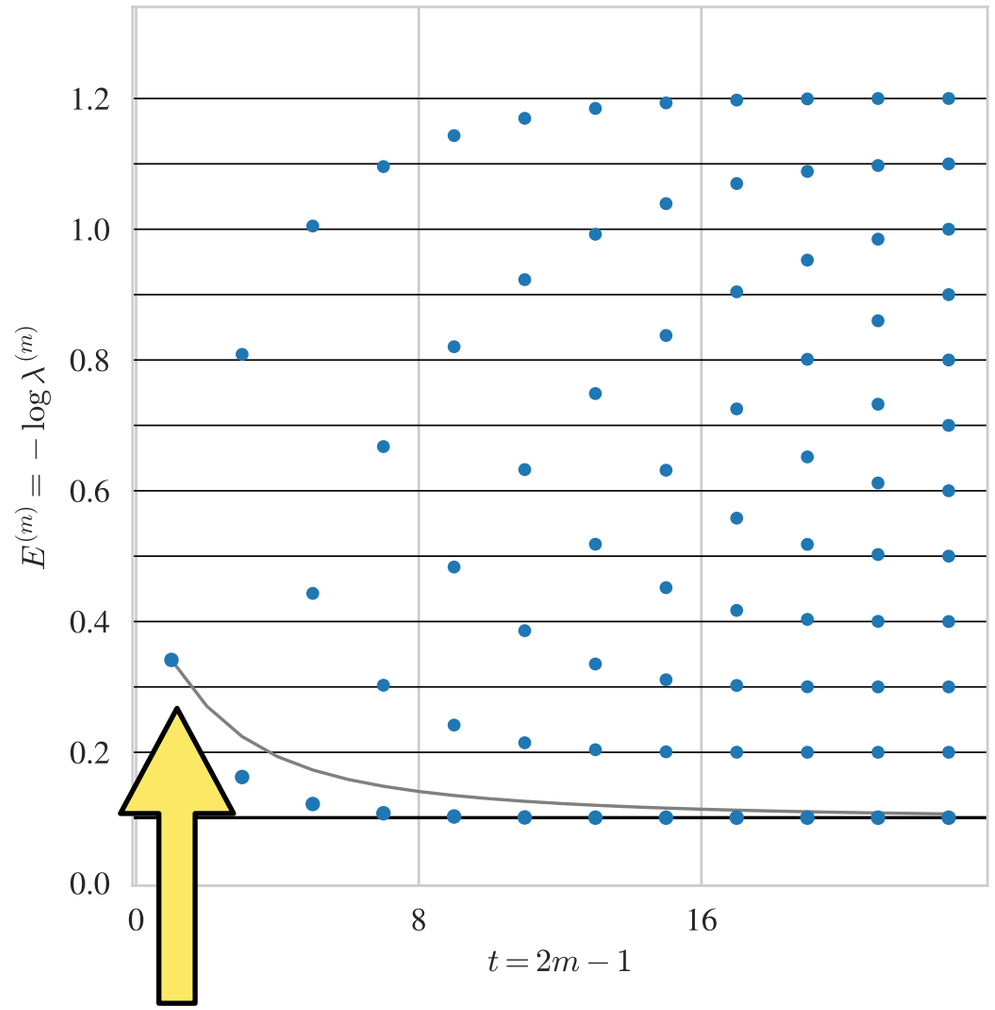
$$\beta_{j+1} = \sqrt{A_j^2 - \alpha_j^2 - \beta_j^2}$$

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

• • •

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 $m > 1$

The residual bound

- 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 $T^{(m)}_{m(m+1)}$

Paige, PhD thesis 1971

Rigorous quantification of excited-state effects!

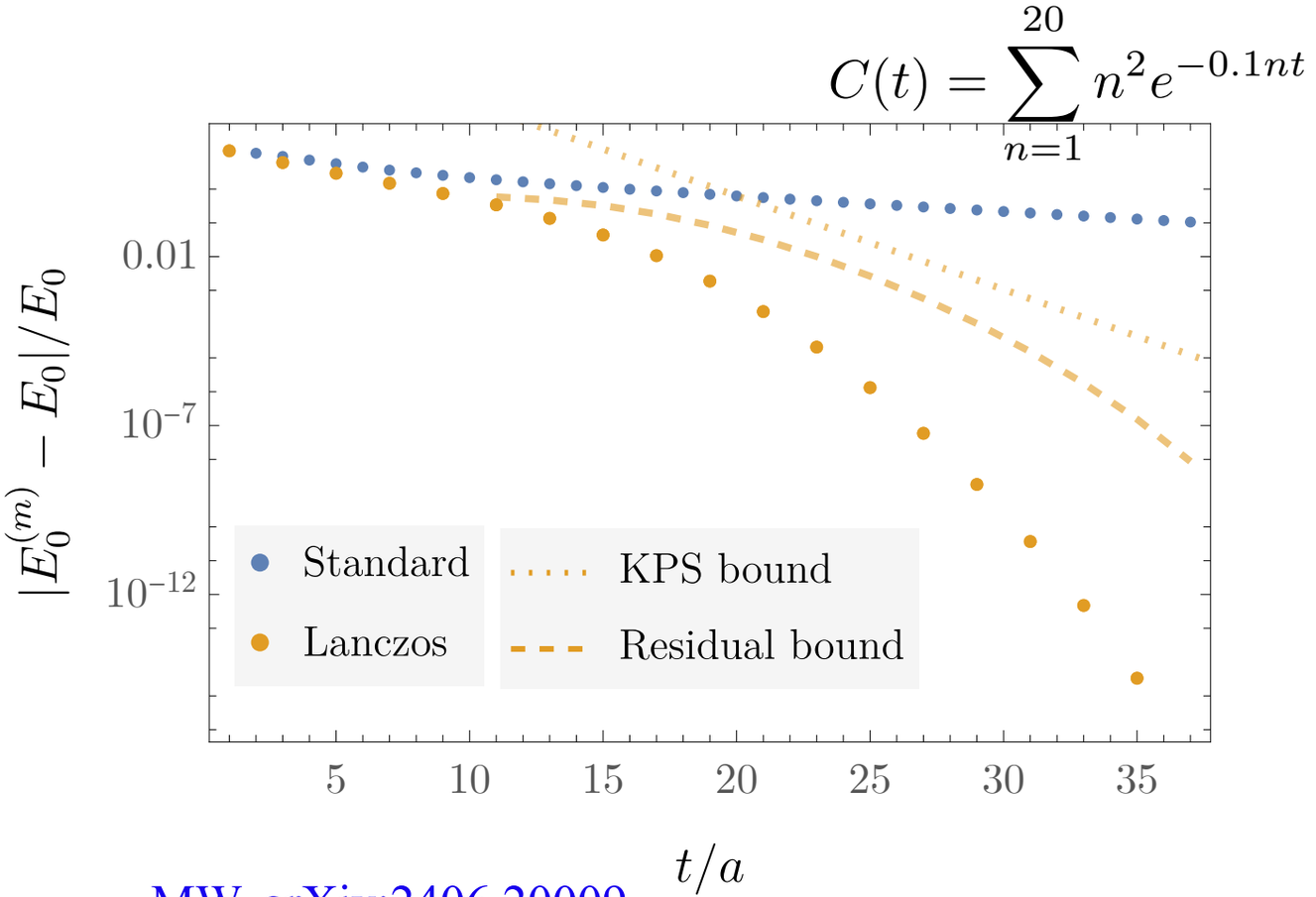
Mock data tests demonstrate

- Lanczos converges exponentially faster than standard effective mass

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

$$|E_0 - E_{\text{eff}}(t)| \propto e^{-\delta(t/a)}$$

- Residual bound provides valid two-sided bound on errors from excited-state effects



MW, arXiv:2406.20009

Will noise destroy Lanczos?

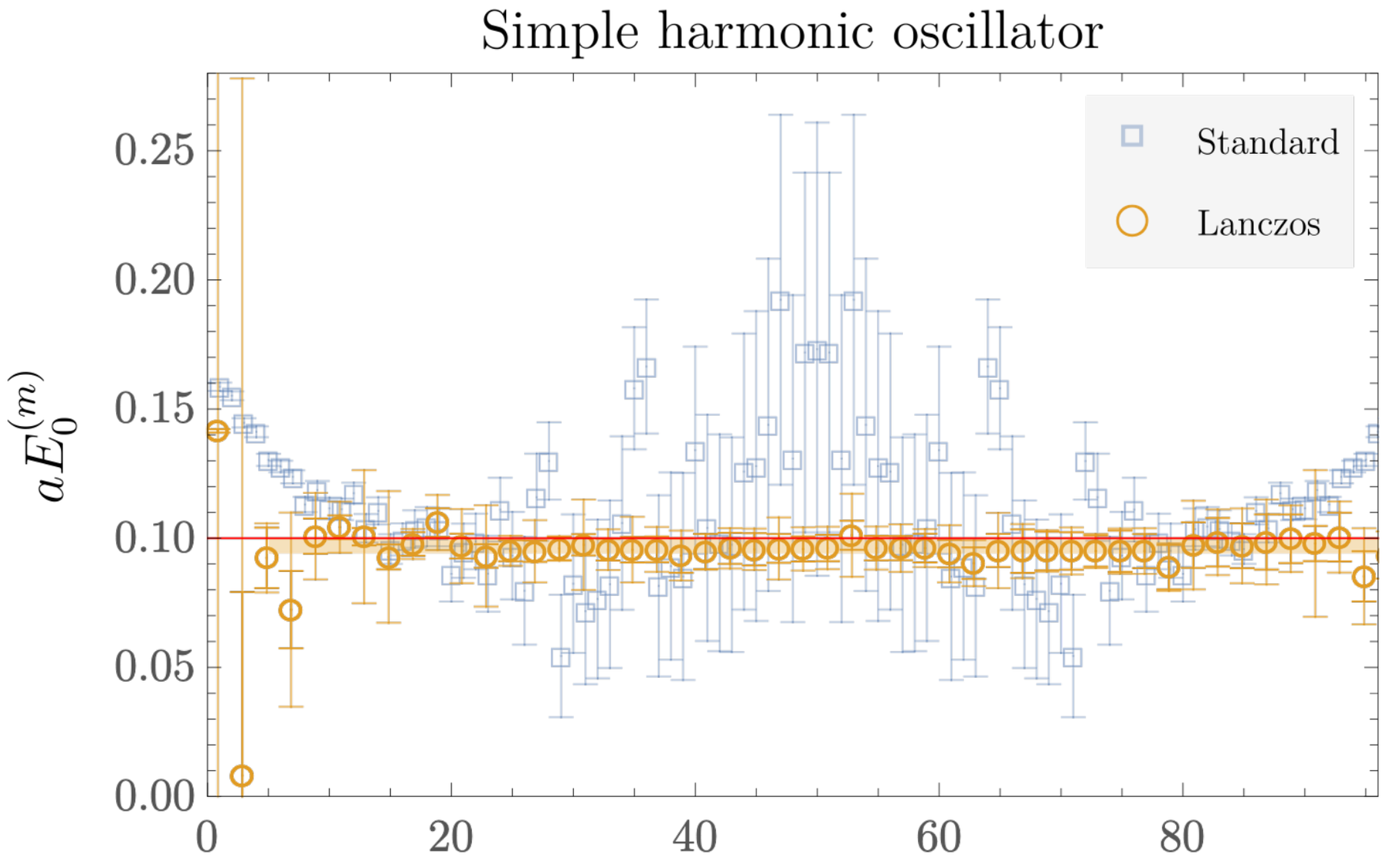
Will noise destroy Lanczos?

- No

Will noise destroy Lanczos?

- No

- Lanczos is surprisingly robust to large-time correlation function noise



MW, arXiv:2406.20009

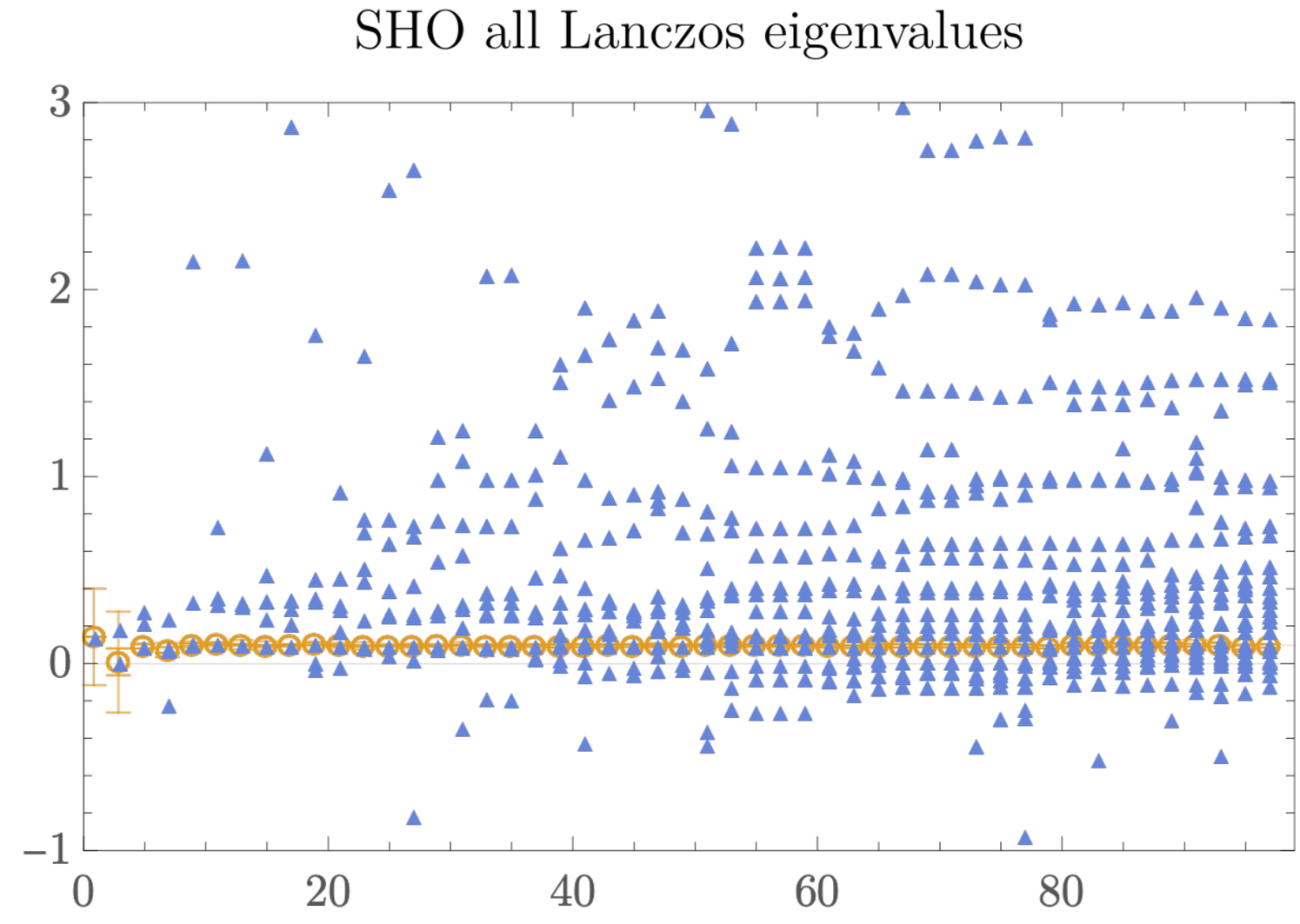
Is it really that easy?

Is it really that easy?

- No

Is it really that easy?

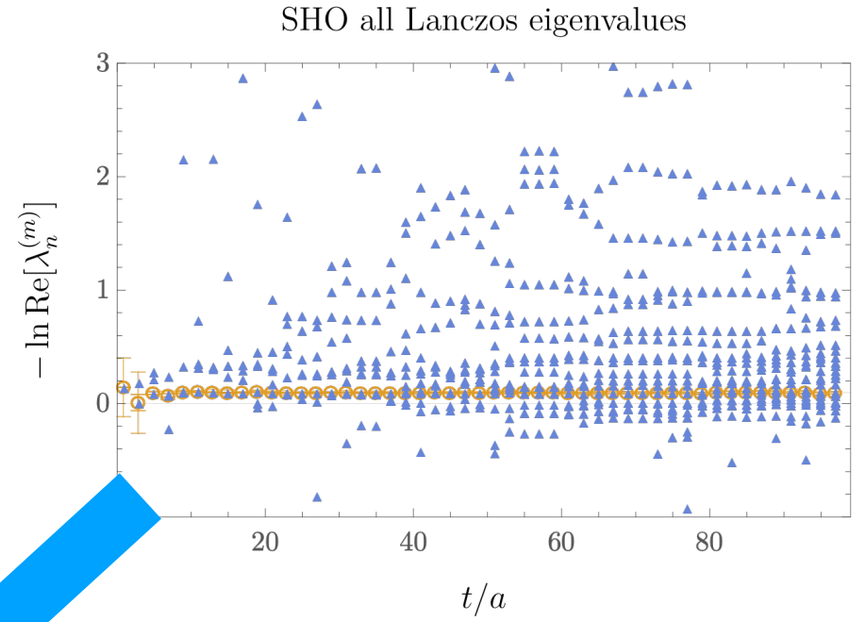
- No
- Lanczos produces an increasingly dense forest of “spurious eigenvalues”

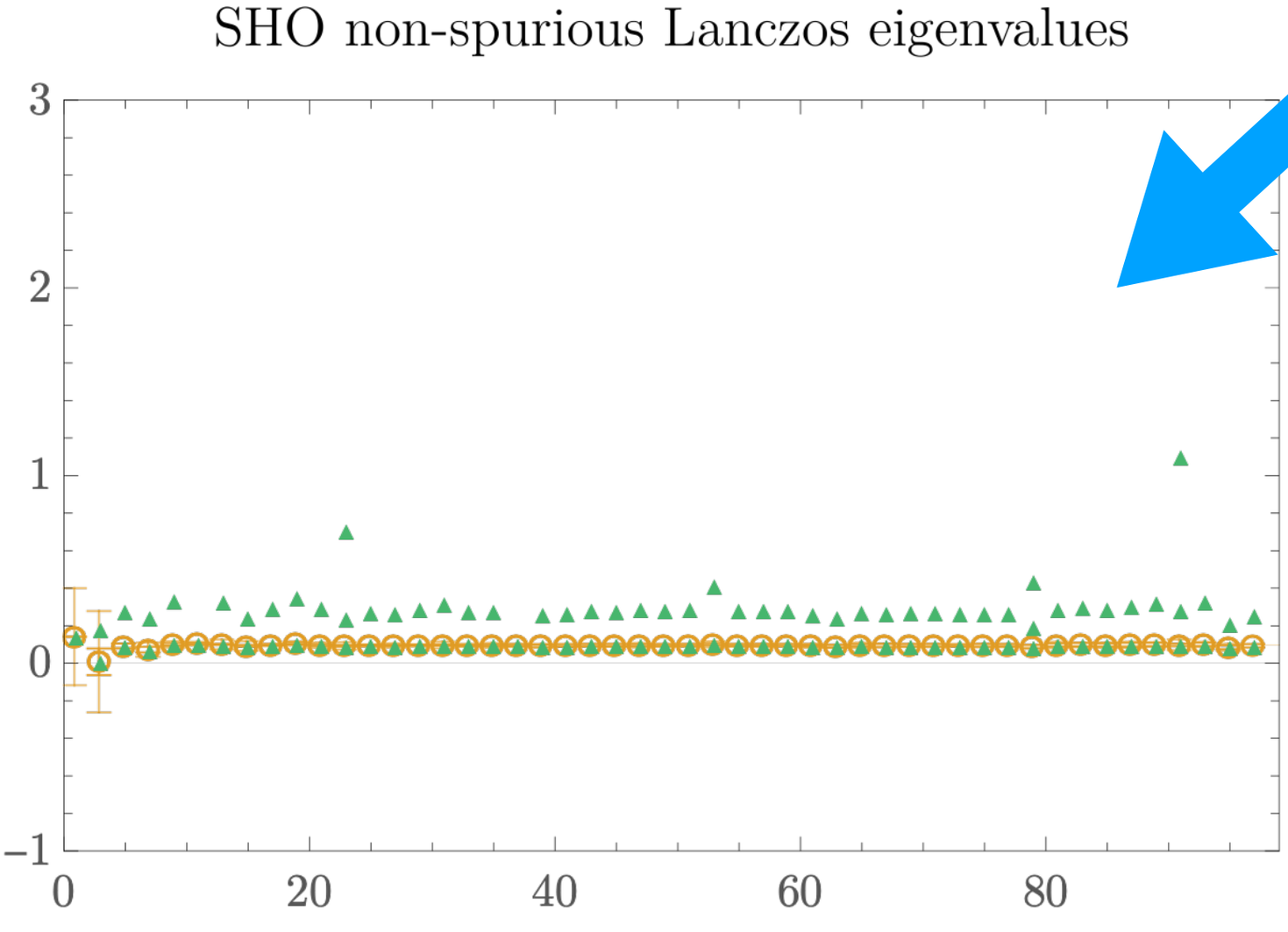


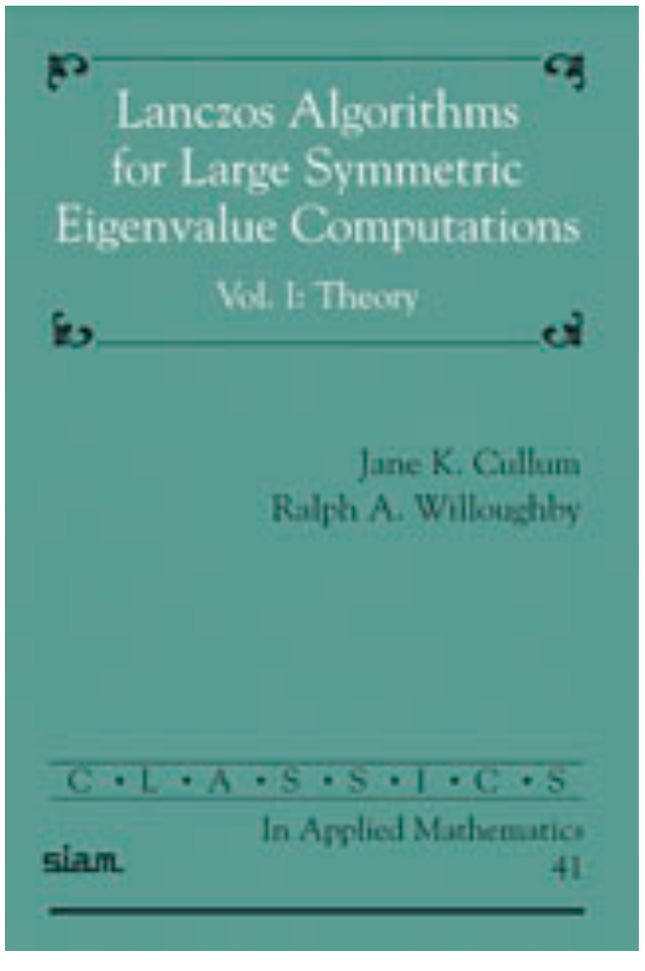
MW, arXiv:2406.20009

Spurious eigenvalues

- We need a way to automatically detect which eigenvalues are spurious and get rid of them







MW, arXiv:2406.20009

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.

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

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

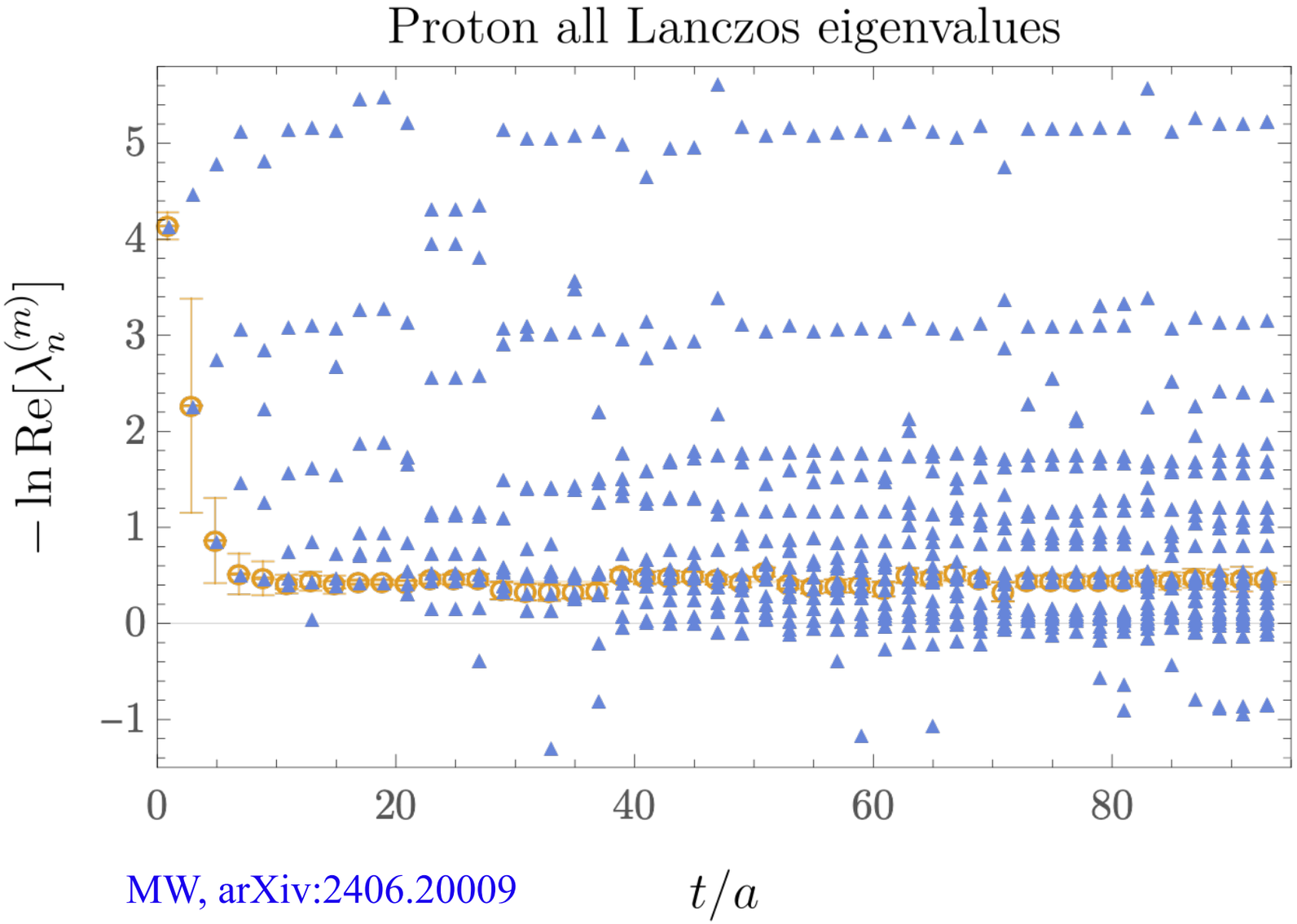
DEFINITION 2. Any simple eigenvalue of T_m that is pathologically close to an eigenvalue of $\hat{T}_2$ will be called "spurious."

Think positive

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

- “Non-zero” can be kept exact even in the presence of noise by adopting oblique Lanczos formalism

Saad, SIAM 19 (1982)



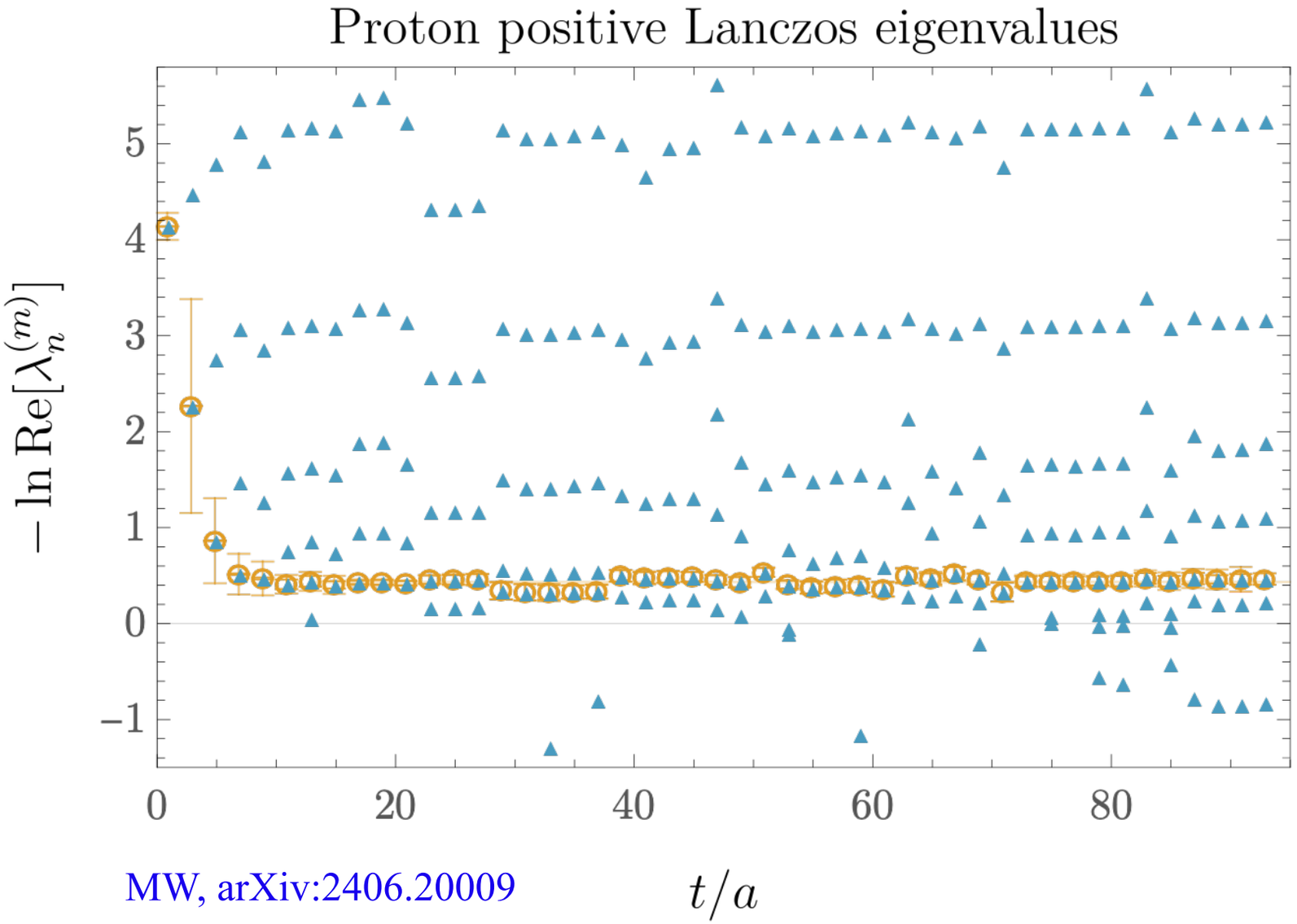
MW, arXiv:2406.20009

Think positive

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

- "Non-zero" can be kept exact even in the presence of noise by adopting oblique Lanczos formalism

Saad, SIAM 19 (1982)



- This gets rid of many spurious eigenvalues but still leaves some that must be wrong because they correspond to $M_N < m_\pi$

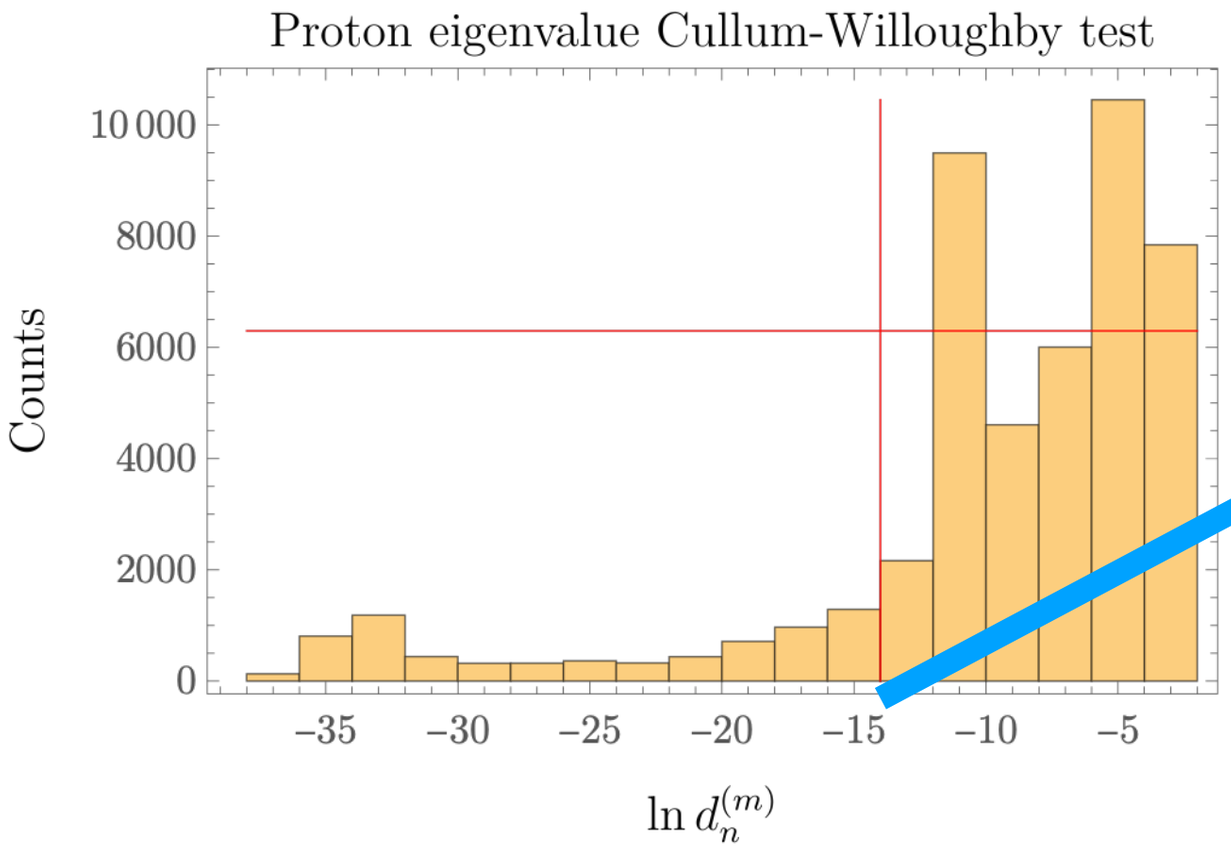
MW, arXiv:2406.20009

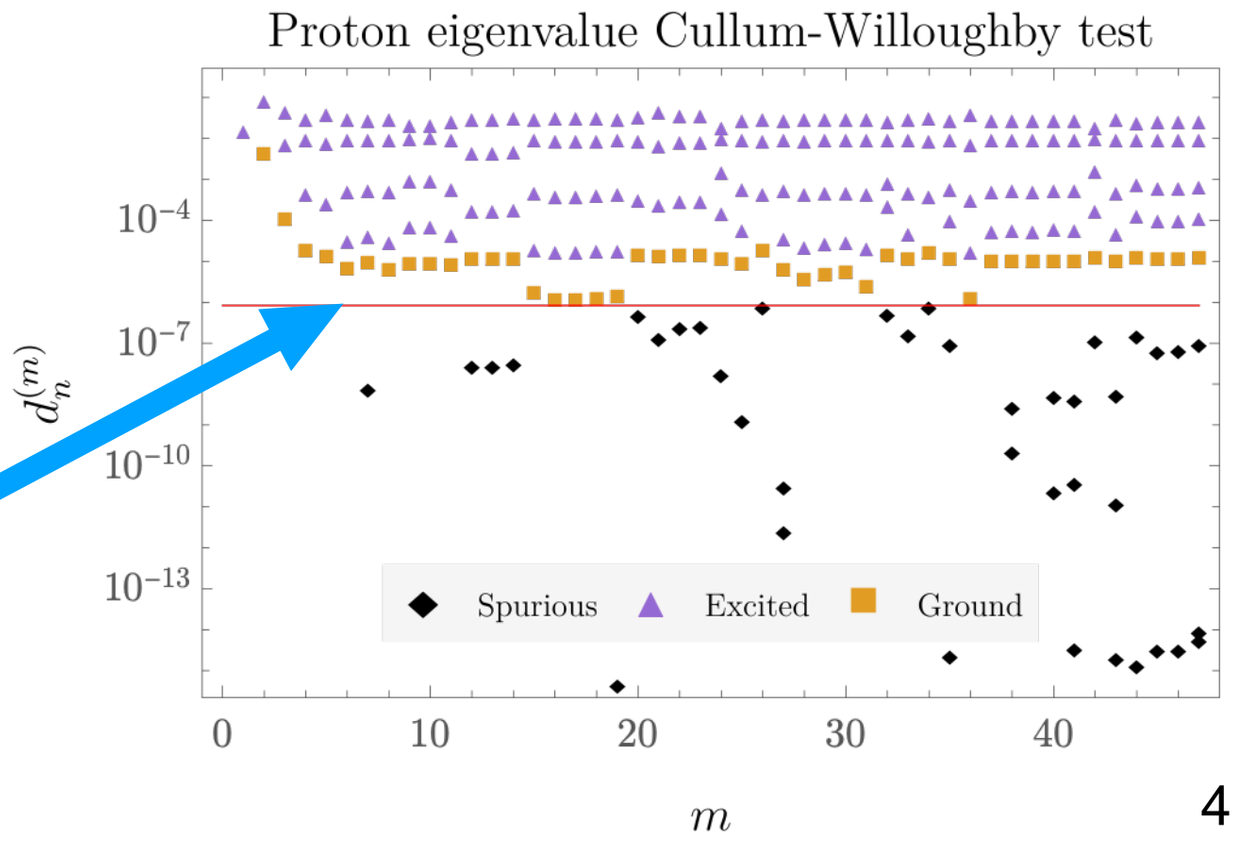
Bootstrapping Cullum-Willoughby

- 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



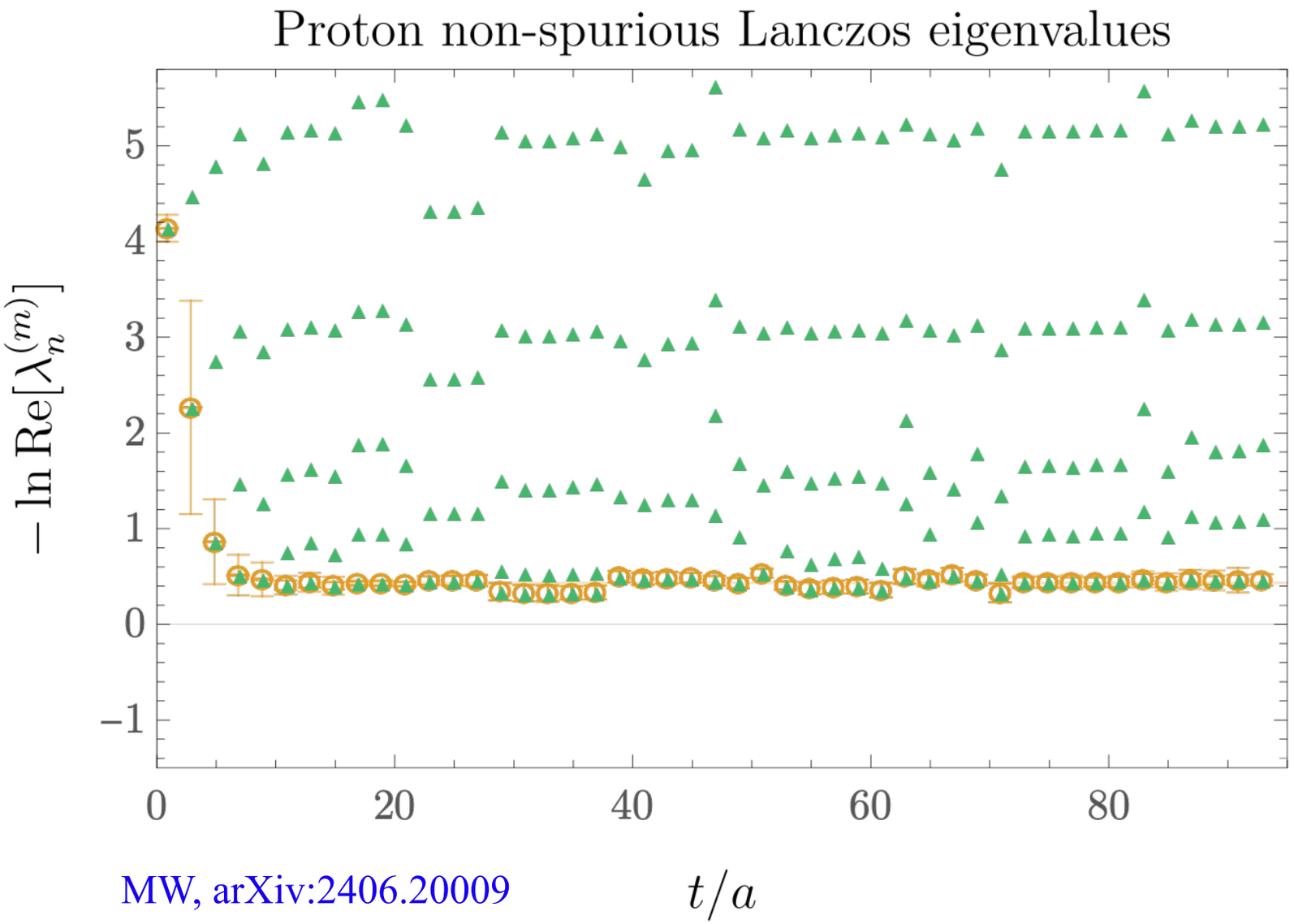


Non-spurious proton energies

- Largest eigenvalue not removed as spurious defines ground-state energy

$$E_0 = -\ln \lambda_0^{(m)}$$

- Excited-state energies also accessible

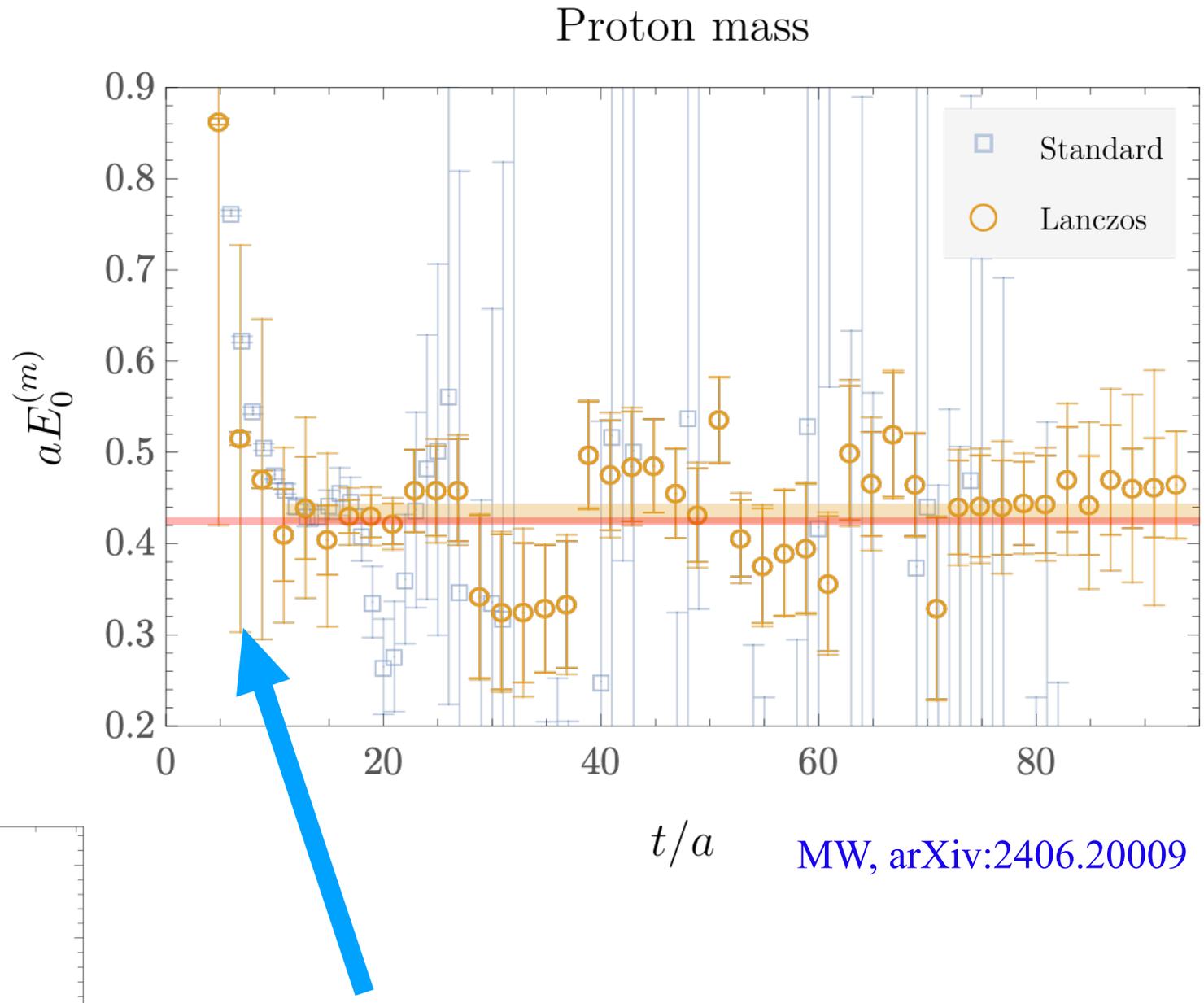


MW, arXiv:2406.20009

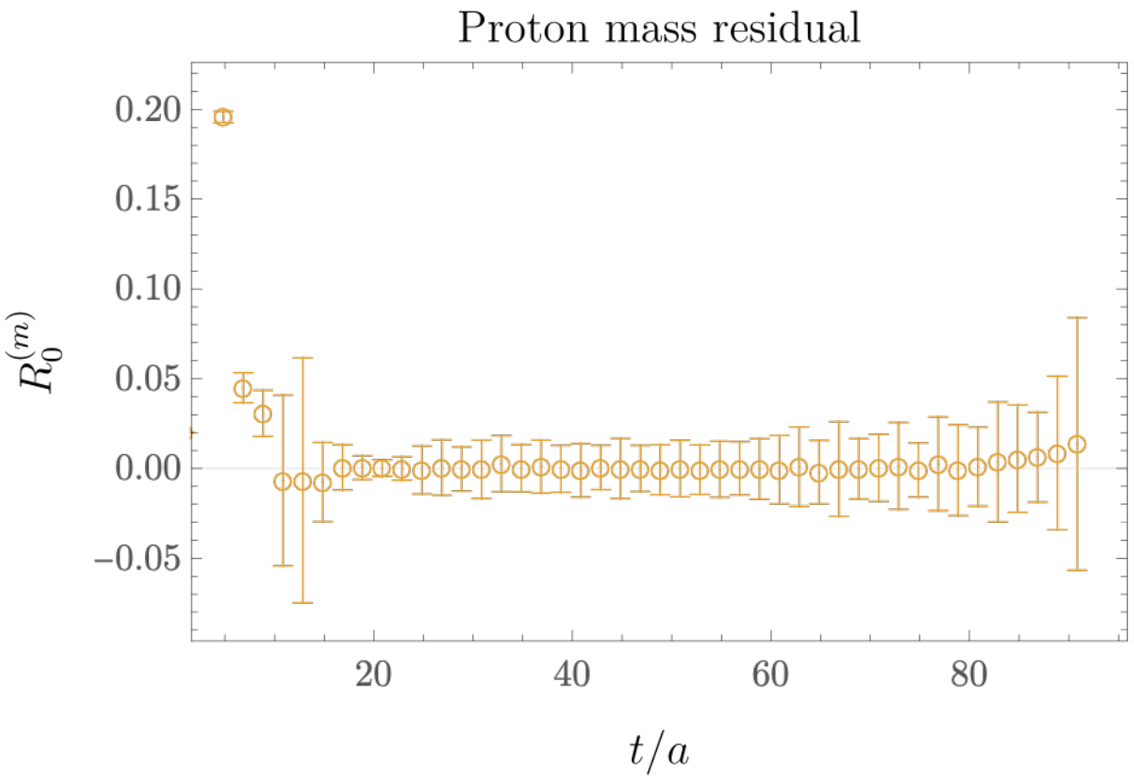
Lanczos proton mass results

- Bootstrap uncertainties complicated by outliers due to spurious eigenvalue misidentification within bootstrap samples

- Robust estimators e.g. based on confidence intervals critical



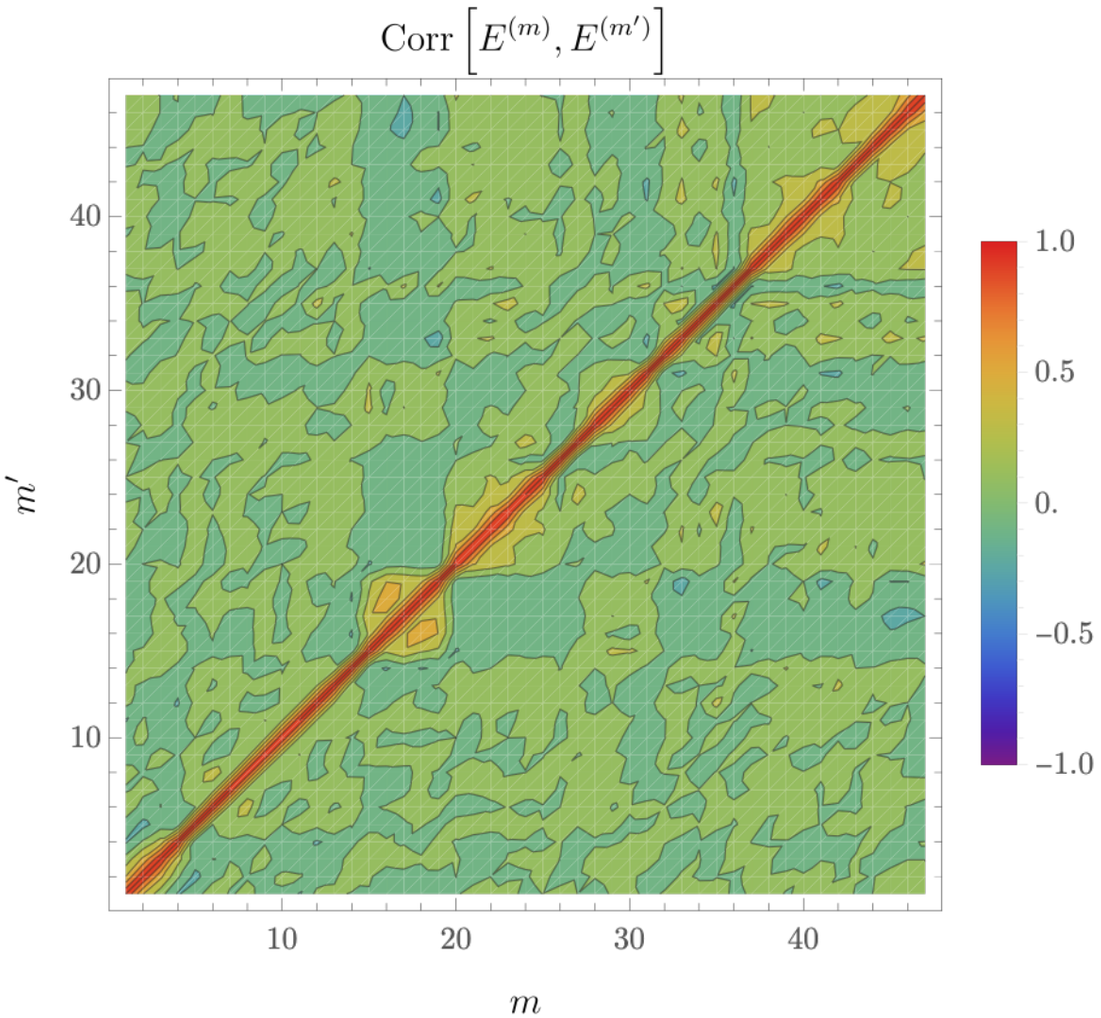
MW, arXiv:2406.20009

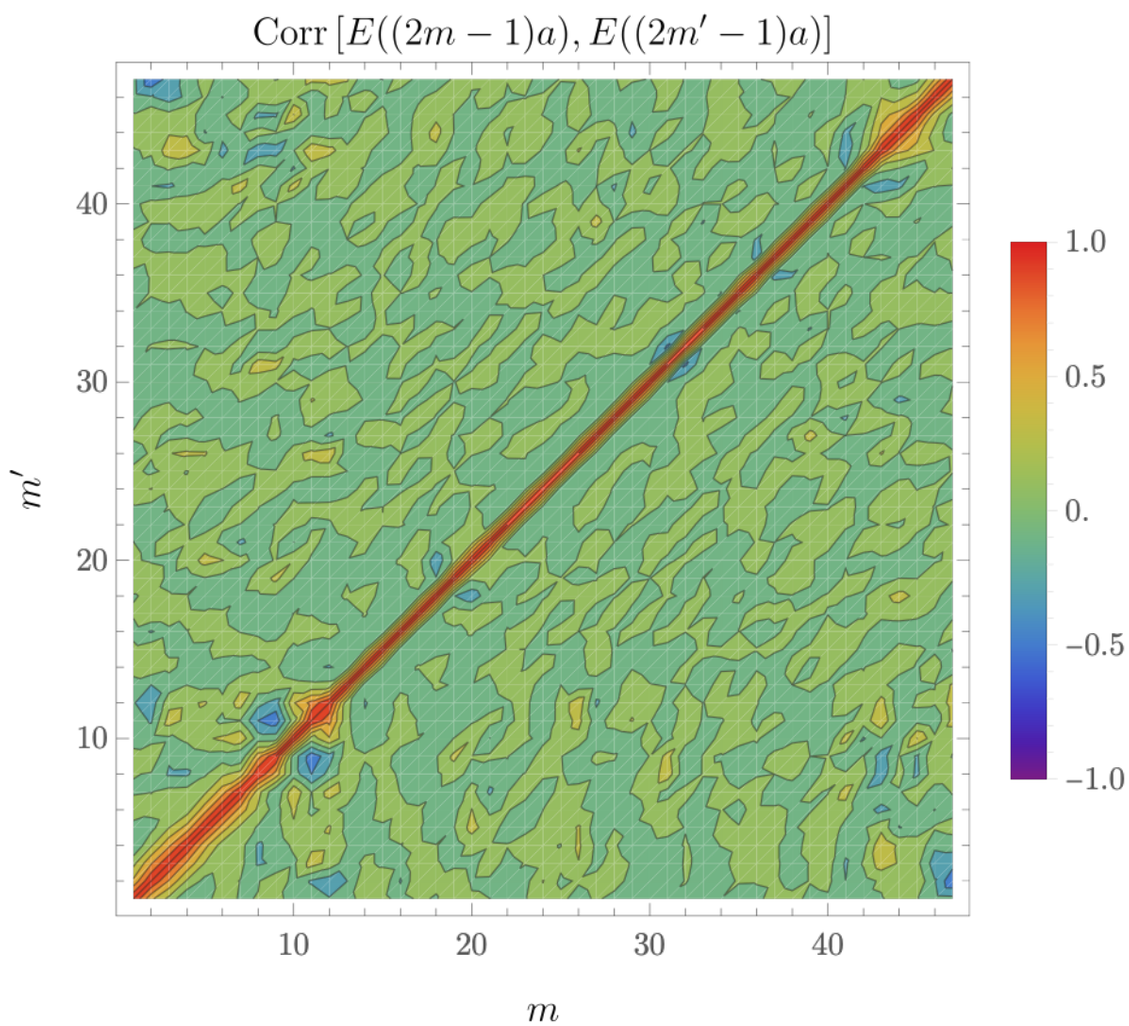


- Residual bound can be used to identify when Lanczos results have converged, provides bound on finite-t approximation errors

Correlations

- Correlations between Lanczos results at different imaginary times fall off rapidly with similar scale to correlations between standard effective mass results



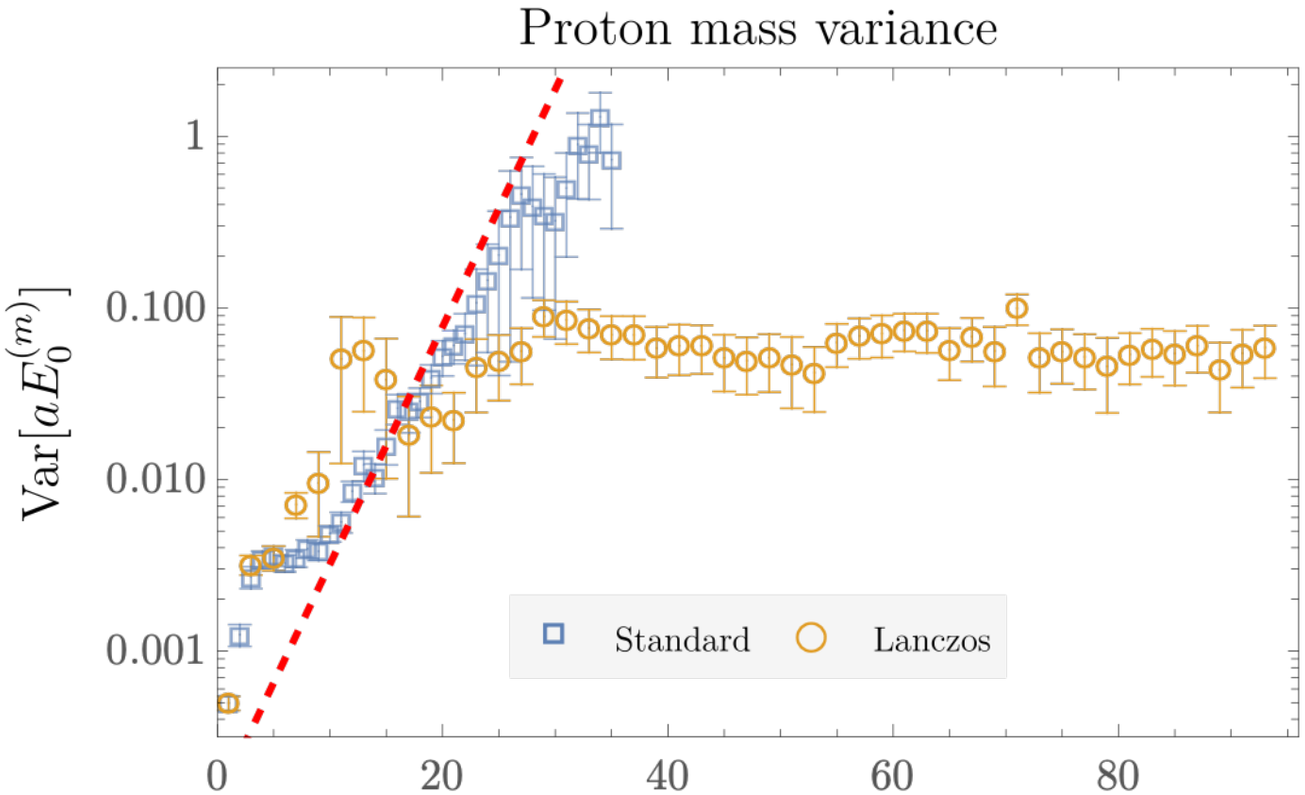


MW, arXiv:2406.20009

Projecting out the noise

- Signal-to-noise of Lanczos results does not degrade exponentially for large t

Why?



- Projection operator solution to signal-to-noise problem:

Della Morte and Giusti, Comp. Phys. Communications 180 (2009)

$$\langle \mathcal{O}(t)\overline{\mathcal{O}}(0)\rangle \longrightarrow \langle \mathcal{O}(t)P\overline{\mathcal{O}}(0)\rangle$$

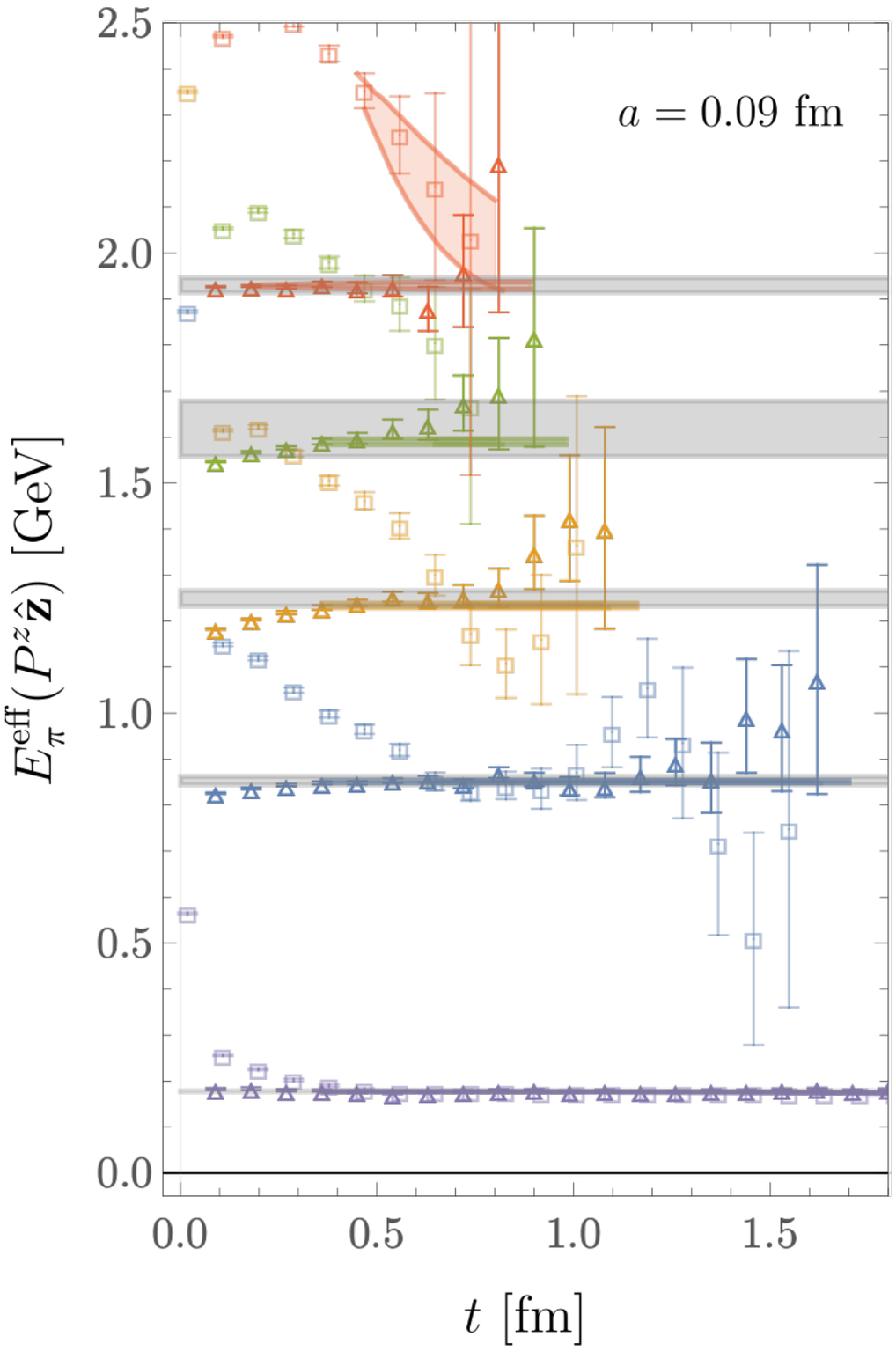
removes states from variance without quantum numbers of "signal squared," e.g. three-pion states in nucleon variance

- Building such projectors is hard — but Lanczos provides Krylov-space approximations

Saad, SIAM 17 (1980) Saad, SIAM 19 (1982)

$$P_n^{(m)} \equiv |y_n^{(m)}\rangle\langle y_n^{(m)}| \approx |n\rangle\langle n|$$

Boosted states are noisy



Exponential signal-to-noise degradation common to all large-momentum correlation functions used for LaMET

e.g. pion state

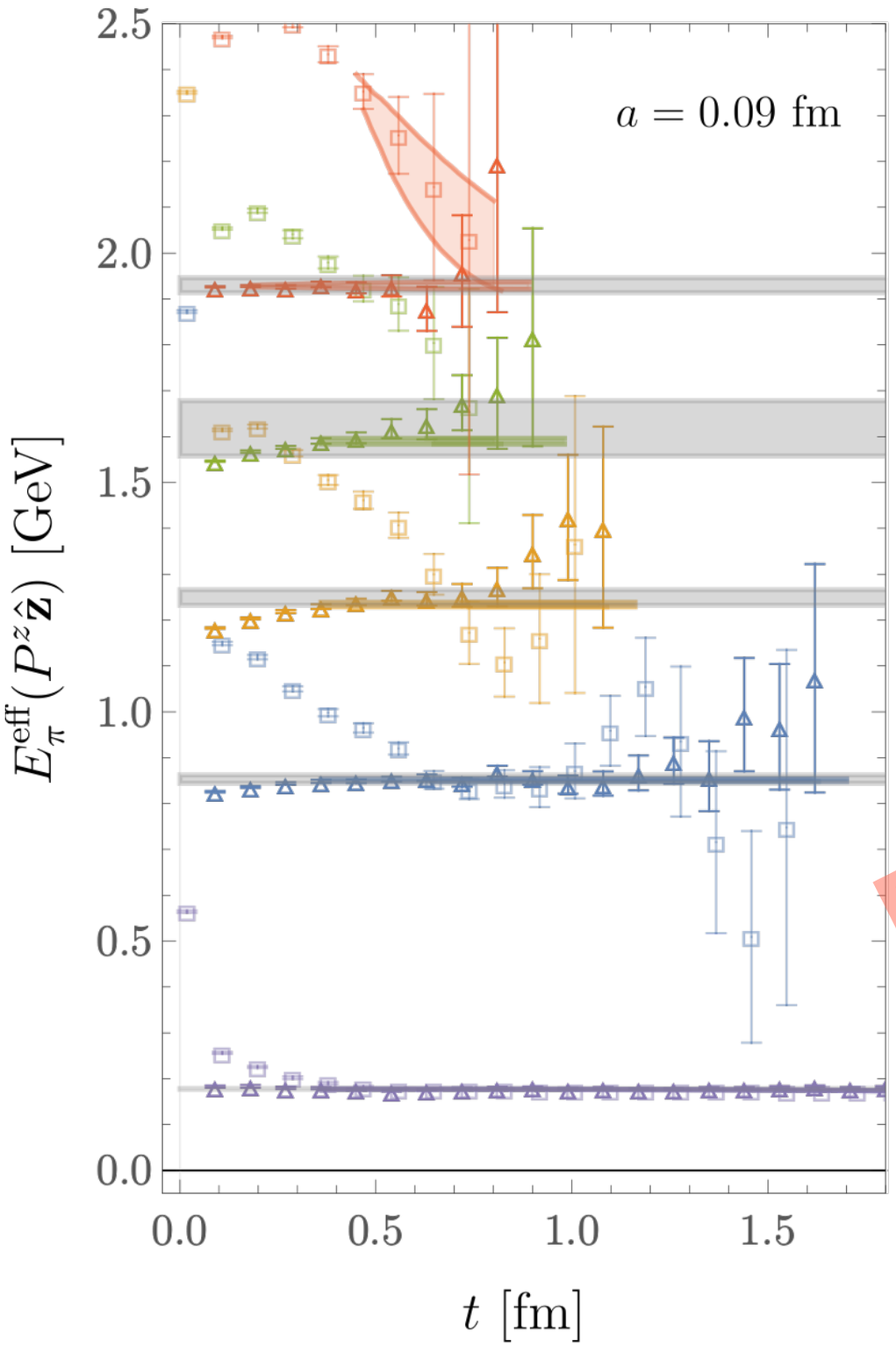
$$\text{signal} \sim e^{-E(\mathbf{P})t}$$

$$\text{variance} \sim e^{-2m_\pi t}$$

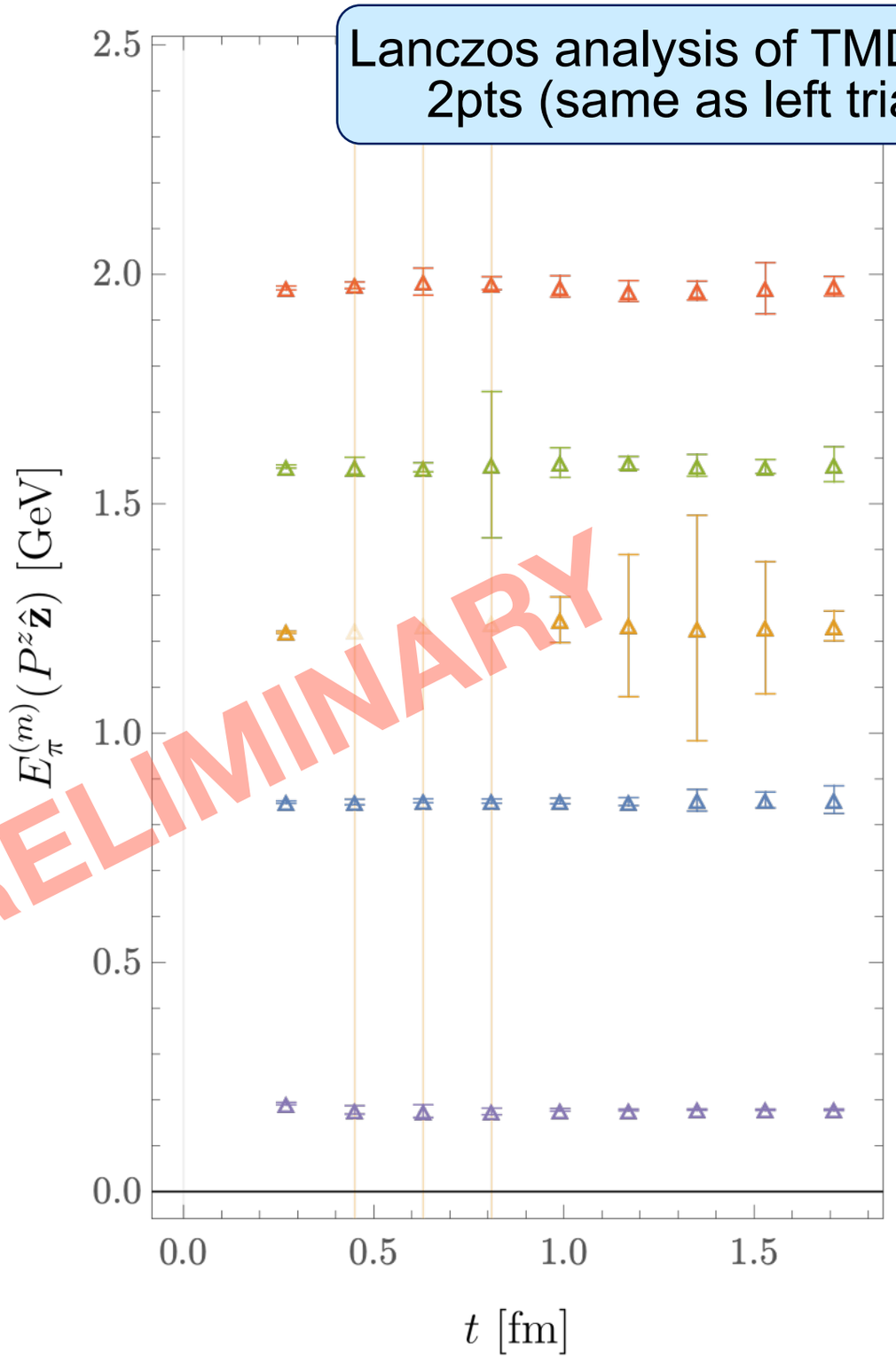
$$\frac{\text{signal}}{\text{noise}} \sim e^{-[E(\mathbf{P})-m_\pi]t}$$

Avkhadiev, Shanahan, MW, Zhao, PRL 132 (2024)

Boosted states are noisy



Lanczos analysis of TMD WF
2pts (same as left triangles)



Avkhadiev, Shanahan, MW, Zhao, PRL 132 (2024)

Excited states are noisy

Example:

Luscher-Weisz gauge action

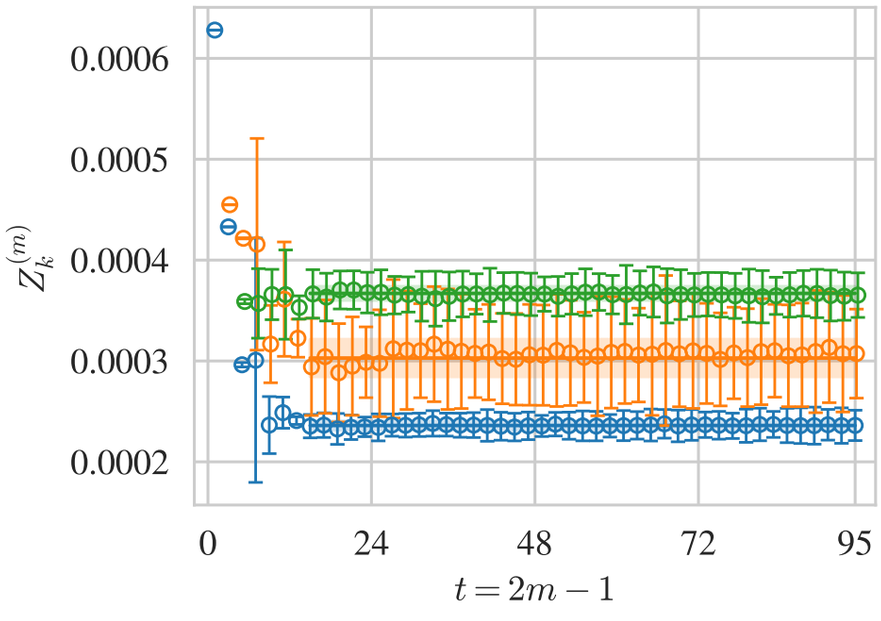
2+1 stout-smeared clover fermions

$M_\pi \approx 170\ \mathrm{MeV} \quad a \approx 0.09\ \mathrm{fm} \quad 48^3 \times 96$

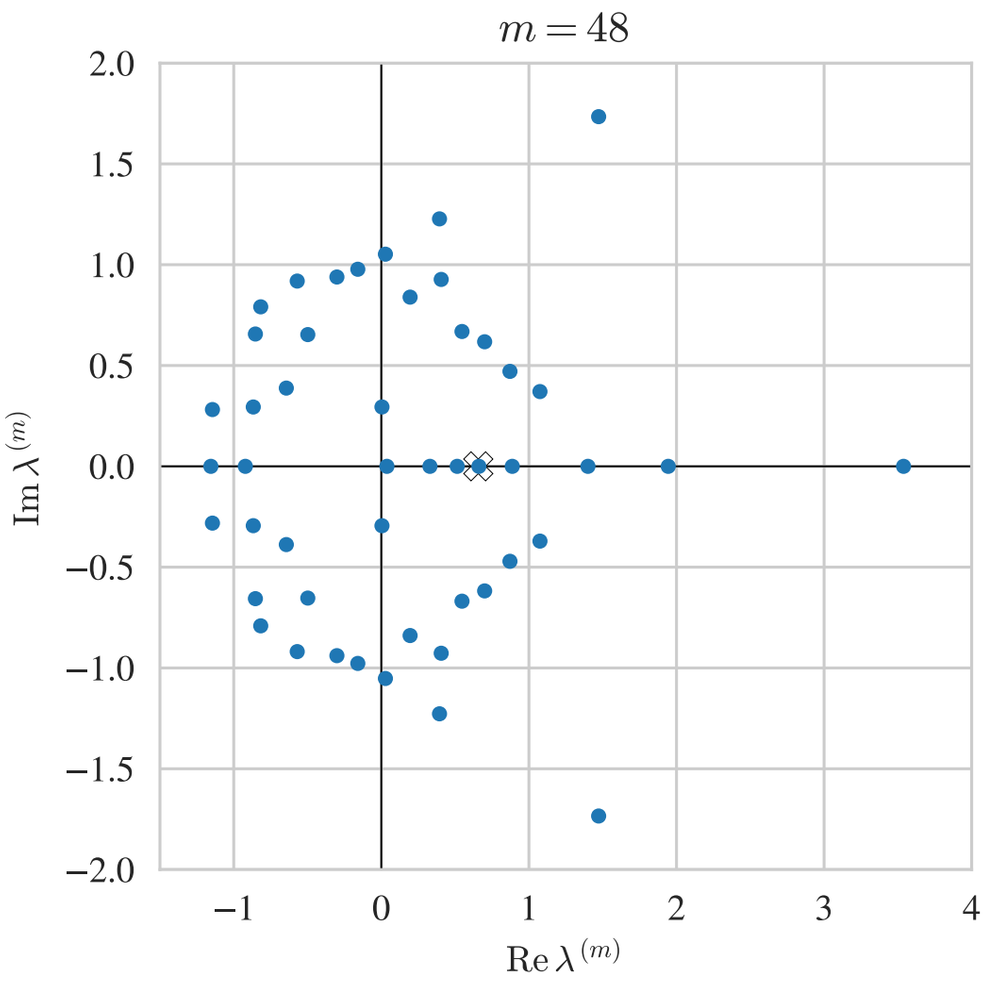
Nucleon $\chi \sim \left(u\, C\, \gamma_5\, d\right) u$

Quarks smeared to $r = 4.5$

3-4 states with physical norms reliably extracted



Ritz values can be filtered into non-Hermitian / spurious and physical subspaces



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

Complex eigenvalues permit exact description of noisy data

Physically interpretable

Excited states are noisy

Example:

Luscher-Weisz gauge action

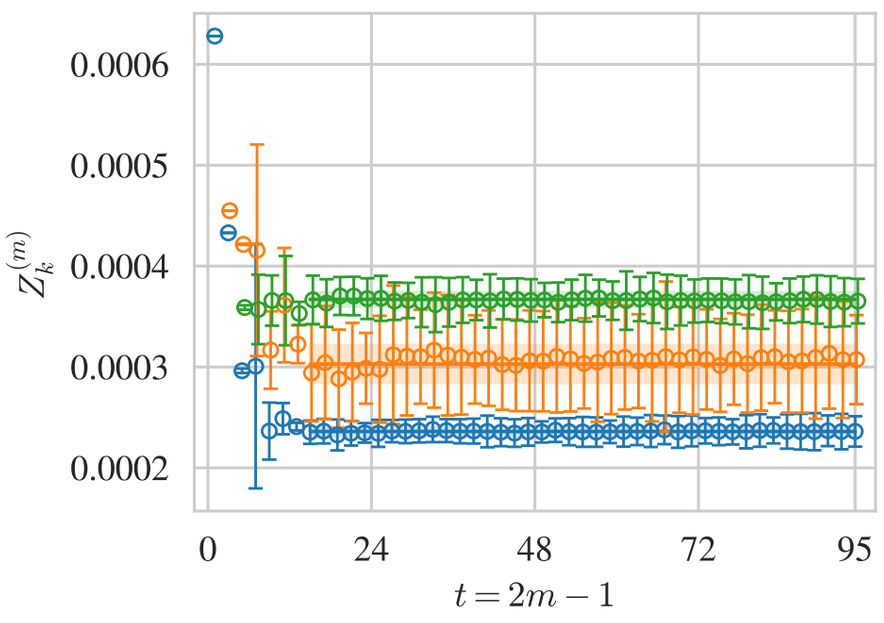
2+1 stout-smeared clover fermions

$M_\pi \approx 170\ \mathrm{MeV} \quad a \approx 0.09\ \mathrm{fm} \quad 48^3 \times 96$

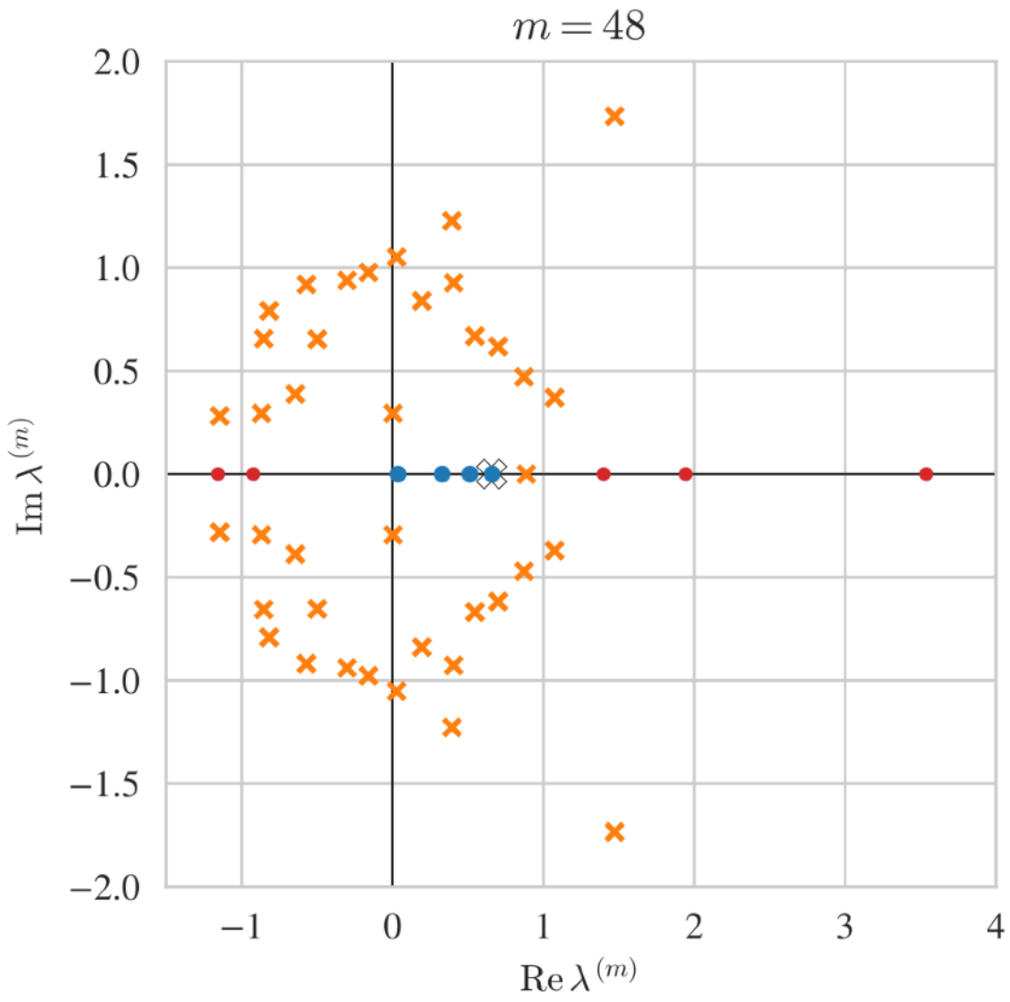
Nucleon $\chi \sim \left(u\, C\, \gamma_5\, d\right) u$

Quarks smeared to $r = 4.5$

3-4 states with physical norms reliably extracted



Ritz values can be filtered into non-Hermitian / spurious and physical subspaces



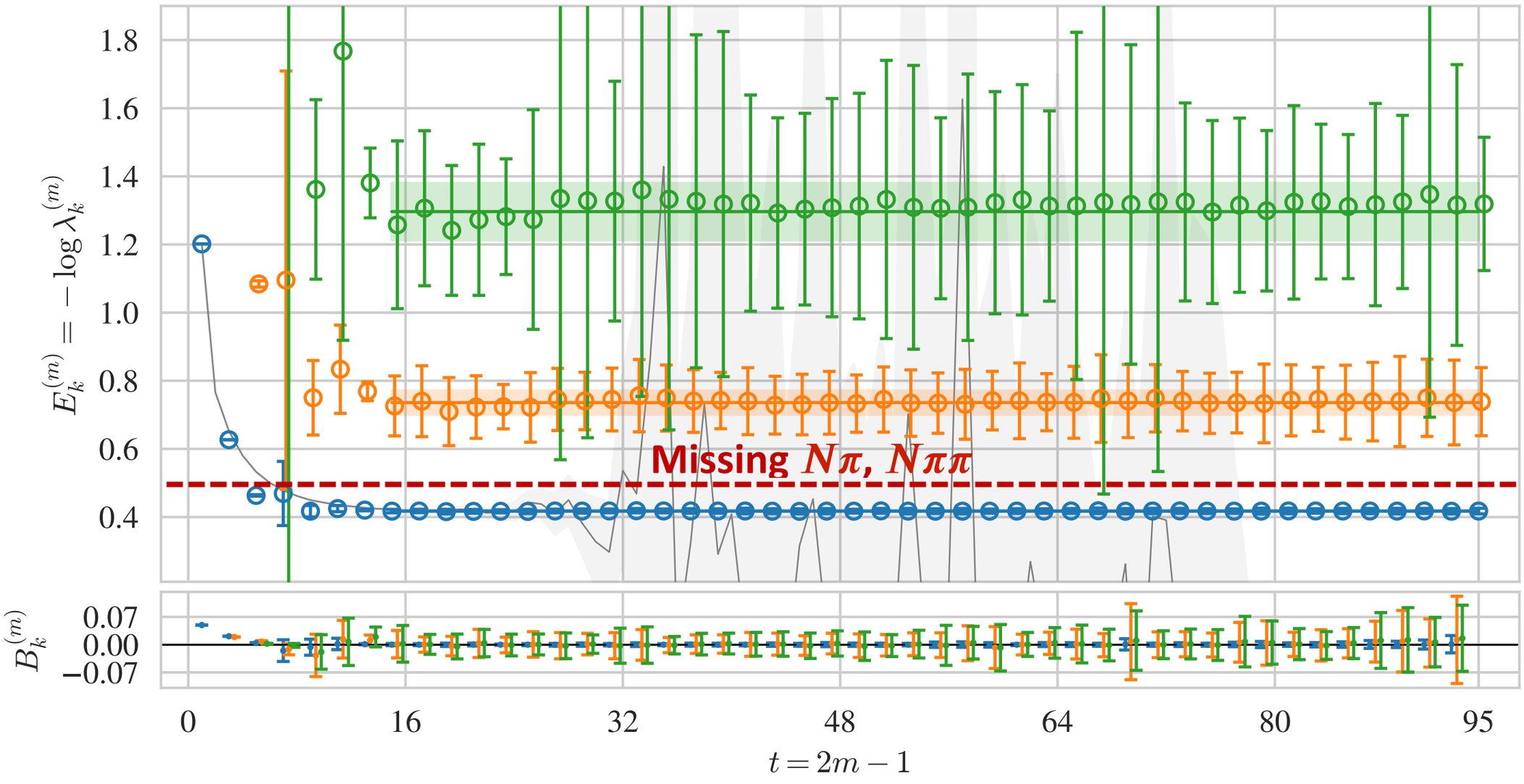
$$T^{(m)} = \sum_{k \in \overline{H}} \left| y_k^{R(m)} \right\rangle \lambda_k^{(m)} \left\langle y_k^{L(m)} \right| + \sum_{k \in H} \left| y_k^{(m)} \right\rangle \lambda_k^{(m)} \left\langle y_k^{(m)} \right|$$

Complex eigenvalues permit exact description of noisy data

Physically interpretable

Excited states are noisy

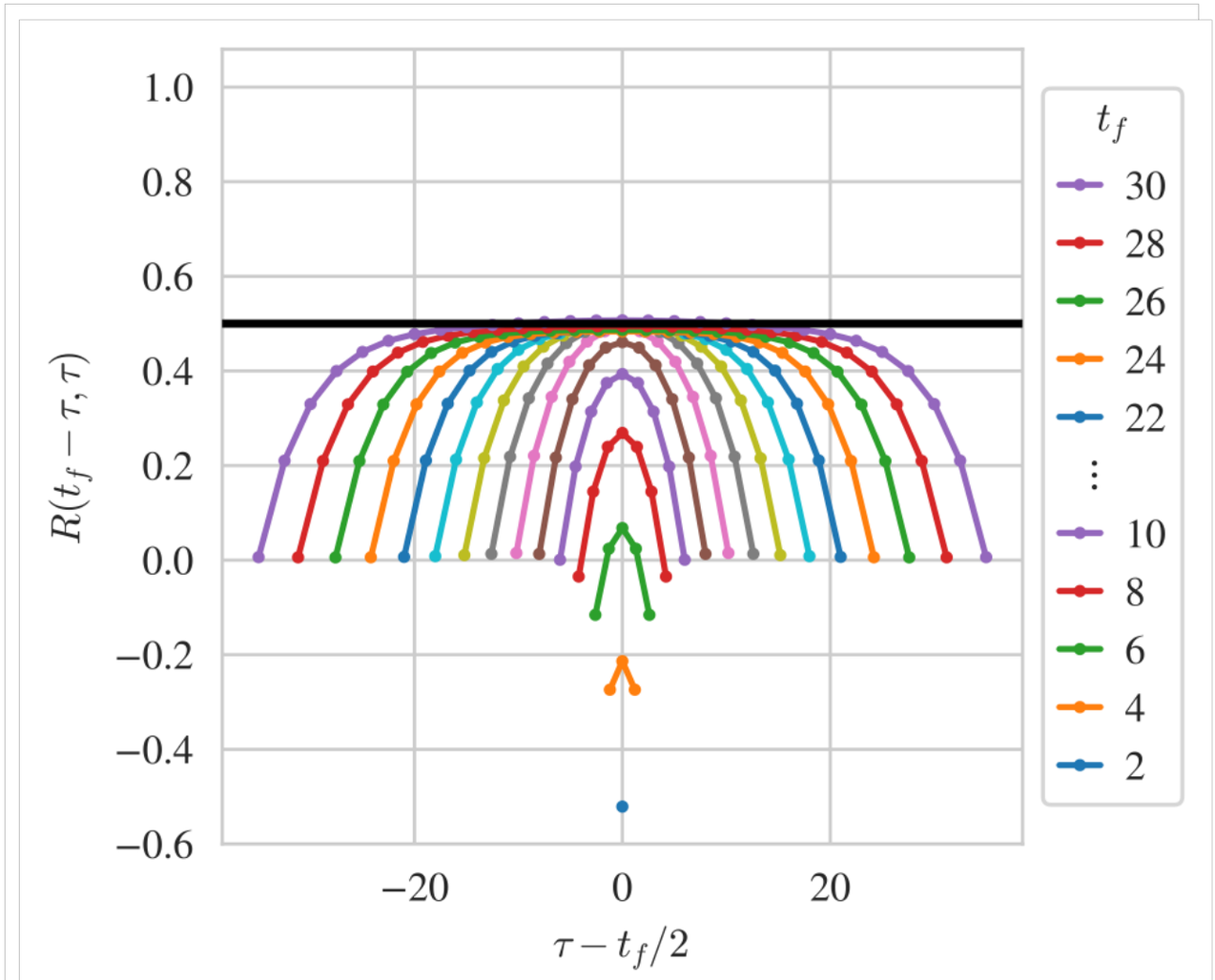
Precise results with no exponential signal-to-noise problem extracted for all states in physical subspace

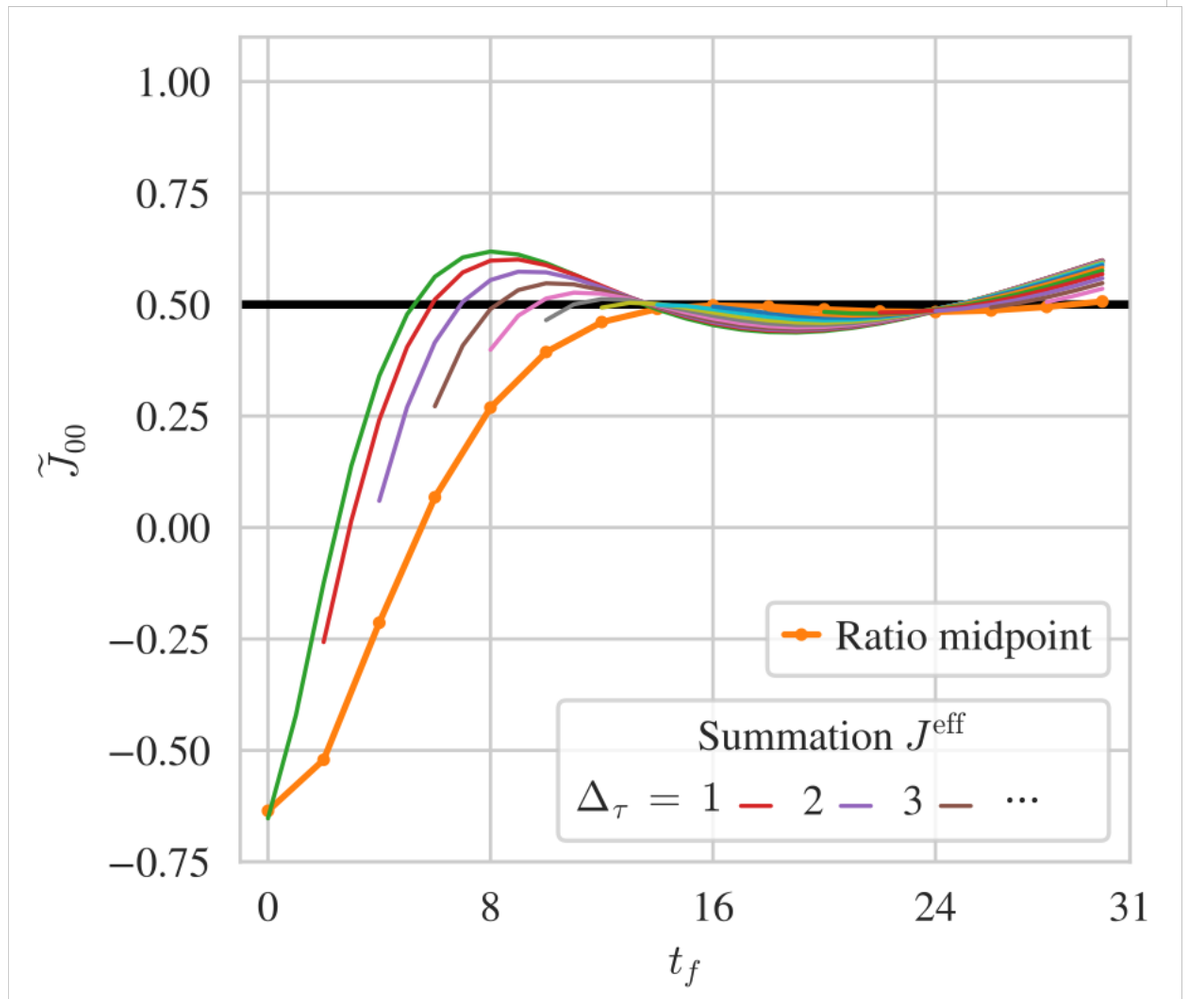


Noteworthy — not all states expected to be in the spectrum are recovered

What about matrix elements?

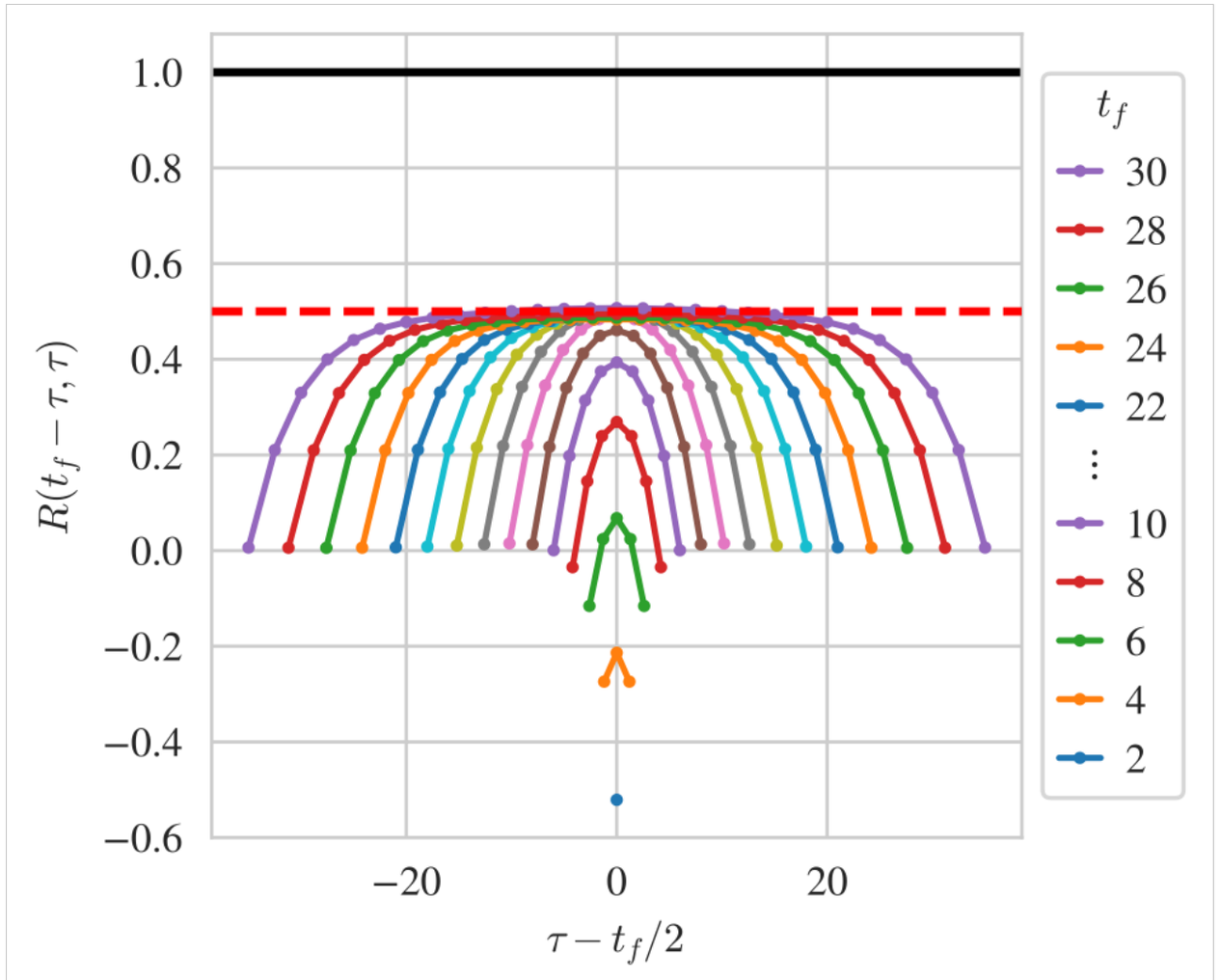
Particularly hard toy example

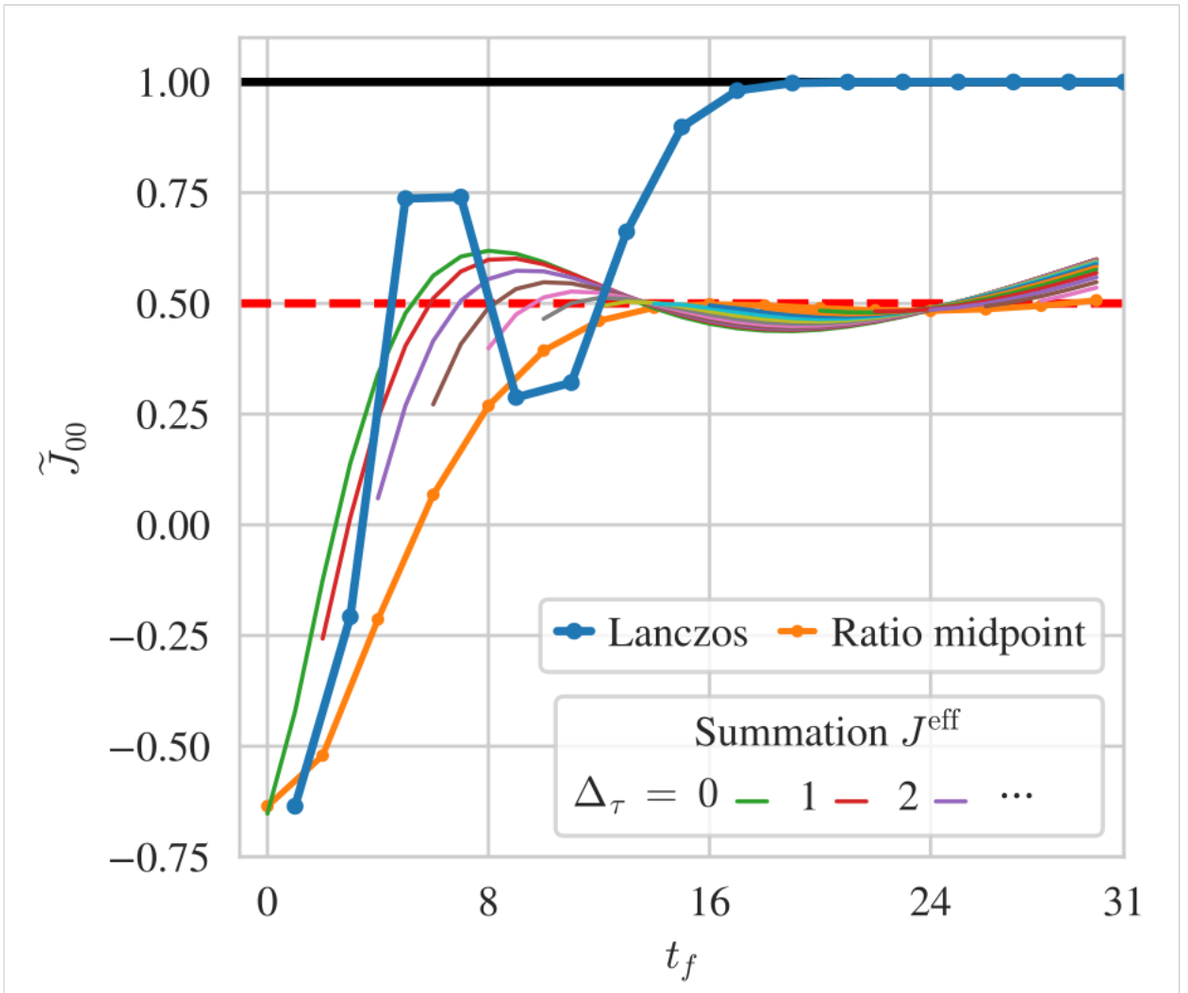




What about matrix elements?

Particularly hard toy example

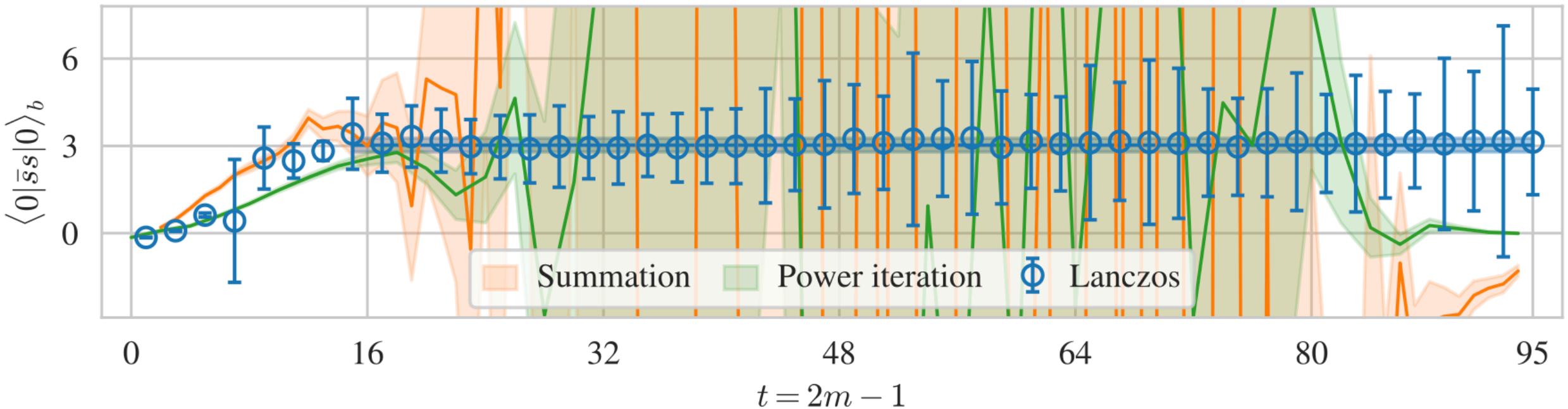




Lanczos for matrix elements

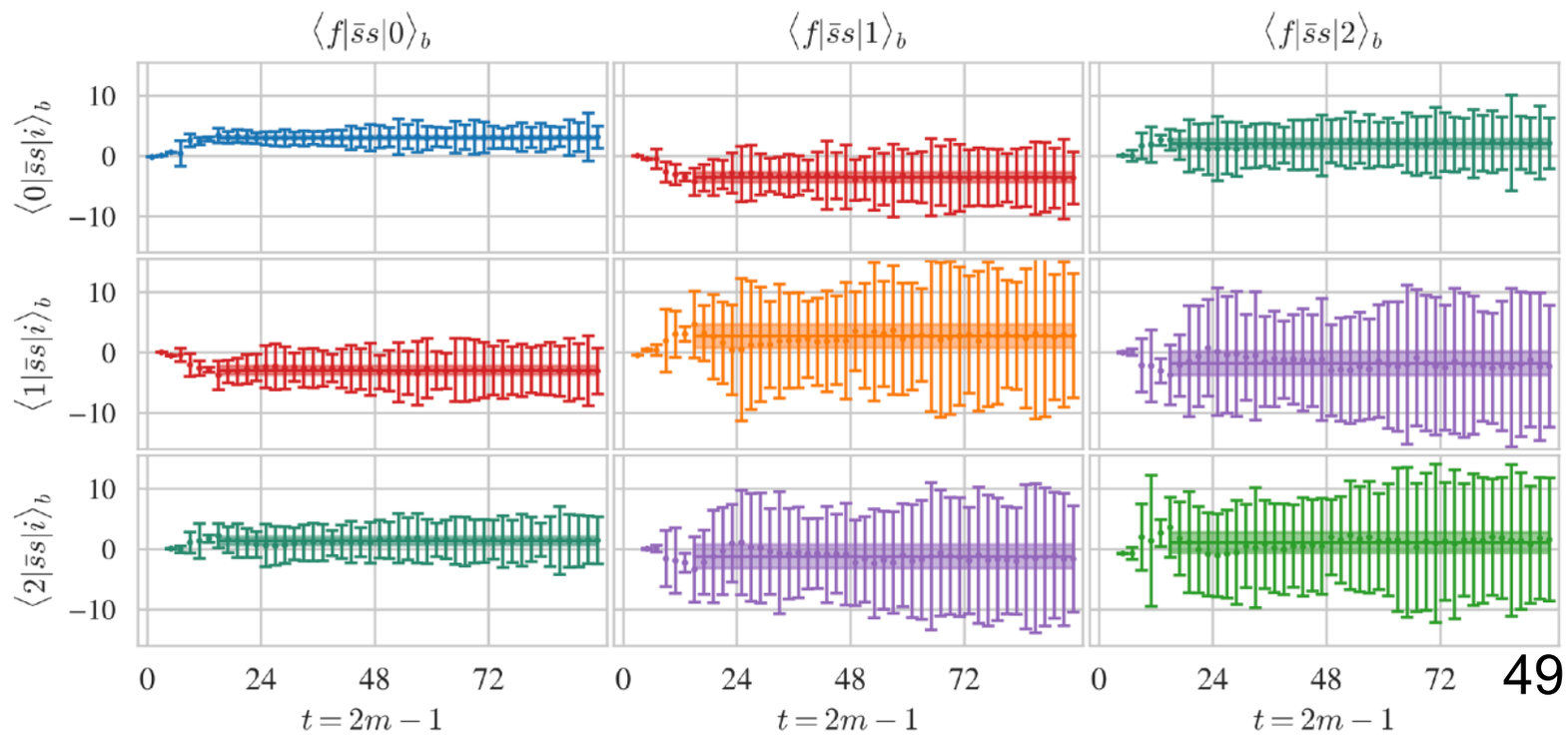
Nucleon strange scalar (bare) matrix element

Hackett, MW, arXiv:2407.21777

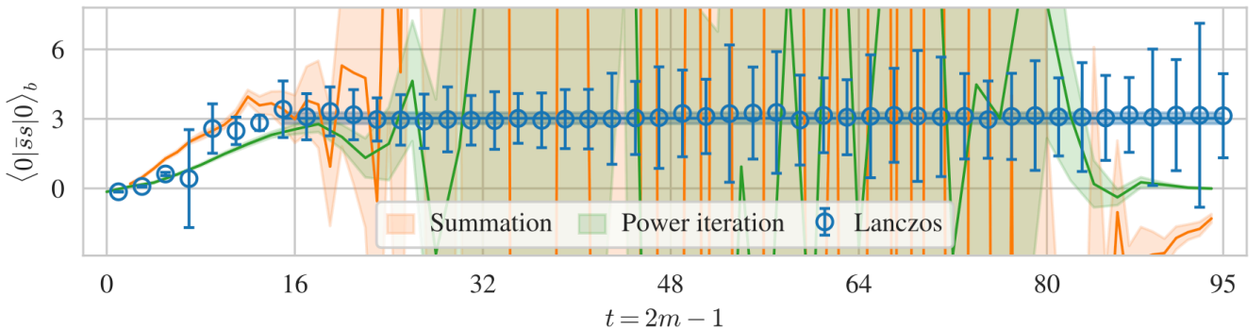


Lanczos eigenvectors provide change of basis allowing matrix elements to be extracted from 3pt functions with simple matrix multiplication

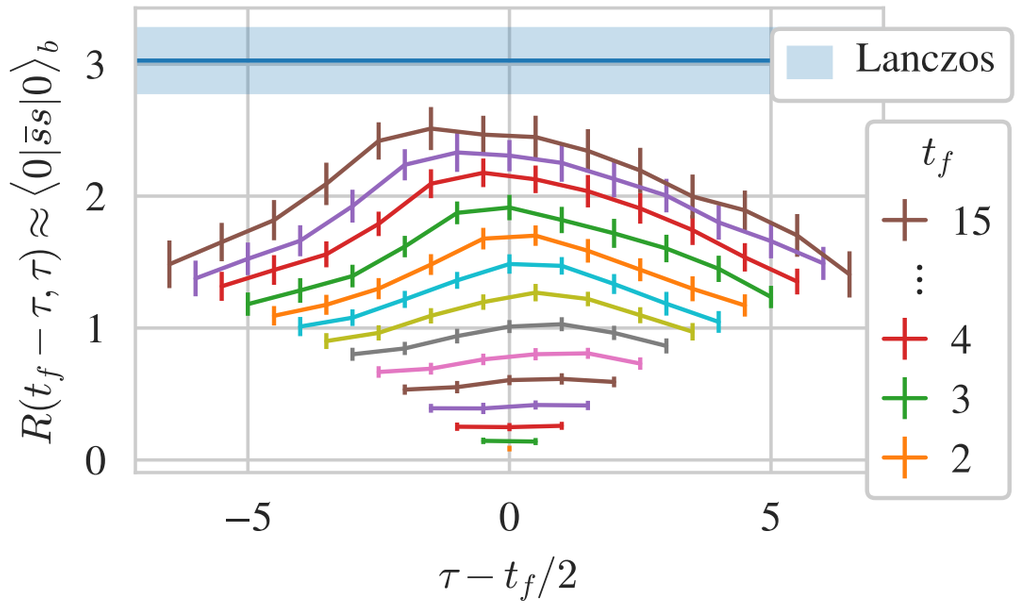
- Excited / transition matrix elements accessible



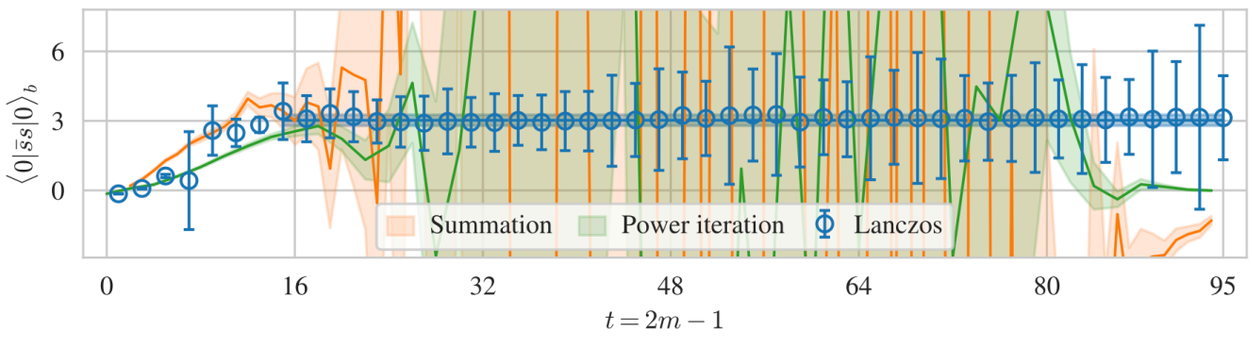
Lanczos for matrix elements



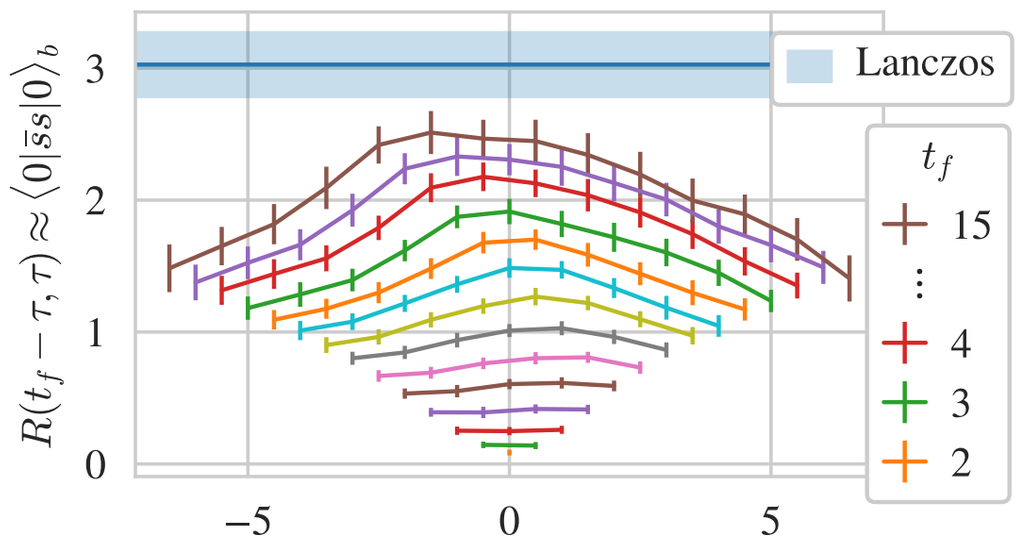
Hackett, MW, arXiv:2407.21777



Lanczos for matrix elements

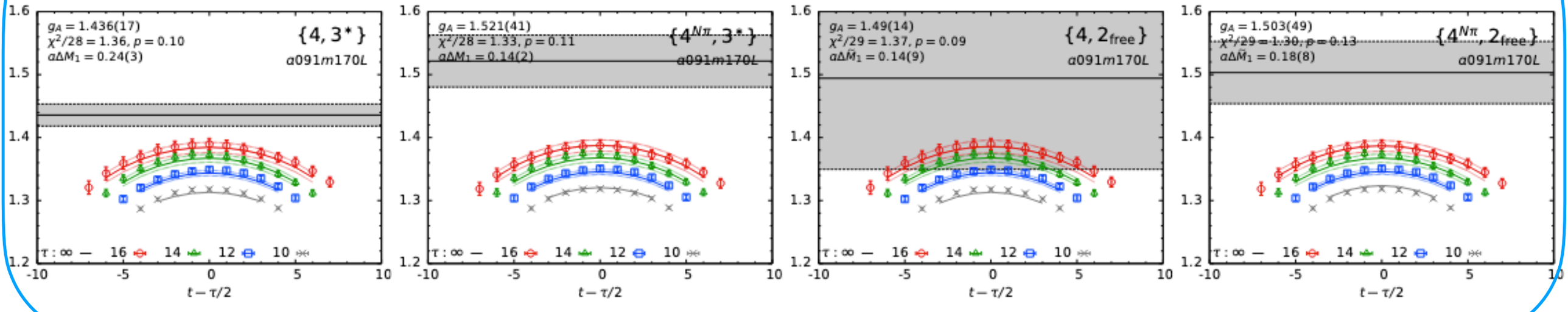


Hackett, MW, arXiv:2407.21777



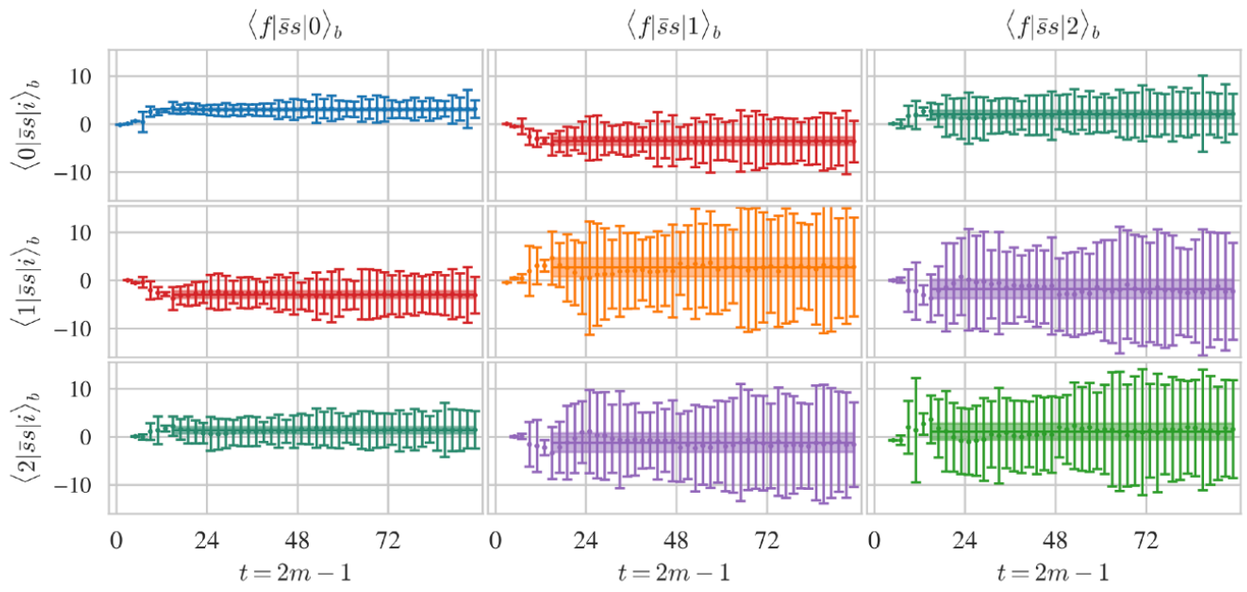
Comparison of Lanczos and standard ratios reminiscent of challenging form factor analyses…

To be continued…



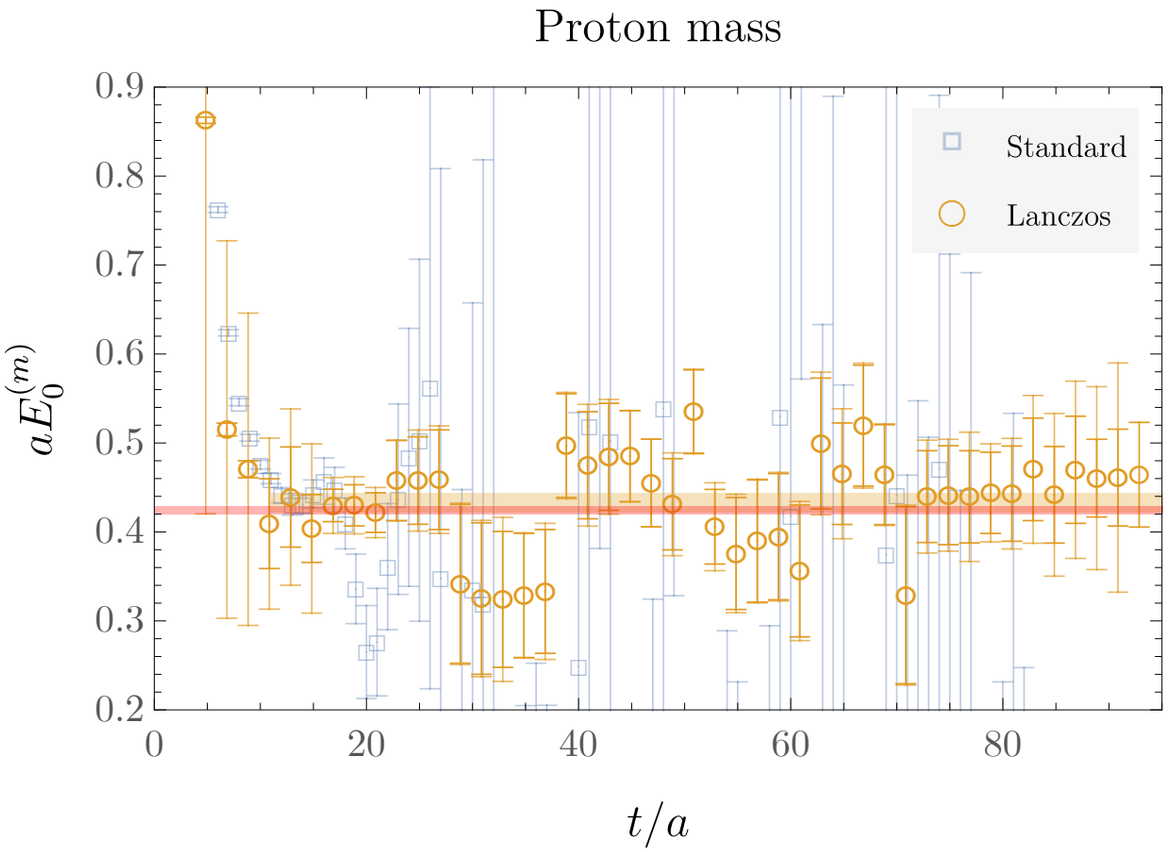
Lanczos for LQCD

- Lanczos enables rapid convergence even with small energy gaps



- Two-sided error bounds allow excited-state effects to be fully quantified

- Lanczos results do not show exponential signal-to-noise degradation



- Spurious eigenvalues lead to challenges: Cullum-Willoughby + bootstrap sufficient?

Lanczos shows promise for LQCD studies of nucleons and nuclei where isolating ground states is challenging; further study needed!