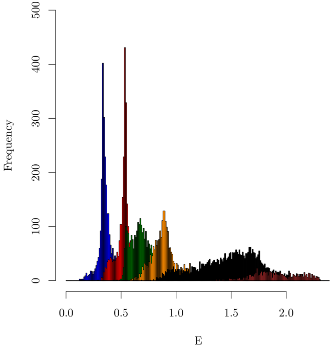
The Prony Generalised Eigenvalue Method

arXiv:2004.10472; arXiv:2411.14981

Johann Ostmeyer, Aniket Sen, Carsten Urbach



Gefördert durch
DFG Deutsche Forschungsgemeinschaft

The (notorious) Challenge

- Euclidean Correlation Functions

$$C(t) = \sum_{l=0}^{N_s} A_l \, e^{-E_l t}$$

$E_0 < E_1 < \ldots$

- obtained from stochastic (MC) simulations
- want to estimate energy levels E_l
- and amplitudes A_l (matrix elements)

The (notorious) Challenge

- Euclidean Correlation Functions

$$C(t) = \sum_{l=0}^{N_s} A_l\, e^{-E_l t}$$

$E_0 < E_1 < \ldots$

- obtained from stochastic (MC) simulations
- want to estimate energy levels E_l
- and amplitudes A_l (matrix elements)

Challenge

- signal-to-noise problem

$$\text{StN} \;\propto\; \exp(-\Delta E t)$$

with in general $\Delta E > 0$

[Lepage (1989)]

- signal deteriorates exponentially
- increasing severity with l

Example: The Effective Mass of the Nucleon

- effective mass

$$M_{\text{eff}}(t) = -\frac{1}{\delta t}\,\log\left(\frac{C(t+\delta t)}{C(t)}\right)$$

- since $E_0 < E_{l\neq 0}$

$$\lim_{t\to\infty} M_{\text{eff}}(t) = E_0$$

- all other contributions to C exponentially suppressed

- but t finite
  $\Rightarrow$ excited state contaminations

Example: The Effective Mass of the Nucleon

- effective mass

$$M_{\text{eff}}(t) = -\frac{1}{\delta t} \log\left(\frac{C(t+\delta t)}{C(t)}\right)$$

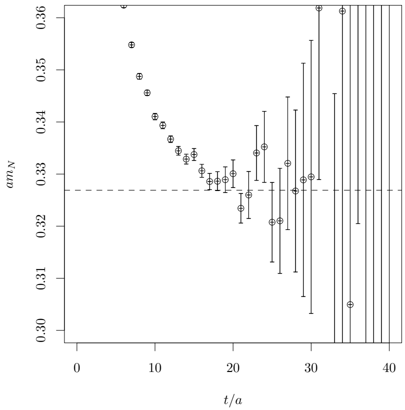
- since $E_0 < E_{l\neq 0}$

$$\lim_{t\to\infty} M_{\text{eff}}(t) = E_0$$

- all other contributions to C exponentially suppressed

- but t finite
  $\Rightarrow$ excited state contaminations

Example Nucleon M_{eff}



[ETMC (2021)]

Example: The Effective Mass of the Nucleon

- effective mass

$$M_{\text{eff}}(t) = -\frac{1}{\delta t} \log\left(\frac{C(t+\delta t)}{C(t)}\right)$$

- since $E_0 < E_{l\neq 0}$

$$\lim_{t\to\infty} M$$

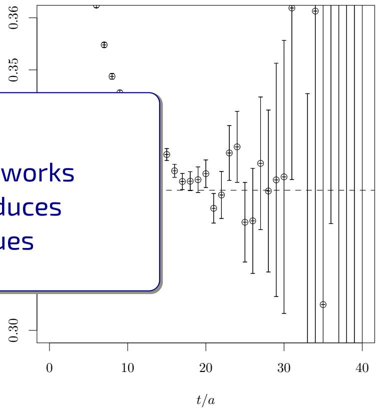
- all other contributions to C exponentially suppressed

- but t finite
  $\Rightarrow$ excited state contaminations

Goal

Find a method that either works at smaller t-values or reduces the noise at large t-values

Example Nucleon M_{eff}



[ETMC (2021)]

Here: Work at smaller t-values

algebraic methods

- generalised eigenvalue method

  [Michael, Teasdale (1983); Lüscher, Wolff (1990), Blossier et al. (2009)]

- Prony method

  [Fleming (2004); Beane et al. (2009); ...]

- Prony generalised eigenvalue method

  [Fischer et al. (2020)]

- generalised pencil of function method

  [Aubin, Orginos (2011); Schiel (2015); Ottnad et al. (2018)]

- Lanczos

  [Wagman (2024)]

Here: Work at smaller t-values

algebraic methods

- generalised eigenvalue method

  [Michael, Teasdale (1983); Lüscher, Wolff (1990), Blossier et al. (2009)]

- Prony method

  [Fleming (2004); Beane et al. (2009); ...]

- Prony generalised eigenvalue method

  [Fischer et al. (2020)]

- generalised pencil of function method

  [Aubin, Orginos (2011); Schiel (2015); Ottnad et al. (2018)]

- Lanczos

  [Wagman (2024)]

all based on the idea:

diagonalise the Euclidean transfer matrix in a larger subspace

e.g. by a “matrix effective mass”

$$M^{-1}(t) \cdot M(t+\delta t)\, v_l = \Lambda_l\, v_l$$

with suitable M

The Prony GEVM

- construct a $n \times n$ Hankel matrix

$$H_{ij}(t) = C(t + i\Delta + j\Delta)$$

for $\Delta > 0$ and $i, j = 0, \ldots, n-1$

The Prony GEVM

- construct a $n \times n$ Hankel matrix

H_i

for $\Delta > 0$ a

- solve the G

H^{-1}

- eigenvalue

with $\delta t = t$

- v_l: access to amplitudes

Example Hankel Matrix for $\Delta = 1$

$$H(t) = \begin{pmatrix} C(t) & C(t+1) & \dots & C(t+n-1) \\ C(t+1) & C(t+2) & \dots & C(t+n) \\ \vdots & \vdots & \ddots & \vdots \\ C(t+n-1) & C(t+n) & \dots & C(t+2n-2) \end{pmatrix}$$

$\Rightarrow$ real-valued and symmetric per construction

$\Rightarrow$ contains $2n - 1$ correlator elements

The Prony GEVM

- construct a $n \times n$ Hankel matrix

$$H_{ij}(t) = C(t + i\Delta + j\Delta)$$

  for $\Delta > 0$ and $i, j = 0, \ldots, n-1$
- solve the GEVP

$$H^{-1}(t_0)\, H(t) v_l = \Lambda_l(t, t_0)\, v_l$$

- eigenvalues $\Lambda_l(t, t_0)$

$$\Lambda_l \;\approx\; e^{-E_l \delta t}$$

  with $\delta t = t - t_0$
- v_l: access to amplitudes

what does $\approx$ mean here?

- $\approx\, \to\, =$ for exactly n states in C
- $n < N_s$: corrections depending on strategy
- strategy here: increase n with $t, \delta t$ fixed
- form of corrections: next slide

The Krylov Point of View: Lanczos

- recast the problem

$$C(t) = \langle\psi|\varphi_t\rangle\,, \quad |\varphi_t\rangle = T^t|\psi\rangle$$

- T: Euclidean transfer matrix
- Krylov space

$$\mathcal{K}_k = \{|\varphi_0\rangle, |\varphi_1\rangle, \ldots, |\varphi_k\rangle\}$$

- Lanczos: based on $\mathcal{K}_{2n}$ find tri-diagonal

$$T_n = V_n^\mathsf{T} \cdot T \cdot V_n$$

[Wagman (2024)]

- T_n similar to T on the sub-space $\mathcal{K}_{2n}$

- eigenvalues of T_n

$$\Lambda_l \;\approx\; e^{-E_l}$$

- Lanczos theory: corrections of order

$$\exp\left(-4n\sqrt{\Delta E \delta t}\right)$$

[Kaniel (1966); Paige (1971); Saad (1980)]

- ΔE splitting to the nearest by state

- Lanczos run-time asymptotically $\mathcal{O}(n^2)$
  PGEVM $\mathcal{O}(n^3)$

PGEVM mathematically equivalent to Lanczos

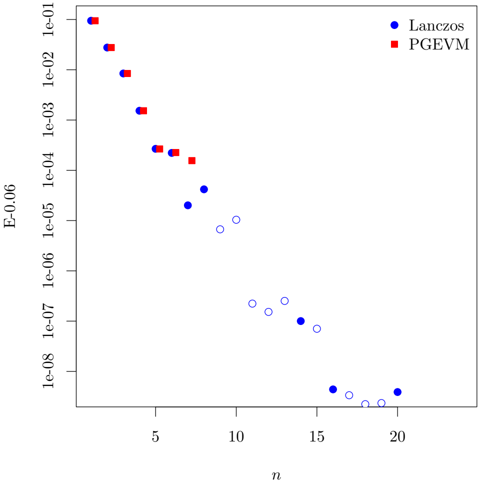
- one can show

$$T_n \quad \sim \quad H^{-1}(t_0)\, H(t)$$

[Ostmeyer, Sen, Urbach (2024)]

⇒ they share all eigenvalues and eigenvectors

- PGEVM and Lanczos identical
- they also share all convergence properties
- clearly observed in artificial data



[Ostmeyer, Sen, Urbach (2024)]

Equivalence goes further...

- correlator C can be replaced by a correlator matrix

$$C_{\alpha\beta}(t) \;=\; \langle O_\alpha(t)\, O_\beta^\dagger(0)\rangle$$

$\Rightarrow$ build a Hankel matrix of correlator matrices $\quad\Rightarrow\quad$ Block PGEVM

- also known as generalised pencil of function (GPOF) method

  [Aubin, Orginos (2011); Schiel, 2015; Ottnad et al. (2018)]

- then:

Block PGEVM $\quad = \quad$ GPOF $\quad \equiv \quad$ Block Lanczos (for $\Delta = 1$ only)

Effective Noise Model

theoretical input:

- $C(t)$ real valued
- spectral decomposition $E_0 < E_1 < \ldots$

$$C(t) = \sum_{l=0}^{N_s} A_l \, e^{-E_l t}$$

PGEVM:

- complex valued E possible
- need to come in pairs E, E^* for C to be real

Effective Noise Model

theoretical input:

- $C(t)$ real valued
- spectral decomposition $E_0 < E_1 < \ldots$

$$C(t) = \sum_{l=0}^{N_s} A_l\, e^{-E_l t}$$

PGEVM:

- complex valued E possible
- need to come in pairs E, E^* for C to be real

consequence

- PGEVM will find approximations to physical energy levels
- effective noise model

$$e^{\mathsf{Re}(E)t} e^{\mathsf{i}\,\mathsf{Im}(E)t}$$

- phoney levels $E \in \mathbb{R}$ possible

Eigenvalue Filtering

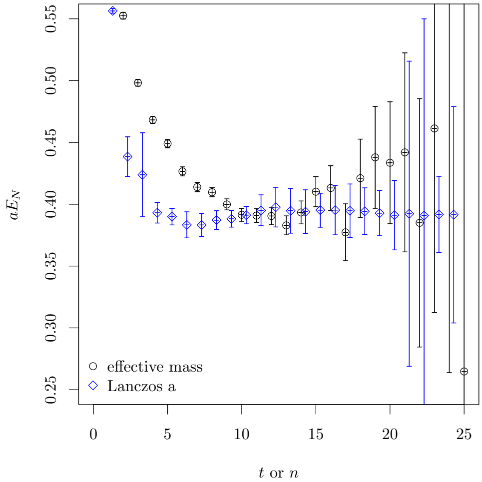
⇒ remove all Λ with $\mathrm{Im}(\Lambda) > \epsilon$

⇒ keep only $\Lambda \in]0, 1]$

- leaves us only with physical and spurious eigenvalues in $]0, 1]$

⇒ highly overestimated errors due to outliers

- even after some outlier removal treatment



[Ostmeyer, Sen, Urbach (2024)]

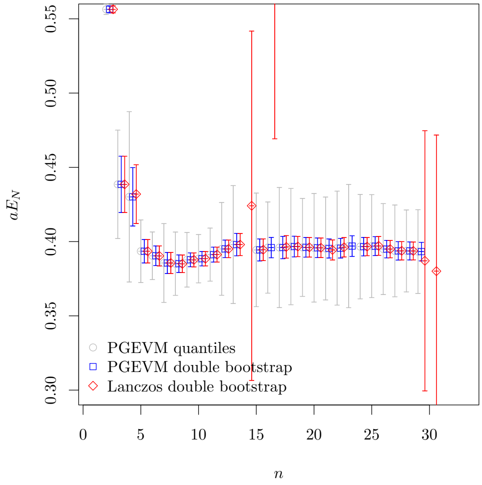
Changing the Estimator for the Expectation Value

- need an outlier-robust estimator
- use the median $\mu^\star$ over bootstrap samples
- for its error apply double bootstrap

  [Wagman (2024) arXiv version 3, Ostmeyer, Sen, Urbach (2024)]

$\Rightarrow$ error estimate reasonable

- only few instabilities left due to misidentification of eigenvalues



[Ostmeyer, Sen, Urbach (2024)]

Eigenvalue Filtering II

more details on filtering

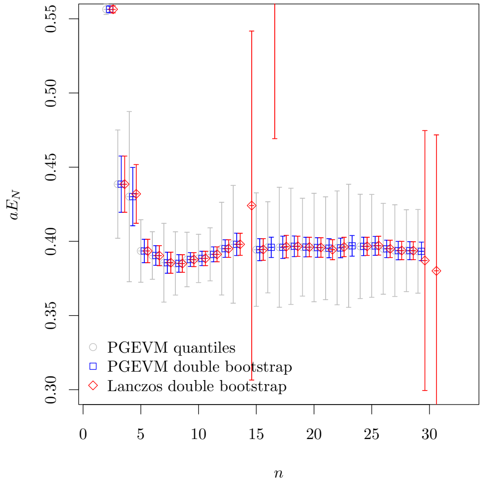
1. given n, chose all $\Lambda \in]0, 1]$
2. sort $\Lambda_0^r > \Lambda_1^r > \ldots, \forall r$
3. compute pivot value

$$L_p = \text{median}_r \Lambda_0^r$$

   over all (bootstrap) samples r

4. pick on all (bootstrap) samples

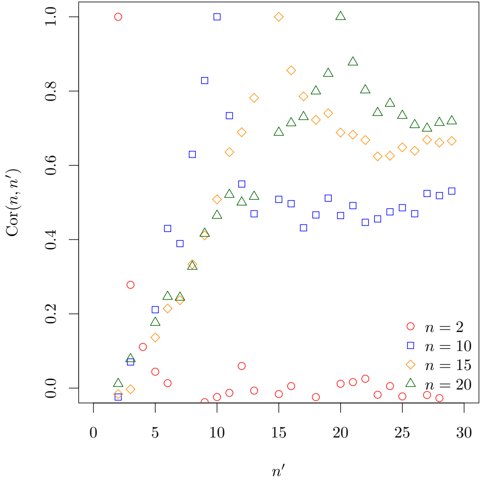
$$\Lambda^r = \min_i (|\Lambda_i^r - L_p|)$$



[Ostmeyer, Sen, Urbach (2024)]

Signal-to-Noise Problem Solved...?

- unfortunately, not!
- results at larger and larger n-values more and more correlated
- flattening out eventually
- consistent with two more (noisy) correlator values added from n to $n+1$



[Ostmeyer, Sen, Urbach (2024)]

So, what's the Point??

Lanczos / PGEVM / GPOF

algebraically deliver estimates for energy lev-
els and amplitudes with minimal assumptions

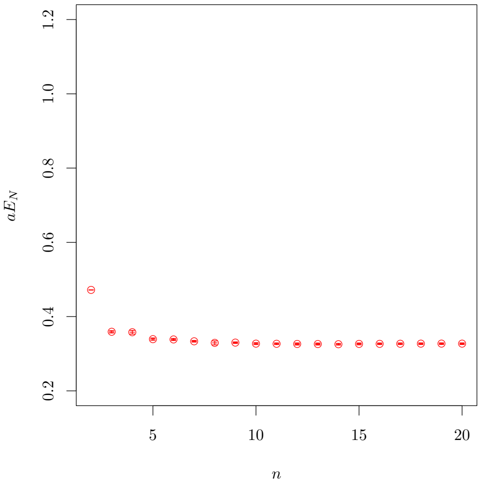
well controlled systematics

no fitting and model averaging needed

Multi-state Fit versus PGEVM: Nucleon

- comparison on C80 ETMC physical point ensemble
- high stat. proton correlator compare to multi-state fit

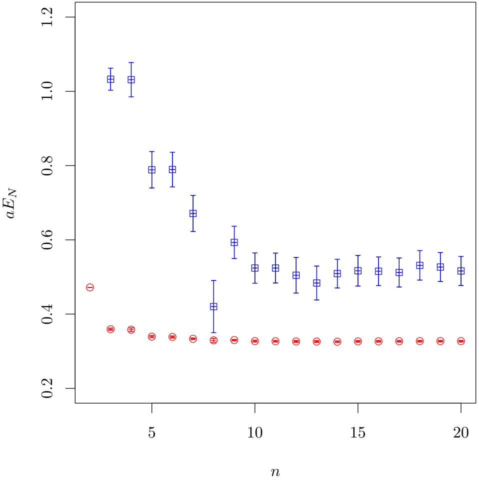
  [ETMC (2024)]



Multi-state Fit versus PGEVM: Nucleon

- comparison on C80 ETMC physical point ensemble
- high stat. proton correlator compare to multi-state fit

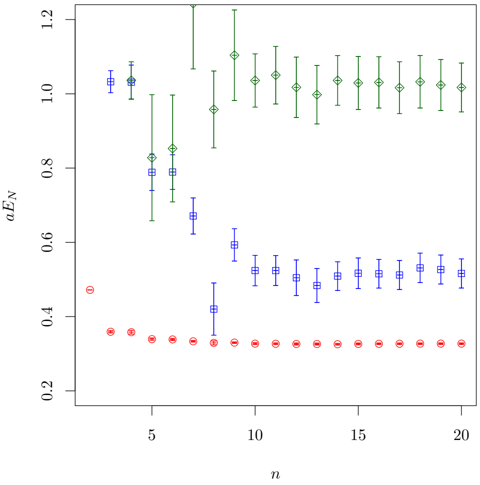
  [ETMC (2024)]
- PGEVM allows to extract ≥ 3 levels
- use result and error at $n = 20$



Multi-state Fit versus PGEVM: Nucleon

- comparison on C80 ETMC physical point ensemble
- high stat. proton correlator compare to multi-state fit

  [ETMC (2024)]
- PGEVM allows to extract ≥ 3 levels
- use result and error at $n = 20$



Multi-state Fit versus PGEVM: Nucleon

- ground state at $n = 20$

$$am_N = 0.3271(22)$$

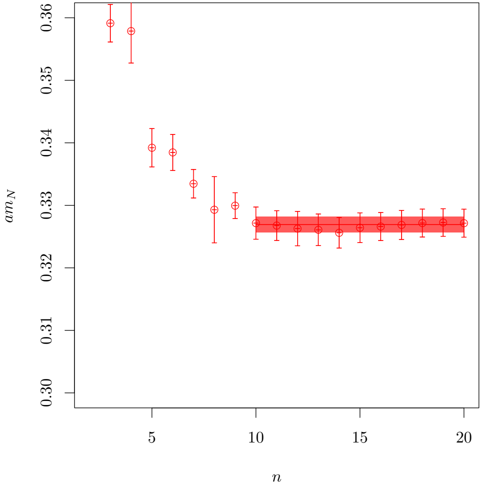
- with fit $(22) \to (12)$,
  but need to chose fit-range

- first excited at $n = 20$

$$aE_N = 0.523(29)$$

- second excited at $n = 20$

$$aE_N = 1.01(5)$$



Multi-state Fit versus PGEVM: Nucleon

- ground state at $n = 20$

$$am_N = 0.3271(22)$$

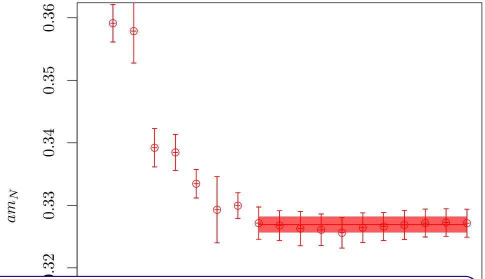
- with fit $(22) \to (12)$, but need to chose fit-range

- first excited at $n = 20$

$$aE_N$$

- second excited a

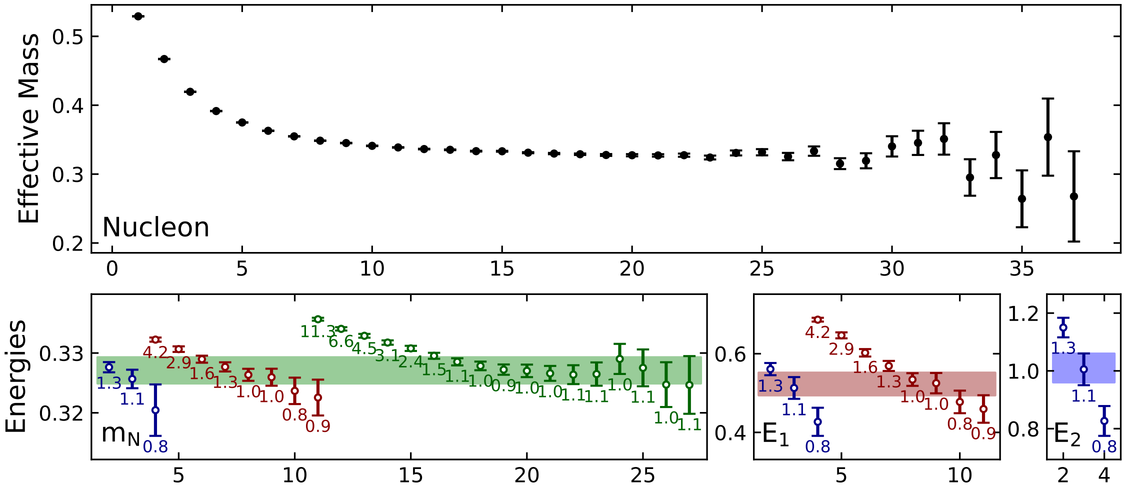
$$aE_N$$



one-state fit			two-state fit			three-state fit		
t_{low}/a	$\bar{\chi}^2$	am_N	t_{low}/a	$\bar{\chi}^2$	am_N	t_{low}/a	$\bar{\chi}^2$	am_N
19	0.9	0.32679(88)	8	1.0	0.3261(11)	3	1.2	0.3253(16)

[ETMC (2021)]

Multi-state Fit versus PGEVM: Nucleon

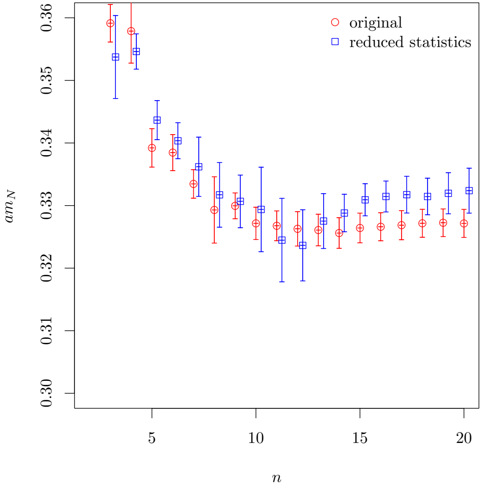


[courtesy: Simone Bacchio]

Comparison to reduced Statistics

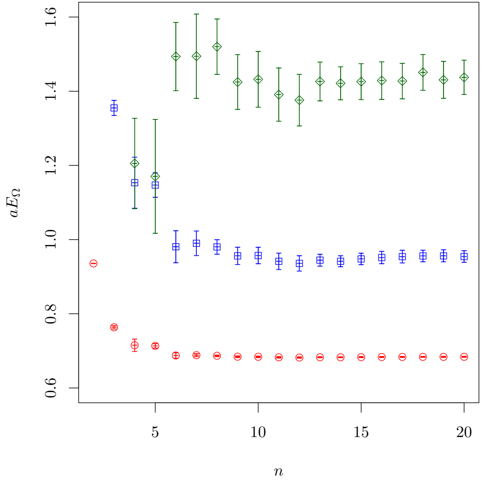
- original statistics 401 configs, 650 sources
- reduced to 101 configs
- compatible results
- error scales roughly as expected

$$am_N = 0.3324(36)$$



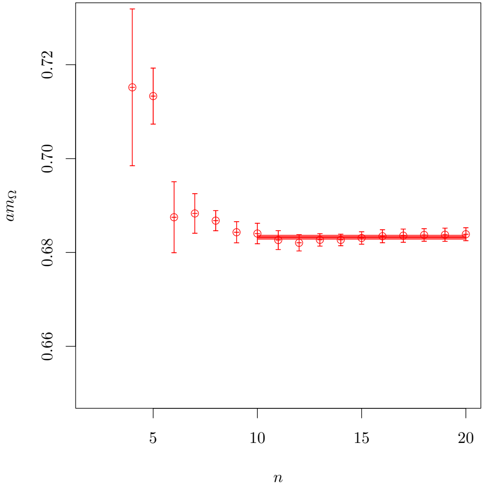
Multi-state Fit versus PGEVM: Ω Baryon

- ETMC B64 physical point ensemble
- high stat. with valence strange
  $a\mu_s = 0.3323817$
- similar conclusion as for the nucleon
- result $am_\Omega = 0.6839(14)$ at $n = 20$
  additional fit: $(14) \rightarrow (5)$
- two-state fit for $t \geq 10$
  $am_\Omega = 0.6825(8)$

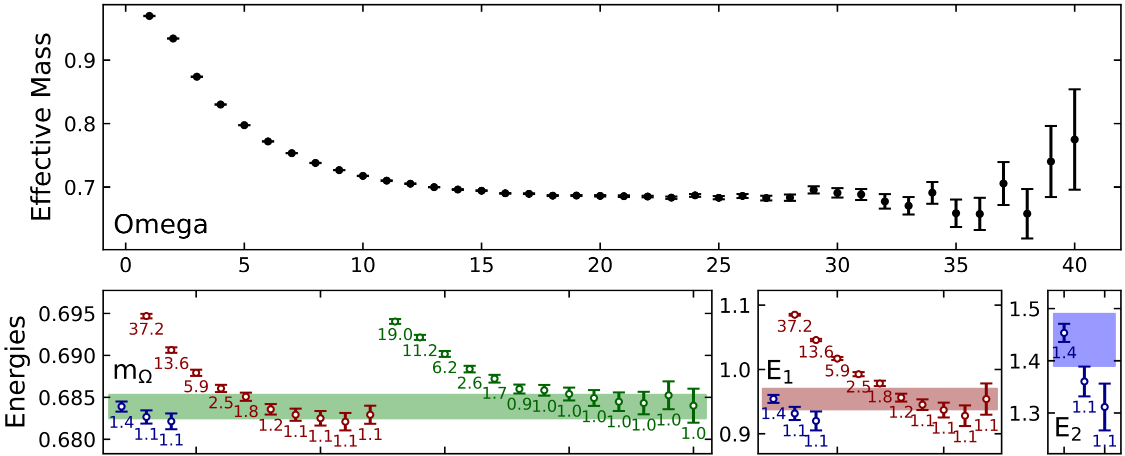


Multi-state Fit versus PGEVM: Ω Baryon

- ETMC B64 physical point ensemble
- high stat. with valence strange $a\mu_s = 0.3323817$
- similar conclusion as for the nucleon
- result $am_\Omega = 0.6839(14)$ at $n = 20$ additional fit: $(14) \rightarrow (5)$
- two-state fit for $t \geq 10$ $am_\Omega = 0.6825(8)$



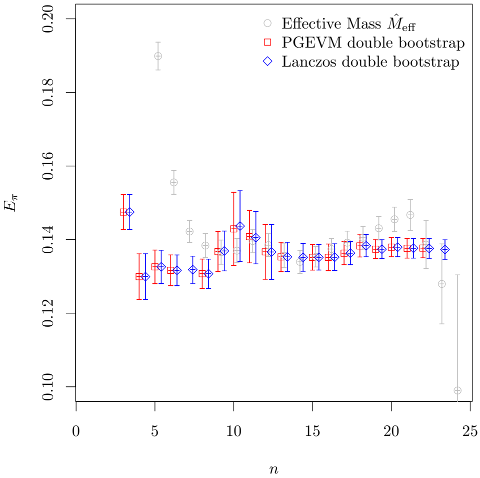
Multi-state Fit versus PGEVM: Ω baryon



[courtesy: Simone Bacchio]

PGEVM for Time-Symmetric Correlators

- PGEVM resolves $\pm E$ separately
- leads to additional systematic effects
- underestimated at smallish n-values?
- of course: could take the largest n result
- but need to understand systematics



A Simple $n = 2$ Model

- consider single energy level E_0 plus noise $\eta(t)$ in symmetric correlator

$$C(t) = \left(e^{-E_0 t} + a\, e^{E_0 t}\right)(1 + b\,\eta(t))\ .$$

- for $n = 2$ analytical solution to PGEVM $t = 0, \delta t = 1$

$$\Lambda_\pm = \frac{C(1)C(2) - C(0)C(3) \pm \sqrt{R}}{2(C(1)^2 - C(0)C(2))}\,,$$

- with

$$\begin{aligned} R =& -3C(1)^2C(2)^2 + 4C(0)C(2)^3 + 4C(1)^3C(3) \\ & -6C(0)C(1)C(2)C(3) + C(0)^2C(3)^2\,. \end{aligned}$$

A Simple $n = 2$ Model

- now expand around $b = 0$
- result for the two eigenvalues

$$\Lambda_- = e^{-E_0}\left(1 + b\,\tilde{R}_- + \mathcal{O}(b^2)\right)\,, \qquad \Lambda_+ = e^{+E_0}\left(1 - \frac{b}{a}\,\tilde{R}_+ + \mathcal{O}(b^2)\right)\,.$$

  $\tilde{R}_\pm$ depending on η, a and E_0

- for $a < b$ the correction to Λ_+ becomes large
- this is the case for $a = \exp(-E_0 T)$

A Simple $n = 2$ Model

solution A

- work where $a \approx 1$
- this is the case when t is chosen around $T/2$
- only feasible for the pion
- otherwise dominated by noise

solution B

- chose Δ maximal for given n
- reason

$$\tilde{R}_+ \propto \exp(-2\Delta E_0)(\ldots)$$

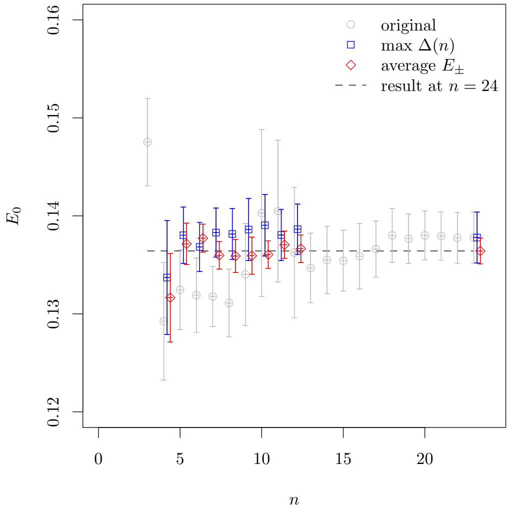
which cancels the $a^{-1} = \exp(E_0 T)$ for $\Delta = T/2$

- in addition: $\tilde{R}_-$ not affected significantly

- there is also a **solution C** (block stratifying H)
- we expect this to still work the same for $n > 2$ and more levels

Pion Example

- fluctuations largely reduced
- stable estimate from $n = 5$ on
- also $+E_0$ estimated well
- averaging reduces uncertainty significantly

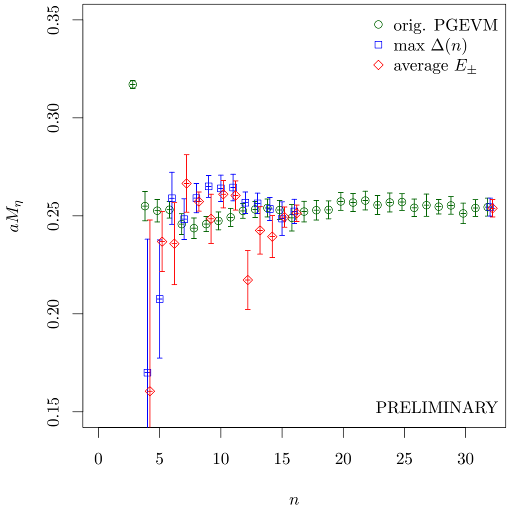


η meson Example

- η/η' significantly affected by noise
- B55 example, $aM_\eta = 0.2481(8)$

  [Ottnad, Urbach (2018)]
- **preliminary** GPOF / Block PGEVM with 2×2 correlator sub-matrices
- original Block PGEVM performs best
- significantly more phoney states

$\Rightarrow$ need to work on state identification

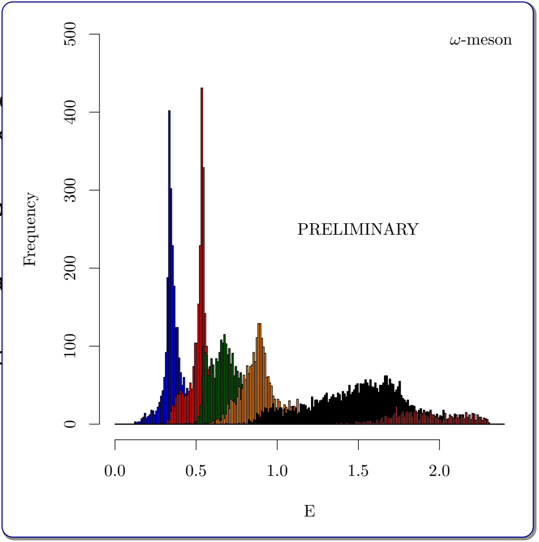


Summary

- Prony GEVM mathematically equivalent to Lanczos
  (more generally GPOF, Block Lanczos and Block PGEVM)
- bootstrap median as estimator is key for reliable error estimates
- full agreement to multi-state fits for nucleon and Ω
- more realistic uncertainty estimate without model averaging and fitting?
- phoney levels become issue for Block PGEVM

Summary

- Prony GEVM mathematically equivalent to Lan
  (more generally GPOF, Block Lanczos and Block
- bootstrap median as estimator is key for reliab
- full agreement to multi-state fits for nucleon a
- more realistic uncertainty estimate without mo
- phoney levels become issue for Block PGEVM



Thanks to...

- **Johann Ostmeyer** and **Aniket Sen**
- **Constantia Alexandrou** and **Simone Bacchio** for the Baryon correlators
- the DFG for the funding
- ... and for your attention