A poor man's attempt at fancy fitting of noisy lattice QCD data with exponentially degrading signal-to-noise ratios

Bayesian Methods in Nuclear Physics
INT June/July 2016

André Walker-Loud

BERKELEY LAB

Fitting Noisy Lattice QCD Correlation Functions

- Lattice QCD and Lattice QCD "data"
  - why "poor man's"?
- Variational Projection
- Input Parameter Free multi-exponential fits
- GPOF
- Matrix Prony

These are all methods I learned working with Kostas Orginos @ W&M/JLab, 2008-2010

Lattice QCD

- *QCD is The fundamental theory of the strong interactions*

$$\mathcal{L}_{QCD} = \bar{q}_{a,\alpha,f}(x)\,[D_\mu\gamma_\mu + m]^{b,\beta,f'}_{a,\alpha,f}\, q_{b,\beta,f'}(x) - \frac{1}{4}G_{\mu\nu}G_{\mu\nu}$$

$q_{b,\beta,f'}(x)$ Quark of *color* b, *spin* β, *flavor* f

flavors = up (u), strange (s), top (t)
down (d), charm (c), bottom (b)

colors = red, green, blue

quarks transform under the fundamental representation of SU(3) color (unitary 3x3)

spin = 4 spin states, 2 particle, 2 anti-particle

Lattice QCD

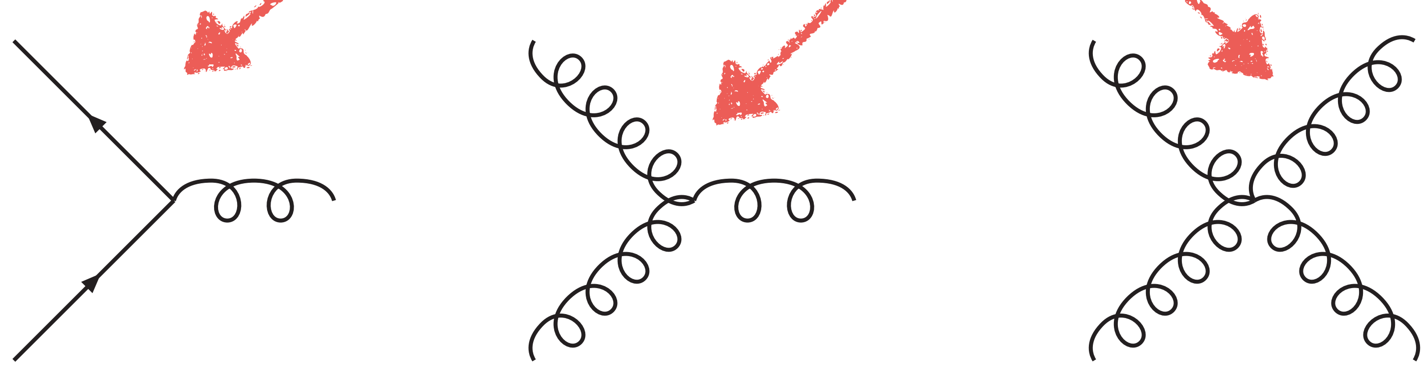
- *QCD is The fundamental theory of the strong interactions*

$$\mathcal{L}_{QCD} = \bar{q}_{a,\alpha,f}(x)\,[D_\mu \gamma_\mu + m]_{a,\alpha,f}^{b,\beta,f'}\, q_{b,\beta,f'}(x) - \frac{1}{4} G_{\mu\nu} G_{\mu\nu}$$

$$[D_\mu]_a^b q_b(x) = \delta_{a,b}\partial_\mu q_b(x) + ig[A_\mu]_a^b q_b(x)$$

gluons adjoint rep. of SU(3) color - 8 gluons

$$\frac{1}{4} G_{\mu\nu} G_{\mu\nu}$$



g g g^2

Lattice QCD

- *QCD is The fundamental theory of the strong interactions*

$$\mathcal{L}_{QCD} = \bar{q}_{a,\alpha,f}(x)\,[D_\mu\gamma_\mu + m]^{b,\beta,f'}_{a,\alpha,f}\, q_{b,\beta,f'}(x) - \frac{1}{4}G_{\mu\nu}G_{\mu\nu}$$

degrees of freedom of QCD are

quarks $q_{b,\beta,f'}(x)$

gluons $[A_\mu]^b_a$

degrees of freedom of nature are protons, neutrons, …

$$M_p = 938.272046 \text{ MeV} \qquad M_n = 939.565379 \text{ MeV}$$

$$M_n - M_p = 1.29333217(42) \text{ MeV}$$

$$M_e = 0.511 \text{ MeV}$$

Lattice QCD

- *QCD is The fundamental theory of the strong interactions*

$$\mathcal{L}_{QCD} = \bar{q}_{a,\alpha,f}(x)\,[D_\mu\gamma_\mu + m]_{a,\alpha,f}^{b,\beta,f'}\,q_{b,\beta,f'}(x) - \frac{1}{4}G_{\mu\nu}G_{\mu\nu}$$

QCD is a remarkably simple theory to write down. At low energies (will define) QCD is a theory of only 3 or 4 parameters:

m_u	mass of the up quark (dimensionfull)
m_d	mass of the down quark (dimensionfull)
(m_s)	mass of the strange quark (dimensionfull)
g	*gauge* coupling between quarks and gluons (dimension-less)

Once these parameters are fixed - everything else is a prediction! - proton mass, He binding energy, neutron star equation of state (maximum neutron star mass), …

Lattice QCD

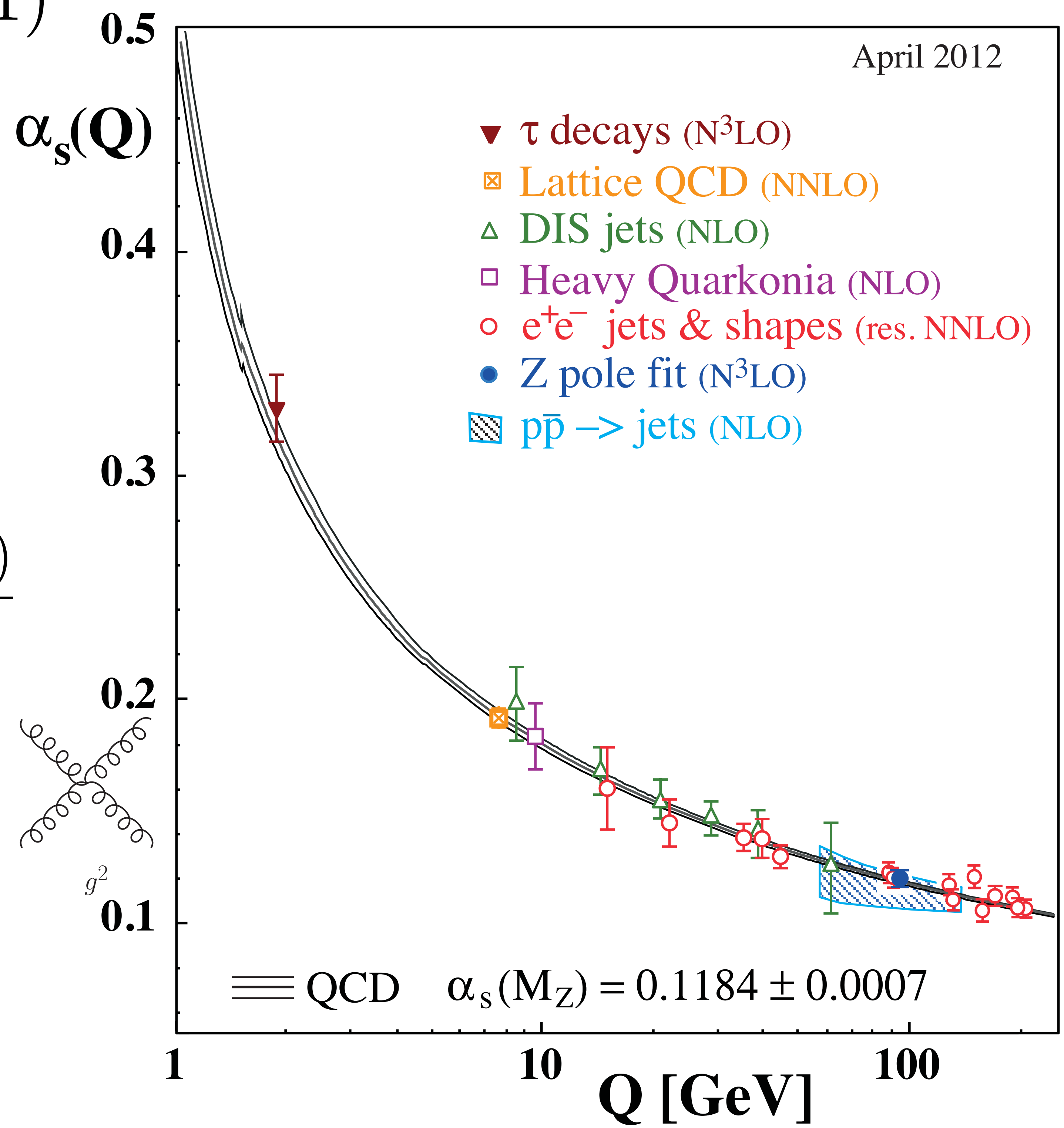
low energy → high energy

$g \sim \mathcal{O}(1)$

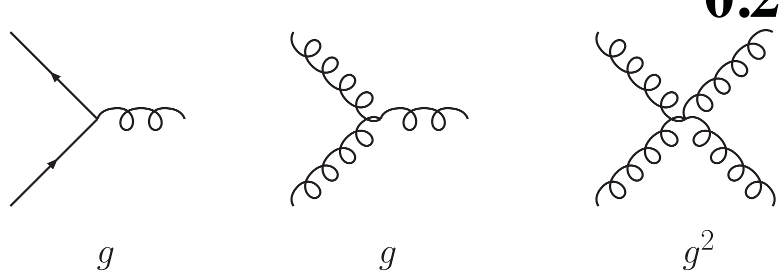
Strong Coupling

$g \to 0$

Asymptotic Freedom



$$\alpha_s(Q) = \frac{g^2(Q)}{4\pi}$$



2004 Nobel Prize
David Gross
David Politzer
Frank Wilczek

Lattice QCD

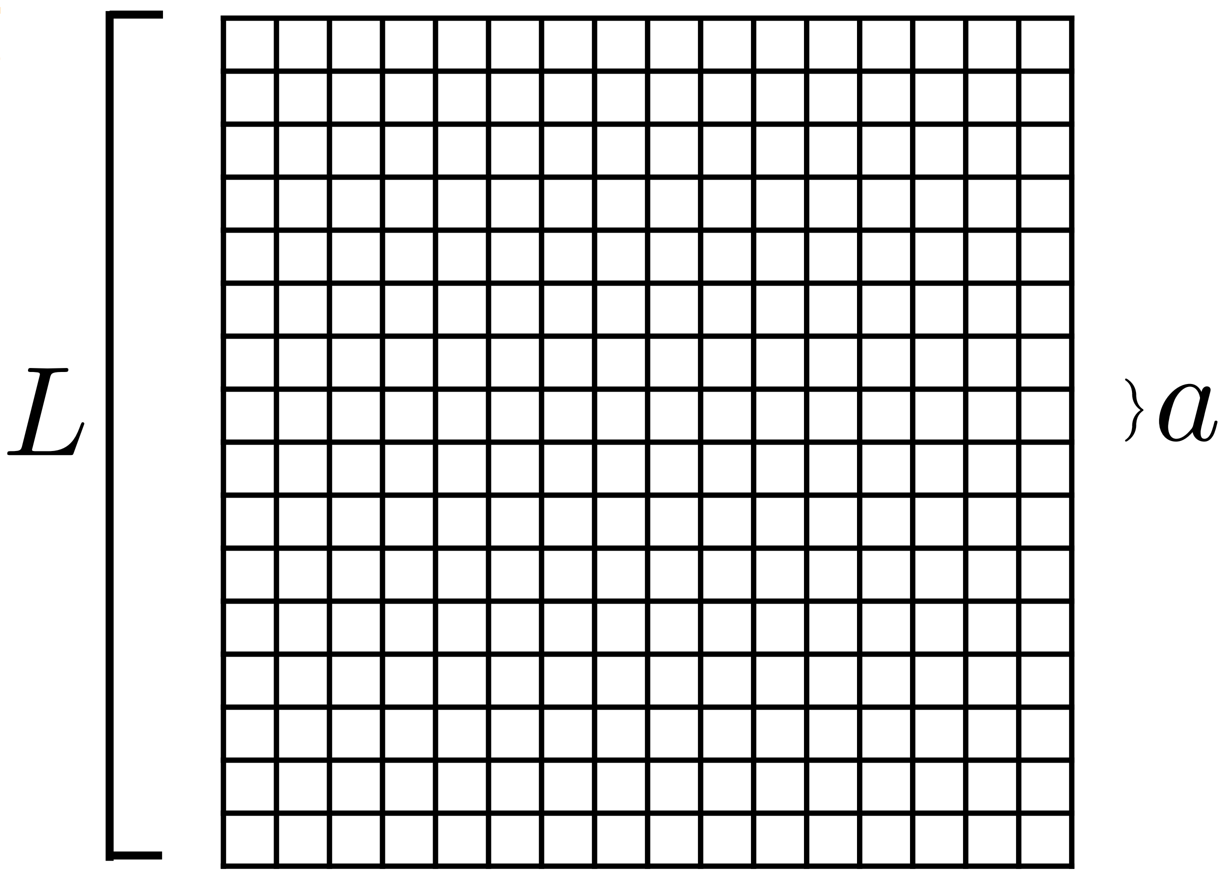
low energy → high energy

$g \sim \mathcal{O}(1)$

Strong Coupling

$g \to 0$

Asymptotic Freedom



$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

small distance
=
large energy

Asymptotic Freedom
Feynman Path Integrals
Wilson Lattice Field Theory
Monte Carlo methods

⇒

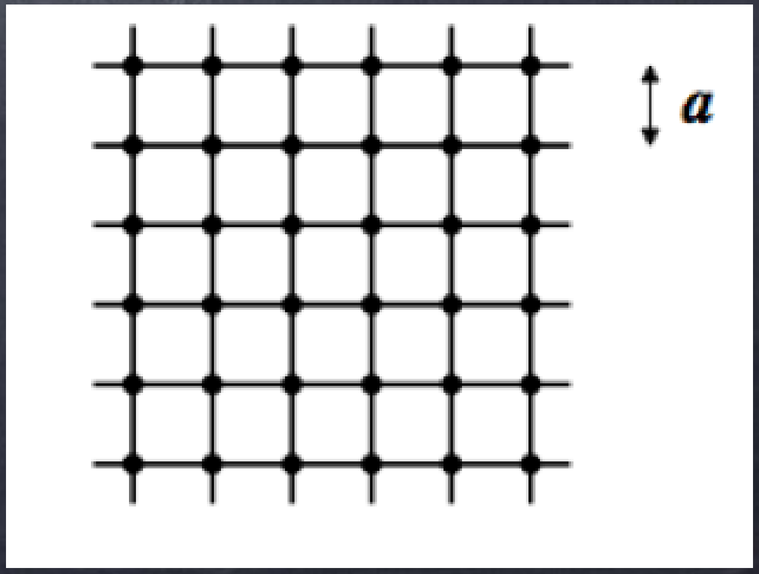
allows numerical solution with exact theory as $a \to 0$ (no uncertainty quantification)

Feynman Path Integrals

$$\mathcal{Z} = \int DA_\mu D\psi D\overline{\psi} e^{iS_{QCD}} \qquad S_{QCD} = \int d^4x \mathcal{L}_{QCD}$$

$$\langle\Omega|\hat{\mathcal{O}}(y)\hat{\mathcal{O}}^\dagger(x)|\Omega\rangle = \frac{1}{\mathcal{Z}}\int DA_\mu D\psi D\overline{\psi} e^{iS_{QCD}}\mathcal{O}(y)\mathcal{O}^\dagger(x)$$

- The path-integral gives us a relation between matrix elements of operators and a high dimensional integral over field configurations.
- We know how to do the integral on the right (in principle at least). The beginning of lattice QFT is to discretize the universe so that we can compute the path-integral representation directly with a computer.
- Suppose we chop the universe into size $32 \times 32 \times 32 \times 64 = 2^{21}$
- our path integral goes over all field configurations on all sites, $n^{2^{21}}$ terms!



Feynman Path Integrals

$$\mathcal{Z} = \int DA_\mu D\psi D\overline{\psi} e^{iS_{QCD}} \qquad S_{QCD} = \int d^4x \mathcal{L}_{QCD}$$

$$\langle\Omega|\hat{\mathcal{O}}(y)\hat{\mathcal{O}}^\dagger(x)|\Omega\rangle = \frac{1}{\mathcal{Z}}\int DA_\mu D\psi D\overline{\psi} e^{iS_{QCD}}\mathcal{O}(y)\mathcal{O}^\dagger(x)$$

- How can we actually perform this integral?
- If we Wick-rotate to Euclidean time, $t \to it_E$, then we have

$$\langle\Omega|\hat{\mathcal{O}}(y_E)\hat{\mathcal{O}}^\dagger(x_E)|\Omega\rangle = \frac{1}{\mathcal{Z}}\int DA_\mu D\psi D\overline{\psi} e^{-S^E_{QCD}}\mathcal{O}(y_E)\mathcal{O}^\dagger(x_E)$$

- For zero quark chemical-potential (zero baryon chemical potential)

$$e^{-S^E_{QCD}} \in \mathbb{R}$$

- We can use this factor as a probability measure to importance sample the integral with Monte-Carlo methods for those field configurations that minimize S^E_{QCD}

Feynman Path Integrals

$$\langle\Omega|\hat{\mathcal{O}}(y_E)\hat{\mathcal{O}}^\dagger(x_E)|\Omega\rangle = \frac{1}{\mathcal{Z}}\int DA_\mu D\psi D\overline{\psi} e^{-S^E_{QCD}}\mathcal{O}(y_E)\mathcal{O}^\dagger(x_E)$$

- We can make N_{cfg} different samples of the field configurations and then our correlation functions are approximated with finite statistics

$$\langle\Omega|\hat{\mathcal{O}}(y_E)\hat{\mathcal{O}}^\dagger(x_E)|\Omega\rangle = \lim_{N_{cfg}\to\infty}\frac{1}{N_{cfg}}\sum_{i=1}^{N_{cfg}}\langle\Omega|\hat{\mathcal{O}}(y_E)[A^i_\mu,\psi_i,\overline{\psi}_i]\hat{\mathcal{O}}^\dagger(x_E)[A^i_\mu,\psi_i,\overline{\psi}_i]|\Omega\rangle$$

$[A^i_\mu,\psi_i,\overline{\psi}_i]$ = the i^{th} value of the fields on "configuration" i

- We really need to compute the mean - not the median (as dictated by the rules of Quantum Field-Theory)

- At finite statistics (finite N_{cfg}) we will have an approximation to the correlation functions with some computable statistical uncertainty that can be systematically improved (with more computing time)

Feynman Path Integrals

$$\langle\Omega|\hat{\mathcal{O}}(y_E)\hat{\mathcal{O}}^\dagger(x_E)|\Omega\rangle = \frac{1}{\mathcal{Z}}\int DA_\mu D\psi D\overline{\psi} e^{-S^E_{QCD}}\mathcal{O}(y_E)\mathcal{O}^\dagger(x_E)$$

- What do we expect our Euclidean spacetime correlation functions to look like? Let us take $x_E=0$ (without loss of generality - translation invariance lets us do this) and $\vec{y}_E = 0$ for simplicity

$$C(t) = \langle\Omega|\hat{\mathcal{O}}(t,\vec{0})\hat{\mathcal{O}}^\dagger(0,\vec{0})|\Omega\rangle$$

- Insert a complete set of states (completeness)

$$1 = \sum_n |n\rangle\langle n|$$

Lattice QCD results are given by a sum of noisy exponentials - a challenging numerical analysis problem

$$C(t) = \sum_n \langle\Omega|\hat{\mathcal{O}}(t)|n\rangle\langle n|\hat{\mathcal{O}}^\dagger(0)|\Omega\rangle$$

$$= \sum_n \langle\Omega|e^{\hat{H}t}\hat{\mathcal{O}}(0)e^{-\hat{H}t}|n\rangle\langle n|\hat{\mathcal{O}}^\dagger(0)|\Omega\rangle$$

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

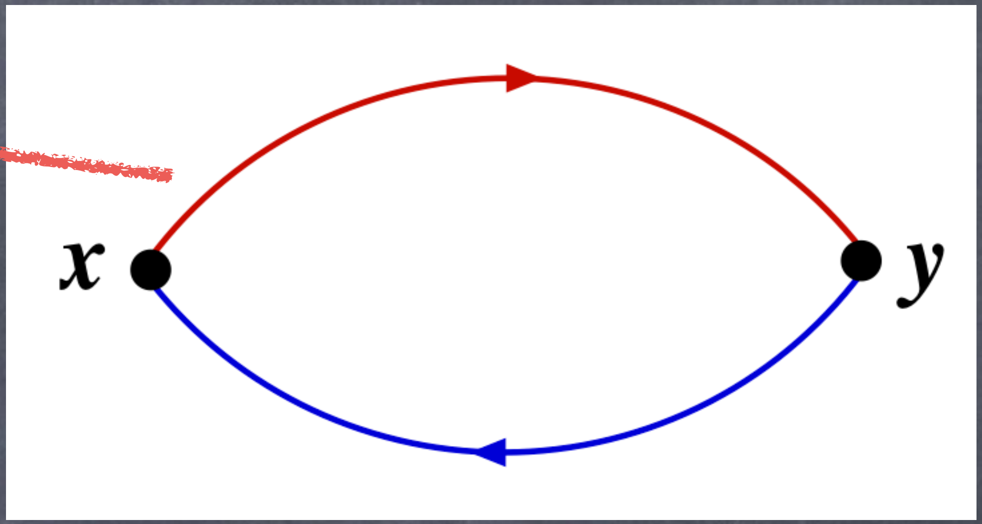
$$Z_n = \langle\Omega|\hat{\mathcal{O}}(0)|n\rangle$$

Quark contractions: Making protons, pions, ...

$$[D_W + M]\, S(x,y;U) = \frac{1}{a^4}\delta_{xy}$$

Quark propagator

Pion correlation function



To solve for the quark propagator, S, we must invert a large sparse matrix

$$[D_W + M]^{-1}$$

Then - we Wick-contract the quarks together to make states of interest: e.g. the pion

$$C(t) = \langle \Omega | \hat{\mathcal{O}}(t,\vec{0}) \hat{\mathcal{O}}^\dagger(0,\vec{0}) | \Omega \rangle$$

$$\hat{\mathcal{O}}^\dagger(y) = \bar{d}(y)\gamma_5 u(y)$$

$$\hat{\mathcal{O}}(x) = \bar{u}(x)\gamma_5 d(x)$$

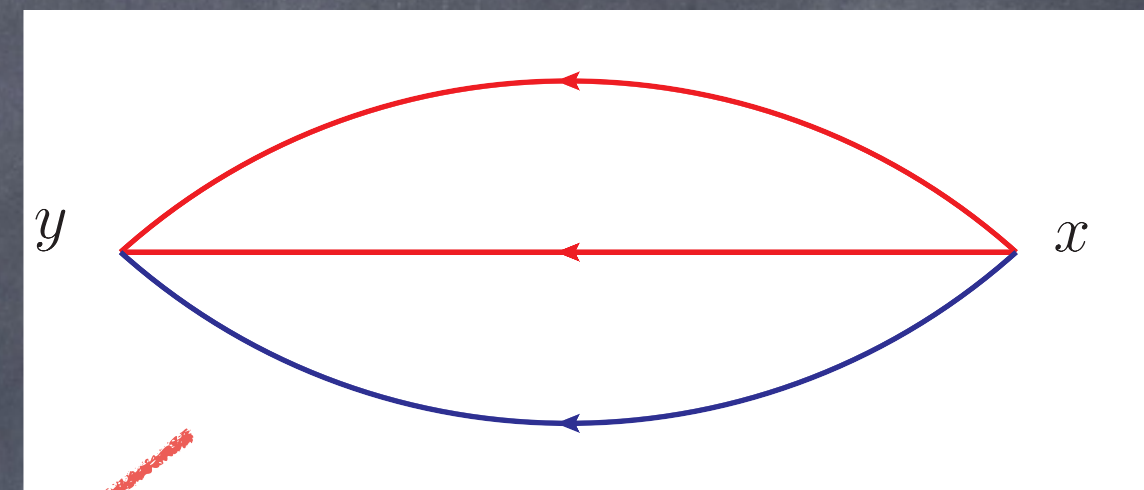
$$\left\langle \overline{u}(x)\gamma_5 d(x) \overline{d}(y)\gamma_5 u(y) \right\rangle_F = -\mathrm{tr}\left\{ \gamma_5 S_{dd}(x,y;U) \gamma_5 S_{uu}(y,x;U) \right\}$$

Quark contractions: Making protons, pions, ...

$$[D_W + M]\, S(x, y; U) = \frac{1}{a^4}\delta_{xy}$$

A proton has 2 u-quarks and 1 d-quark. The contractions are slightly more complex - we need to keep track of which u-quark from x goes to which u-quark at y - 2 contractions (Nu! * Nd!)

Proton correlation function



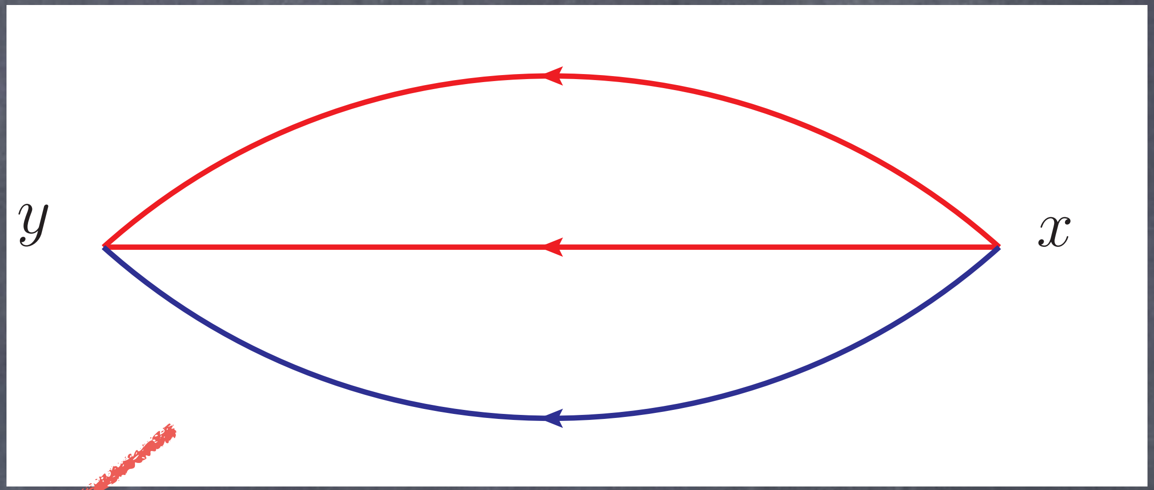
2 protons (proton-proton scattering) has 4! * 2! = 48 contractions
He3 (ppn) has 5! * 4! = 2880 contractions
He4 (ppnn) has 6! * 6! = 518400 contractions!
...

Symmetries can be used to largely reduce this growth
Yamazaki, Kuramashi, Ukawa - Phys.Rev. D81 (2010)
But this Wick-contraction cost can be dominant for multi-nucleon lattice QCD calculations

Quark contractions: Making protons, pions, ...

$$[D_W + M]\, S(x, y; U) = \frac{1}{a^4}\delta_{xy}$$

Proton correlation function



A proton has 2 u-quarks and 1 d-quark. The contractions are slightly more complex - we need to keep track of which u-quark from x goes to which u-quark at y - 2 contractions (Nu! * Nd!)

2 protons (proton-proton scattering) has 4! * 2! = 48 contractions

He3 (ppn) has 5! * 4! = 2880 contractions

He4 (ppnn) has 6! * 6! = 518400 contractions!

...

The cancellations (- signs) in these contractions give rise to a signal-to-noise problem

Signal-to-Noise

Calculations involving nucleons suffer a severe signal-to-noise problem

$$Signal = Ze^{-m_N t}\left[1 + \delta Z_n e^{-(E_n - m_N)t}\right]$$

Signal for a proton correlation function

$$\frac{Signal}{Noise} \sim \sqrt{N_{sample}}\; e^{-(m_N - \frac{3}{2}m_\pi)t}$$

Signal-to-noise for a proton correlation function

$$m_N \simeq 939 \text{ MeV}$$

$$m_\pi \simeq 135 \text{ MeV}$$

$$\frac{Signal}{Noise} \sim \sqrt{N_{sample}}\; e^{-A(m_N - \frac{3}{2}m_\pi)t}$$

Signal-to-noise for A nucleons

Solving this problem requires solutions at early Euclidean-time before the Noise becomes large - but this requires sophisticated "wave-functions" for the proton, which compounds the Wick-contraction cost mentioned above

Lattice QCD: Recap

- Lattice QCD is a stochastic approximation to the path-integral formulation of QCD in imaginary (Euclidean) time
  - LQCD applications to NP require Peta- to Exa-scale computing - want to get the most out of our cycles
- Numerical Results exactly modeled as a sum of noisy exponentials with exponentially degrading signal-to-noise ratios

$$C_N(t) = \sum_n \tilde{Z}_n Z_n^\dagger e^{-E_n t}$$

$$\lim_{t\to\infty} C_{AN}(t) = \tilde{Z}_0 Z_0^\dagger e^{-(A m_N + \Delta E_A)t}$$

$$\Delta E_A \simeq 10 \text{ MeV}$$

$$m_N \simeq 10^3 \text{ MeV}$$

$$A = 1, 2, 3, 4, \ldots$$

- gaps to excited states can also be in the 10-100 MeV range
  - to resolve both energy levels, need $t \sim 1/(E_1 - E_0)$ which is precisely where the noise is growing unwieldy
- current calculations typically use just 1 or 2 Markov chains - due to computational costs - though this may change soon

Lattice QCD: Recap

- Why "poor man's"?

$$C(t) = \sum_n \tilde{Z}_n Z_n^\dagger e^{-E_n t}$$

$$\rightarrow C_{ij}(t) = \sum_n \tilde{Z}_{i,n} Z_{j,n}^\dagger e^{-E_n t}$$

a "rich man" would create a large basis of operators that all couple to the same states - so that a diagonalization of this basis can be performed via a generalized eigenvalue problem leaving one with linear combinations of operators that couple predominantly to single states, n.

- This idea works extremely well for mesons (quark—anti-quark states) but is prohibitively costly for 2 or more nucleons

Lattice QCD: Recap

- Standard Tool: Effective Mass

$$C(t) = \sum_n \tilde{Z}_n Z_n^\dagger e^{-E_n t}$$

$$m_{eff}(t,\tau) \equiv \frac{1}{\tau} \ln\left(\frac{C(t)}{C(t+\tau)}\right)$$

$$\lim_{t\to\infty} m_{eff}(t,\tau) = E_0$$

Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

Very simple idea:
if a fit function has mixed linear/non-linear dependence on the fit parameters, one does not need to perform a numerical minimization on all the parameters - one can first perform a linear least squares on the linear parameters, solving as a function of the non-linear ones

Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

Very simple idea: first perform linear-least squares

in lattice QFT, correlation functions fit with

$$\chi^2 = \sum_{t,t'} [y(t) - f(\lambda,t)]\, C^{-1}_{t,t'}\, [y(t') - f(\lambda,t')]$$

$$f(\lambda,t) = \sum_n Z_n \lambda_n^t \qquad \lambda_n = e^{-E_n} \qquad \lambda = \{Z_n, \lambda_n\}$$

$$\frac{\partial \chi^2}{\partial Z_n} = 0$$

(repeated indices summed over)

$$= \left[-\lambda_n^t C^{-1}_{t,t'} (y(t') - Z_m \lambda_m^{t'}) - (y(t) - Z_m \lambda_m^t) C^{-1}_{t,t'} \lambda_n^{t'} \right]$$

Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

Very simple idea: first perform linear-least squares

symmetry in $t \leftrightarrow t'$

$$\frac{\partial \chi^2}{\partial Z_n} = 0$$

$$= \left[-\lambda_n^t C_{t,t'}^{-1} (y(t') - Z_m \lambda_m^{t'}) - (y(t) - Z_m \lambda_m^t) C_{t,t'}^{-1} \lambda_n^{t'} \right]$$

$$\Rightarrow \quad Z_m \lambda_m^t C_{t,t'}^{-1} \lambda_n^{t'} = y(t) C_{t,t'}^{-1} \lambda_n^{t'}$$

$$Z_m \Lambda_{m,n} = y(t) C_{t,t'}^{-1} \lambda_n^{t'} \qquad \Lambda_{m,n} = \lambda_m^t C_{t,t'}^{-1} \lambda_n^{t'}$$

Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

Very simple idea: first perform linear-least squares

$$\frac{\partial \chi^2}{\partial Z_n} = 0 \quad \Rightarrow \quad Z_m = \Lambda^{-1}_{m,n}\, y(t)\, C^{-1}_{t,t'}\, \lambda_n^{t'}$$

we have solved for the overlap factors as functions of the eigenvalues

plug these solutions back into χ^2 and perform numerical minimization on just the non-linear parameters (E_n)

Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

Very simple idea: first perform linear-least squares

$$\chi^2 = \sum_{t,t'} [y(t) - f(\lambda, t)]\, C^{-1}_{t,t'}\, [y(t') - f(\lambda, t')]$$

$$f(\lambda, t) = \sum_n Z_n \lambda_n^t \qquad \lambda_n = e^{-E_n}$$

$$Z_m = \Lambda^{-1}_{m,n}\, y(t)\, C^{-1}_{t,t'}\, \lambda_n^{t'} \qquad \Lambda_{m,n} = \lambda_m^t C^{-1}_{t,t'} \lambda_n^{t'}$$

When counting the degrees of freedom in the fit DON'T forget to count the overlap factors you have also determined!

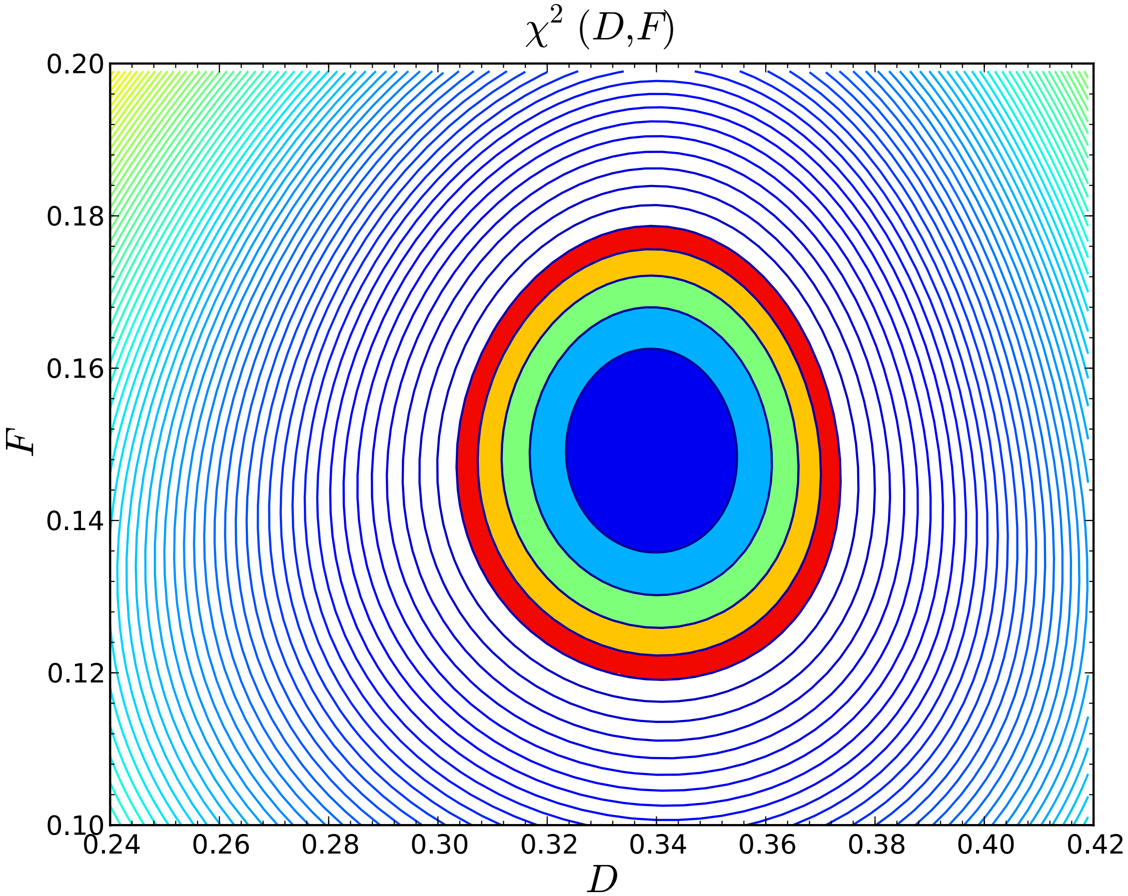
readily extend this to matrix of correlation functions

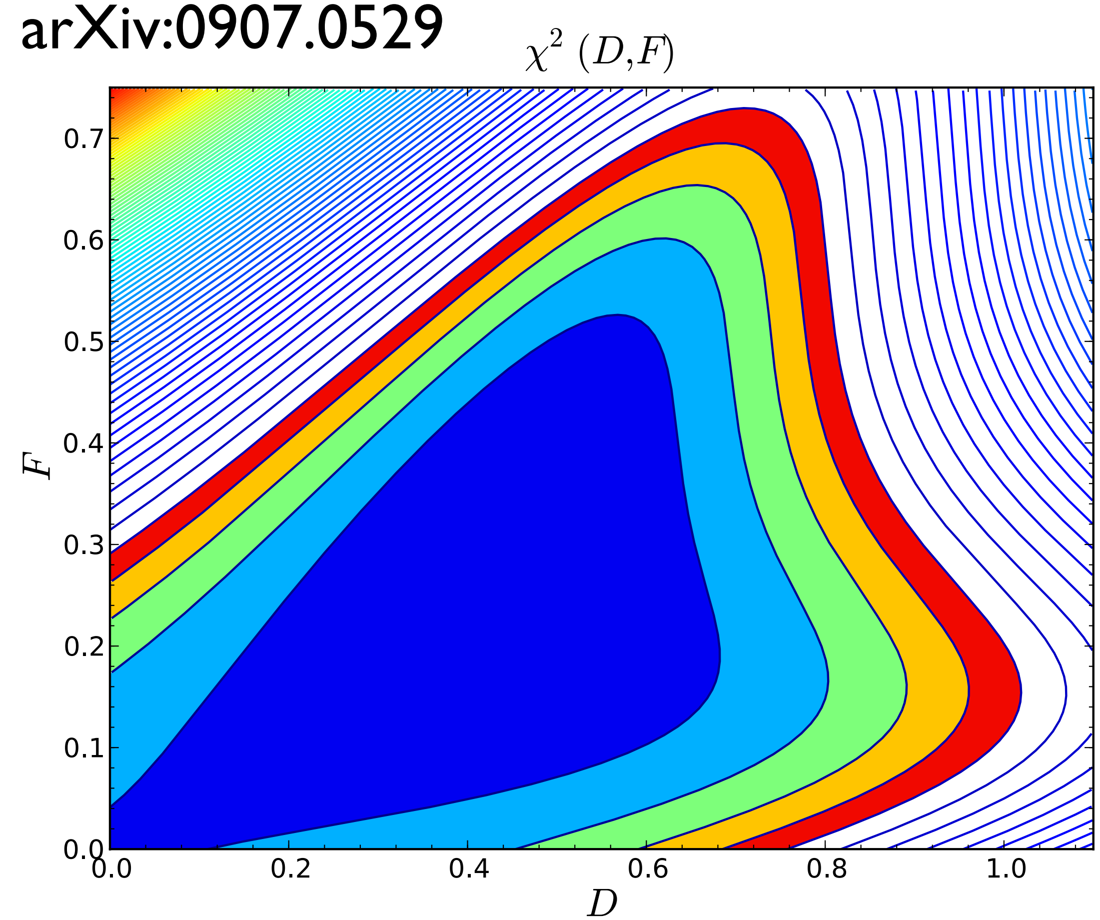
Variational Projection

Golub & Pereyra Inverse Problems 19 (2003)
arXiv:0907.0529

can also apply the same idea to chiral extrapolation formulae which usually have mixed linear/non-linear dependence on LECs

E. Jenkins, A. Manohar, J. Negele and AWL
arXiv:0907.0529





NLO: 11 LECs

NNLO: 30 LECs

Fit to Octet and Decuplet baryon mass results with SU(3) Baryon Chiral Perturbation Theory

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

"The fit of a sum of exponentials to noisy data"
(Note: typos in paper)

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

Given the outputs of a non-degenerate n-dimensional linear system

$$y_t = c^T x_t \qquad c_i \neq 0, i = 1, \ldots, n$$

$$\{x_t, c\} \in \mathbb{R}^n$$

$$x_{t+1} = T x_t \qquad x_0^T = (1, \ldots, 1)$$

$$y_t = c^T T^t x_0$$

construct the p x q Hankel Matrix with p,q > n

$$H_{ij} = y_{i+j-1}$$

$$H = \begin{pmatrix} y_0 & \cdots & y_{q-1} \\ \vdots & \ddots & \vdots \\ y_{p-1} & \cdots & y_{q+p-2} \end{pmatrix}$$

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

The Hankel Matrix can be factorized

$$H = \begin{pmatrix} y_0 & \cdots & y_{q-1} \\ \vdots & \ddots & \vdots \\ y_{p-1} & \cdots & y_{q+p-2} \end{pmatrix}$$

$$y_t = c^T T^t x_0$$

$$H = FG = \begin{pmatrix} c^T \\ c^T T \\ \vdots \\ c^T T^{p-1} \end{pmatrix} \begin{pmatrix} x_0 & Tx_0 & \cdots & T^{q-1} x_0 \end{pmatrix}$$

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

The factorization matrices can be inverted $H = FG$

$F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ $G : \mathbb{R}^q \rightarrow \mathbb{R}^n$ (H is $p \times q$ matrix)

$$\mathbb{1}^n = \tilde{F}^{-1} H \tilde{G}^{-1}$$

$$\{\tilde{F}^{-1}, \tilde{G}^{-1}\} = \text{pseudo-inverse}\{F, G\}$$

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

The shifted Hankel Matrix can be factorized

$$H_{ij} = y_{i+j-1} \qquad H^+_{ij} = y_{i+j} \qquad y_t = c^T T^t x_0$$

$$H = FG \qquad H^+ = FTG$$

We can construct a matrix similar to T

$$T_s = \tilde{F}^{-1} H^+ \tilde{G}^{-1} = \tilde{F}^{-1} F T G \tilde{G}^{-1}$$

How do we determine $\{F, G\}$?

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

The shifted Hankel Matrix can be factorized

$$H_{ij} = y_{i+j-1} \qquad H^{+}_{ij} = y_{i+j} \qquad y_t = c^T T^t x_0$$

$$H = FG \qquad H^{+} = FTG$$

How do we determine $\{F, G\}$?

- First - pick a window in time of interest you want to analyze, t_i, t_f
- One can (should) shift with $\Delta t = \tau > 1$

$$H^{\tau}_{ij} = y_{i+j-1+\tau}$$

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

- First - pick a window in time of interest you want to analyze, t_i, t_f
- One can (should) shift with $\Delta t = \tau > 1$

$$H^{\tau}_{ij} = y_{i+j-1+\tau}$$

$$H = \begin{pmatrix} y_0 & \cdots & y_{q-1} \\ \vdots & \ddots & \vdots \\ y_{p-1} & \cdots & y_{q+p-2} \end{pmatrix}$$

```python
# written in python language
import numpy as np

corr = # data array in form array([ncfg,nt,data[ncfg,nt]])
ncfg = corr.shape[0]
nt = corr.shape[1]
dt = 4 #shift you chose

h_tmp = np.zeros([nt/2 -dt, nt/2 -dt])
h_tmp_shift = np.zeros_like(h_tmp)
y_t = np.mean(corr,axis=0)
for t1 in range(nt/2 - dt):
    for t2 in range(nt/2 - dt):
        h_tmp[t1,t2] = y_t[t1+t2]
        h_tmp_shift[t1,t2] = y_t[t1+t2+dt]

hankel = h_tmp[t0:tf:dt,0:tf-t0:dt]
hankel_shift = h_tmp_shift[t0:tf:dt,0:tf-t0:dt]
```

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

Perform a singular value decomposition on H

$$H = U\Sigma V^T \qquad \Sigma = \mathrm{diag}(\mathrm{singular\ values})$$

$$T_s = \Sigma^{-1/2}\, U^T\, H^+\, V\, \Sigma^{-1/2} \qquad H^+ = FTG$$

One very important point: need to truncate the singular values to a reasonable number

- If you desire just the ground state - chose range of time suitably (contamination from just a single state) and take two singular values
- For determining 2 states (gs + 1st excited) pick 3 singular values etc.
- study the singular values to see how much information you can get from correlation function

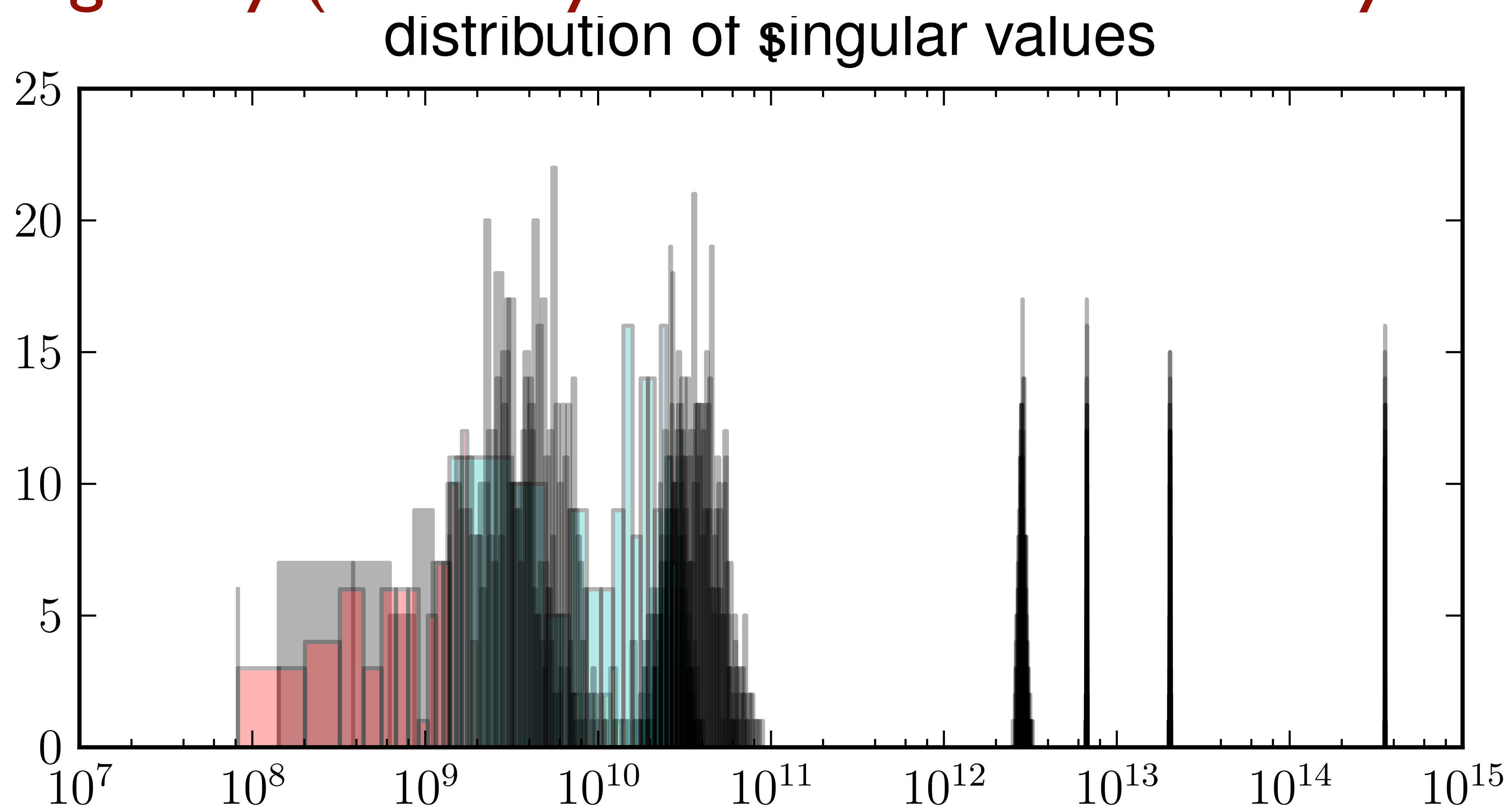
Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

Perform a singular value decomposition on H

$$H = U\Sigma V^T \qquad \Sigma = \mathrm{diag}(\text{singular values})$$

Don't get greedy (don't try and determine too many excited states)



in this case, at most, one can determine 4 energies and only 3 with confidence (the 4th absorbs the "slop")

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

Perform a singular value decomposition on H

$$H = U\Sigma V^T \qquad \Sigma = \mathrm{diag}(\text{singular values})$$

take n_s largest singular values

$$T_s = \Sigma^{-1/2}[0:n_s]\, U^T[0:ns,:]\, H^+\, V[:,0:n_s]\, \Sigma^{-1/2}[0:n_s]$$

$$\lambda_n, O_n = \mathrm{eig}(T_s) \qquad H^+ = FTG$$

$$E_n = -\ln(\lambda_n)$$

Now we have the eigen-energies - how do we get the overlap factors?

See VarPro!

Input Parameter Free Multi-Exp Fits

De Groen & De Moor
Comp. Appl. Math. 20 (1987)

$$H_{ij} = y_{i+j-1} \quad H^{+}_{ij} = y_{i+j} \qquad H = U \Sigma V^T \qquad n_s$$

$$T_s = \Sigma^{-1/2}[0:n_s]\, U^T[0:ns,:]\, H^+\, V[:,0:n_s]\, \Sigma^{-1/2}[0:n_s]$$

$$\lambda_n, O_n = \mathrm{eig}(T_s) \qquad E_n = -\ln(\lambda_n)$$

$$Z_m = \Lambda^{-1}_{m,n}\, y(t)\, C^{-1}_{t,t'}\, \lambda^{t'}_n \qquad \Lambda_{m,n} = \lambda^t_m C^{-1}_{t,t'} \lambda^{t'}_n$$

Use these values to seed the multi-exponential fit!

GPOF:

K. Orginos Latt2010 (unpublished) but on web
C. Aubin & K. Orginos Latt2011 (on PoS)
C. Aubin, K. Orginos & AWL private work

Lattice Correlation functions:

$$C_{ij}(t) = \langle O_i(t)\tilde{O}_j^\dagger(0)\rangle$$
$$= \sum_n Z_i^n \tilde{Z}_j^{n,\dagger} e^{-E_n t}$$

Ideally, $\tilde{O}_i = O_i$ then one can solve a Generalized EigenValue Problem - Blossier et. al. JHEP 0904 (2009)

But this is often prohibitively expensive (requiring all-to-all propagators)

Can we find a solution that handles non-symmetric, non-positive definite correlation functions? Of Course!

GPOF:

K. Orginos Latt2010 (unpublished) but on web
C. Aubin & K. Orginos Latt2011 (on PoS)
C. Aubin, K. Orginos & AWL private work

Consider

$$K_{ij}^{pq}(t,\tau) = \langle O_i(t+p\tau)\tilde{O}_j^\dagger(-q\tau)\rangle = C_{ij}(t+(p+q)\tau)$$

$$K(t) = \begin{pmatrix} C(t) & C(t+\tau) & C(t+2\tau) & \dots \\ C(t+\tau) & C(t+2\tau) & C(t+3\tau) & \dots \\ C(t+2\tau) & C(t+3\tau) & C(t+4\tau) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Time-evolution of operators produces orthogonal correlators $O(t+p\tau) = e^{p\tau H}O(t)e^{-p\tau H}$

One can then proceed with a generalized eigenvalue-like solution: a non-symmetric C can be shifted into a square-matrix K

GPOF:

K. Orginos Latt2010 (unpublished) but on web
C.Aubin & K. Orginos Latt2011 (on PoS)
C.Aubin, K. Orginos & AWL private work

$$K_{ij}^{pq}(t,\tau) = \langle O_i(t+p\tau)\tilde{O}_j^\dagger(-q\tau)\rangle = C_{ij}(t+(p+q)\tau)$$

But over shifting (too many p's and q's) leads to linearly dependent information

one must perform a singular-value decomposition (SVD) on K(t) and cut the number of allowed singular values

$$K(t_0) = U\Sigma(t_0)V^\dagger \qquad \Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots \\ 0 & \sigma_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

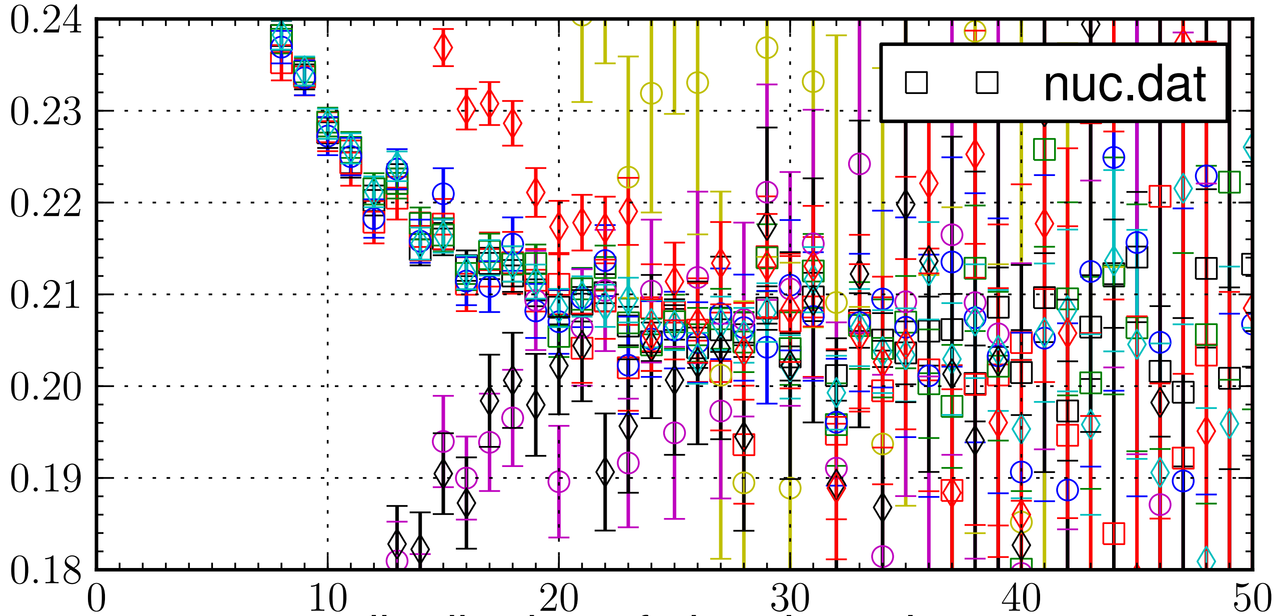
and $A(t) = \Sigma^{-1/2} U K(t) V^\dagger \Sigma^{-1/2}$

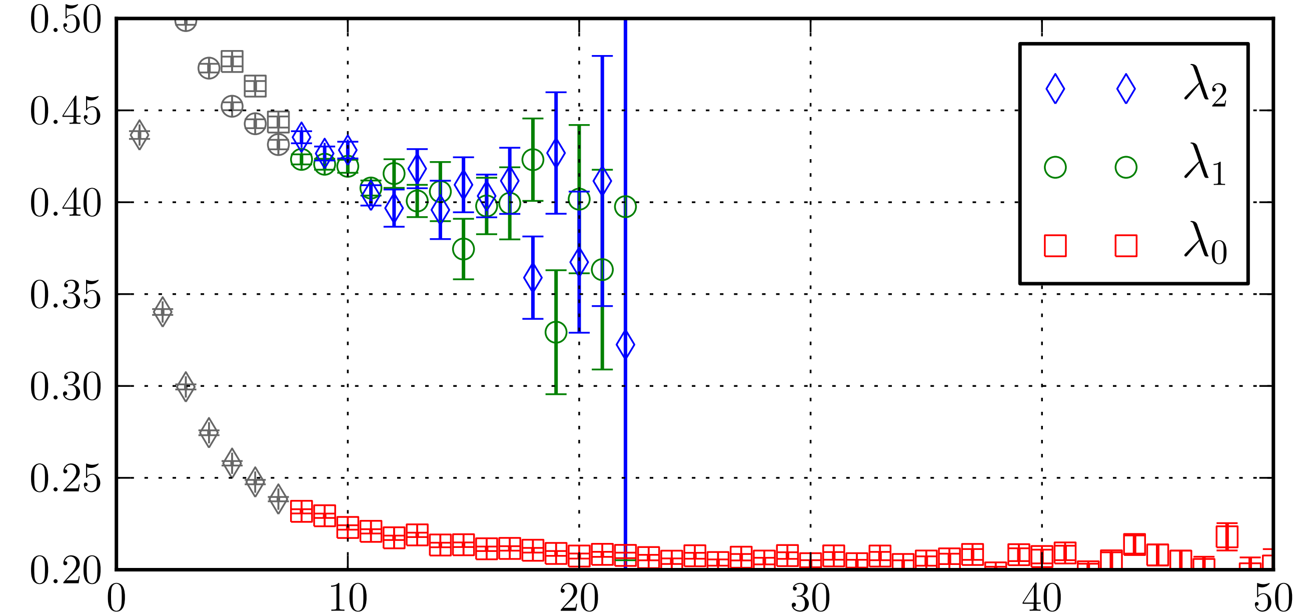
gives an eigenvalue problem for a non-symmetric matrix

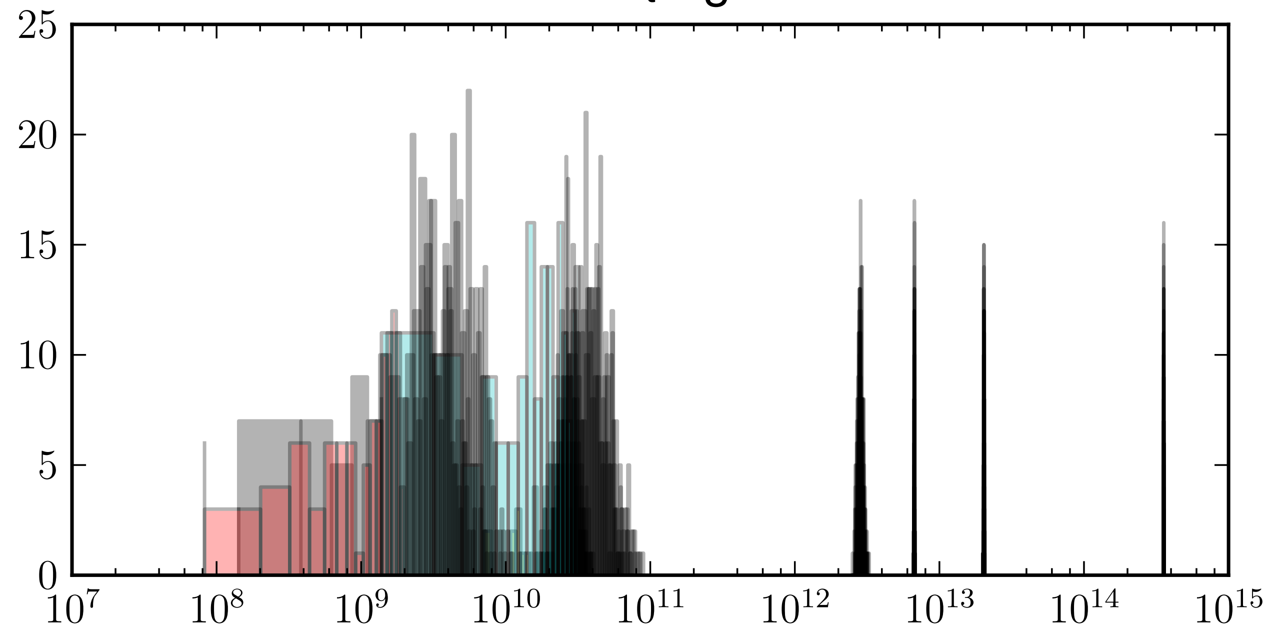
$$A\tilde{q}^n = \lambda_n \tilde{q}^n \qquad q^{n\dagger} A = \lambda_n q^{n\dagger} \qquad q^{n\dagger}\tilde{q}^m = \delta_{nm}$$

GPOF:

K. Orginos Latt2010 (unpublished) but on web
C. Aubin & K. Orginos Latt2011 (on PoS)
C. Aubin, K. Orginos & AWL private work





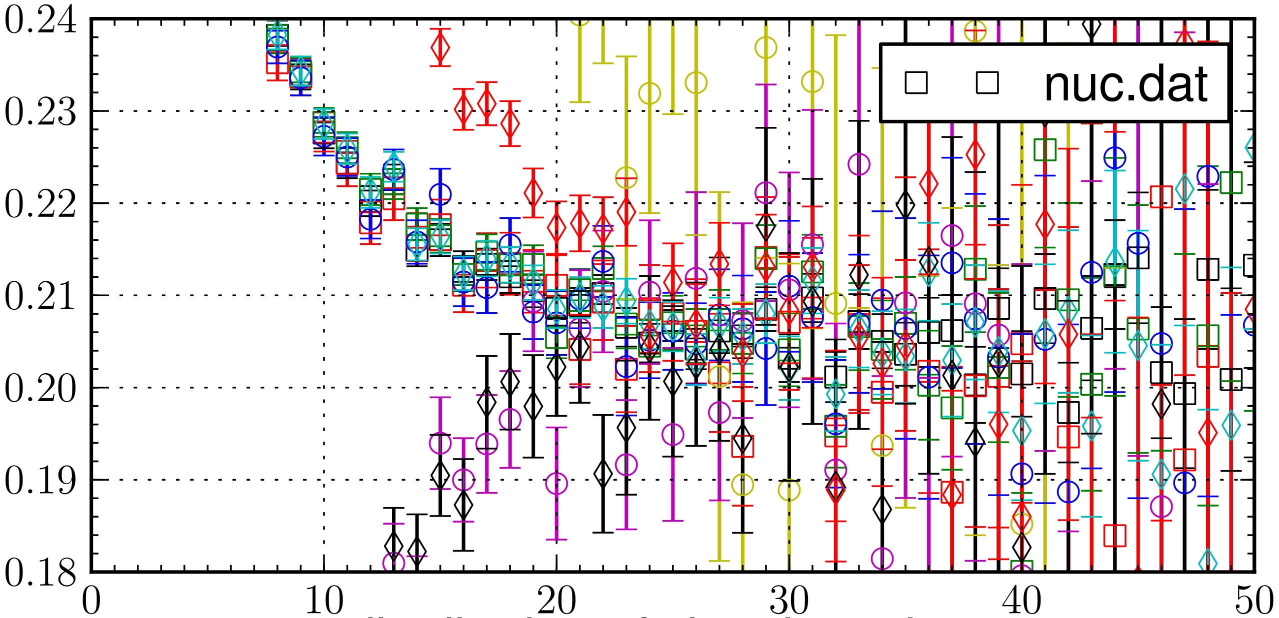


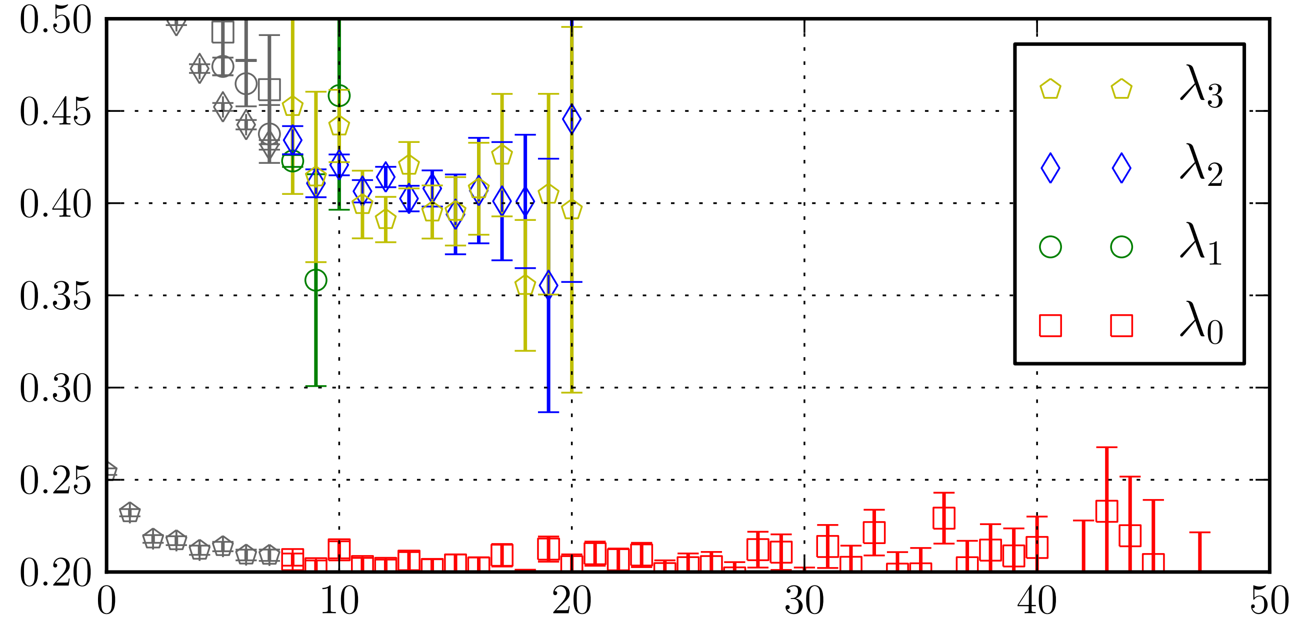
keeping 3 singular values

$$\lambda_n(t) = (1-\delta)e^{-E_n t} + \delta e^{-E_{N+1} t}$$

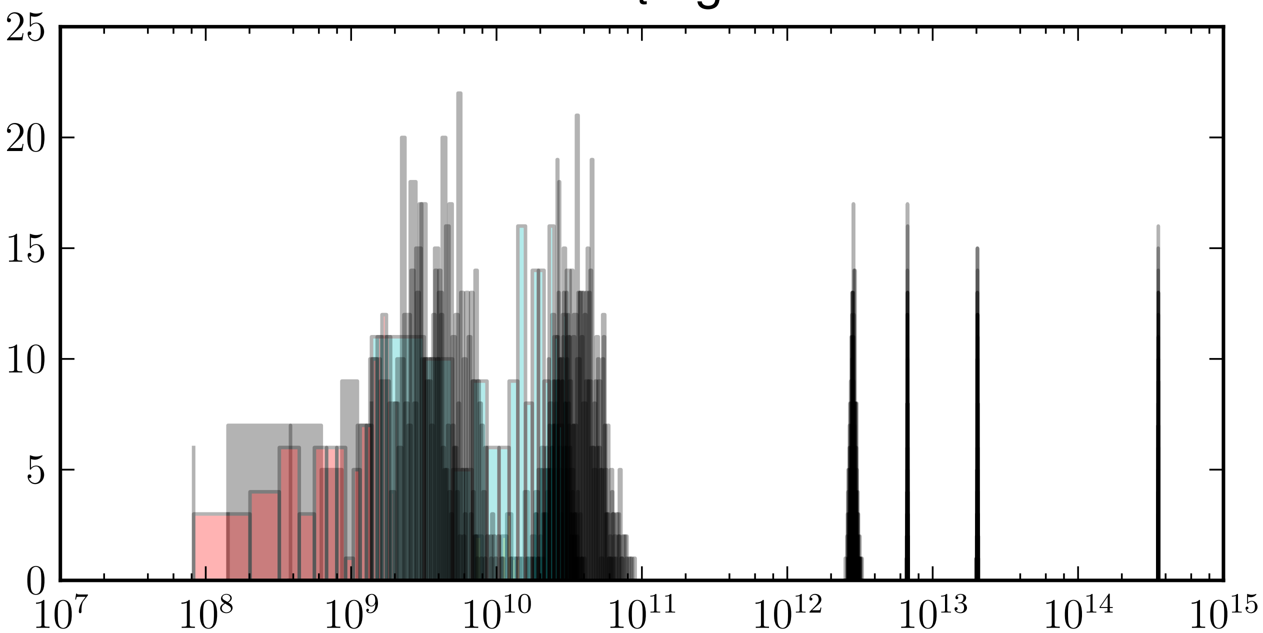
GPOF:

K. Orginos Latt2010 (unpublished) but on web
C.Aubin & K. Orginos Latt2011 (on PoS)
C.Aubin, K. Orginos & AWL private work





distribution of singular values

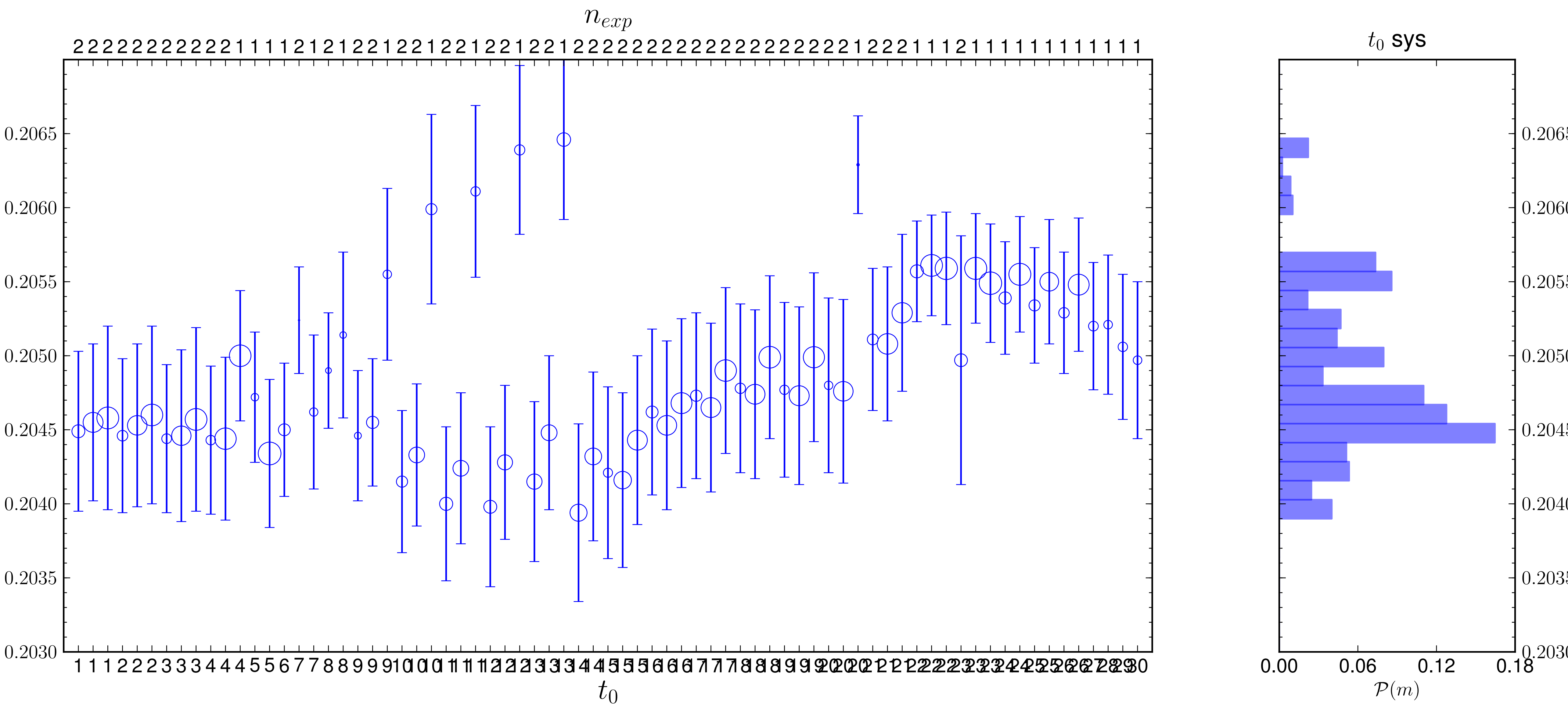


keeping 4 singular values

$$\lambda_n(t) = (1-\delta)e^{-E_n t} + \delta e^{-E_{N+1} t}$$

GPOF:

K. Orginos Latt2010 (unpublished) but on web
C.Aubin & K. Orginos Latt2011 (on PoS)
C.Aubin, K. Orginos & AWL private work



sweeping over choices or parameters to determine systematics

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

Prony = some guys name

G. R. de Prony Journal de l'cole Polytechnique, volume 1, cahier 22, 24-76 (1795)

Matrix = matrix

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

Start with a vector of correlation functions
perhaps a single src and multiple sinks

$$y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_N(t) \end{pmatrix}$$

We would like to construct an operator such that

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

$$\hat{T}(\tau) = \text{ Transfer Operator}$$

Matrix Prony

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

$$y(t+\tau) y^T(t) = \hat{T}(\tau) y(t) y^T(t)$$

Transpose

$$\hat{M}(\tau) y(t+\tau) y^T(t) = \hat{V}(\tau) y(t) y^T(t) \qquad \hat{T} = \hat{M}^{-1}\hat{V}$$

Nothing special about t

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

⬇

$$\hat{M}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) = \hat{V}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t) y^T(t)$$

We have assumed T (and M and V) are independent of t

but this is precisely what we desire from a “transfer matrix”

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$\hat{M}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) = \hat{V}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t) y^T(t)$$

a solution

$$\hat{M}(\tau) = \left[\sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) \right]^{-1}, \quad \hat{V}(\tau) = \hat{V} = \left[\sum_{t_0}^{t_0+\Delta t} y(t) y^T(t) \right]^{-1}$$

must sum over sufficient number of time slices to make matrices "full rank" for two-components, must sum over at least two time slices

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$\hat{M}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) = \hat{V}(\tau) \sum_{t_0}^{t_0+\Delta t} y(t) y^T(t)$$

a solution

$$\hat{M}(\tau) = \left[\sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) \right]^{-1}, \quad \hat{V}(\tau) = \hat{V} = \left[\sum_{t_0}^{t_0+\Delta t} y(t) y^T(t) \right]^{-1}$$

most robust results come from maximizing Δt

this requires that over a large range of time our ansatz is satisfied - only N states contribute in $t_0 \to t_0 + \Delta t$

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{M}^{-1}(\tau)\hat{V}y(t) = \hat{T}(\tau)y(t)$$

$$\hat{M}(\tau) = \left[\sum_{t_0}^{t_0+\Delta t} y(t+\tau)y^T(t)\right]^{-1}, \quad \hat{V}(\tau) = \hat{V} = \left[\sum_{t_0}^{t_0+\Delta t} y(t)y^T(t)\right]^{-1}$$

given M and V one then solves the eigenvalue equation

$$\hat{T}(\tau)q_n = (\lambda_n)^\tau q_n \qquad \lambda_n = e^{-E_n} \qquad \hat{T}(\tau) = \hat{M}^{-1}(\tau)\hat{V}$$

given λ_n check to see if ansatz is satisfied - over range of $t_0 \to t_0 + \Delta t$ there should be no significant evidence of excited state contamination

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

in Python

```
''' MATRIX PRONY '''
import numpy as np
from scipy import linalg as spla
def matrix_prony(corr,ti,tau,td=1):
    corr = np.array(corr)
    n_cfg,n_t,n_corr = corr.shape
    if tau < n_corr:
        print 'Matrix Prony Error:'
        print '  must chose at least %d time slices' %n_corr
        sys.exit(-1)
    y_avg = np.mean(corr,axis=0)
    Minvt = np.array(map(np.outer,np.roll(y_avg,-td,0),y_avg))
    Vinvt = np.array(map(np.outer,y_avg,y_avg))
    Minv = np.sum(Minvt[ti:ti+tau],axis=0)
    Vinv = np.sum(Vinvt[ti:ti+tau],axis=0)
    Transfer = np.dot(Minv,spla.inv(Vinv))
    le,Oe = spla.eig(Transfer)
    ind = np.argsort(le); indr = ind[::-1]

    return le, Oe, indr
```

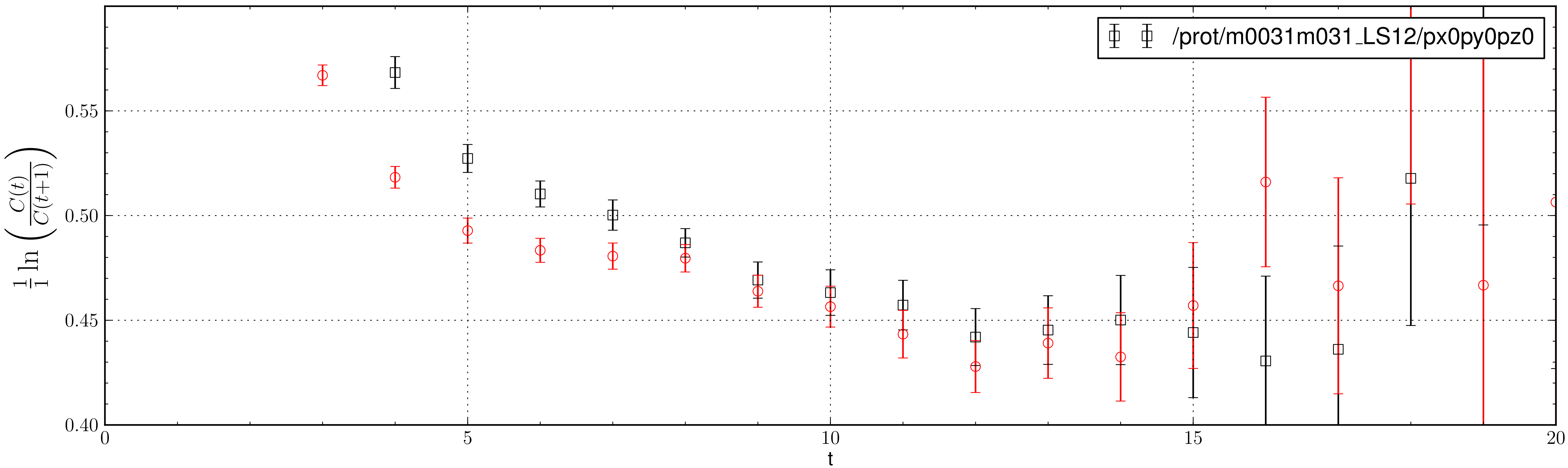
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau)y(t)$$

take correlation function with
one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=5 and t=15 there appear to be two states

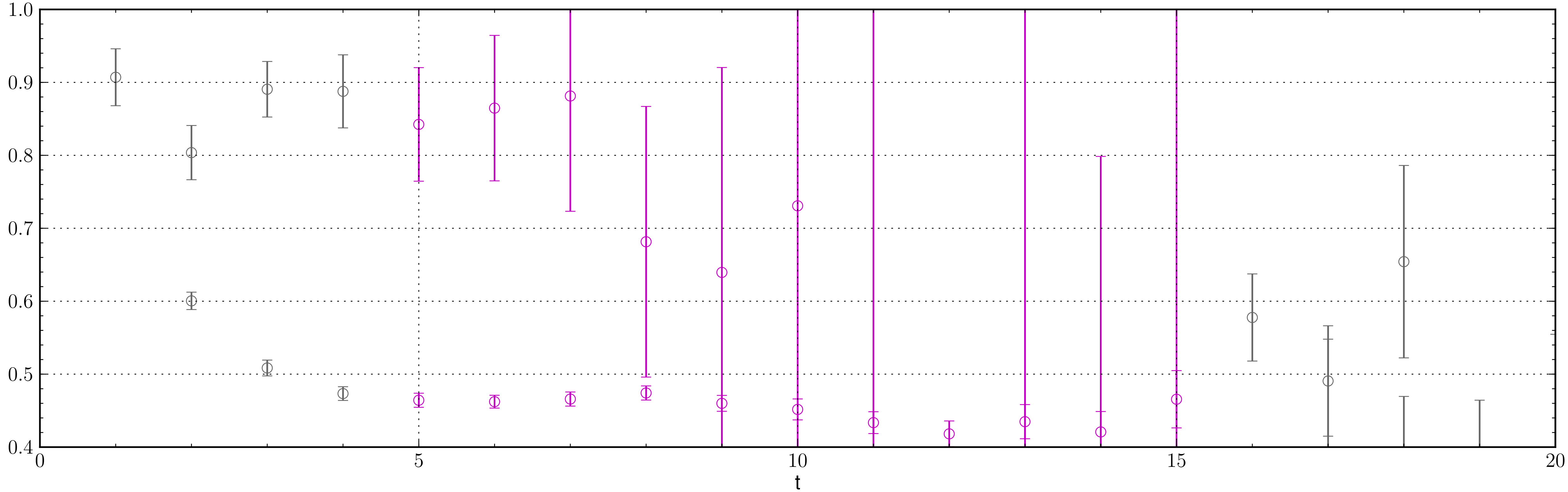
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

take correlation function with
one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=5 and t=15 there appear to be two states

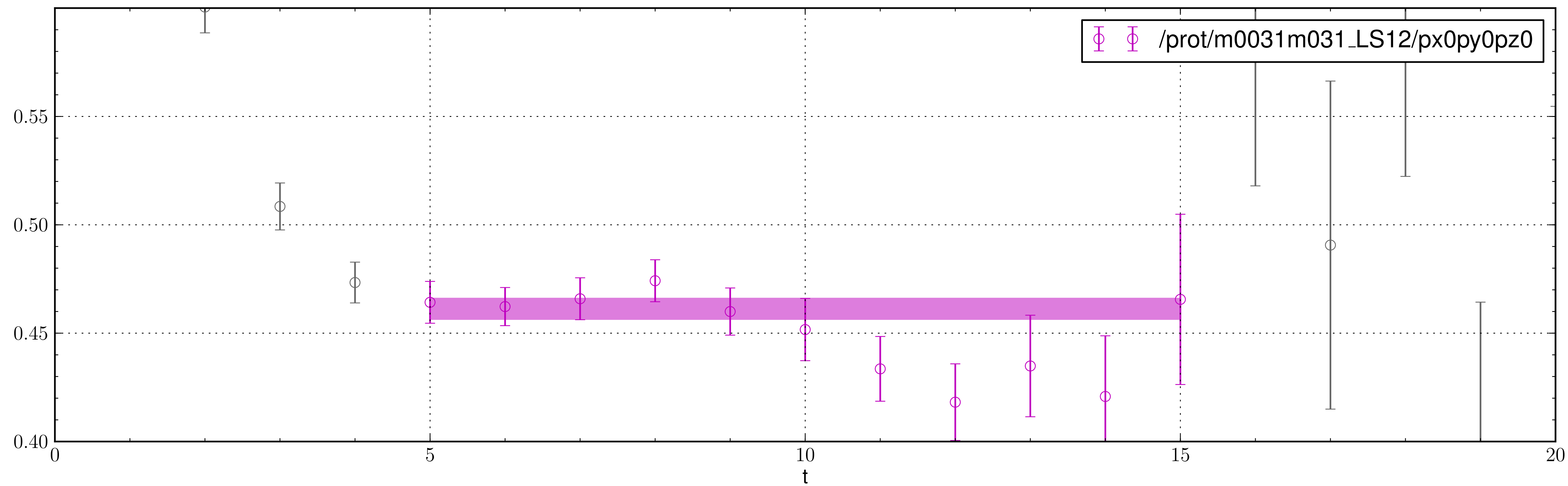
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

take correlation function with
one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=5 and t=15 there appear to be two states

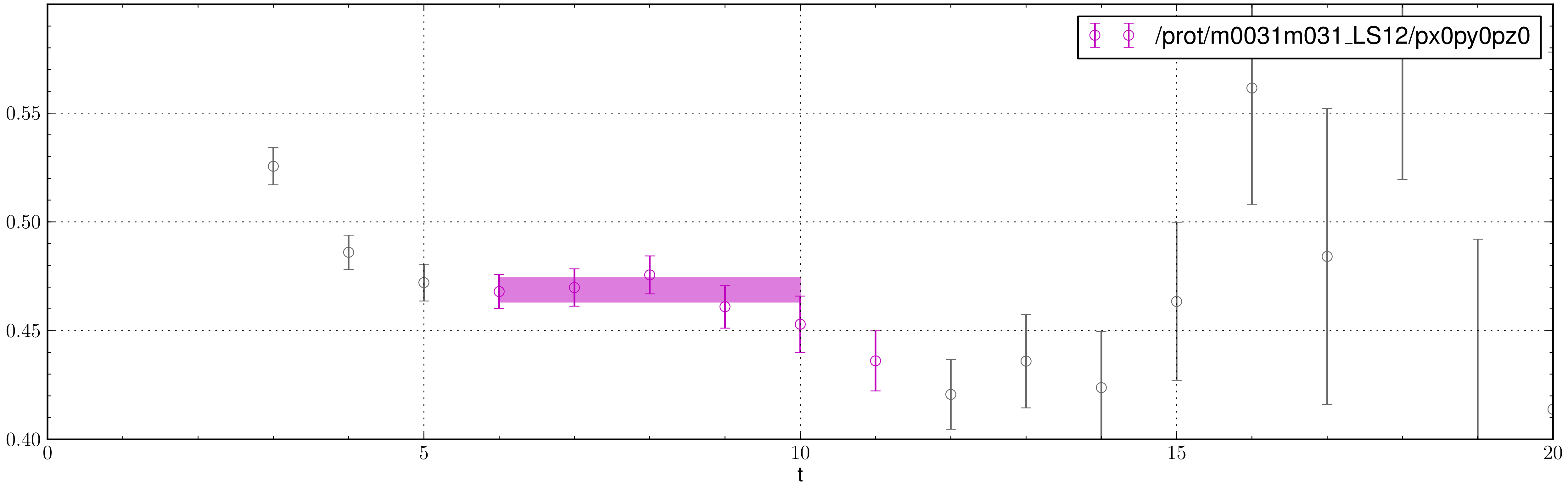
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau)y(t)$$

take correlation function with
one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=6 and t=11 also good

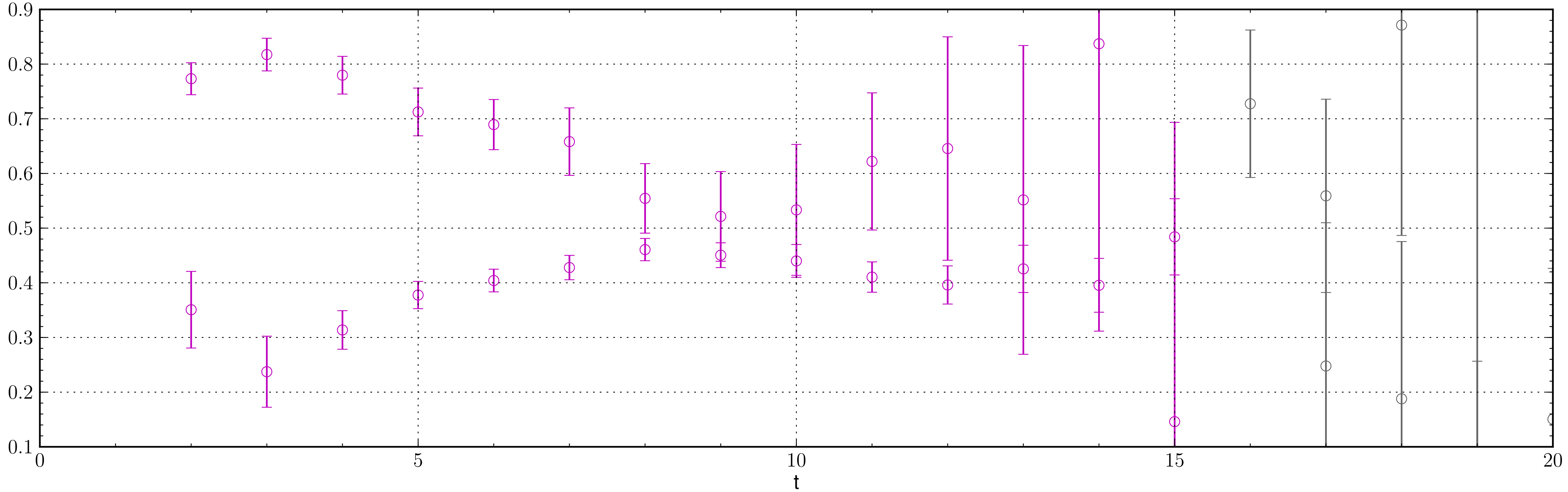
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

take correlation function with
one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=2 and t=15 too aggressive

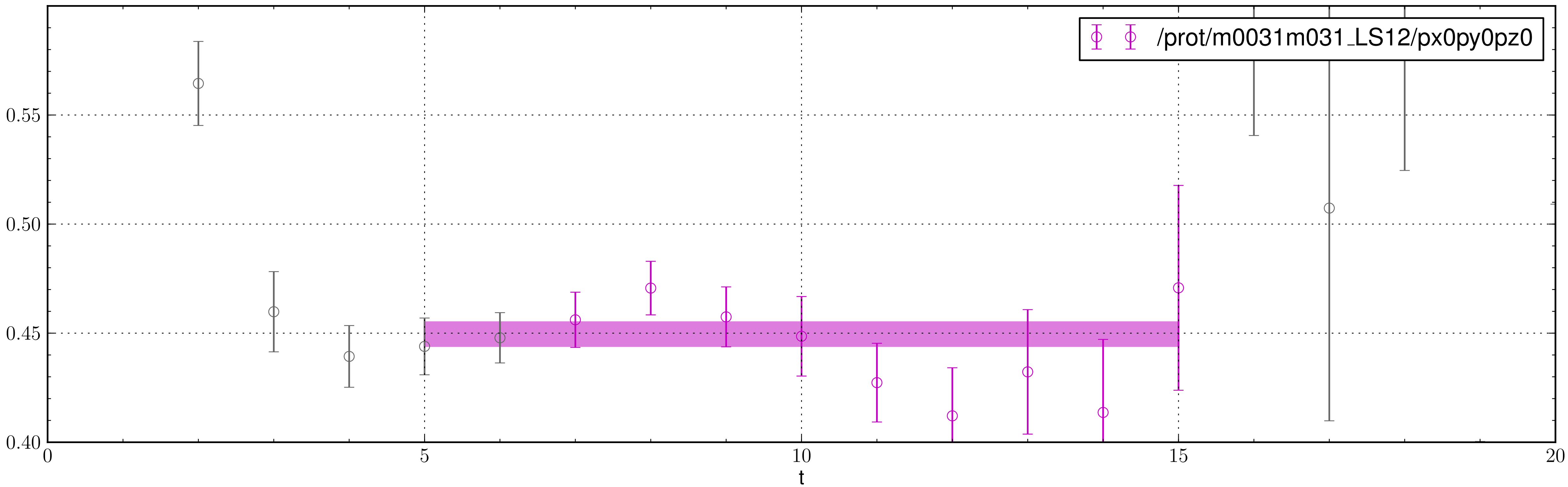
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau)y(t)$$

take correlation function with one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



between t=7 and t=15 also good

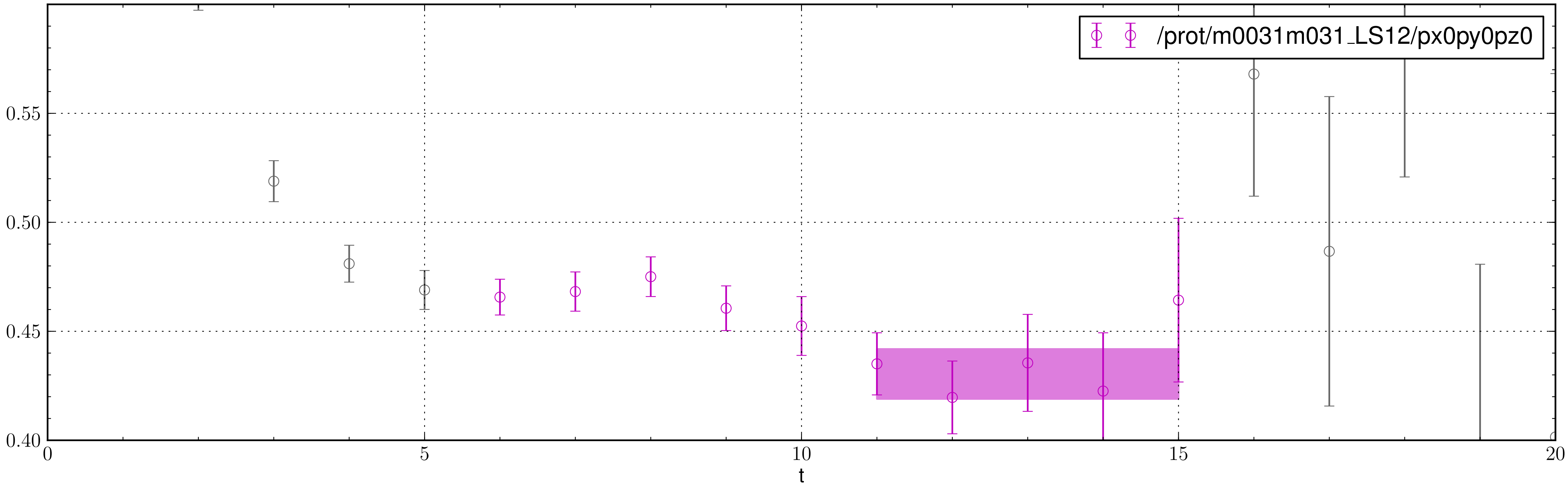
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau)y(t)$$

take correlation function with one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$

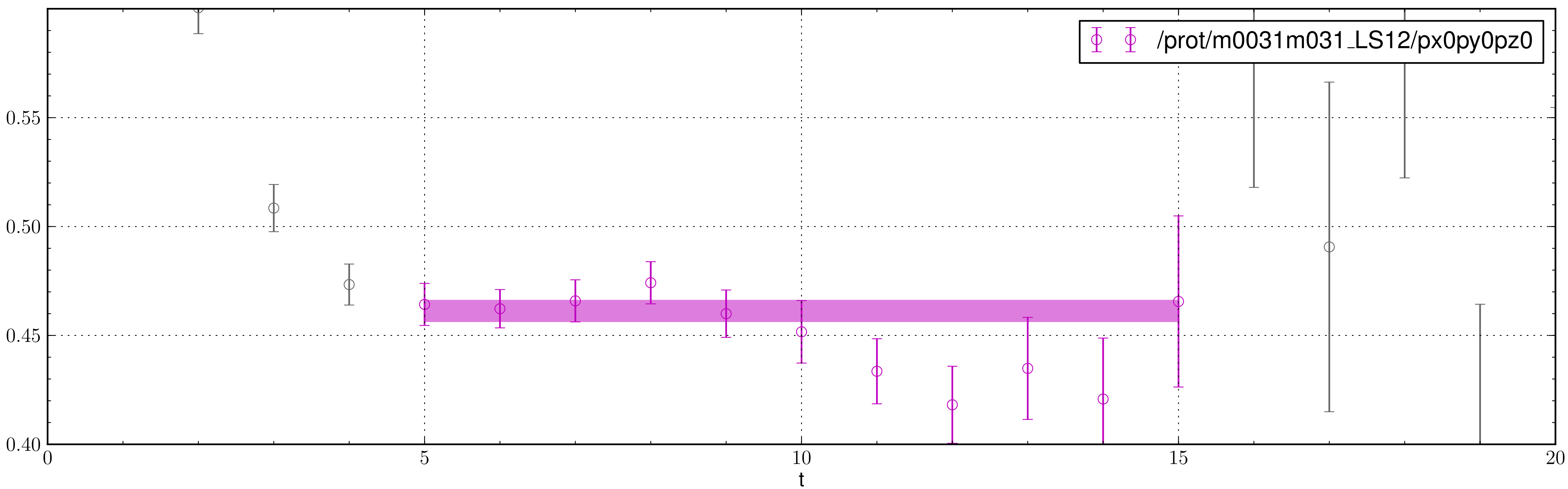


between t=7 and t=15

different range of fit ok

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466



exponential decay of signal-to-noise for baryons challenging

- is late-time dip of effective mass true ground state value?
- lack of positive-definite correlation functions allow "false plateaus"
- results much more sensitive to choices of fit-range, MP-range
- want algorithm to weight all possible reasonable choices

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

The algorithm I currently use
(systematics are not uniquely defined)

- pick minimum Δt_{MP}

$$\hat{M}(\tau) = \left[\sum_{t_0}^{t_0+\Delta t} y(t+\tau) y^T(t) \right]^{-1}$$

- pick minimum Δt_{plat} in exp fit $t_f = t_i + \Delta t_{plat}$

  Δt_{plat} choice guided by excited state masses

- pick minimum and maximum t you want to consider (correlator dependent)

- loop over independently chosen $t_0, \Delta t_{MP}, t_i, \Delta t_{plat}$

Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

The algorithm I currently use
(systematics are not uniquely defined)

- specific for baryons, for each fit, pick a weight factor

$$\omega_i = \frac{Q_i}{\sigma_i^2} \qquad Q = \int_{\chi^2_{min}}^{\infty} d\chi^2 \mathcal{P}(\chi^2, dof) \qquad Q \in [0, 1]$$

$\sigma_i =$ statistical uncertainty for a given fit

this choice allows for late time fluctuations being real, but suppresses them by their larger statistical uncertainty

$$\bar{m} = \frac{\sum_i w_i m_i}{\sum_j w_j}$$

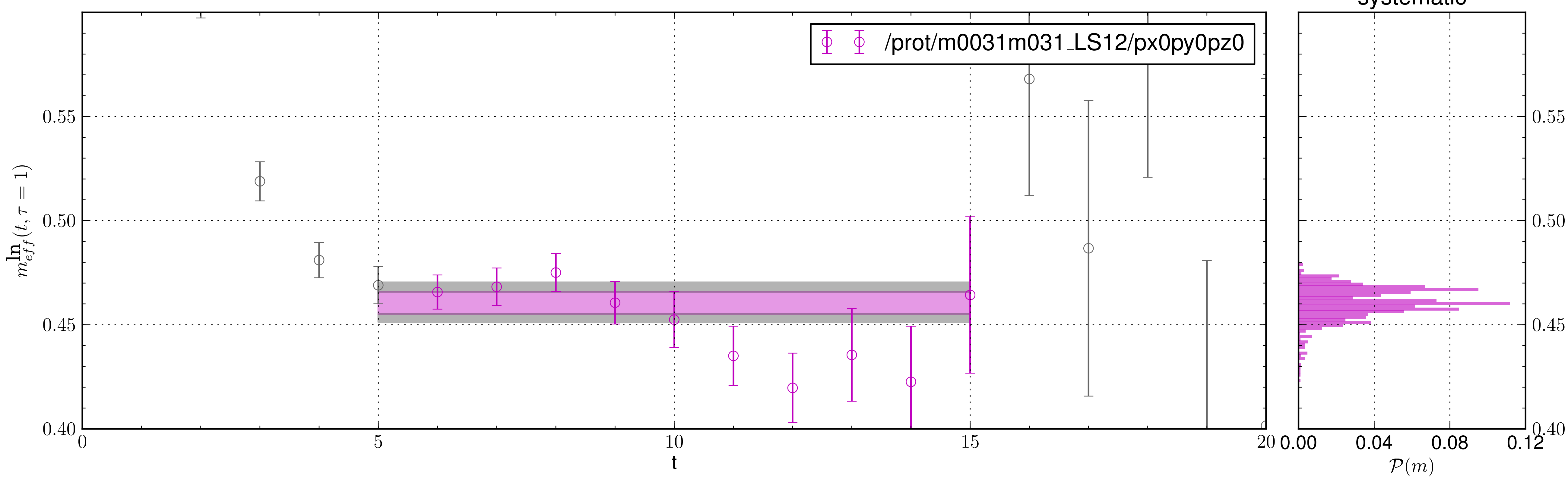
Matrix Prony

arXiv:1301.1114
arXiv:0905.0466

$$y(t+\tau) = \hat{T}(\tau) y(t)$$

take correlation function with one source and two sinks

$$y(t) = \begin{pmatrix} y_{SP}(t) \\ y_{SS}(t) \end{pmatrix}$$



16% and 84% quantiles chosen for systematic uncertainty

inner band statistical outer band stat+sys added in quadrature

$$\mathcal{P}(m) = \text{ mass probability distribution function}$$

$$\mathcal{P}(m_i) = \frac{w_i}{\sum_j w_j}$$

Questions

- Can the noise of our numerical results be used to rigorously chose a number of singular values in the Hankel Matrix/GPOF methods?

- The growth of computing power and algorithms means that TODAY is the beginning of a renaissance in nuclear physics where these exciting things are just becoming possible!

Thank You