Exponential Time Series Analysis in Lattice QCD

George T. Fleming

Yale University

QCDNA IV 02 May 2007

Outline

- Definition of the problem
- Estimating ground state energies
  - Limitations due to variance growth
- Estimating excited state energies
  - Bayesian Methods
    - Constrained Fitting
    - Maximum Entropy Method
  - Black Box Methods
  - Variational Methods

Exponential Time Series (I)

- What we heard yesterday:
  - Generating ensembles of gauge fields $\{U\}$.
  - Solving linear systems $[D(U) + m]\,\chi = \delta(x - x_0)$ for propagation of quarks from a single point x_0 to anywhere x.
- Using these pieces we construct correlation functions $C(t)$ to compute physical observables. A typical model function:
  $C_{ij}(\mathbf{p}; t, t_0) = \sum_{k=1}^{K} Z^*_{ik}(\mathbf{p}, t_0)\ Z_{kj}(\mathbf{p}, t)\ \exp\left[-E_k(\mathbf{p})(t - t_0)\right]$
  - i, j label arrangements of quarks and gluons at times t_0, t, same irrep of $O_h \rtimes T^3$.
  - $Z_{kj}(\mathbf{p}, t)$ are model parameters that indicate how likely an arrangement j will "look like" a physical state k at time t. Should be independent of t on average.
  - Energies $E_k(\mathbf{p})$ can be compared to real experimental data.
  - Energies are ordered: $0 \leq E_1(\mathbf{p}) \leq E_2(\mathbf{p}) \leq \cdots$
  - Easy to Extract lowest energy $E_1(\mathbf{p})$ at large times: $t \gg t_0$

Exponential Time Series (II)

- A meson correlation function in some more detail:
$$C_{ij}(\mathbf{p}; t, t_0) = \sum_{\mathbf{x}} e^{i\mathbf{p}\cdot\mathbf{x}} \mathrm{Tr}\left[\Gamma_i M^{-1}(\mathbf{x}_0, t_0; \mathbf{x}, t)\Gamma_j M^{-1}(\mathbf{x}, t; \mathbf{x}_0, t_0)\right]$$
- Solving the linear system for a single site $\mathbf{x}_0$ is sufficient to compute the correlation function.
- Fourier transform at the source $\mathbf{x}_0$ would be prohibitively expensive. Physics/symmetry saves the day.
- Sometimes there are disconnected contributions:
$$C_{ij}^{\mathrm{disc}}(\mathbf{p}; t, t_0) = \sum_{\mathbf{x}} e^{i\mathbf{p}\cdot\mathbf{x}} \mathrm{Tr}\left[\Gamma_i M^{-1}(\mathbf{x}_0, t_0; \mathbf{x}_0, t_0)\right] \mathrm{Tr}\left[\Gamma_j M^{-1}(\mathbf{x}, t; \mathbf{x}, t)\right]$$
- Really expensive. Trace estimation needs hundreds or thousands of linear solves. (A. Stathopoulos, J. Osborn)

Exponential Time Series (III)

- Returning to our typical model function:
  $$C_{ij}(\mathbf{p}; t, t_0) = \sum_{k=1}^{K} Z^*_{ik}(\mathbf{p})\ Z_{kj}(\mathbf{p})\ \exp\left[-E_k(\mathbf{p})(t - t_0)\right]$$
- Treating i, j, k as matrix indices gives the form: (A. Lichtl)
  $$\mathbf{C}(\mathbf{p}; t, t_0) = \mathbf{Z}^\dagger(\mathbf{p})\ \mathbf{\Lambda}(\mathbf{p})^{t-t_0}\ \mathbf{Z}(\mathbf{p})$$
- A matrix element of $\mathbf{C}$ gives a $T \times K$ Vandermonde system:
  $$\mathbf{y} = \mathbf{\Phi a} : \quad y_t = C_{ij}(t), \alpha_k = e^{-E_k}, \phi_{tk} = \alpha_k^t, a_k = Z^*_{ik} Z_{kj} e^{-E_k t_0}$$
- The Vandermonde system is useful for black box methods. (H.-W. Lin)
- Vandermonde form useful starting point for implementing variable projections: $\boldsymbol{\alpha}$ nonlinear, $\mathbf{a}$ linear. (V. Pereyra)
- Sometimes, we'd rather work with the spectral function even if ill-posed (P. Petreczky)
  $$C(t) = \int_0^\infty e^{-\omega t} \sigma(\omega)$$

Estimating ground state energies

- For $t \gg t_0$, the correlator projects onto the ground state:
$$C_{ij}(t) \approx Z_{i1}^* Z_{1j} \exp[-E_1(t - t_0)]$$
- A non-linear least squares fit over $t_{\max} \geq t \geq t_{\min} \gg t_0$ is most common method. Data is highly correlated, so covariance matrices must be used in χ^2 minimization.
$$\chi^2 = [\mathbf{y} - \mathbf{\Phi}(\alpha)\mathbf{a}]^\dagger \mathcal{C}^{-1} [\mathbf{y} - \mathbf{\Phi}(\alpha)\mathbf{a}]$$
- The simplest black box method, called *effective mass*, gives a non-optimal estimate:
$$E_1 = -\log\left[C_{ij}(t+1)/C_{ij}(t)\right]$$

Limitations due to variance growth

- Quick review of hadron spectrum: $m_\pi \ll m_\rho \lesssim m_N$
- Dominant variance of the correlation function

  $$\mathrm{Var}\ C_{ij}(t, t_0) =$$

  $$\mathrm{Tr}\left[M^{\dagger -1}(t, t_0)\Gamma_j^\dagger M^{\dagger -1}(t_0, 0)\Gamma_i^\dagger \Gamma_i M^{-1}(t_0, t)\Gamma_j M^{-1}(t, t_0)\right]$$

- $\Gamma_i^\dagger \Gamma_i$ contracts quarks in irrep of the vacuum. Lowest energy state in that channel is two pions at rest.

  $$\mathrm{Var}\, C_{ii}(\mathbf{p}; t, t_0) \approx e^{-2m_\pi(t-t_0)}, \quad t \gg t_0$$

- Signal-to-noise falls exponentially for all correlators except pion at rest:

  $$\mathrm{SNR}\ \approx\ \frac{\exp[-E_1(t - t_0)]}{\sqrt{\exp[-2m_\pi(t - t_0)]}} = \exp[-(E_1 - m_\pi)(t - t_0)]$$

Estimating Excited State Energies

- Least squares minimization can be notoriously difficult
  - Very sensitive to initial guesses. Variable projections (VARPRO) needs guesses only for nonlinear parameters and it does improve fit stability.
  - Can estimate E_1 on $t_{\mathrm{min}1} \leq t \leq t_{\mathrm{max}}$, freezing it and then estimating E_2 on $t_{\mathrm{min}2} \leq t \leq t_{\mathrm{min}1}$. Can be very sensitive to choice of $t_{\mathrm{min}1}$.
- Three general approaches are being pursued to address this problem:
  - Bayesian Methods
  - Black Box Methods
  - Variational Methods

Bayesian Methods

- Augment likelihood functions with prior probabilities.
- Priors should encode *only* our understanding of the model without any reference to data. Otherwise, predictability is lost.
- Constrained Fitting
  - A popular method for choosing Gaussian-distributed priors. Prevents minimizer from running away along flat directions. Difficult to understand how to relate these Gaussian parameters to a theoretical calculation.
  - VARPRO may be useful in the constrained fitting context in that priors need to be given only for non-linear parameters.
- Maximum Entropy Method
  - Uses the Shannon-James entropy to ensure that there is a price to pay when maximizing the likelihood function for a spectral function solution that has too many bumps and wiggles.
  - The computed spectral function approaches the spectral function of perturbative QCD at high frequencies: $\omega \rightarrow \infty$.

Black Box Methods (I)

- The effective mass is a simple example. Starting from the $T = 2, K = 1$ Vandermonde system

$$\begin{pmatrix} y_t \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} \alpha_1^t \\ \alpha_1^{t+1} \end{pmatrix} (a_1) \quad \Rightarrow \quad \alpha_1 = \frac{y_{t+1}}{y_t}, \quad a_1 = \frac{y_t}{\alpha_1^t}$$

- The effective mass has been generalized to find two mass estimates. Starting from the $T = 4, K = 2$ Vandermonde system

$$\begin{pmatrix} y_t \\ y_{t+1} \\ y_{t+2} \\ y_{t+3} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \\ \alpha_1^2 & \alpha_2^2 \\ \alpha_1^3 & \alpha_2^3 \end{pmatrix} \begin{pmatrix} a_1 \alpha_1^t \\ a_2 \alpha_2^t \end{pmatrix}.$$

Black Box Methods (II)

- Three quantities are computed from the data:

$$A = y_{t+1}^2 - y_t y_{t+2}, \; B = y_t y_{t+3} - y_{t+1} y_{t+2}, \; C = y_{t+2}^2 - y_{t+1} y_{t+3}$$

- The two solutions for the non-linear parameters α_k come from the familiar quadratic equation

$$\alpha_{1,2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

- It is plausible that there are algebraic solutions to the three-state and four-state effective mass problems when properly reduced to cubic and quartic equations after separation of variables.
- (H.-W. Lin)

Variational Methods

- The variational method is the generalization of effective mass to correlator matrices.
- For a $N \times N$ correlator matrix $\mathbf{C}$ the N effective masses are solutions to the generalized eigenvalue problem:

$$\mathbf{C}(t, t_0)\mathbf{z} = \lambda \mathbf{C}(t+1, t_0)\mathbf{z}$$

- The variational method involves only two time slices. A generalized method to extract masses from multiple time slices would involve solving a higher order matrix polynomial equation, *e.g.*

$$\mathbf{C}(t, t_0)\mathbf{z} - 2\lambda \mathbf{C}(t+1, t_0)\mathbf{z} + \lambda^2 \mathbf{C}(t+2, t_0)\mathbf{z} = 0$$

- There are twice as many solutions, but not all of them need to be physical. Open question: Does the benefit of having more solutions outweigh the burden of eliminating unphysical ones?
- (A. Lichtl)

Conclusions

- Lattice QCD calculations are getting so advanced that it is no longer sufficient to simply measure the easiest things, e.g., ground state masses of hadrons.
- Many possible directions to try:
  - Bayesian Methods
    - Constrained Fitting
    - Maximum Entropy Method
  - Black Box Methods
  - Variational Methods
- Many possible pitfalls!
  - Biased Priors
  - Unphysical Solutions
  - Sub-optimal Solutions