Three t GEVP

Dan Hoying

September 24, 2019

Outline

Physics Motivation: Excited States

- Excited state contamination on the lattice is intrinsic
- $K \to \pi\pi$: $\pi\pi$ phase shift puzzle: solution: Excited States
- PCAC Puzzle (Y.C. Jang): solution: Excited States
- Had Spec, matrix elements, etc.
- The dual problem: late time noise
- Excited states are a common part of most error budgets.

GEVP: Solving Excited States

The GEVP allows us to project our energy eigenspace (from lattice Hamiltonian) to the operator basis.
Define positive definite, hermitian GEVP matrix $C(t)$

$$
\begin{aligned}
C_{ij}(t) &= \left\langle 0|O_i(t)O_j^\dagger(0)|0\right\rangle \\
&= \sum_n^{\infty} \left\langle 0|O_i(t)\,|n\rangle\,\langle n|\,O_j^\dagger(0)|0\right\rangle \\
&= \sum_n^{\infty} e^{-E_n t}\psi_{in}\psi_{jn}^*
\end{aligned}
$$

ψ_{in} is overlap factor, E_n is the lattice energy desired.

Warnings

1. This is a math talk.
2. I have no lattice results.
3. I have numerical results from a toy model, but I won't share them now.
4. This work is very preliminary, unpublished, and certainly not peer reviewed.
5. No new $\pi\pi \to \pi\pi$ results, sorry. (see my Lattice 2019 Wuhan talk).

 You've been warned! 

Outline

Generalized Eigenvalue Problem

Problem: find E_n given lattice data. We can restate that problem using the Generalized Eigenvalue Problem (GEVP).

$$
\begin{gathered}
C(t)\left|v(t,t_0)\right\rangle = \lambda(t,t_0)C(t_0)\left|v(t,t_0)\right\rangle \\
\Rightarrow\ (e.g.)\ C(t')\left|v(t',t_0)\right\rangle = \lambda(t',t_0)C(t_0)\left|v(t',t_0)\right\rangle \\
t_0 < t' < t
\end{gathered}
$$

Excited State Contamination (Systematic Error)

Split C into two parts: $C = C^0 + C^1$. Define as follows:

$$C^0(t) = \sum_{n=0}^{N} e^{-E_n t} \psi_{in} \psi^*_{jn}$$

$$C^1(t) = \sum_{n=N+1}^{\infty} e^{-E_n t} \psi_{in} \psi^*_{jn}$$

- C^1 is the perturbing matrix, or the perturbation.
- (Naive) power counting: use parameter α to count powers of C^1:
- Count exponents. (e.g.)
  $O(\alpha^2) = \langle v_n^1(t, t_0) | C^1(t) | v_n^0 \rangle \to 1 + 1 + 0 = 2$
- (Any item without a superscript is something we can find from lattice data by solving the appropriate GEVP.)
- (Subscripts denote energy. No subscript means the energy we want to solve for at the moment.)

First Big Question

First Big Question:

- How does this systematic error behave asymptotically as a function of $t_0, t(, t+1)$?

Answer: Alpha Collab Theorem [1]

Systematic error in E_n is $O(e^{-(E_{N+1}-E_n)t})$
Assumptions:

- Single contaminating state: We are at late enough t, t_0 that we can neglect $E_n, n > N + 1$.
- GEVP derivative: We must take a derivative with respect to t to find the effective mass.
- (Only if we need the asymptotic error bound.)
- GEVP derivative $\Rightarrow$ We need three (!) time separations worth of lattice data: $t_0, t, t + 1$.

I refer to this theorem as the alpha theorem, the method as the alpha method, and its assumptions as the alpha assumptions.

Time Dependence Saga

- Initial GEVP analysis only looked at t' dependence (fixed t_0).
- Alpha Collab extended analysis to two independently varying time slices t_0, t' (fixed $t = t' + 1$).
- Thus, the natural extension of this line of reasoning: How does the systematic error depend on three independent varying time slices t_0, t', t?

Second Big Question: What can we do in general with three time slices worth of lattice data?

Outline

Examining the Perturbation

Start with standard GEVP. Expand E_n to second order in α. We find this mess:

$$
\begin{aligned}
\frac{\lambda}{\lambda_n^0} = 1 &+ \frac{\langle v^0 | C^1(t) | v^0 \rangle}{\langle v^0 | C^0(t) | v^0 \rangle} - \frac{\langle v^0 | C^1(t_0) | v^0 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} + \left(\frac{\langle v^0 | C^1(t_0) | v^0 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} \right)^2 \\
&- \frac{\langle v^0 | C^1(t_0) | v^0 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} \frac{\langle v^0 | C^1(t) | v^0 \rangle}{\langle v^0 | C^0(t) | v^0 \rangle} + \\
&- \frac{\langle v^0 | C^1(t_0) | v^1 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} - \frac{\langle v^1 | C^1(t_0) | v^0 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} + \frac{\langle v^0 | C^1(t) | v^1 \rangle}{\langle v^0 | C^0(t) | v^0 \rangle} + \frac{\langle v^1 | C^1(t) | v^0 \rangle}{\langle v^0 | C^0(t) | v^0 \rangle} + \\
&+ \frac{\langle v^1 | C^0(t) | v^1 \rangle}{\langle v^0 | C^0(t) | v^0 \rangle} - \frac{\langle v^1 | C^0(t_0) | v^1 \rangle}{\langle v^0 | C^0(t_0) | v^0 \rangle} + O(\alpha^3)
\end{aligned}
$$

Error Types

Split the errors into two types:

1. Diagonal Error: Terms with brackets which involve only v^0.
2. Eigenvector (Off-Diagonal) Error: Terms with brackets with $v^r, r \neq 0$.

The diagonal nomenclature come from the matrix elements of C^1 in the v^0 basis.

This distinction is, however, only meaningful if we can find some way to actually separate the two contributions.

Computing the Off-Diagonal Error

- We must separate this error somehow from the diagonal contributions.
- How to separate? Two facts: Generalized orthogonality relation (GOR; eq. (1)), and the time independence of v_n^0.
- A useful side observation: since the relation vanishes, it must vanish order by order in the perturbation series.

$$
\begin{aligned}
0 &= \langle v_n | C(t) | v_m \rangle & (1)\\
\Rightarrow 0 &= \left\langle v_n^0 \middle| C^0(t) \middle| v_m^0 \right\rangle \\
\Rightarrow 0 &= \left\langle v_n^0 \middle| C^1(t) \middle| v_m^0 \right\rangle + \left\langle v_n^1(t,t_0) \middle| C^0(t) \middle| v_m^0 \right\rangle + \left\langle v_n^0 \middle| C^0(t) \middle| v_m^1(t,t_0) \right\rangle
\end{aligned}
$$

$(etc.)$

Exploiting the GOR

- We can exploit the GOR on three time slices.
- We find a pure systematic: $\langle v_n(t, t')|C(t)|v_m(t, t_0)\rangle$ (e.g.)
- This object is only non-zero if there is off-diagonal error.
- We thus gain insight into the off-diagonal perturbation series.

Computing the Entire Off-Diagonal Series (Prospects)

I form the following construction (an $O(\alpha)$ systematic)

$$\left\langle v_m(t,t')\middle|C(t_0)\middle|v_n(t,t')\right\rangle = \left\langle v_m^0\middle|C^0(t_0)\middle|v_n^1(t,t')\right\rangle + \\ + \left\langle v_m^1(t,t')\middle|C^0(t_0)\middle|v_n^0\right\rangle + \left\langle v_m^0\middle|C^1(t_0)\middle|v_n^0\right\rangle + O(\alpha^2)$$

(Analogous equations exist for $C(t), C(t')$.) Expand v^1:

$$\left|v_n^1(t,t_0)\right\rangle = \sum_{m\neq n}\left|v_m^0\right\rangle \frac{1}{\lambda_m^0-\lambda_n^0}(\lambda_n^0(t,t_0)\frac{\color{red}{\left\langle v_m^0\middle|C^1(t_0)\middle|v_n^0\right\rangle}}{\left\langle v_m^0\middle|C^0(t_0)\middle|v_m^0\right\rangle}+ \\ - \frac{\color{blue}{\left\langle v_m^0\middle|C^1(t)\middle|v_n^0\right\rangle}}{\left\langle v_m^0\middle|C^0(t_0)\middle|v_m^0\right\rangle})$$

(unknowns to solve for are highlighted)

Systems of Equations

- We thus obtain have $3 * \binom{N}{2}$ equations and the same number of unknowns.
- $\Rightarrow$ We can solve for $\left\langle v_m^0 | C^1 | v_n^0 \right\rangle$ to $O(\alpha^2)$.
- We thus (nominally) have the pieces of the perturbation series to $O(\alpha^2)$.
- When we compute the series with imprecise pieces, we will make all mistakes at a higher order. If we keep track of these (small) mistakes, we can sum the series indefinitely!
- Do we have all the (imprecise) pieces of the Off-Diagonal series yet? No.

Summary of Remaining Steps

Due to time constraints I'll summarize the remaining Off-Diagonal steps. General scheme: Follow analogous steps of Rayleigh-Schrödinger (RS) Perturbation Theory.

- Impose normalization on the eigenvectors, but with a $C(t)$ in the middle (e.g. $\left\langle v_n^1(t, t_0) | v_n^0 \right\rangle = 0 \rightarrow \left\langle v_n^1(t, t_0) | C(t) | v_n^0 \right\rangle = 0$.
- Use available software for the RS case, and modify accordingly (still to do, but the pattern is similar, and the differences look straightforward to program).
- Remaining piece of the perturbation series: $\left\langle v_n^0 | C^1(t) | v_n^0 \right\rangle$. It turns out these can be found via

$$
\begin{aligned}
O(\alpha^2) = & \left\langle v_m(t, t') | C(t_0) | v_n(t, t') \right\rangle + \\
& - \left\langle v_m(t', t_0) | C(t_0) | v_n(t, t') \right\rangle - \left\langle v_m(t, t') | C(t_0) | v_n(t', t_0) \right\rangle
\end{aligned}
$$

Aside: The Commutator

Bracket - C^1 (Off-Diagonal) Connection: Commutator

- If we obtain v_n^0, it turns out we can find C^1 (and, of course, vice-versa).
- To see how, take GEVP$\rightarrow$ EVP: $C^{-1}(t_0)C(t)\,|v(t,t_0)\rangle = \lambda(t,t_0)\,|v(t,t_0)\rangle$
- $\Rightarrow \left[C^{-1}(t_0)C(t'), C^{-1}(t')C(t)\right] = 0$ if $|v_n\rangle = \left|v_n^0\right\rangle$. (sim. diag.)
- Rotate to $\left|v_n^0\right\rangle$ basis; solve for (off-diagonal) C^1.

Diagonal Perturbations

- Defined as perturbations, which, on three time slices, do not appreciably change the eigenbasis.
- After removing Off-diagonal perturbations in v_n and λ_n, these remain.
- Our system of equations is under-determined (by one variable).
- Without model constraints, we can only hope to bound the remaining systematic error.
- (In progress, nearly complete.)

Diagonal Perturbations: Results

- I've obtained an upper bound on E_n which is (likely) tight on three time slices (without model assumptions) (optimization still in progress).
- Lower bound on E has a (mathematical) critical point, and is slightly more involved to calculate (optimization still in progress).
- However, (asymptotic) bounds on E are now explicit (no unknown coefficient)!
- $\Rightarrow$ We can explicitly bound the systematic error from excited states.

Outline

Toy Model: 3x3 GEVP +1 Extra State

(Assume off-diagonal perturbations are negligible.)
Procedure:

1. Generate the GEVP from hand picked energies and (pseudo)-random $U(1)$ overlap factors.
2. Subtract the energy via a guess from the alpha method
3. Minimize a quadratic cost function (sum of squares of GEVP's) over our known contamination amplitude (a) and energy difference (Δ).
4. Shift the results by the amount we subtracted in our guess.

Toy Model: Results (1/2)

- The cost function is not convex (no guarantee that we will converge to a global minimum).
- However, for the default starting point, we obtain (empirically), an improved energy
- The new energy is always (again, empirically) an improvement, but the percent improvement varies a lot.
- (mostly, it seems, because the absolute size of the improvement does not vary much)

Toy Model: Results (2/2)

- Varying the energy does not affect the results.
- Different starting points for the minimizer, though, yield very inconsistent results.
- (a, Δ vary a lot)
- (Rarely), a starting point can yield a worse energy!

Toy Model: Future Directions

Questions:

1. Can the cost function be made convex?
2. Or, can we find a bounded domain where the cost function is convex?
3. Off-diagonal errors vary in their directional effect on the energy, can we still include them in this cost function?
4. Is this minimization procedure complementary (or can we make it complementary) with our other error reduction strategies?

Outline

Conclusions (1/2)

1. I extend the GEVP systematic error analysis to include the third time slice of lattice data.
2. I illustrate a (plausible) procedure for removing off-diagonal errors (needs testing; especially for convergence).
3. I indicate bounds then exist for the remaining diagonal systematic error.
4. My results from a toy model show promise. Can we make improvements in this direction?

Open Questions (2/2)

- A challenge: Code up the (off-diagonal) perturbation series, including some way to track higher order mistakes made in the sum procedure.
- What should we expect about the relative sizes of the off-diagonal and diagonal errors?
- What time slice choices will yield the smallest errors?

Thanks

Thanks!

References

Benoit Blossier et al. "On the generalized eigenvalue method for energies and matrix elements in lattice field theory". In: *JHEP* 04 (2009), p. 094. DOI: 10.1088/1126-6708/2009/04/094. arXiv: 0902.1265 [hep-lat].