Bayesian Interpretation of Backus Gilbert methods

with L Del Debbio, M Panero, N Tantalo

Alessandro Lupo

CERN Lattice ☕

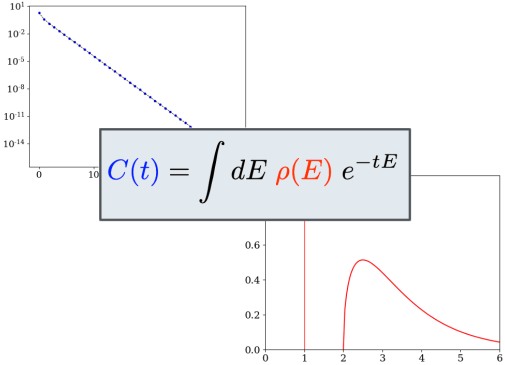
June 4, 2024

The inverse problem

- We are concerned with computing the spectral density $\rho(E)$ associated to a lattice correlator $C(t)$
- Ill-posed in presence of a finite set of noisy data
- There are ways to regularise the problem. Different assumptions, one way to express the result

$$\rho_\sigma(E) = \sum_t g_t(\sigma; E)\, C(t)$$

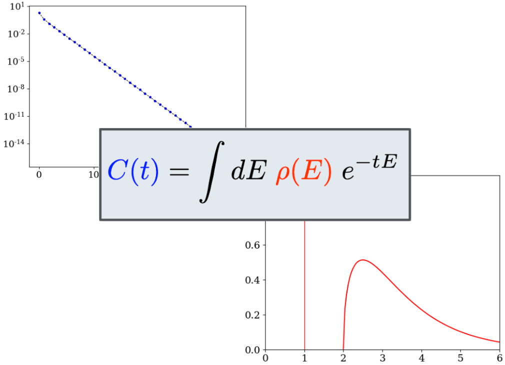
$$\rho(E) = \lim_{\sigma \to 0} \rho_\sigma(E)$$



$$C(t) = \int dE\; \rho(E)\; e^{-tE}$$

Difficulties

- Finite set of measurements vs function with potentially continuous support
- Target function is a distribution
- Information is suppressed by $\exp(-tE)$
- We work we data that is affected by errors



$$C(t) = \int dE \; \rho(E) \; e^{-tE}$$

Smearing

- Smearing must be introduced to have a function that is smooth even in a finite volume

$$\rho_\sigma(\omega) = \int dE\, \mathcal{S}_\sigma(E,\omega)\, \rho(E)$$

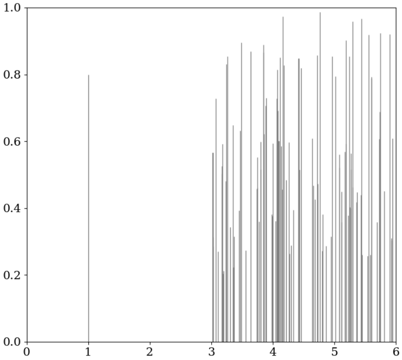
- Linear combinations of correlators automatically produce a smeared SD

$$\begin{aligned}\rho_\sigma(\omega) &= \sum_t g_t(\sigma;\omega)\; C(t) \\ &= \sum_t g_t(\sigma;\omega) \int dE\; e^{-tE} \rho(E)\end{aligned}$$

- We can now take the infinite volume limit

$$\lim_{L\to\infty} \rho_L(E) = 🚫$$

$$\lim_{\sigma\to 0} \lim_{L\to\infty} \rho_L(\sigma;E) = \rho(E)$$



Bayesian Inference with Gaussian Processes

- Aim for a probability distribution over a functional space of possible spectral densities

- Consider the stochastic field $\mathcal{R}(E)$ Gaussian-distributed around the prior value $\rho^{\text{prior}}(E)$ with covariance $\mathcal{K}^{\text{prior}}(E, E')$.

$$\mathcal{GP}\left(\rho^{\text{prior}}(E), \mathcal{K}^{\text{prior}}(E, E')\right)$$

- Similarly, assume that observational noise is Gaussian: $\eta(t)$

$$\mathbb{G}\left(\eta, \text{Cov}_d\right) = \exp\left(-\frac{1}{2}\vec{\eta}^T \ \text{Cov}_d^{-1} \ \vec{\eta}\right)$$

- The stochastic variable associated to the correlator, $\mathcal{C}$, is related to $\mathcal{R}$ and η via

$$\mathcal{C}(t) = \int dE \, e^{-tE}\mathcal{R}(E) + \eta(t)$$

- Incomplete list of references:
  FASTSUM collab. , Valentine, Sambridge 19 , Horak, Pawlowski, Rodríguez-Quintero, Turnwald, Urban 21
  Del Debbio, Giani, Wilson 21

- The joint, posterior distribution is again Gaussian, centred around ρ^{post} centre and variance:

$$\rho^{\text{post}}(\omega) = \rho^{\text{prior}}(\omega) + \sum_{t=1}^{t_{\max}} g_t^{\text{GP}}(\omega) \left(C(t) - \int_0^\infty dE \, e^{-tE} \rho^{\text{prior}}(E) \right)$$

$$\mathcal{K}^{\text{post}}(\omega,\omega) = \left(\mathcal{K}^{\text{prior}}(\omega,\omega) - \sum_{t=1}^{t_{\max}} g_t^{\text{GP}}(\omega) \, f_t^{\,\text{GP}}(\omega) \right)$$

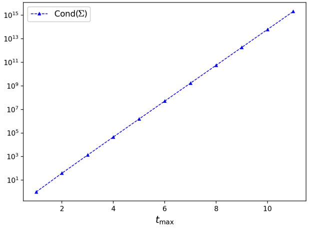
- The coefficients can be written as

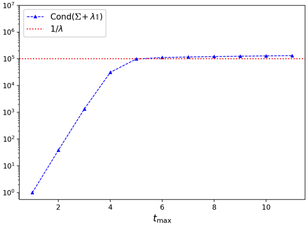
$$\vec{g}^{\text{GP}}(\omega) = (\Sigma^{GP} + \lambda \text{Cov}_{\text{d}})^{-1} \, \vec{f}^{\,\text{GP}}$$

- With the following ingredients:

$$\Sigma^{\text{GP}}{}_{tr} = \int dE_1 \int dE_2 \; e^{-tE_1} \, \mathcal{K}^{\text{prior}}(E_1, E_2) \, e^{-rE_2} \quad \text{ill cond}$$

$$f_t^{\,\text{GP}}(\omega) = \int dE \; \mathcal{K}^{\text{prior}}(\omega, E) \, e^{-tE}$$





Backus-Gilbert methods: ideal world

- {Hansen Lupo Tantalo 19} Choose an appropriate smearing kernel such that when $\sigma \to 0$ we recover $S_\sigma(E,\omega) \to \delta(E-\omega)$

- We need to find the set of coefficients spanning $S_\sigma(E,\omega)$:

$$\sum_{\tau=1}^{\infty} g_\tau^{\text{true}}(\sigma, E)\, e^{-a\tau\omega} = S_\sigma(E,\omega)$$

- We can find the coefficients by minimising

$$A[g] = \int_{E_0}^{\infty} dE\; e^{\alpha E} \left| \sum_{\tau=1}^{\infty} g_\tau(\sigma, E)\, e^{-a\tau\omega} - S_\sigma(E,\omega) \right|^2$$

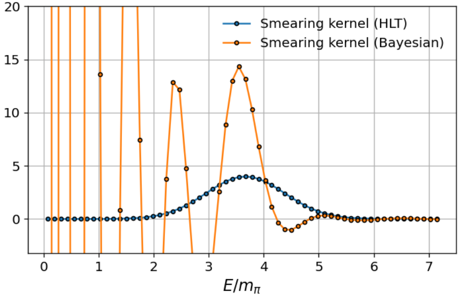
- Without errors on $C(t)$ and infinitely many points, this is the solution.

Backus-Gilbert methods: less ideal world

- In reality, the correlator is known at a finite number of points. This translates into a systematic error in the reconstructed kernel and therefore in the reconstructed SD

$$\sum_{\tau=1}^{\tau_{\max}} g_\tau(\sigma, E)\, C(a\tau) = \rho_\sigma(E) + r(\tau_{\max}, \sigma; E)$$

- The sum truncated to $\tau_{\max}$ is however well-defined and define unambiguously a given smearing kernel

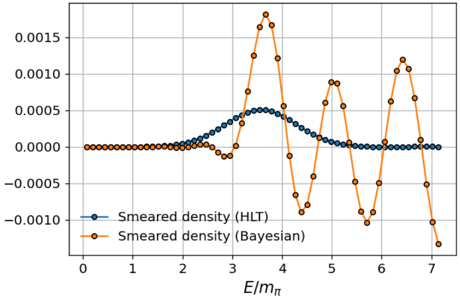


Backus-Gilbert methods: less ideal world

- In reality, the correlator is known at a finite number of points. This translates into a systematic error in the reconstructed kernel and therefore in the reconstructed SD

$$\sum_{\tau=1}^{\tau_{\max}} g_\tau(\sigma, E) C(a\tau) = \rho_\sigma(E) + r(\tau_{\max}, \sigma; E)$$

- The sum truncated to $\tau_{\max}$ is however well-defined and define unambiguously a given smearing kernel



Backus-Gilbert methods: real world

- The main complication is that noisy data severely hinder this approach. Minimising $A[g]$ amounts to solve a massively ill-conditioned linear system

$$\vec{g} = \Sigma^{-1}\vec{f}$$

$$\Sigma_{tr} = \int dE_1 \; e^{-tE_1} \; e^{-rE_1}$$

- Backus-Gilbert regularisation:

$$\int_0^\infty dE \; e^{\alpha E} \; \left| \sum_{t=1}^{t_{\max}} g_t e^{-tE} - \mathcal{S}_\sigma(\omega, E) \right|^2 + \lambda \; \vec{g} \cdot \mathrm{Cov}_d \cdot \vec{g}$$

- The linear system is now

$$\vec{g} = (\Sigma + \lambda \mathrm{Cov_d})^{-1} \; \vec{f}$$

Comparing equations

- In both cases the coefficients that generate the solution are written as:

$$\vec{g}^{\,\mathrm{GP}}(\omega) = \left(\Sigma^{\mathrm{GP}} + \lambda \mathrm{Cov_d}\right)^{-1} \vec{f}^{\,\mathrm{GP}}$$

$$\Sigma^{\mathrm{GP}}{}_{tr} = \int dE_1 \int dE_2 \, e^{-tE_1} \, \mathcal{K}^{\mathrm{prior}}(E_1, E_2) \, e^{-rE_2}$$

$$f_t{}^{\mathrm{GP}}(\omega) = \int dE \, \mathcal{K}^{\mathrm{prior}}(\omega, E) \, e^{-tE}$$

$$\Sigma_{tr} = \int dE_1 \, e^{-tE_1} \, e^{-rE_1}$$

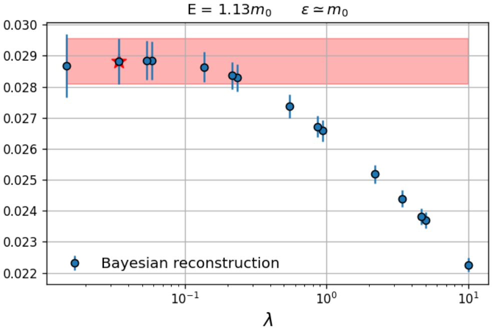
$$f_t(\omega) = \int dE \, S_\sigma(\omega, E) \, e^{-tE}$$

- They can be mapped into one another only at $\sigma = 0$.
- They regularise the problem in the very same way.
- What about λ?

Unphysical parameters & physical results

- λ introduces a bias. Recent application of BG methods perform a "stability analysis" {Bulava et al. 21 }
- We could do the same with the Bayesian reconstruction. Let us pick a prior:

$$\mathcal{K}_{\epsilon}^{\text{prior}}(E, E') = \frac{e^{-(E-E')^2/2\epsilon^2}}{\lambda}\ , \qquad \rho^{\text{prior}} = 0$$

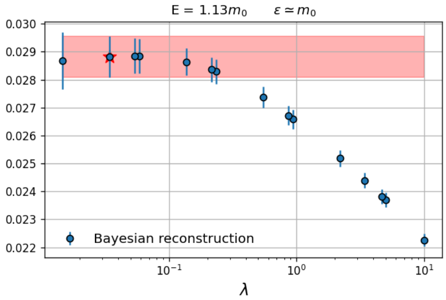


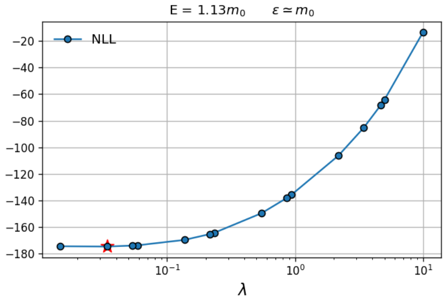
Unphysical parameters & physical results

- In the Bayesian literature, hyperparameters are determined by minimising the negative log likelihood (NLL)

$$-\log P(\text{data}|\text{parameters})$$

- The methods seem compatible





Bayesian formulation of BG

- Compute the posterior probability distribution for a spectral density smeared with a **fixed kernel** $G_\sigma(E, E') = \exp^{-(E-E')^2/2\sigma^2}$

- Diagonal model covariance:

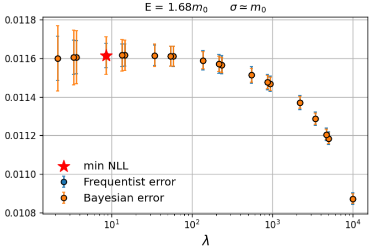
$$\mathcal{K}(E, E') = \frac{\delta(E - E')}{\lambda} \quad ,$$

- The solution is now given by the same coefficients as HLT19

$$g^{\mathrm{GP}}(\sigma; \omega) = g(\sigma; \omega) \quad \text{even at finite } \sigma$$

  - The only difference is in the error (bootstrap for Backus-Gilbert methods)

$$\mathcal{K}^{\sigma}_{\mathrm{post}}(\omega, \omega)^2 = \frac{1}{2} \int dE \left(\sum_t g_t(\sigma, \omega) e^{-tE} - G_\sigma(E, \omega) \right) G_\sigma(E, \omega)$$



More on the bias

- Generate toys for spectral densities / correlators distributed with the covariance measured on the lattice.

$$C(t) = \sum_{n=0}^{n_{\max}-1} w_n \, e^{-|t|E_n} \, , \qquad E_0 < E_1 \leq \ldots \, ,$$

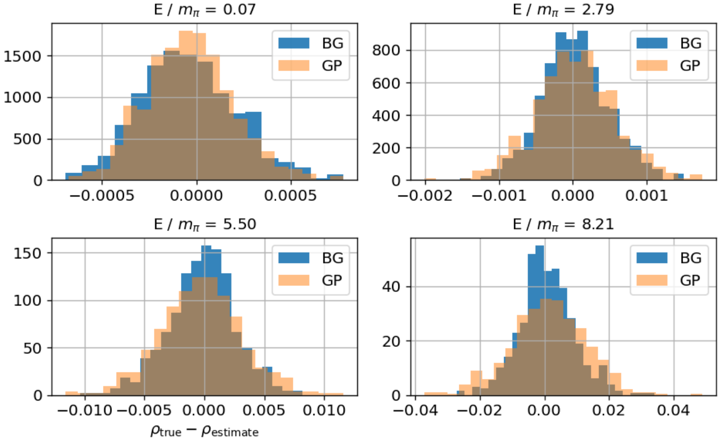
- Solving for each can give an idea of the size of the bias, if any
- Example: generate weights w_n with a GP, centred around the Gounaris-Sakurai parametrisation of the R-ratio, and covariance:

$$K_{\text{weights}}(n, n') = \kappa \, \exp\left(- \frac{(E_n - E_{n'})^2}{2\epsilon^2} \right) \, ,$$

- For the corresponding correlators, we inject noise from a covariance matrix measured on the lattice.

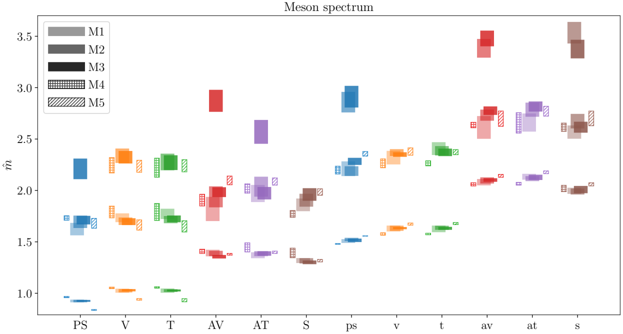
Preliminary results

- Results for ρ_σ (true) - ρ_σ (estimate)
- Same plots for the pull are being analysed. Stay tuned!



Updates about spectroscopy

- In a previous paper [2211.09581] we explored the possibility to perform finite-volume spectroscopy using smeared spectral densities
- Recent developments in [2405.01388]



work with:

E. Bennett, L. Del Debbio, N. Forzano, R.C. Hill, D. K. Hong, H. Hsiao, J.-W. Lee, C.-J. D. Lin, B. Lucini, AL, M. Piai, D. Vadacchino, F. Zierler