Bayesian Interpretation of Backus-Gilbert methods

Alessandro Lupo

Based on work with L. Del Debbio, M. Panero, N. Tantalo
arXiv:2308.xxxx

LATTICE 2023, Fermilab

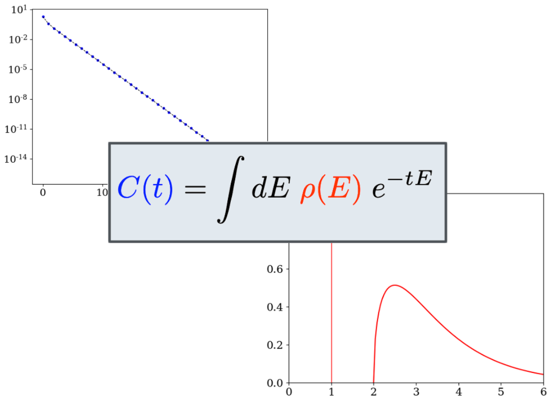
Inverse problem: generalities

- Computing the spectral density $\rho(E)$ associated to a lattice correlator $C(t)$

- Ill-posed in presence of a finite set of noisy data.

- Regularisations are available: Backus-Gilbert & Bayesian methods have different philosophies but share similarities

$$\rho(E) = \lim_{\sigma\to 0} \sum_t g_t(\sigma; E)\, C(t)$$



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

Wish list

- To obtain a function that is smooth even at finite volume:

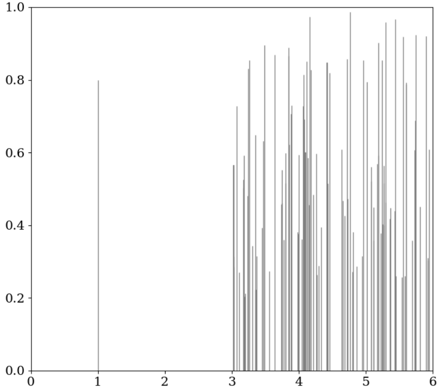
$$\rho_\sigma(\omega) = \int dE \, \mathcal{S}_\sigma(E,\omega) \, \rho(E)$$

- For some applications, a fixed smearing kernel across lattice spacings, volumes, ... to control systematics of fits & extrapolations

- Understand dependence of the result on algorithmic inputs, parameters, priors ...

- Remark: linear combination of $C(t)$ is always smeared

$$\rho_\sigma(E) = \sum_t g_t(\sigma; E) \, C(t)$$
$$= \sum_t g_t(\sigma; E) \int dE \, e^{-tE} \rho(E)$$



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 $\rho^{\mathrm{prior}}(E)$ with covariance $\mathcal{K}^{\mathrm{prior}}(E, E')$.

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

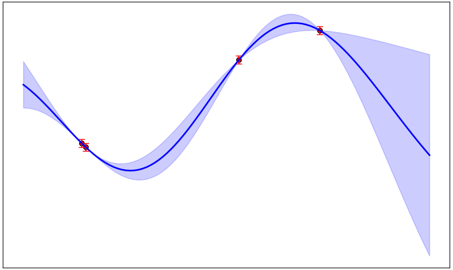
- Similarly, assume that observational noise is Gaussian: $\eta(t)$

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

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

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

Valentine, Sambridge 19
Horak, Pawlowski, Rodríguez-Quintero, Turnwald, Urban 21
Del Debbio, Giani, Wilson 21



Bayesian Inference with Gaussian Processes

- The joint, posterior distribution for ρ^{post} at some energy has centre and variance (set $\rho^{\mathrm{prior}} = 0$)

$$\rho^{\mathrm{post}}(\omega) = \sum_{t=1}^{t_{\max}} g_t^{\mathrm{GP}}(\omega)\, C(t)$$

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

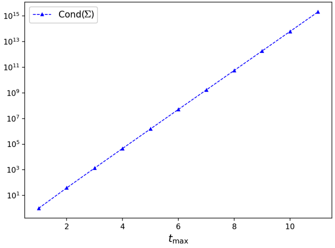
- The coefficients are

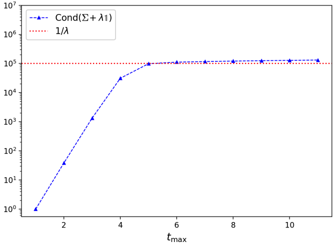
$$g_t^{\mathrm{GP}}(\omega) = \sum_{r=1}^{t_{\max}} \left(\frac{1}{\Sigma + \mathrm{Cov}_d} \right)_{tr} F_r(\omega)$$

- Ingredients:

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

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





Bayesian Inference with Gaussian Processes

- Let the model prior for the spectral density have covariance

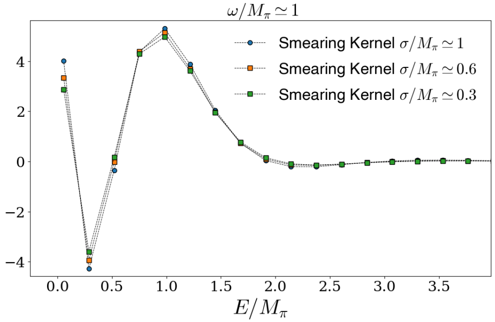
$$\mathcal{K}_{\sigma}(E,E') = \frac{e^{-(E-E')^2/2\sigma^2}}{\lambda} , \qquad \rho^{\rm prior} = 0$$

- The model covariance acts as a smearing kernel,

$$\mathcal{S}_{\sigma}^{\rm GP}(E,\omega) = \sum_{t=1}^{t_{\rm max}} g_t^{\rm GP}(\sigma;\omega) e^{-tE}$$

- If we are interested in removing the smearing, there should be a limit in which $\mathcal{S}$ approaches a δ-function

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



Prior dependence

- Allow the model prior to vary

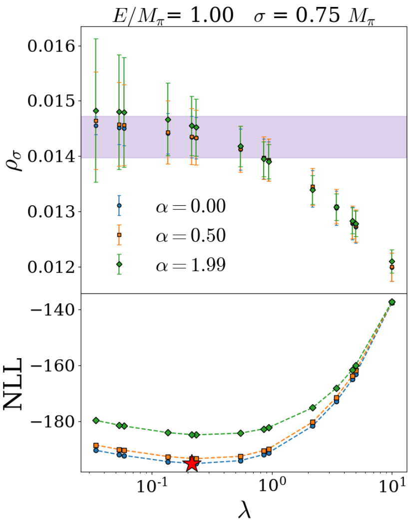
$$\mathcal{K}_{\sigma}^{\mathrm{prior}}(E, E') = \frac{e^{-(E-E')^2/2\sigma^2}}{\lambda} e^{\alpha E} \, ,$$

- By changing α and λ we can explore the dependence on the posterior from the prior

- There is a region in which the dependence on the prior is absorbed in the statistical error

- Choice of the parameters: minimise Negative Log Likelihood

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



Backus Gilbert

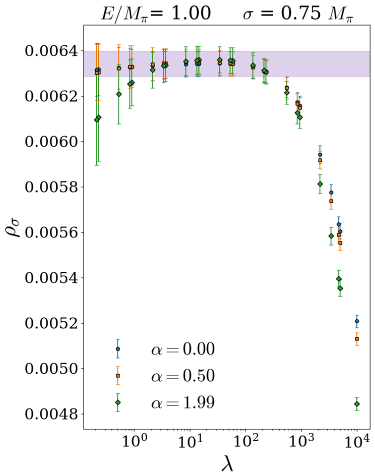
- (Hansen Lupo Tantalo 19) Target a spectral density smeared with a chosen kernel
- Recipe for g_t: make them such that

$$\sum_{t=1}^{\infty} g_t(\omega)\, e^{-tE} = \mathcal{S}_\sigma(E-\omega) \implies \sum_{t=1}^{\infty} g_t(\omega)\, c(t) = \rho_\sigma(\omega)$$

- How? By minimising

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

- $\lambda \in (0,1)$ and $\alpha < 2$ are algorithmical input parameters.



Backus Gilbert

- Our prescription leads to

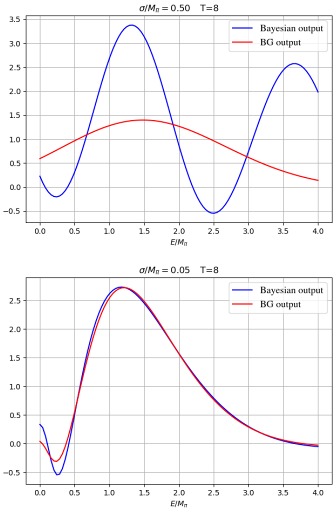
$$g^{\mathrm{BG}}(\sigma;\omega) = \left(\frac{1}{\Sigma^0 + \lambda' \, \mathrm{Cov}_d} \right)_{tr} F_r(\sigma;\omega) \,, \qquad \lambda' = \lambda/(1-\lambda)$$

- The ingredients are

$$\Sigma^0_{tr} = \int dE \, e^{-tE} e^{-rE} e^{\alpha E} \,, \quad \text{ill-cond}$$

$$F_t(\sigma;\omega) = \int dE \, e^{-tE} \, e^{\alpha E} \mathcal{S}_\sigma(E,\omega)$$

- When $\sigma \to 0$, $g^{\mathrm{GP}} \to g^{\mathrm{BG}}$ but for $\sigma \neq 0$ they differ



Backus-Gilbert as a Gaussian Process

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

- Let the model prior for the spectral density have **diagonal** covariance

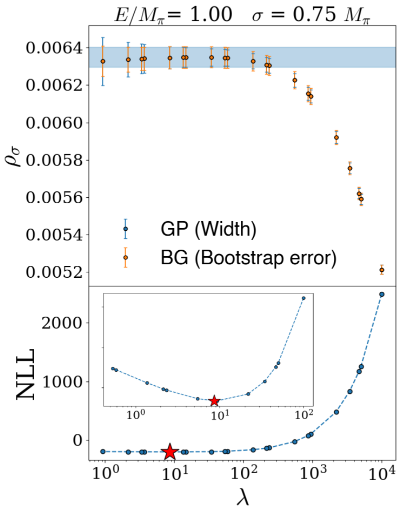
$$\mathcal{K}(E, E') = \frac{\delta(E - E')}{\lambda} e^{\alpha E} ,$$

- By computing the posterior probability, one gets coefficients g^{GPsmr}

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

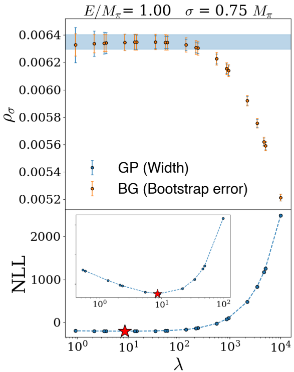
- The only difference is in the error (bootstrap for Backus-Gilbert methods)

$$\Gamma^{\mathrm{GPsmr}}(\sigma;\omega)^2 = \frac{1}{2}\int dE \left(\sum_t g_t^{\mathrm{BG}}(\sigma,\omega)e^{-tE} - G_\sigma(E,\omega)\right) G_\sigma(E,\omega)$$



Conclusions

- Same regularisation of the problem via Cov_d
- Algorithmic parameters of BG can be understood as Bayesian priors, and vice versa
- There is a region in which the result does not depend on the inputs / priors within statistical error. In the same region, the NLL finds its minimum.
- The statistical error of BG is of the same order of magnitude of the Bayesian error Γ^{GPsmr}



This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 813942.