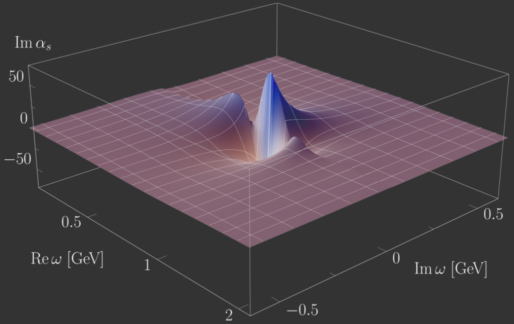
Fredholm inversion with Gaussian processes

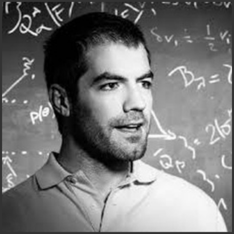


Julian M. Urban lettucefield.org

MIT / IAIFI

Collaborators

J. M. Pawlowski



S. Zafeiropoulos

J. Rodríguez-Quintero

N. Wink

C. S. Schneider

J. Horak

J. Turnwald

J. M. Urban

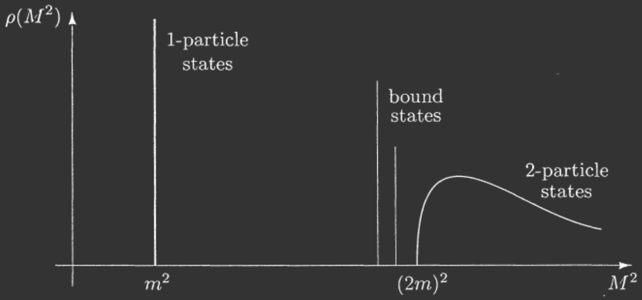


Figure 7.2. The spectral function $\rho(M^2)$ for a typical interacting field theory. The one-particle states contribute a delta function at m^2 (the square of the particle's mass). Multiparticle states have a continuous spectrum beginning at $(2m)^2$. There may also be bound states.

Analogous expressions hold for the case $y^0 > x^0$. Both cases can be summarized in the following general representation of the two-point function (the *Källén-Lehmann spectral representation*):

$$\langle\Omega| T\phi(x)\phi(y) |\Omega\rangle = \int_0^\infty \frac{dM^2}{2\pi}\, \rho(M^2)\, D_F(x-y; M^2), \tag{7.6}$$

where $\rho(M^2)$ is a positive *spectral density* function,

$$\rho(M^2) = \sum_\lambda (2\pi)\delta(M^2 - m_\lambda^2)\big|\langle\Omega| \phi(0) |\lambda_0\rangle\big|^2. \tag{7.7}$$

From imaginary to real time

- Many non-perturbative approaches to strongly interacting QFTs only provide numerical access to Euclidean correlators

From imaginary to real time

- Many non-perturbative approaches to strongly interacting QFTs only provide numerical access to Euclidean correlators
- Direct analytic continuation of imaginary-time *data* problematic

From imaginary to real time

- Many non-perturbative approaches to strongly interacting QFTs only provide numerical access to Euclidean correlators
- Direct analytic continuation of imaginary-time *data* problematic
- Indirect access via Källén-Lehmann spectral representation:

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\,\rho(\omega)}{p^2 + \omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

From imaginary to real time

- Many non-perturbative approaches to strongly interacting QFTs only provide numerical access to Euclidean correlators
- Direct analytic continuation of imaginary-time *data* problematic
- Indirect access via Källén-Lehmann spectral representation:

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\,\rho(\omega)}{p^2+\omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

- Computing spectral function $\rho(\omega)$ from noisy data $\mathbf{G}$ is an ill-posed linear inverse problem (Fredholm inversion)

Outline

- Motivation for probabilistic ansatz
- Gaussian process regression (GPR)
- Fredholm inversion with GPR
- Related & original work
- Application 1: bound state masses
- Incorporating asymptotic constraints
- Application 2: S-matrix building blocks

Motivation for probabilistic ansatz

Motivation for probabilistic ansatz

- A *problem* is a function $f^{-1} : \mathcal{G} \longrightarrow \mathcal{R}$ from a space of initial conditions $\mathcal{G}$ to a space of solutions $\mathcal{R}$

Motivation for probabilistic ansatz

- A *problem* is a function $f^{-1} : \mathcal{G} \longrightarrow \mathcal{R}$ from a space of initial conditions $\mathcal{G}$ to a space of solutions $\mathcal{R}$

- Usually interested in a solution $\rho \in \mathcal{R}$ for a point $G \in \mathcal{G}$

Motivation for probabilistic ansatz

- A *problem* is a function $f^{-1} : \mathcal{G} \longrightarrow \mathcal{R}$ from a space of initial conditions $\mathcal{G}$ to a space of solutions $\mathcal{R}$

- Usually interested in a solution $\rho \in \mathcal{R}$ for a point $G \in \mathcal{G}$

- The combination $\{f^{-1}, G\}$ is a *problem instance*

Motivation for probabilistic ansatz

- A *problem* is a function $f^{-1}: \mathcal{G} \longrightarrow \mathcal{R}$ from a space of initial conditions $\mathcal{G}$ to a space of solutions $\mathcal{R}$

- Usually interested in a solution $\rho \in \mathcal{R}$ for a point $G \in \mathcal{G}$

- The combination $\{f^{-1}, G\}$ is a *problem instance*

- For $f[\rho] \equiv \displaystyle\int \mathrm{d}\omega\, K(p,\omega)\rho(\omega) = G(p),\ \mathcal{G}, \mathcal{R}$ are *function* spaces

Motivation for probabilistic ansatz

- A *problem* is a function $f^{-1} : \mathcal{G} \longrightarrow \mathcal{R}$ from a space of initial conditions $\mathcal{G}$ to a space of solutions $\mathcal{R}$
- Usually interested in a solution $\rho \in \mathcal{R}$ for a point $G \in \mathcal{G}$
- The combination $\{f^{-1}, G\}$ is a *problem instance*
- For $f[\rho] \equiv \int \mathrm{d}\omega\, K(p,\omega)\rho(\omega) = G(p),\ \mathcal{G}, \mathcal{R}$ are *function* spaces
- Determination of $\rho(\omega)$ from $G(p)$ is well-posed in principle

Motivation for probabilistic ansatz

- In practice: discrete *data* $\mathbf{G}$ with finite measurement error

Motivation for probabilistic ansatz

- In practice: discrete *data* $\mathbf{G}$ with finite measurement error
- Deterministic computation of solution *function* $\rho(\omega)$ is ill-posed due to violation of uniqueness (cf. Hadamard)

Motivation for probabilistic ansatz

- In practice: discrete *data* $\mathbf{G}$ with finite measurement error
- Deterministic computation of solution *function* $\rho(\omega)$ is ill-posed due to violation of uniqueness (cf. Hadamard)
- Probabilistic ansatz: model initial & solution functions *jointly* for simultaneous inference from available data

$$\longrightarrow \text{ Objective: } P(\{G(p), \rho(\omega)\} | f, \mathbf{G})$$

Motivation for probabilistic ansatz

- In practice: discrete *data* $\mathbf{G}$ with finite measurement error
- Deterministic computation of solution *function* $\rho(\omega)$ is ill-posed due to violation of uniqueness (cf. Hadamard)
- Probabilistic ansatz: model initial & solution functions *jointly* for simultaneous inference from available data

$$\longrightarrow \text{ Objective: } P(\{G(p), \rho(\omega)\}|f, \mathbf{G})$$

- Desirable: incorporation of additional prior information in the form of constraints and biases

Gaussian processes

Gaussian processes

- Consider $\omega, \rho(\omega) \in \mathbb{R}$ for concreteness

Gaussian processes

- Consider $\omega, \rho(\omega) \in \mathbb{R}$ for concreteness
- Define prior distribution with a Gaussian process (GP):

$$\rho(\omega) \sim \mathcal{GP}\left(\mu(\omega), C(\omega, \omega')\right),\; C(\omega, \omega') = C(\omega', \omega)$$

Gaussian processes

- Consider $\omega, \rho(\omega) \in \mathbb{R}$ for concreteness
- Define prior distribution with a Gaussian process (GP):

$$\rho(\omega) \sim \mathcal{GP}\left(\mu(\omega), C(\omega, \omega')\right),\ C(\omega, \omega') = C(\omega', \omega)$$

- GPs are normal distributions in *function space*

Gaussian processes

- Consider $\omega, \rho(\omega) \in \mathbb{R}$ for concreteness
- Define prior distribution with a Gaussian process (GP):

$$\rho(\omega) \sim \mathcal{GP}\left(\mu(\omega), C(\omega, \omega')\right),\ C(\omega, \omega') = C(\omega', \omega)$$

- GPs are normal distributions in *function space*
- Analogy: free scalar field theory $\sim$ ∞-dim. multivariate normal

$\longrightarrow$ mean $\mu(\omega)$ $\sim$ vacuum expectation value

$\longrightarrow$ covariance $C(\omega, \omega')$ $\sim$ propagator

Gaussian processes

- Finite-dim. multivariate normal for any finite set of points:

$$\begin{pmatrix} \rho(\omega_1) \\ \vdots \\ \rho(\omega_N) \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu(\omega_1) & & C(\omega_1, \omega_1) & \dots & C(\omega_1, \omega_N) \\ \vdots & , & \vdots & \ddots & \vdots \\ \mu(\omega_N) & & C(\omega_N, \omega_1) & \dots & C(\omega_N, \omega_N) \end{pmatrix}$$

$$= (2\pi)^{-\frac{N}{2}} \det(\mathbf{C})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\rho} - \boldsymbol{\mu})^T \mathbf{C}^{-1} (\boldsymbol{\rho} - \boldsymbol{\mu})\right)$$

Gaussian processes

- Finite-dim. multivariate normal for any finite set of points:

$$\begin{pmatrix} \rho(\omega_1) \\ \vdots \\ \rho(\omega_N) \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu(\omega_1) & C(\omega_1,\omega_1) & \dots & C(\omega_1,\omega_N) \\ \vdots & , \quad \vdots & \ddots & \vdots \\ \mu(\omega_N) & C(\omega_N,\omega_1) & \dots & C(\omega_N,\omega_N) \end{pmatrix}$$

$$= (2\pi)^{-\frac{N}{2}} \det(\mathbf{C})^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\rho}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\boldsymbol{\rho}-\boldsymbol{\mu})\right)$$

- Sampling similar to pseudofermion generation in LQCD:

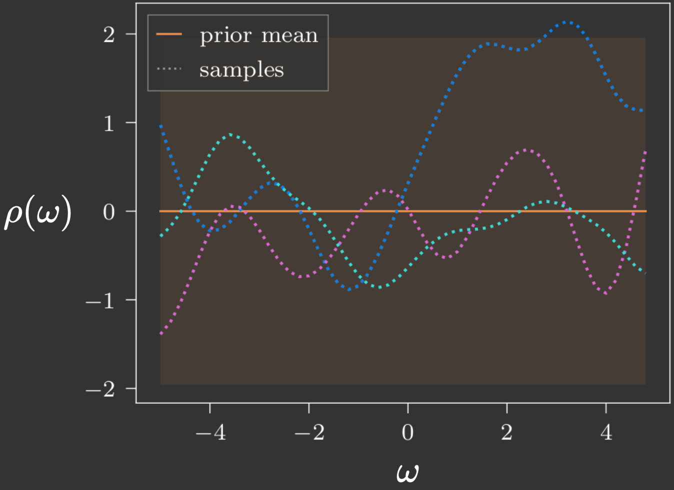
  1. Cholesky decomposition: $\mathbf{C} = \mathbf{A}\mathbf{A}^T$
  2. $\boldsymbol{\rho} = \mathbf{A}\boldsymbol{\eta}$ with $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbb{1})$

Gaussian processes

$$\mu(\omega) = 0,\ C(\omega, \omega') = \sigma^2 \exp\left(-\frac{(\omega - \omega')^2}{2\xi^2}\right)$$

Gaussian processes

$$\mu(\omega) = 0,\ C(\omega, \omega') = \sigma^2 \exp\left(-\frac{(\omega - \omega')^2}{2\xi^2}\right)$$



Gaussian process regression

Gaussian process regression

- Joint distribution of *observed* $\boldsymbol{\rho}$ and *unobserved* $\rho(\omega)$:

$$\begin{pmatrix} \rho(\omega) \\ \boldsymbol{\rho} \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \mu(\omega) \\ \boldsymbol{\mu} \end{matrix} \, , \, \begin{matrix} C(\omega,\omega') & \mathbf{C}^T(\omega) \\ \mathbf{C}(\omega') & \mathbf{C} + \sigma_n^2 \mathbb{1} \end{matrix} \right)$$

$$\boldsymbol{\mu}_i = \mu(\omega_i),\ \mathbf{C}_i(\omega) = C(\omega_i, \omega),\ \mathbf{C}_{ij} = C(\omega_i, \omega_j)$$

Gaussian process regression

- Joint distribution of *observed* $\boldsymbol{\rho}$ and *unobserved* $\rho(\omega)$:

$$\begin{pmatrix} \rho(\omega) \\ \boldsymbol{\rho} \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \mu(\omega) \\ \boldsymbol{\mu} \end{matrix} , \begin{matrix} C(\omega,\omega') & \mathbf{C}^T(\omega) \\ \mathbf{C}(\omega') & \mathbf{C} + \sigma_n^2 \mathbb{1} \end{matrix} \right)$$

$$\boldsymbol{\mu}_i = \mu(\omega_i),\ \mathbf{C}_i(\omega) = C(\omega_i,\omega),\ \mathbf{C}_{ij} = C(\omega_i,\omega_j)$$

- Closed-form expression for conditional posterior:

$$\rho(\omega)|\boldsymbol{\rho} \sim \mathcal{N}\left(\mu(\omega) + \mathbf{C}^T(\omega)\left(\mathbf{C} + \sigma_n^2 \cdot \mathbb{1}\right)^{-1}(\boldsymbol{\rho} - \boldsymbol{\mu}),\ C(\omega,\omega') - \mathbf{C}^T(\omega)\left(\mathbf{C} + \sigma_n^2 \cdot \mathbb{1}\right)^{-1}\mathbf{C}(\omega') \right)$$

Gaussian process regression

- Joint distribution of *observed* $\boldsymbol{\rho}$ and *unobserved* $\rho(\omega)$:

$$\begin{pmatrix} \rho(\omega) \\ \boldsymbol{\rho} \end{pmatrix} \sim \mathcal{N}\left(\begin{matrix} \mu(\omega) \\ \boldsymbol{\mu} \end{matrix} , \begin{matrix} C(\omega,\omega') & \mathbf{C}^T(\omega) \\ \mathbf{C}(\omega') & \mathbf{C} + \sigma_n^2 \mathbb{1} \end{matrix} \right)$$

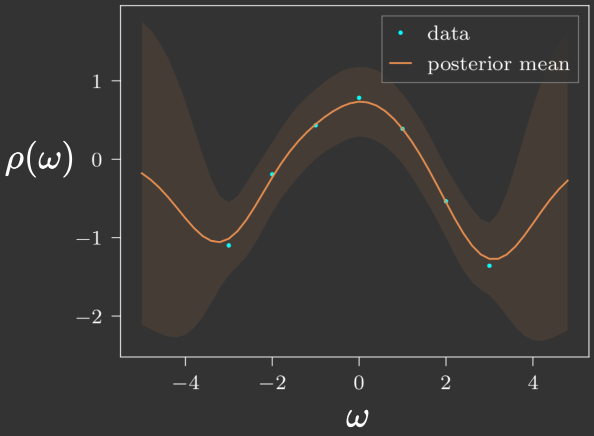
$$\boldsymbol{\mu}_i = \mu(\omega_i),\ \mathbf{C}_i(\omega) = C(\omega_i,\omega),\ \mathbf{C}_{ij} = C(\omega_i,\omega_j)$$

- Closed-form expression for conditional posterior:

$$\rho(\omega)|\boldsymbol{\rho} \sim \mathcal{N}\Big(\underbrace{\overset{\text{prior}}{\mu(\omega)} + \overset{\text{prior}}{\mathbf{C}^T(\omega)} \big(\mathbf{C} + \sigma_n^2 \cdot \mathbb{1}\big)^{-1} \overset{\text{data}}{(\boldsymbol{\rho} - \boldsymbol{\mu})}}_{\text{prediction}},\ \underbrace{C(\omega,\omega') - \mathbf{C}^T(\omega)\big(\mathbf{C} + \overset{\text{error}}{\sigma_n^2} \cdot \mathbb{1}\big)^{-1} \mathbf{C}(\omega')}_{\text{uncertainty}} \Big)$$

Gaussian process regression

$$\rho(\omega)|\boldsymbol{\rho} \sim \mathcal{N}\left(\underbrace{\mu(\omega) + \mathbf{C}^T(\omega)\left(\mathbf{C} + \sigma_n^2 \cdot \mathbb{1}\right)^{-1}(\boldsymbol{\rho} - \boldsymbol{\mu})}, \underbrace{C(\omega,\omega') - \mathbf{C}^T(\omega)\left(\mathbf{C} + \sigma_n^2 \cdot \mathbb{1}\right)^{-1}\mathbf{C}(\omega')}\right)$$



Hyperparameter optimization

- GPR: *non-parametric* ansatz with *hyperparameters* $\boldsymbol{\alpha}$

Hyperparameter optimization

- GPR: *non-parametric* ansatz with *hyperparameters* $\boldsymbol{\alpha}$
- Common selection approach: maximize likelihood

$$p(\boldsymbol{\rho}|\boldsymbol{\alpha}) = \left((2\pi)^N \det\left(\mathbf{C}_\alpha + \sigma_n^2 \mathbb{1}\right)\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\rho}-\boldsymbol{\mu})^T \left(\mathbf{C}_{\boldsymbol{\alpha}} + \sigma_n^2 \mathbb{1}\right)^{-1} (\boldsymbol{\rho}-\boldsymbol{\mu})\right)$$

or, conventionally, minimize negative log-likelihood (NLL)

$$-\log p(\boldsymbol{\rho}|\boldsymbol{\alpha}) = \frac{1}{2}(\boldsymbol{\rho}-\boldsymbol{\mu})^T \left(\mathbf{C}_{\boldsymbol{\alpha}} + \sigma_n^2 \mathbb{1}\right)^{-1} (\boldsymbol{\rho}-\boldsymbol{\mu}) + \frac{1}{2}\log\det\left(\mathbf{C}_\alpha + \sigma_n^2 \mathbb{1}\right) + \frac{N}{2}\log 2\pi$$

Hyperparameter optimization

- GPR: *non-parametric* ansatz with *hyperparameters* $\boldsymbol{\alpha}$
- Common selection approach: maximize likelihood

$$p(\boldsymbol{\rho}|\boldsymbol{\alpha}) = \left((2\pi)^N \det\left(\mathbf{C}_\alpha + \sigma_n^2 \mathbb{1}\right)\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{\rho} - \boldsymbol{\mu})^T \left(\mathbf{C}_{\boldsymbol{\alpha}} + \sigma_n^2 \mathbb{1}\right)^{-1} (\boldsymbol{\rho} - \boldsymbol{\mu})\right)$$

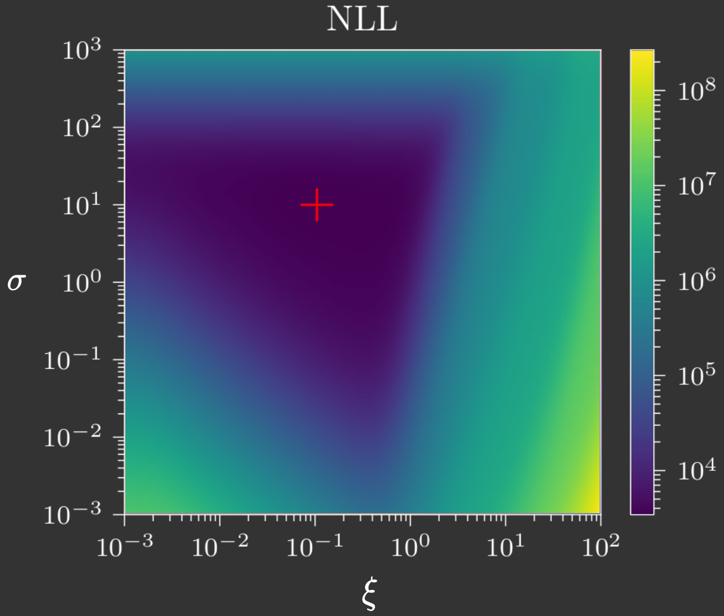
or, conventionally, minimize negative log-likelihood (NLL)

$$-\log p(\boldsymbol{\rho}|\boldsymbol{\alpha}) = \frac{1}{2}(\boldsymbol{\rho} - \boldsymbol{\mu})^T \left(\mathbf{C}_{\boldsymbol{\alpha}} + \sigma_n^2 \mathbb{1}\right)^{-1} (\boldsymbol{\rho} - \boldsymbol{\mu}) + \frac{1}{2}\log\det\left(\mathbf{C}_\alpha + \sigma_n^2 \mathbb{1}\right) + \frac{N}{2}\log 2\pi$$

- Ignores structure & unstable directions of posterior landscape

  $\longrightarrow$ stabilize or integrate out with suitable *hyperprior* $q(\boldsymbol{\alpha})$

Hyperparameter optimization



sklearn to the rescue

```
>>> from sklearn.datasets import make_friedman2
>>> from sklearn.gaussian_process import GaussianProcessRegressor
>>> from sklearn.gaussian_process.kernels import DotProduct, WhiteKernel
>>> X, y = make_friedman2(n_samples=500, noise=0, random_state=0)
>>> kernel = DotProduct() + WhiteKernel()
>>> gpr = GaussianProcessRegressor(kernel=kernel,
...         random_state=0).fit(X, y)
>>> gpr.score(X, y)
0.3680...
>>> gpr.predict(X[:2,:], return_std=True)
(array([653.0..., 592.1...]), array([316.6..., 316.6...]))
```
>>>

Fredholm inversion with GPR

Fredholm inversion with GPR

- Linear transformations $G = \mathcal{K} \circ \rho$ preserve Gaussian statistics

Fredholm inversion with GPR

- Linear transformations $G = \mathcal{K} \circ \rho$ preserve Gaussian statistics
- $\rho \sim \mathcal{GP}(\mu, C) \longrightarrow G \sim \mathcal{GP}(\mathcal{K} \circ \mu, \mathcal{K} \circ C \circ \mathcal{K}^T)$

Fredholm inversion with GPR

- Linear transformations $G = \mathcal{K} \circ \rho$ preserve Gaussian statistics
- $\rho \sim \mathcal{GP}(\mu, C) \longrightarrow G \sim \mathcal{GP}(\mathcal{K} \circ \mu, \mathcal{K} \circ C \circ \mathcal{K}^T)$
- Joint prior over *indirect observations* $\mathbf{G}$ and prediction ρ :

$$\begin{pmatrix} \rho \\ \mathbf{G} \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & C & C \circ \mathcal{K}^T \\ \mathcal{K} \circ \mu ' & \mathcal{K} \circ C & \mathcal{K} \circ C \circ \mathcal{K}^T + \sigma_n^2 \mathbb{1} \end{pmatrix}$$

Fredholm inversion with GPR

- Linear transformations $G = \mathcal{K} \circ \rho$ preserve Gaussian statistics
- $\rho \sim \mathcal{GP}(\mu, C) \longrightarrow G \sim \mathcal{GP}(\mathcal{K} \circ \mu, \mathcal{K} \circ C \circ \mathcal{K}^T)$
- Joint prior over *indirect observations* $\mathbf{G}$ and prediction ρ :

$$\begin{pmatrix} \rho \\ \mathbf{G} \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & C & C \circ \mathcal{K}^T \\ \mathcal{K} \circ \mu, & \mathcal{K} \circ C & \mathcal{K} \circ C \circ \mathcal{K}^T + \sigma_n^2 \mathbb{1} \end{pmatrix}$$

- Closed-form posterior like before, with $\mathcal{K} \circ$ where necessary

Fredholm inversion with GPR

$$G(p) \sim \mathcal{GP}\left(\int \mathrm{d}\omega\, K(p,\omega)\mu(\omega), \int \mathrm{d}\omega\, \mathrm{d}\omega' K(p,\omega)C(\omega,\omega')K(p',\omega')\right)$$

$$P(\rho(\omega)|\mathbf{G}) = \mathcal{N}\Bigg(\mu(\omega) + \int \mathrm{d}\eta' C(\omega,\eta')K(\mathbf{p}^T,\eta')$$

$$\left(\int \mathrm{d}\zeta\, \mathrm{d}\zeta' K(\mathbf{p},\zeta)C(\zeta,\zeta')K(\mathbf{p}^T,\zeta') + \sigma_n^2 \mathbb{1}\right)^{-1} \left(\mathbf{G} - \int \mathrm{d}\eta\, K(\mathbf{p},\eta)\mu(\eta)\right),$$

$$C(\omega,\omega') - \int \mathrm{d}\eta' C(\omega,\eta')K(\mathbf{p}^T,\eta')$$

$$\left(\int \mathrm{d}\zeta\, \mathrm{d}\zeta' K(\mathbf{p},\zeta)C(\zeta,\zeta')K(\mathbf{p}^T,\zeta') + \sigma_n^2 \mathbb{1}\right)^{-1} \int \mathrm{d}\eta\, K(\mathbf{p},\eta)C(\eta,\omega')\Bigg)$$

Fredholm inversion with GPR

- Linear transformations $G = \mathcal{K} \circ \rho$ preserve Gaussian statistics
- $\rho \sim \mathcal{GP}(\mu, C) \longrightarrow G \sim \mathcal{GP}(\mathcal{K} \circ \mu, \mathcal{K} \circ C \circ \mathcal{K}^T)$
- Joint prior over *indirect observations* $\mathbf{G}$ and prediction ρ :

$$\begin{pmatrix} \rho \\ \mathbf{G} \end{pmatrix} \sim \mathcal{N}\begin{pmatrix} \mu & C & C \circ \mathcal{K}^T \\ \mathcal{K} \circ \mu' & \mathcal{K} \circ C & \mathcal{K} \circ C \circ \mathcal{K}^T + \sigma_n^2 \mathbb{1} \end{pmatrix}$$

- Closed-form posterior like before, with $\mathcal{K} \circ$ where necessary
- Generalizable to *arbitrary* derivative & integral constraints or known values of ρ by treating them as observations

Related/original work

- Krige, Danie G. (1951): *A statistical approach to some basic mine valuation problems on the Witwatersrand*. J. of the Chem., Metal. and Mining Soc. of South Africa. 52 (6): 119–139.

- Valentine, Andrew P. and Sambridge, Malcolm (2019): *Gaussian process models—I. A framework for probabilistic continuous inverse theory*. Geophysical Journal International, Volume 220, Issue 3, March 2020, Pages 1632–1647

- Del Debbio, Luigi and Giani, Tommaso and Wilson, Michael (2021): *Bayesian approach to inverse problems: an application to NNPDF closure testing.* arXiv:2111.05787 [hep-ph]

- Candido, Alessandro and Del Debbio, Luigi and Giani, Tommaso and Petrillo, Giacomo (2023): *Inverse problems in PDF determinations.* arXiv:2302.14731 [hep-lat]

Numerical Demo

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\,\rho(\omega)}{p^2+\omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

Numerical Demo

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\, \rho(\omega)}{p^2 + \omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

- Example with analytic solution: relativistic Breit-Wigner dist.

$$\rho(\omega) = \frac{4A\Gamma\omega}{4\Gamma^2\omega^2 + (M^2 + \Gamma^2 - \omega^2)^2},\ G(p) = \frac{A}{(p+\Gamma)^2 + M^2}$$

Numerical Demo

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\,\rho(\omega)}{p^2+\omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

- Example with analytic solution: relativistic Breit-Wigner dist.

$$\rho(\omega) = \frac{4A\Gamma\omega}{4\Gamma^2\omega^2 + (M^2+\Gamma^2-\omega^2)^2}, \; G(p) = \frac{A}{(p+\Gamma)^2+M^2}$$

- Common approach: discretize integral, solve system of linear eq.

$$G(p_i) \approx \sum_{\omega_j} \Delta\omega\, K(p_i,\omega_j)\rho(\omega_j) \quad \longleftrightarrow \quad \mathbf{G} \approx \tilde{\mathbf{K}}\boldsymbol{\rho}$$

Numerical Demo

$$G(p) = \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \frac{2\omega\,\rho(\omega)}{p^2+\omega^2} \equiv \int_0^\infty \mathrm{d}\omega\, K(p,\omega)\rho(\omega)$$

- Example with analytic solution: relativistic Breit-Wigner dist.

$$\rho(\omega) = \frac{4A\Gamma\omega}{4\Gamma^2\omega^2 + (M^2+\Gamma^2-\omega^2)^2}, \quad G(p) = \frac{A}{(p+\Gamma)^2+M^2}$$

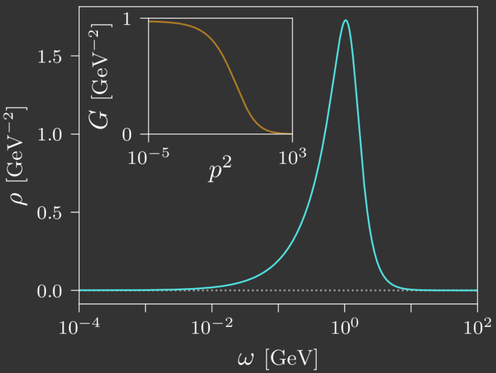
- Common approach: discretize integral, solve system of linear eq.

$$G(p_i) \approx \sum_{\omega_j} \Delta\omega\, K(p_i,\omega_j)\rho(\omega_j) \quad \longleftrightarrow \quad \mathbf{G} \approx \tilde{\mathbf{K}}\boldsymbol{\rho}$$

- Additive Gaussian noise: $\mathbf{G} + \boldsymbol{\epsilon},\ \epsilon \sim \mathcal{N}(0,\sigma^2),\ \sigma = 10^{-4}$

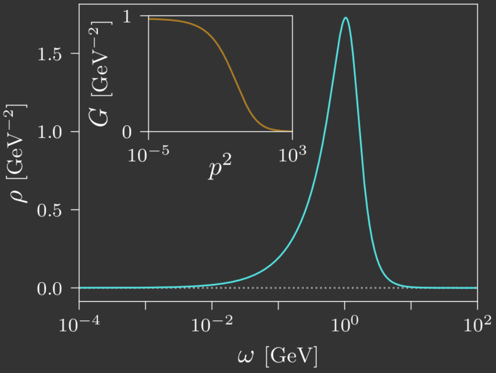
Numerical Demo

- Perfect data: $\boldsymbol{\rho} \approx \tilde{\mathbf{K}}^{-1}\mathbf{G}$

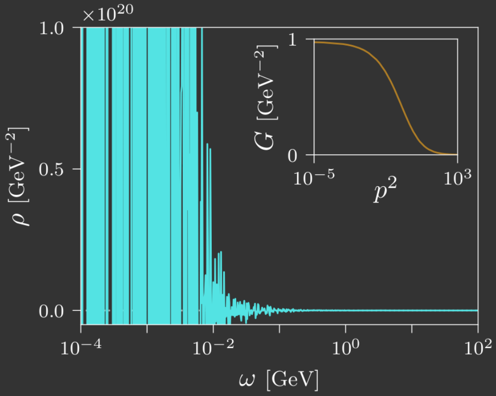


Numerical Demo

- Perfect data: $\boldsymbol{\rho} \approx \tilde{\mathbf{K}}^{-1}\mathbf{G}$

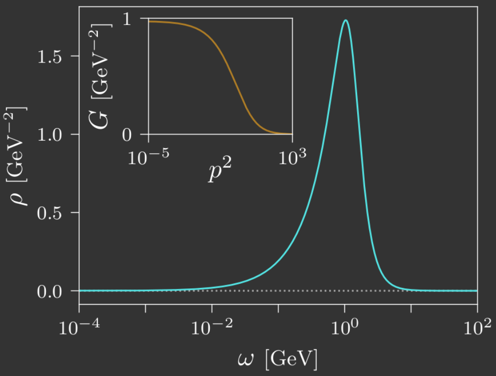


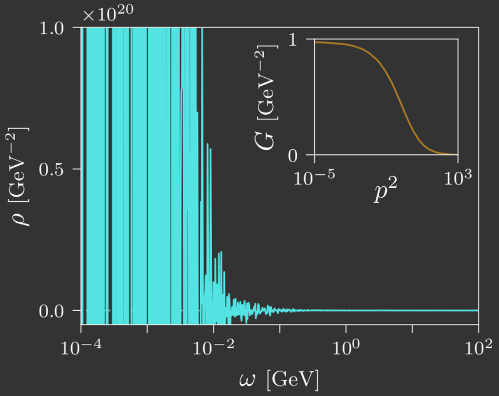
- Noisy data: $\boldsymbol{\rho} \approx \tilde{\mathbf{K}}^{-1}(\mathbf{G} + \boldsymbol{\epsilon})$



Numerical Demo

- Perfect data: $\boldsymbol{\rho} \approx \tilde{\mathbf{K}}^{-1}\mathbf{G}$
- Noisy data: $\boldsymbol{\rho} \approx \tilde{\mathbf{K}}^{-1}(\mathbf{G} + \boldsymbol{\epsilon})$





- Condition number $\kappa(\tilde{\mathbf{K}}) \approx \mathcal{O}(10^{33})$

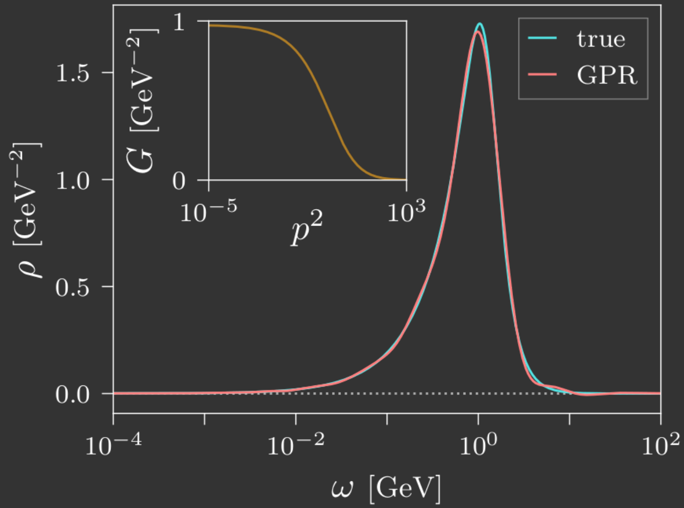
  $\longrightarrow$ expect to lose $\sim$ 33 digits of accuracy with naive inversion

Numerical Demo

- GPR: $\boldsymbol{\rho} \approx \boldsymbol{\mu} + \mathbf{C}\tilde{\mathbf{K}}^T \left(\tilde{\mathbf{K}}\mathbf{C}\tilde{\mathbf{K}}^T + \sigma_n^2 \mathbb{1} \right)^{-1} \left((\mathbf{G} + \boldsymbol{\epsilon}) - \tilde{\mathbf{K}}\boldsymbol{\mu} \right)$

Numerical Demo

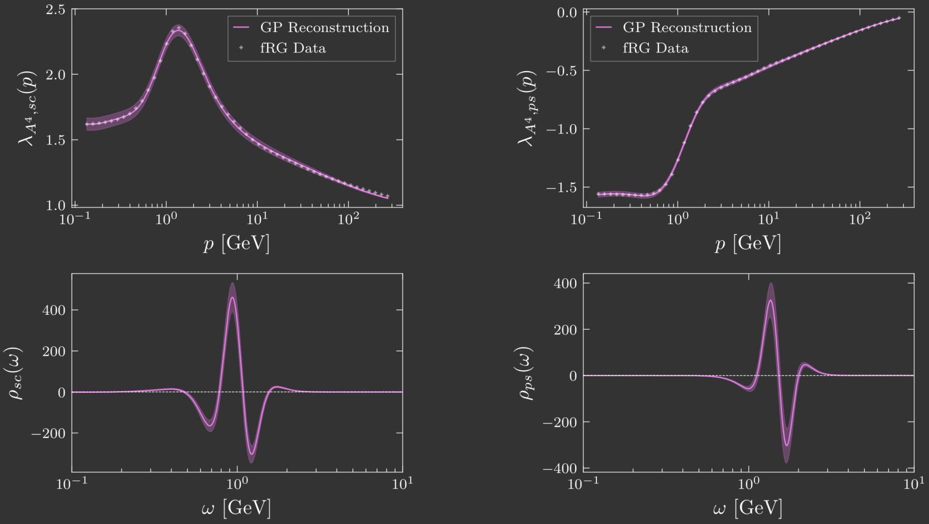
- GPR: $\boldsymbol{\rho} \approx \boldsymbol{\mu} + \mathbf{C}\tilde{\mathbf{K}}^T \left(\tilde{\mathbf{K}}\mathbf{C}\tilde{\mathbf{K}}^T + \sigma_n^2 \mathbb{1} \right)^{-1} \left((\mathbf{G} + \boldsymbol{\epsilon}) - \tilde{\mathbf{K}}\boldsymbol{\mu} \right)$



Application 1: bound state masses

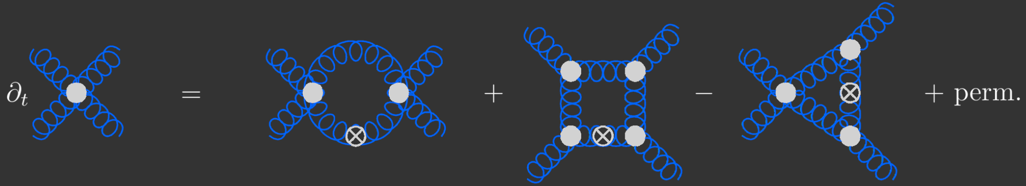
2212.01113 [hep-ph]

- Scalar glueball: $J^{PC} = 0^{++}$
- Pseudo-sc. glueball: $J^{PC} = 0^{-+}$



Application 1: bound state masses

2212.01113 [hep-ph]



J^{PC}	lattice		DSE-BSE		this work
0^{++}	1760(70)	[9]	1850(130)	[18]	1870(75)
	1740(70)	[11]	1640	[15]	
	1651(23)	[12]			
	1618(26)(25)	[13]			
0^{-+}	2650(60)	[9]	2580(180)	[18]	2700(120)
	2610(70)	[11]	4530	[15]	
	2600(40)	[12]			
	2483(61)(55)	[13]			

Exact asymptotic constraints

Exact asymptotic constraints

- Standard GPR: equality constraints $\mathcal{A} \circ \rho = b$ (with $\mathcal{A}$ any linear operator) by treating b as an indirect observation

Exact asymptotic constraints

- Standard GPR: equality constraints $\mathcal{A} \circ \rho = b$ (with $\mathcal{A}$ any linear operator) by treating b as an indirect observation
- Asymptotic constraints $\mathcal{A} \circ \rho \overset{\omega \to \infty}{\propto} \phi(\omega)$

Exact asymptotic constraints

- Standard GPR: equality constraints $\mathcal{A} \circ \rho = b$ (with $\mathcal{A}$ any linear operator) by treating b as an indirect observation
- Asymptotic constraints $\mathcal{A} \circ \rho \overset{\omega \to \infty}{\propto} \phi(\omega)$ by modifying kernel via

Mercer's theorem: $\forall$ continuous symmetric PSD kernels

$$C \ : \ [a, b] \times [a, b] \longrightarrow \mathbb{R}, \ C(x, y) = C(y, x) \, \forall \, x, y \in [a, b]$$

$\exists$ orthonormal basis of eigenfunctions $e_i(x)$ with non-negative eigenvalues λ_i s.t. $C(x, y) = \sum_{i=1} \lambda_i e_i(x) e_i(y)$.

Exact asymptotic constraints

- Standard GPR: equality constraints $\mathcal{A} \circ \rho = b$ (with $\mathcal{A}$ any linear operator) by treating b as an indirect observation
- Asymptotic constraints $\mathcal{A} \circ \rho \overset{\omega \to \infty}{\propto} \phi(\omega)$ by modifying kernel via

Mercer's theorem: $\forall$ continuous symmetric PSD kernels

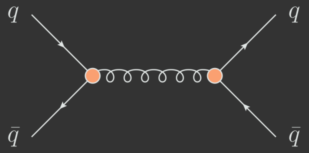
$$C \;:\; [a,b] \times [a,b] \longrightarrow \mathbb{R},\; C(x,y) = C(y,x) \;\forall\, x,y \in [a,b]$$

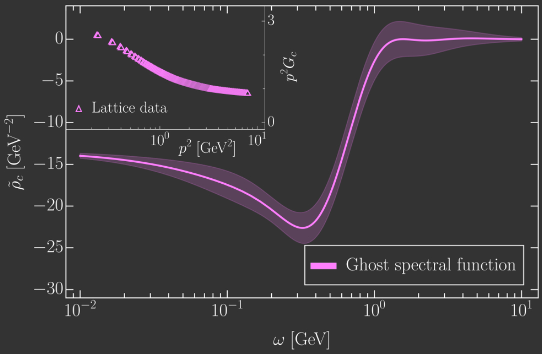
$\exists$ orthonormal basis of eigenfunctions $e_i(x)$ with non-negative eigenvalues λ_i s.t. $C(x,y) = \sum_{i=1} \lambda_i e_i(x) e_i(y)$.

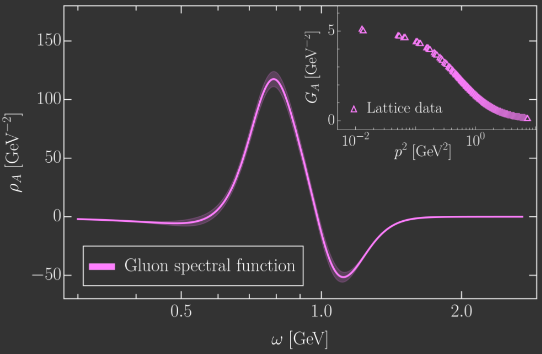
- E.g.: smooth transition from perturbative to non-pert. region

Application 2: S-matrix building blocks

2107.13464 [hep-ph], 2301.07785 [hep-ph]

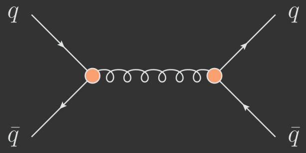


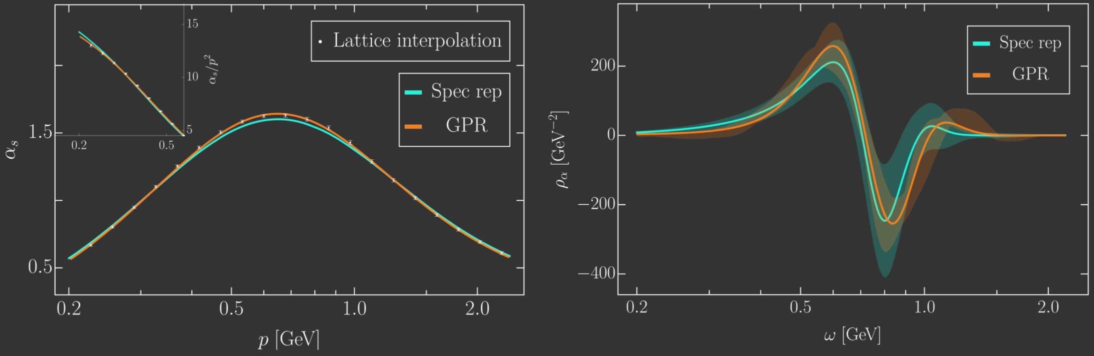




Application 2: S-matrix building blocks

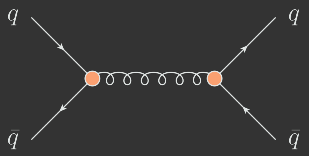
2107.13464 [hep-ph], 2301.07785 [hep-ph]

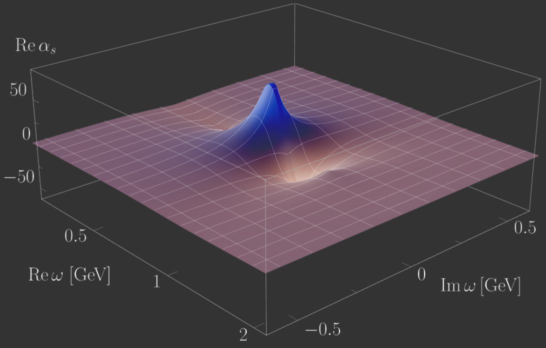


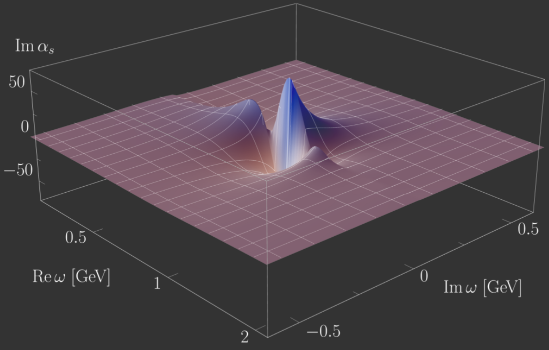


Application 2: S-matrix building blocks

2107.13464 [hep-ph], 2301.07785 [hep-ph]







Outlook

- Algorithm:
  - Integrate out hyperparams via hyperpriors & HMC
  - Automatic kernel selection, hyperkernels, deep kernel learning
  - Inequality constraints like positivity & monotonicity
  - Smeared spectral densities from finite-volume correlators
- Data:
  - Increased resolution/precision of lattice & fRG results
  - Improved real-time constraints from spectral DSE
- Software:
  - **`fredipy`**: python package for Fredholm inversion, coming soon!
  - Couple to **`STAN-HMC`** for faster hyperparameter integration

Thanks!