Reconstruction of bottomonium spectral functions in thermal QCD using Kernel Ridge Regression

Sam Offler

38th International Symposium on Lattice Field Theory

July 30, 2021

In collaboration with Gert Aarts, Chris Allton, Benjamin Jaeger, Seyong Kim, Maria Paola Lombardo, Ben Page, Sinead Ryan, Jon-Ivar Skullerud, Thomas Spriggs

Gen 2L Setup

- N_f = 2+1 Wilson quarks
- a_s = 0.1136 fm and a_τ^{-1} = 5.997 GeV
- M_π = 236 MeV
- Anisotropy, ζ = 3.45
- N_s = 32
- 47 MeV $\leq T \leq$ 375 MeV

See arXiv:2007.04188 (hep-lat)

NRQCD

- An effective field theory by expanding dispersion relation in powers of v.
- Terms up to v^4 are included.
- Correlators are generated from an initial value problem.
- All physics is contained within the spectral functions and correlators but former is harder to access.

$$G(\tau) = \int_{\omega_{min}}^{\omega_{max}} \mathrm{d}\omega\, K(\tau,\omega)\rho(\omega)$$

where $K(\tau,\omega) = e^{-\omega\tau}$

- Inversion is an ill-posed problem

Kernel Ridge Regression

- Linear case: $y = \mathbf{w}x \rightarrow$ Kernel case : $y = \mathbf{C}(x, x')\alpha$
- $\mathbf{w} \rightarrow \alpha$
- Relate this to lattice quantities by setting $y \rightarrow \rho(\omega)$ and $x \rightarrow G(\tau)$
- $\rho(\omega)$ can be parameterised by a vector $\mathbf{p}$.
- Kernel function contains "distance" between input data (L1, L2 norms, etc)
- α is determined by minimizing a cost function.

KRR

$$E = (\mathbf{p} - \mathbf{C}\alpha)^2 + \lambda|\mathrm{I}\alpha|^2$$

- λ is regularisation constant to prevent overfitting.
- We can determine α using:

$$\alpha = (\mathbf{C} + \mathrm{I}\lambda)^{-1}\mathbf{p}$$

where

$$C_{ij}(x_i, x_j) = \exp\left\{-\gamma \sum_{\tau} \left[\log(G_i(\tau)) - \log(G_j(\tau))\right]^2\right\}$$

- γ is scale length, determines "strength" of correlations.
- Finally predictions are made using $\mathbf{p} = C(x_{\mathrm{train}}, x_{\mathrm{NRQCD}})\alpha$

For a more thorough explanation see arXiv:1912.12900 [hep-lat]

Considerations

- We have two issues that need considering
- Is the training data appropriate?
- How do we optimize hyperparameters?

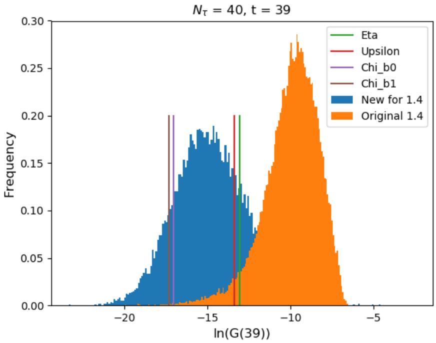
Data Generation

- To train the model we require mock data (spectral functions and corresponding correlators).
- The mock ρ were made from 5 Gaussians over a range of $a_\tau(\omega_{max} - \omega_{min}) = 1.4$.
- 20000 ρ were generated, each with 15 associated parameters.
- Corresponding correlators were calculated from

$$G(\tau) = \Delta\omega \sum_{\omega=\omega_{\min}}^{\omega_{\max}} K(\tau,\omega)\rho(\omega)$$

Data Generation

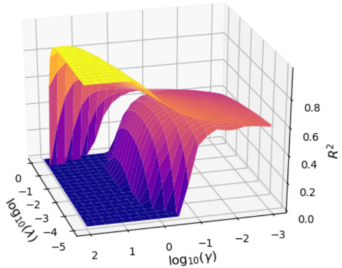
- Does this need improving and can it be improved?
- We compared these mock correlators to NRQCD data and selected the most similar.
- This was done for 4 channels and 10 values of N_τ to obtain a set of 'optimal' parameters for generating ρ.



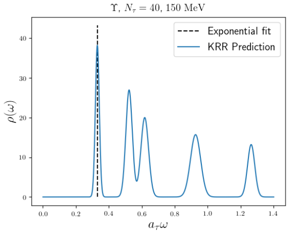
Hyperparameters.

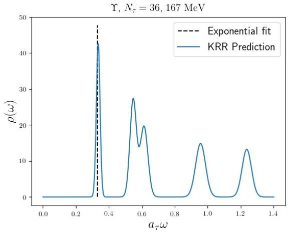
- KRR models are trained using a subset of the mock data.
- A second smaller test set is used to select the hyperparameters
- It is common to minimize the test set error to do this.
- Instead (λ, γ) were chosen by maximising R^2 for the test set, with $|R^2_{train} - R^2_{test}| < 0.01$.

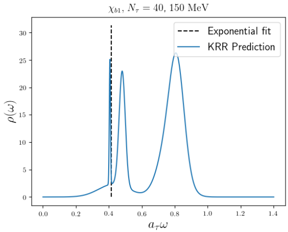
$$R^2 = 1 - \frac{\sum_i |y_i^{true} - y_i^{pred}|^2}{\sum_i |y_i^{true} - \bar{y}^{true}|^2}$$

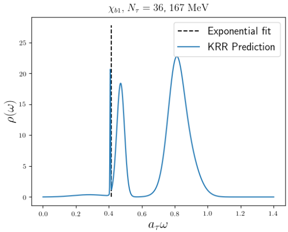


Results (Preliminary)









Summary

- We have compared mock data to actual NRQCD correlators and used this to generate more appropriate training functions.
- An alternate method for optimizing KRR hyperparameters has been discussed
- Groundstate state mass of Υ and χ_{b1} show agreement with exponential fit.

Further work

- Alternative forms for training data (Breit-Wigner functions, continuum, etc.).
- Introducing noise to the training step.

Thank you for listening