



News from bottomonium spectral functions in thermal QCD

Sam Offler*, **Gert Aarts**, **Chris Allton**, **Jonas Glesaaen**

Department of Physics, Swansea University, Swansea SA2 8PP, United Kingdom

E-mail: {s.p.offler.967106, g.aarts, c.allton}@swansea.ac.uk, jonas@glesaaen.com

Benjamin Jäger

CP3-Origins & Danish IAS, Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense M, Denmark

E-mail: jaeger@cp3.sdu.dk

Seyong Kim

Department of Physics, Sejong University, Seoul 143-747, Korea

E-mail: skim@sejong.ac.kr

Maria Paola Lombardo

INFN, Sezione di Firenze, 50019 Sesto Fiorentino (FI), Italy

E-mail: Mariapaola.Lombardo@lnf.infn.it

Sinead M. Ryan

School of Mathematics, Trinity College, Dublin 2, Ireland

E-mail: ryan@maths.tcd.ie

Jon-Ivar Skullerud

Dept. of Theoretical Physics, National University of Ireland Maynooth, County Kildare, Ireland

E-mail: jonivar@thphys.nuim.ie

New results on bottomonium at nonzero temperature are presented, using the FASTSUM Generation 2L ensembles. Preliminary results for spectral function reconstruction using Kernel Ridge Regression, a machine learning technique, are shown as well and compared to results from the Maximum Entropy Method.

37th International Symposium on Lattice Field Theory - Lattice2019

16-22 June 2019

Wuhan, China

*Speaker.

1. Introduction

Quarkonia, bound states of a heavy quark and anti-quark, are among the most-studied probes of the quark-gluon plasma, both experimentally and theoretically, see e.g. the reviews [1, 2]. In this contribution, we report on the latest preliminary results of the FASTSUM collaboration [3] for bottomonium, using lattice NRQCD on our new *Generation 2L* thermal ensembles. These calculations, with $N_f = 2 + 1$ flavours of Wilson-clover quark, are performed on anisotropic lattices, with $\xi = a_s/a_\tau = 3.453(6)$ and a fixed cutoff of $a_\tau^{-1} = 5.997(34)$ GeV, a pion mass of $m_\pi = 236(2)$ MeV, and a physical strange quark. Some details on the ensembles are given in Table 1 and a full discussion of Gen 2L is given in a companion contribution [5]. In particular, an estimate of the pseudocritical temperature from the renormalised chiral condensate is given by $T_{pc} = 162(1)$ MeV, which we will use as an indication for the crossover temperature.

The current study extends our previous Generation 1 [6–9] and especially Generation 2 [10] analysis, which differs from the Gen 2L ensembles by having a heavier pion, $m_\pi = 384(4)$ MeV. Hence one of the objectives of this study is to investigate the role of the light quarks. A second objective is to explore new methods for extracting spectral information from Euclidean correlators and we will present some very preliminary results obtained using a machine learning approach, namely Kernel Ridge Regression.

2. NRQCD correlation functions

We follow the approach discussed in detail in Refs. [7, 10]. Bottomonium correlators are generated by solving the NRQCD evolution equations for the b quark, whose mass is tuned via the non-relativistic dispersion relation of the spin-averaged $1S$ groundstate. Hadronic spectral quantities are determined up to an overall energy constant E_0 , which is fixed via the experimental value of the $\Upsilon(1S)$ mass. We find that $E_0 = 7464.5$ MeV. Some results in the Υ (S wave) and χ_{b1} (P wave) channels are shown here. In Fig. 1 we present the ratio of Euclidean correlators at a given temperature T to the correlator at $T_0 = 47$ MeV, the lowest temperature on which the bottomonium correlators were computed, allowing for a first indication of thermal effects. Fig. 2 contains temperature-dependent effective masses. In agreement with earlier observations [7–11], we note significant temperature dependence in P waves, not present in S waves. In fact, there is quantitative agreement with the Gen 2 results [10], indicating only a weak pion mass dependence.

N_τ	256*	128	64	56	48	40	36
T [MeV]	23	47	94	107	125	150	167
N_{cfg}	750	306	1041	1042	1123	1102	1119
N_τ	32	28	24	20	16	12	8
T [MeV]	187	214	250	300	375	500	750
N_{cfg}	1090	1031	1016	1030	1102	1267	1048

Table 1: Generation 2L ensembles, $m_\pi = 236(2)$ MeV, lattice size $32^3 \times N_\tau$, spatial lattice spacing $a_s = 0.1136(6)$ fm, temporal lattice spacing $a_\tau^{-1} = 5.997(34)$ GeV, anisotropy $a_s/a_\tau = 3.453(6)$. The choice of parameters and the ensemble at the lowest temperature are courtesy of the HadSpec collaboration [4].

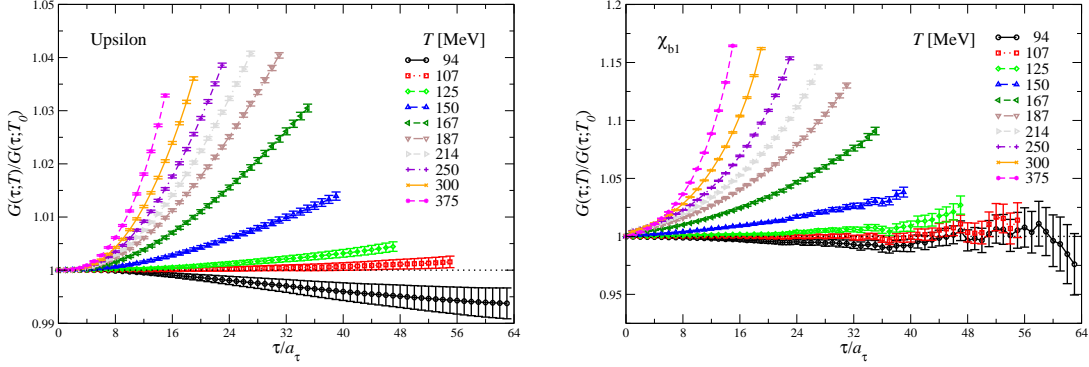


Figure 1: Thermal modification, $G(\tau;T)/G(\tau;T_0)$ with $T_0 = 47$ MeV, of the correlation functions in the Υ (left) and χ_{b1} (right) channels. Note the different vertical scale.

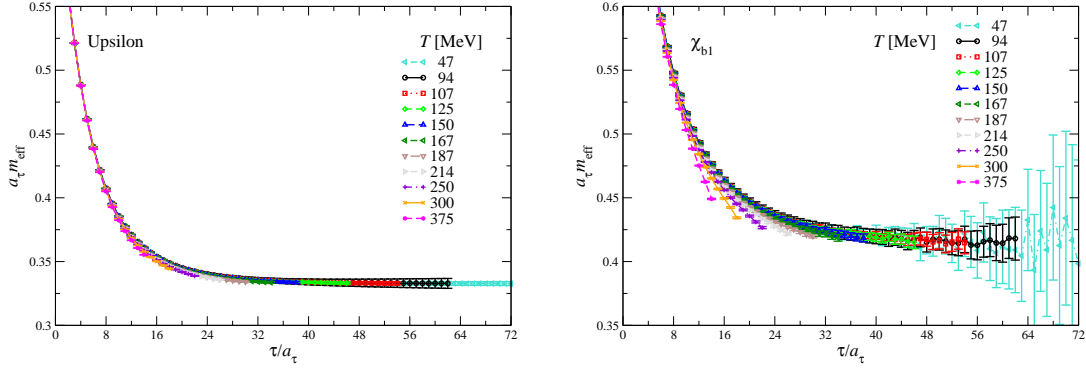


Figure 2: Effective mass in the Υ (left) and χ_{b1} (right) channels, for a range of temperatures. The resulting masses at the lowest temperature are $m(1S) = 9.460$ GeV and $m(1P) = 10.005$ GeV respectively. The latter is somewhat on the heavy side, which requires further investigation.

3. Kernel Ridge Regression

In order to gain a more detailed understanding of spectral modification in a thermal medium, it is of interest to extract the spectral function $\rho(\omega)$, related to the Euclidean correlator $G(\tau)$ via¹

$$G(\tau) = \int_{\omega_{\min}}^{\omega_{\max}} \frac{d\omega}{2\pi} K(\tau, \omega) \rho(\omega), \quad K(\tau, \omega) = e^{-\omega\tau}. \quad (3.1)$$

As is well known, determining $\rho(\omega)$ from a numerically computed correlator $G(\tau)$ at a finite number of temporal lattice points is an ill-posed inversion problem. This has been addressed using the Maximum Entropy Method (MEM) and other approaches (see Ref. [2] and references therein), but these methods are not yet entirely satisfactory as they suffer from systematic uncertainties.

¹The form of the kernel $K(\tau, \omega)$ depends on the problem; for NRQCD it is the simple exponential given here [6, 7].

In the last decade the field of machine learning has undergone significant development, with considerable success in image classification and reconstruction, given incomplete and noisy input data. Hence it makes sense to investigate whether machine learning methods can be applied to spectral reconstruction in the context of lattice QCD. If successful, this alternative approach can then be used to support the results of other methods such as MEM, with the potential to surpass it in the future. While there are many methods available, here we will focus on Kernel Ridge Regression (KRR), which has been used in Ref. [12] to address the analytic continuation problem in quantum many-body physics, using mock data only. An alternative Deep Neural Network approach has been tested on mock data in Ref. [13]. In this contribution, we will apply KRR to actual lattice data.

Kernel Ridge Regression, like linear regression, is used to determine a mapping between two datasets. However, KRR uses sets of functions as opposed to sets of values as in linear regression. In our context the KRR model is used to map a set of correlators to a set of spectral functions, i.e. to establish the relation (G, ρ) . The method for training a KRR model can be separated into three steps: generation of training data, determination of a parameter matrix α which relates the input and output data, and lastly selecting the optimal KRR hyperparameters. The final step is optional, but is done to improve predictions. We note here that the generation of training data is straightforward: one can construct an essentially unlimited set of spectral functions by combining multiple (narrow and broad) peaks and a continuum contribution at large energies, while imposing an upper limit $\omega < \omega_{\max}$ arising from the finite lattice cutoff [7]. In practice we generated the training set by combining Gaussians of various widths, while enforcing the constraints $\rho(\omega = \omega_{\min}) = 0$, $\rho(\omega) > 0$, and

$$\int_{\omega_{\min}}^{\omega_{\max}} \frac{d\omega}{2\pi} \rho(\omega) = G(\tau = 0) = \int d^3x S(\mathbf{x}), \quad (3.2)$$

where $S(\mathbf{x})$ is the source used in the NRQCD formulation, see Ref. [7] for details. Given these spectral functions, the associated Euclidean correlators follow from the easy-to-evaluate integral (3.1), which is performed either by integration or by summation, after discretising the ω interval using N_ω points.

One issue with spectral reconstruction, i.e. the inversion of Eq. (3.1), is the mismatch in the number of points N_ω needed to describe the spectral function $\rho(\omega)$, say $\mathcal{O}(1000)$, and the number of points available from the correlator $G(\tau)$, which is $\mathcal{O}(N_\tau)$. This can partly be avoided by expressing $\rho(\omega)$ in terms of an incomplete set of basis functions, for instance by employing the decomposition familiar from MEM [14],

$$\rho(\omega) = m(\omega) \exp \sum_{k=1}^{N_a} a_k f_k(\omega). \quad (3.3)$$

Here $m(\omega)$ is a polynomial in ω , a_k ($k = 1, \dots, N_a$) are a set of coefficients, and $f_k(\omega)$ form the set of basis functions, satisfying

$$\int_{\omega_{\min}}^{\omega_{\max}} d\omega f_k(\omega) f_l(\omega) = \delta_{kl}. \quad (3.4)$$

For the results shown here, we use a constant $m(\omega) = m_0$. The choice of basis functions is inspired by the construction used in MEM [14]. For this we consider the kernel $K(\tau_k, \omega_n)$, after discretising both τ and ω , as an $N_a \times N_\omega$ matrix, and generate the basis functions by taking the SVD of $K(\tau_k, \omega_n)$. The crucial difference with MEM is that N_a does not depend on the number of time

slices available at a given temperature, N_τ , but is a free parameter chosen such that $N_\tau \ll N_a \ll N_\omega$. Changing N_a allows for checks of stability. Given a spectral function in the training set and a constant $m(\omega) = m_0$, the coefficients a_k are then determined by

$$a_k = \int_{\omega_{\min}}^{\omega_{\max}} d\omega f_k(\omega) \log \left[\frac{\rho(\omega)}{m_0} \right]. \quad (3.5)$$

KRR will now establish the map between $G(\tau)$ and the set $\{a_k\}$.

To train the model, we note that KRR is a combination of two techniques. The first of these, the kernel method, can be treated as a generalised form of linear regression. In the linear case $y = w^T \phi(x)$, where y are the target data, $\phi(x)$ is a vector of functions of the input data x , and w is a vector of parameters [15]. The cost function to be minimized is $E = \frac{1}{2}(y - w^T \phi(x))^2$. In our application, the input data are a training set of Euclidean correlators $G_i(\tau)$ ($i = 1, \dots, N_{\text{train}}$). Rather than using these directly, they are combined in a $N_{\text{train}} \times N_{\text{train}}$ matrix \mathbf{C} , with

$$C_{ij} = \exp \left(-\frac{1}{\sigma^2} \sum_n [G_i(\tau_n) - G_j(\tau_n)]^2 \right), \quad (3.6)$$

where the sum goes over all time slices.² The quantity σ is a hyperparameter, which sets a correlation length in the space of correlators. The target data, i.e. the spectral functions, are encoded in a $N_{\text{train}} \times N_a$ matrix \mathbf{Y} , in which each row contains the N_a coefficients determining $\rho_i(\omega)$ according to Eq. (3.3). Finally, the input and target data are assumed to be related according to

$$\mathbf{Y} = \mathbf{C}\alpha, \quad (3.7)$$

which defines the $N_{\text{train}} \times N_a$ matrix α as the equivalent of w in linear regression. The aim of the training stage is to determine this α .

The second ingredient in KRR is ridge regression, which adds an additional term in the cost function, proportional to the square of the parameters, to prevent overfitting. This gives the following cost function

$$E(\mathbf{Y}, \mathbf{C}, \alpha) = \frac{1}{2}(\mathbf{Y} - \mathbf{C}\alpha)^2 + \frac{1}{2}\lambda \alpha^T \mathbf{C}\alpha, \quad (3.8)$$

where λ is the second hyperparameter, used to regularise the influence of the additional term. Minimising this cost function with respect to α then determines the optimal parameter matrix, for given σ and λ , as

$$\alpha_{\text{opt}} = (\mathbf{C} + \lambda \mathbf{I})^{-1} \mathbf{Y}. \quad (3.9)$$

The hyperparameters are determined via a cross-validation procedure. After this training stage, it is possible to make predictions for a spectral function (\mathbf{Y}') given an actual Euclidean correlator using

$$\mathbf{Y}' = \mathbf{C}' \alpha_{\text{opt}}, \quad (3.10)$$

where \mathbf{C}' is determined from the squared difference between the NRQCD correlator and the training correlators, i.e. it is a matrix of size $1 \times N_{\text{train}}$.

²The actual form of the kernel function can be varied with the only restriction being that it determines a difference between the input data [15]. It should not be confused with the integral kernel appearing in Eq. (3.1).

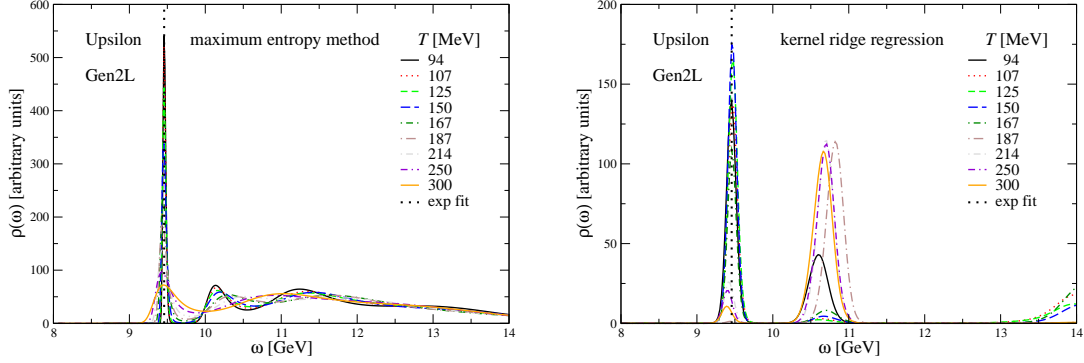


Figure 3: Spectral functions, obtained with the Maximum Entropy Method (MEM, left) and Kernel Ridge Regression (KRR, right) using $N_a = 80$ basis functions, in the Υ channel for a number of temperatures, in Gen 2L. The dotted line indicates the mass at the lowest temperature, obtained used a conventional fit.

First preliminary results of this approach are presented in Fig. 3, in the Υ channel, comparing MEM (left) and KRR (right). The latter are obtained using $N_a = 80$ coefficients to parametrise the spectral functions, at all temperatures. The size of the training set is $N_{\text{train}} = 9000$. Training is repeated at every temperature, using the same training set $\{\rho_i(\omega)\}$, but for correlators appropriate at that temperature, i.e. with given N_τ . We note a clear first peak, whose position coincides with the mass obtained using a standard exponential fit at the lowest temperature, indicated with the dotted line. As the temperature increases, the size of the first peak drops above $T = 187$ MeV, in both MEM and KRR. Given the difference between the two approaches, this is an interesting observation. The area under the spectral function is the same at all temperature, see Eq. (3.2); hence a reduction of the size of the first peak leads to more spectral weight at higher energies, both in MEM and KRR. In the case of KRR this leads to a more pronounced second peak. However, the second peak is at this moment not yet reliable and requires further study.

4. Summary

We have shown new results for bottomonium on the FASTSUM Gen 2L ensembles, finding quantitative agreement with previous Gen 2 results. Further analysis will be carried out shortly, including the application of the methods developed in Ref. [16]. We have presented preliminary results for spectral reconstruction using a machine learning technique, namely Kernel Ridge Regression, applied to actual lattice QCD data for the first time. The position of the first peak was seen to coincide with the one obtained from conventional fitting and from MEM, with a drop of the peak size at higher temperature seen both with KRR and MEM. Even though much remains to be done, we take these first results as a positive stimulus for the future.

Acknowledgments. We are grateful for support from STFC via grants ST/L000369/1 and ST/P00-055X/1, the Swansea Academy for Advanced Computing, SNF, ICHEC, COST Action CA15213

THOR, and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 813942. Computing resources were made available by HPC Wales, Supercomputing Wales, and PRACE via access to the Marconi-KNL system hosted by CINECA, Italy. This work used the DiRAC Extreme Scaling service and the DiRAC Blue Gene Q Shared Petaflop system at the University of Edinburgh, operated by the Edinburgh Parallel Computing Centre on behalf of the STFC DiRAC HPC Facility. This equipment was funded by by BIS National E-infrastructure capital grant ST/K000411/1, STFC capital grant ST/H008845/1, and STFC DiRAC Operations grants ST/K005804/1 and ST/K005790/1, and also through BEIS capital funding via STFC capital grant ST/R00238X/1 and STFC DiRAC Operations grant ST/R001006/1. DiRAC is part of the National e-Infrastructure.

References

- [1] G. Aarts *et al.*, Eur. Phys. J. A **53** (2017) no.5, 93 [arXiv:1612.08032 [nucl-th]].
- [2] A. Rothkopf, arXiv:1912.02253 [hep-ph].
- [3] FASTSUM collaboration, fastsum.gitlab.io/
- [4] D. J. Wilson, R. A. Bričeno, J. J. Dudek, R. G. Edwards and C. E. Thomas, Phys. Rev. D **92** (2015) no.9, 094502 [arXiv:1507.02599 [hep-ph]]; G. K. C. Cheung *et al.* [Hadron Spectrum Collaboration], JHEP **1612** (2016) 089 [arXiv:1610.01073 [hep-lat]].
- [5] G. Aarts *et al.*, PoS LATTICE **2019** (2019) 075 [arXiv:1912.09827 [hep-lat]].
- [6] G. Aarts, S. Kim, M. P. Lombardo, M. B. Oktay, S. M. Ryan, D. K. Sinclair and J.-I. Skullerud, Phys. Rev. Lett. **106** (2011) 061602 [arXiv:1010.3725 [hep-lat]].
- [7] G. Aarts, C. Allton, S. Kim, M. P. Lombardo, M. B. Oktay, S. M. Ryan, D. K. Sinclair and J. I. Skullerud, JHEP **1111** (2011) 103 [arXiv:1109.4496 [hep-lat]].
- [8] G. Aarts, C. Allton, S. Kim, M. P. Lombardo, M. B. Oktay, S. M. Ryan, D. K. Sinclair and J. I. Skullerud, JHEP **1303** (2013) 084 [arXiv:1210.2903 [hep-lat]].
- [9] G. Aarts, C. Allton, S. Kim, M. P. Lombardo, S. M. Ryan and J.-I. Skullerud, JHEP **1312** (2013) 064 [arXiv:1310.5467 [hep-lat]].
- [10] G. Aarts, C. Allton, T. Harris, S. Kim, M. P. Lombardo, S. M. Ryan and J. I. Skullerud, JHEP **1407** (2014) 097 [arXiv:1402.6210 [hep-lat]].
- [11] S. Kim, P. Petreczky and A. Rothkopf, Phys. Rev. D **91** (2015) 054511 [arXiv:1409.3630 [hep-lat]]; JHEP **1811** (2018) 088 [arXiv:1808.08781 [hep-lat]].
- [12] L.-F. Arsenault, R. Neuberg, L.A. Hannah and A.J. Millis, Inverse Problems **33** (2017) 115007 [arXiv:1612.04895 [cond-mat]].
- [13] L. Kades, J. M. Pawłowski, A. Rothkopf, M. Scherzer, J. M. Urban, S. J. Wetzel, N. Wink and F. Ziegler, arXiv:1905.04305 [physics.comp-ph].
- [14] G. Aarts, C. Allton, J. Foley, S. Hands and S. Kim, Phys. Rev. Lett. **99** (2007) 022002 [hep-lat/0703008 [HEP-LAT]].
- [15] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag New York Inc.
- [16] R. Larsen, S. Meinel, S. Mukherjee and P. Petreczky, Phys. Rev. D **100** (2019) no.7, 074506 [arXiv:1908.08437 [hep-lat]]; Phys. Lett. B **800** (2020) 135119 [arXiv:1910.07374 [hep-lat]].