Sparse modeling study to extract spectral functions from lattice QCD data

Junichi Takahashi

Meteorological College, Japan Meteorological Agency

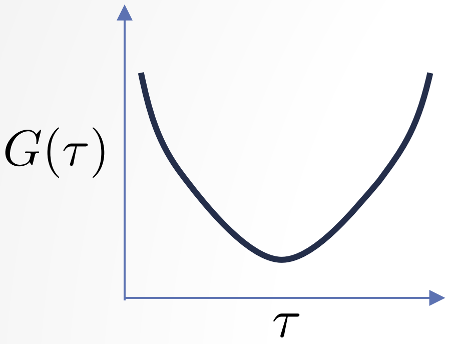
Hiroshi Ohno (University of Tsukuba), Akio Tomiya (TWCU)

Lattice 2024

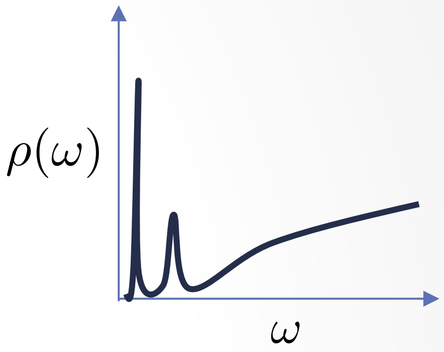
Liverpool, UK, July 29th, 2024

Imaginary time correlation function and spectral function

Imaginary time correlation function



Spectral function



Analytical continuation

$$G(\tau) = \int_0^\infty d\omega K(\tau,\omega)\rho(\omega)$$

Discretized & schematically

$$N_\tau \sim O(10) \left\{ \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_{N_\tau} \end{pmatrix} \right. = \begin{pmatrix} K_{11} & K_{12} & \cdots & \cdots & K_{1N_\omega} \\ K_{21} & K_{22} & \cdots & \cdots & K_{2N_\omega} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ K_{N_\tau 1} & \cdots & \cdots & \cdots & K_{N_\tau N_\omega} \end{pmatrix} \left. \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_{N_\omega} \end{pmatrix} \right\} N_\omega \sim O(1000)$$

Extracting spectral functions is an ill-posed inverse problem.

Spectral function extracted from lattice QCD data

- Previous works (not inclusive)
  - ✓ Maximum entropy method (MEM)
    [M. Asakawa, T. Hatsuda and Y. Nakahara, Prog. Part. Nucl. Phys. 46 (2001) 459-508]
  - ✓ Stochastic method
    [H.-T. Ding, et al., Phys. Rev. D 97, 094503 (2018)]
  - ✓ Backus Gilbert method
    [B. B. Brandt, A. Francis, H. B. Meyer, and D. Robaina, Phys. Rev. D 92, 094510 (2015)]
  - ✓ Sparse modeling (SpM)
    [E. Itou, Y. Nagai, J. High Energ. Phys. 2020, 7 (2020)]

> It is important to check the spectral function with each other in various ways and to properly estimate the systematic error.

- Our presentation @ Lattice2023
  - ➢ We checked the applicability of SpM by performing mock data tests, which had never been done before for the calculation of spectral functions using SpM.
  - ➢ We tried to extract spectral functions from mean values of charmonium correlation functions in the vector channel obtained from lattice QCD.

Update from our previous study

- Mock data test
  - ➢ Using correlation functions reconstructed from mock charmonium spectral functions.
  - ➢ Manually adding noises corresponding to the errors to the correlation functions and preparing several of them.

    ⇧

    The mean value and covariance matrix of the correlation functions can be considered as in actual lattice calculations.

- Spectral function from lattice QCD data
  - ➢ Extracting the spectral functions in the pseudoscalar channel in addition to the vector channel.
  - ➢ Estimating errors of the spectral functions by Jackknife analyses.

Sparse modeling

- Extracting spectral functions by using SpM has been proposed in condensed matter physics. [H. Shinaoka, J. Otsuki, M. Ohzeki, K. Yoshimi, Phys. Rev. B 96 (2017) 035147] [J. Otsuki, M. Ohzeki, H. Shinaoka, K. Yoshimi, Phys. Rev. E 95 (2017) 061302]

  - Fitting with different regularization in MEM (we do not use default models.)

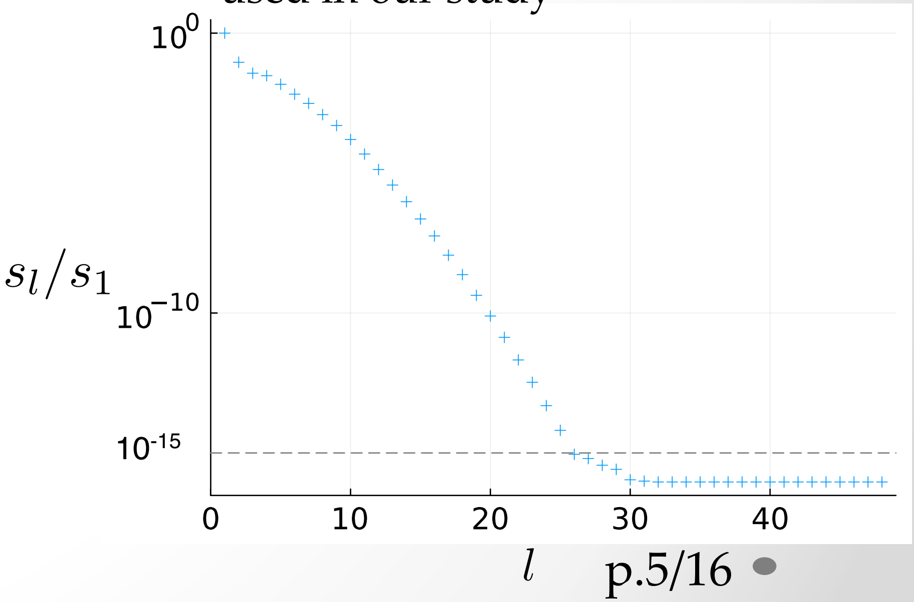
- The ranks of the spectral function and the correlation function are reduced by dropping the contribution of small singular values s_l.

Singular value decomposition of kernel

$$K = USV^{\mathrm{t}} \longrightarrow \boldsymbol{\rho}' \equiv V^{\mathrm{t}}\boldsymbol{\rho} \quad \boldsymbol{G}' \equiv U^{\mathrm{t}}\boldsymbol{G}$$

➢ In our study, the components of **ρ**' and **G**' corresponding to small singular values satisfied with $s_l/s_1 < 10^{-15}$ dropped.

✓ An example of singular values of K used in our study



Sparse modeling

- Cost function: χ^2 term + L1 regularization term (LASSO form problem)

$$F(\boldsymbol{\rho}') = \frac{1}{2}(\boldsymbol{G}' - S\boldsymbol{\rho}')^2 + \lambda||\boldsymbol{\rho}'||_1$$

L1 norm: $\|\boldsymbol{\rho}'\|_1 \equiv \sum_l |\rho'_l|$

λ is a positive hyperparameter.

- This optimization problem is solved iteratively by alternating direction method of multipliers (ADMM). [S. Boyd, et al., Foundations and Trends R in Machine Learning 3, 1 (2011)]
  - The problem is solved at various λ and the most likely spectral function ρ can be found at the optimal λ.
  - In our study, estimation of the optimal λ is same as the previous study.

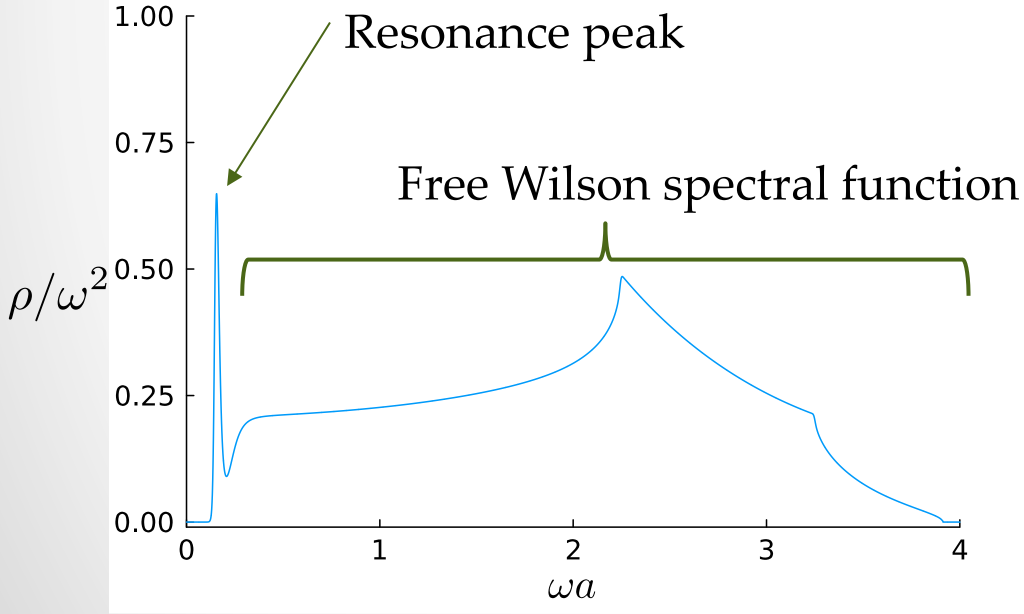
[E. Itou, Y. Nagai, J. High Energ. Phys. 2020, 7 (2020)]

See backup slides for detail.

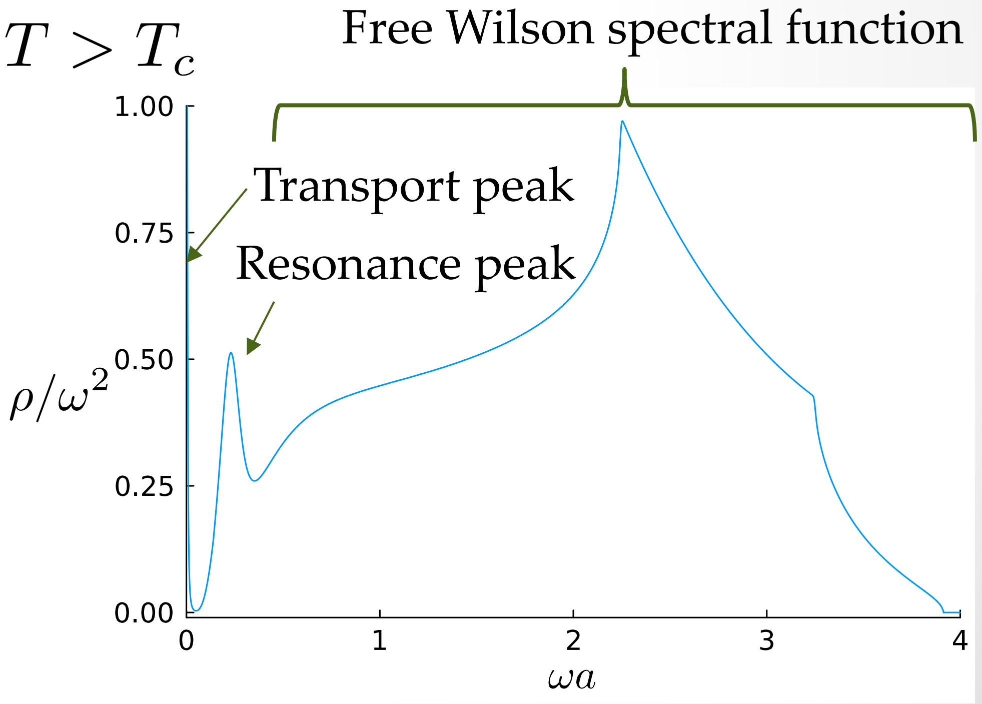
Mock data test

- Mock spectral functions: [H.-T. Ding, et al., Phys. Rev. D 97, 094503 (2018)]
  - ➢ $T < T_c$: resonance peak (J/ψ meson mass) + free Wilson spectral function
  - ➢ $T > T_c$: transport peak + broader resonance peak + free Wilson spectral function
- The range of ω: $0 \leq \omega a \leq 4$ (a: lattice spacing), # of ω points: N_ω=8001

$T < T_c$



$T > T_c$



Mock data test

- The central values of correlation function G(τ)

$$G(\tau) = \int d\omega \rho(\omega) K(\omega, \tau)$$

  - # of τ points: $N_\tau = 48,\ 64,\ 96$

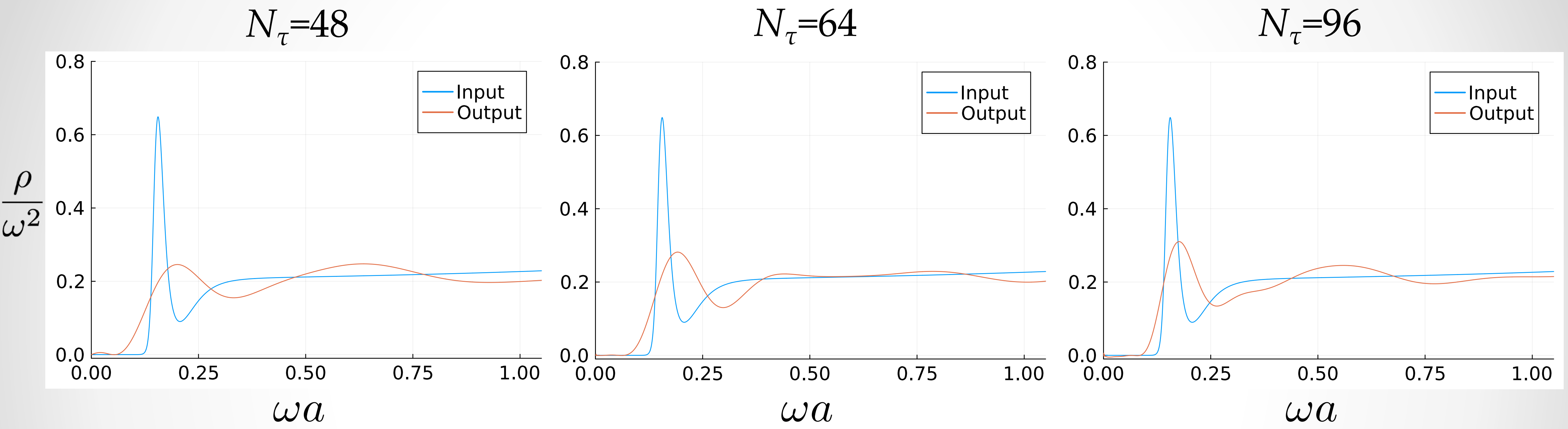
- Errors of G(τ) are generated by gaussian random numbers
  - Variance: $\sigma(\tau) = \epsilon \cdot \tau \cdot G(\tau)$
  - Consider three types of noise level: $\epsilon = 10^{-2},\ 5 \times 10^{-3},\ 10^{-5}$
  - Prepare G(τ) with errors ($N_{\rm conf}$=300)

    ⇨ calculate covariance matrix

$$C_{ij} = \frac{1}{N_{\rm conf}(N_{\rm conf} - 1)} \sum_{n=1}^{N_{\rm conf}} \left(G(\tau_i) - G^{(n)}(\tau_i)\right) \left(G(\tau_j) - G^{(n)}(\tau_j)\right)$$

$$G(\tau_i) = \frac{1}{N_{\rm conf}} \sum_{n=1}^{N_{\rm conf}} G^{(n)}(\tau_i),$$

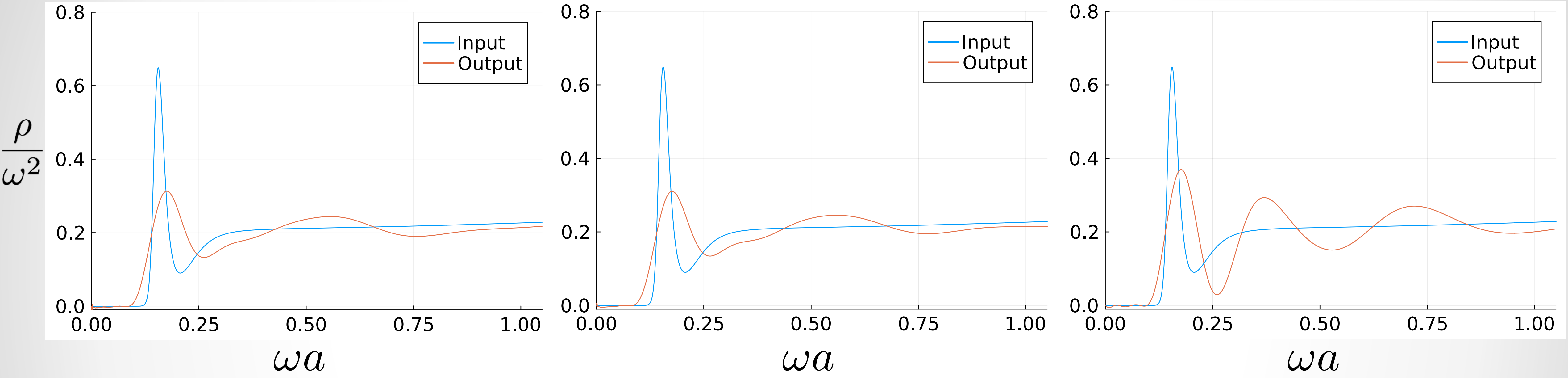
Results of mock data test (T<T$_c$, noise level $\varepsilon=5\times10^{-3}$)



- The resonance peak becomes sharper as N_τ increases.
- Similar results are obtained for other noise levels.

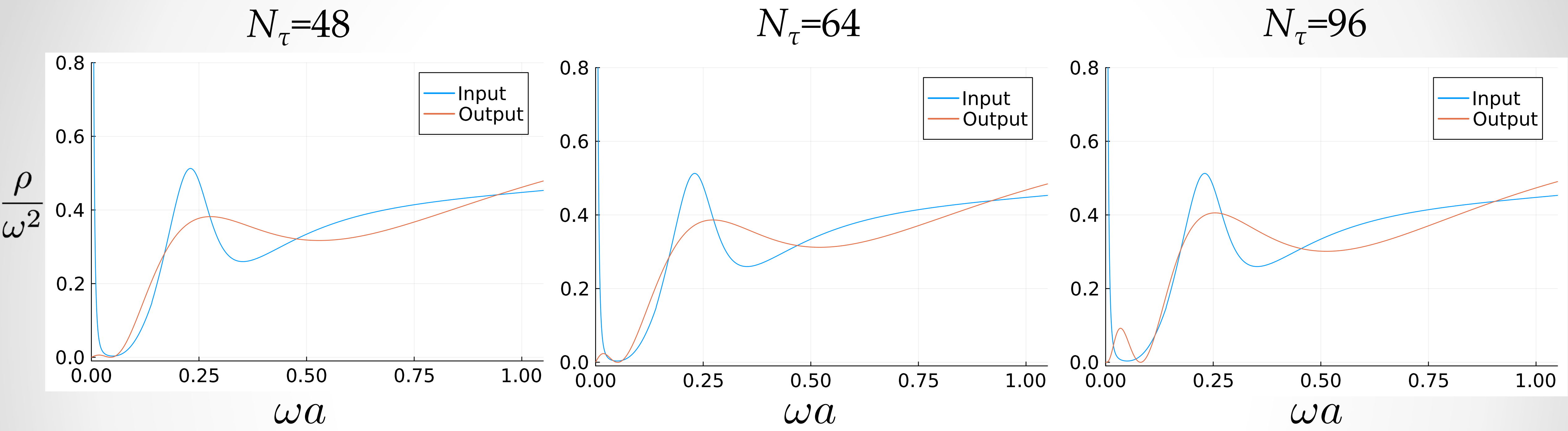
Results of mock data test (T<T$_c$, N_τ=96)

Noise level $\varepsilon = 10^{-2}$ $\varepsilon = 5\times10^{-3}$ $\varepsilon = 10^{-5}$



- The resonance peak becomes sharper as noise level ε decreases.
- Similar results are obtained for other N_τ.

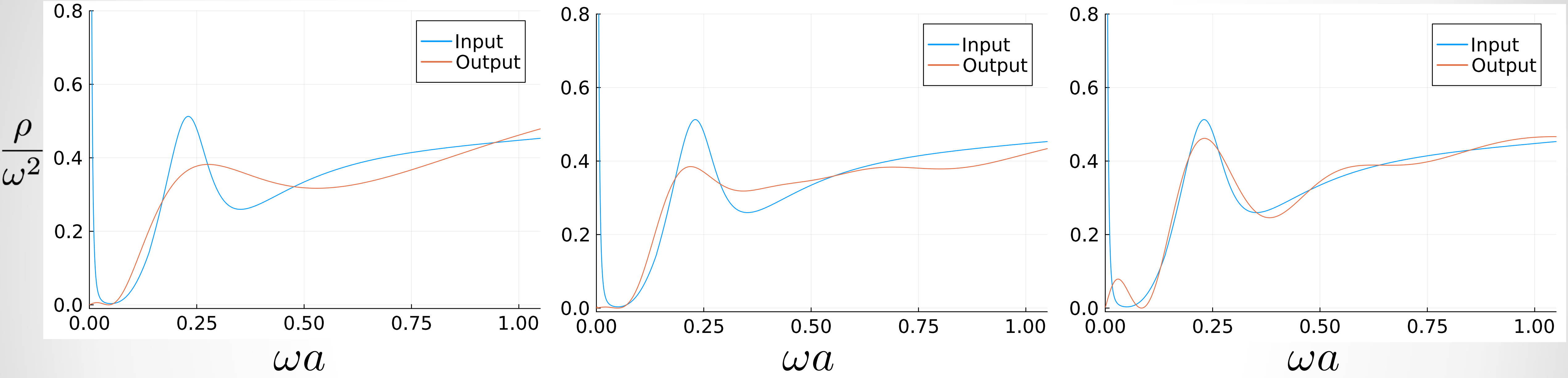
Results of mock data test (T>T$_c$, noise level ε=10^{-2})



- The resonance peak becomes sharper as N_τ increases.
- The transport peaks are not reproduced.
- Similar results are obtained for other noise levels.

Results of mock data test (T>T_c, N_τ=48)

Noise level $\varepsilon=10^{-2}$ $\quad$ $\varepsilon=5\times10^{-3}$ $\quad$ $\varepsilon=10^{-5}$



- The resonance peak becomes sharper as noise level ε decreases.
- Similar results are obtained for other N_τ.

Lattice QCD data

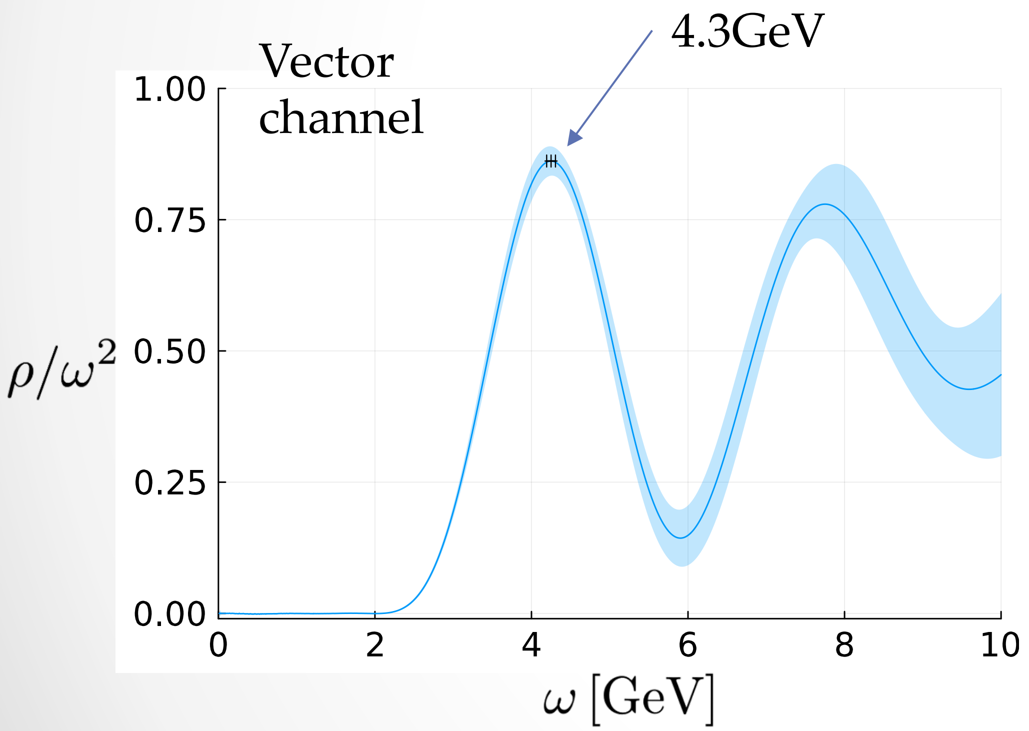
- Standard plaquette gauge + O(a)-improved Wilson quark action
- In the quenched approximation
- Lattice spacing: $a = 0.010\,\mathrm{fm},\ a^{-1} \simeq 18.97\,\mathrm{GeV}$
- Spatial and temporal extents: $N_\sigma = 128,\ N_\tau = 96,\ 48$ for $T/T_c = 0.73,\ 1.46$
- Charmonium correlation functions in pseudoscalar channel and vector channel
- # of conf.: 234 $(N_\tau = 96)$, 461 $(N_\tau = 48)$

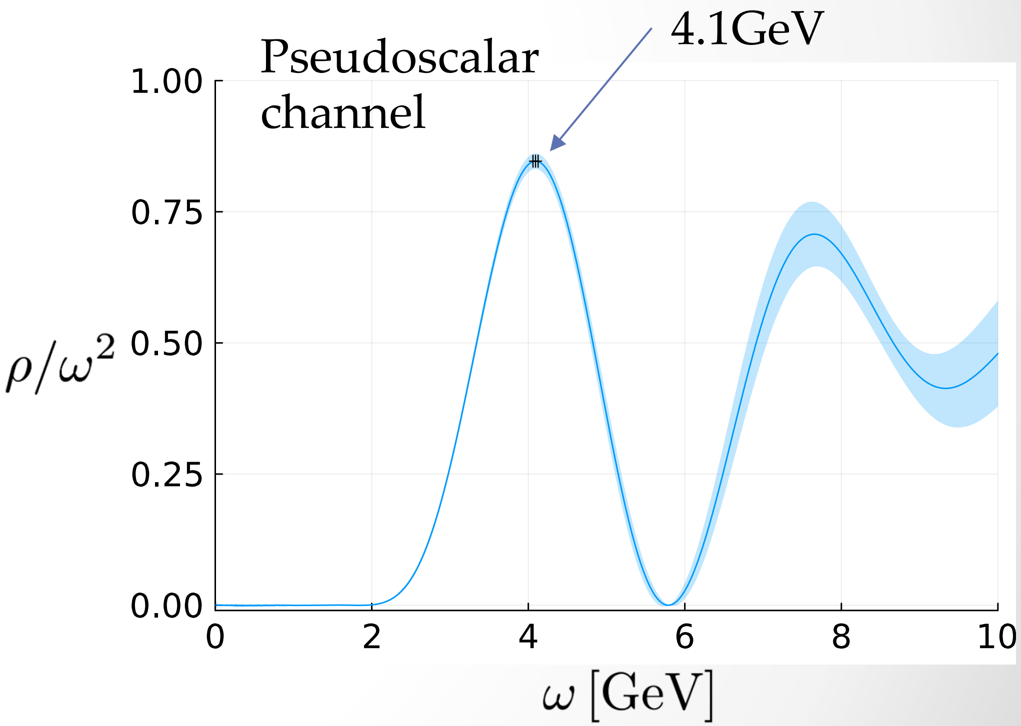
  [H.-T. Ding, A. Francis, O. Kaczmarek, F. Karsch, H. Satz, W. Soeldner, Phys. Rev. D 86, 014509 (2012)]

- Statistical errors: Jackknife analyses

Results from lattice QCD data ($T=0.73T_c$)

- Broad peak about 4GeV.
  - ➢ Results of MEM: 3.48GeV (vector), 3.31GeV (pseudoscalar) [H.-T. Ding, A. Francis, O. Kaczmarek, F. Karsch, H. Satz, W. Soeldner, Phys. Rev. D 86, 014509 (2012)]





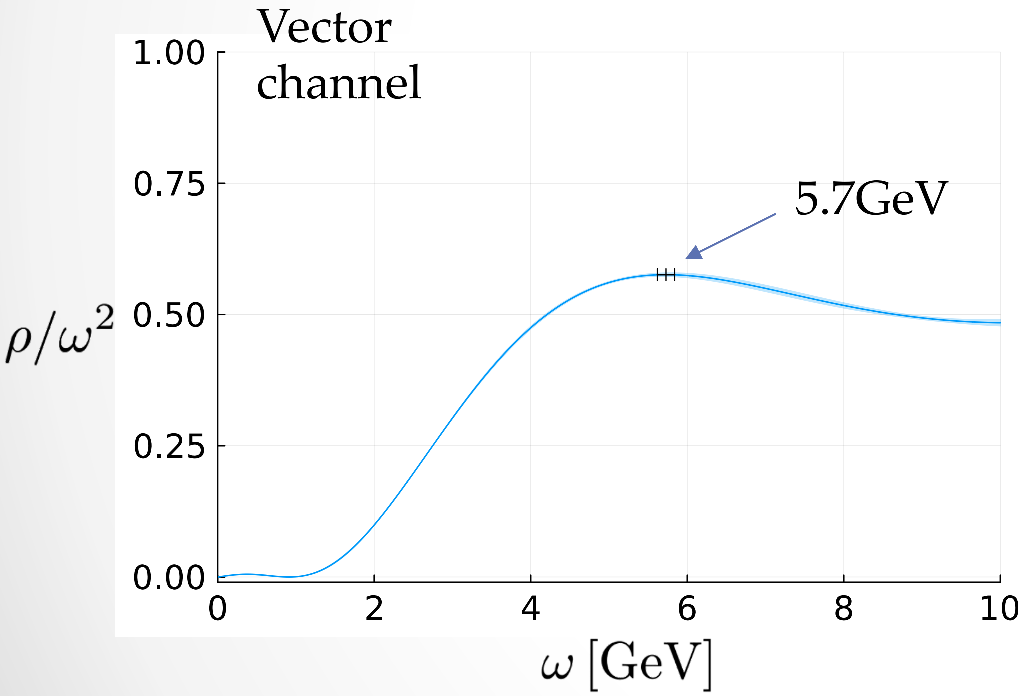
Shaded areas: statistical errors, solid lines: mean values,
horizontal error bars: statistical uncertainties of the peak location

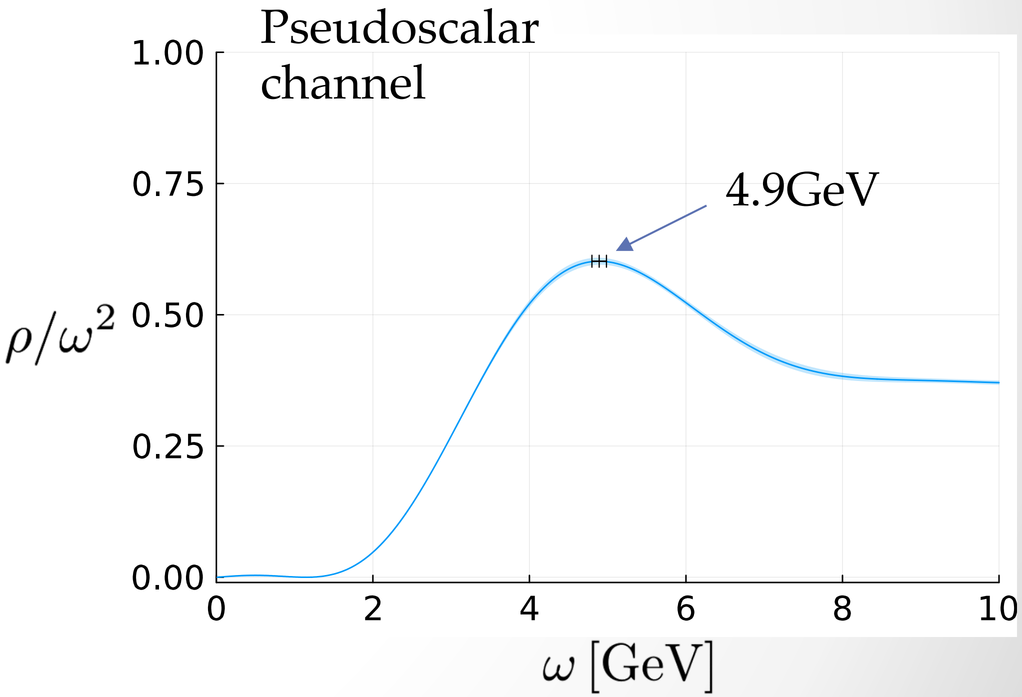
Results from lattice QCD data ($T=1.46T_c$)

- The resonance peaks become much broader and are shifted higher energy (~5GeV).
  - ➢ Results of MEM: 4.7GeV (vector), 4.1GeV (pseudoscalar)

[H.-T. Ding, A. Francis, O. Kaczmarek, F. Karsch, H. Satz, W. Soeldner, Phys. Rev. D 86, 014509 (2012)]

- The transport peaks are not appeared.





Shaded areas: statistical errors, solid lines: mean values,
horizontal error bars: statistical uncertainties of the peak location

Summary

- Sparse modeling (SpM) is a useful method for obtaining a reasonable solution to an ill-posed inverse problem such as the extraction of spectral functions.
- We applied SpM to extract spectral functions from charmonium correlation functions.
- Mock data test:
  - ➢ The longer N_τ or the smaller the noise level ε, the better the spectral function can be reproduced.
- Results from lattice QCD data:
  - ➢ Although the position and width of the peaks are not the same as the MEM results, the qualitative behavior is similar. ⇨ Reflecting model-independent properties.
- To estimate the transport peak, assumptions beyond SpM, such as the shape of the transport peak, might be needed.

Sparse modeling in our study

1. Calculate the covariance matrix C and carry out the Cholesky decomposition:

$$C^{-1} = W^{\mathrm{t}} W$$

2. Transform G and K by W: $\boldsymbol{G}_W = W\boldsymbol{G},\ K_W = WK$

3. Carry out the singular value decomposition of K_W: $K_W = USV^{\mathrm{t}}$

S is an Nτ×Nω diagonal matrix. U and V are Nτ×Nτ and Nω×Nω orthogonal matrices, respectively.

Nτ and Nω are the # of points of G(τ) and ρ(ω), respectively.

4. Transform the basis of G_W and ρ by U^t and V^t, respectively:

$$\boldsymbol{\rho}'_W = V^{\mathrm{t}}\boldsymbol{\rho},\ \boldsymbol{G}'_W = U^{\mathrm{t}}\boldsymbol{G}_W$$

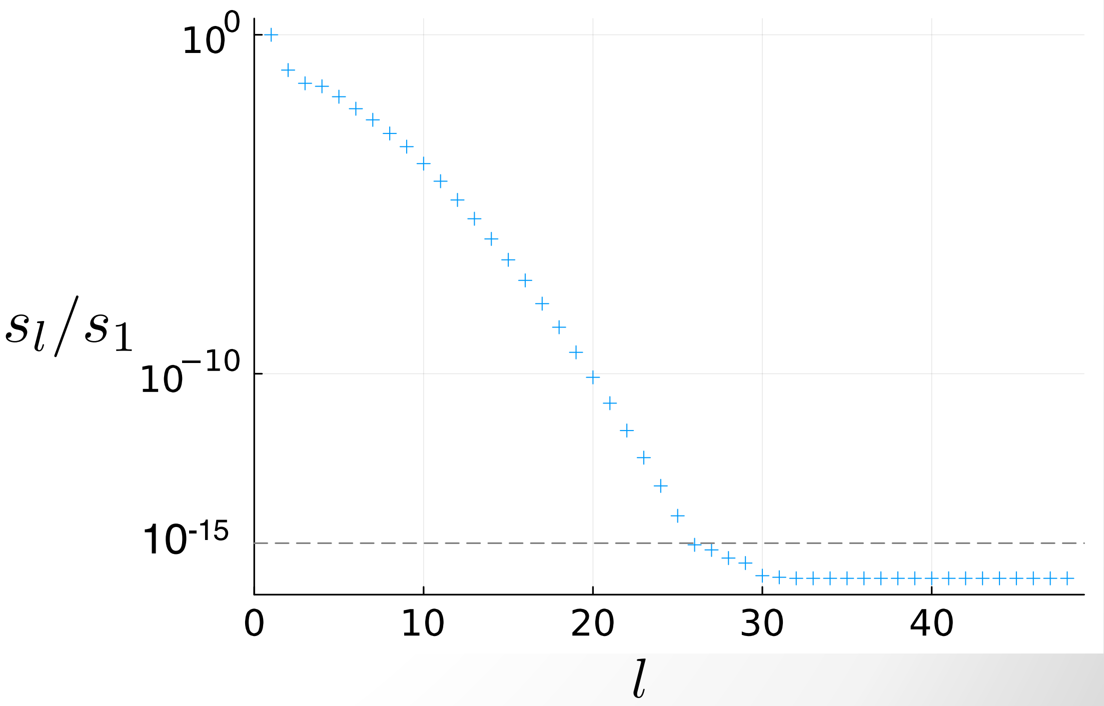
Sparse modeling in our study

5. The components of $\boldsymbol{\rho}_W$' and $\mathbf{G}_W$' corresponding to small singular values satisfied with $s_l/s_1 < 10^{-15}$ are dropped.

   - At the same time, the sizes of U, S and V are reduced.

$$\begin{cases} S : N_\tau \times N_\omega \rightarrow L \times L, \\ U : N_\tau \times N_\tau \rightarrow N_\tau \times L, \\ V : N_\omega \times N_\omega \rightarrow N_\omega \times L \end{cases}$$

L is the number of components of singular values satisfied with $s_l/s_1 < 10^{-15}$.

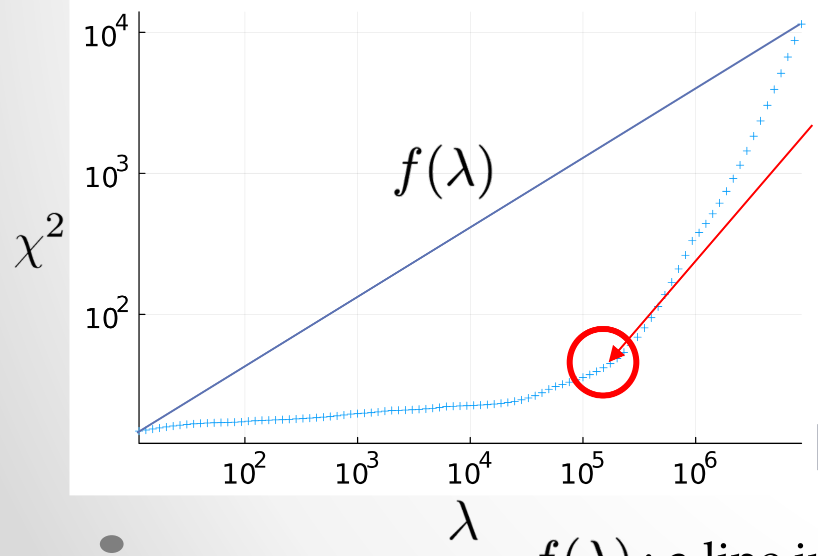


Sparse modeling in our study

6. The cost function F(ρ’) consists of the square error and the L1 regularization term.

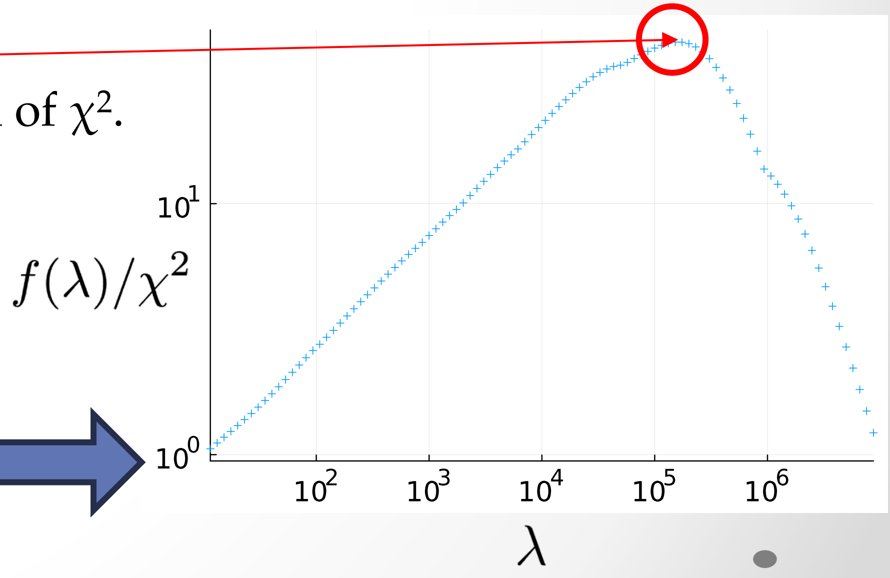
$$F(\boldsymbol{\rho}'_W) = \frac{1}{2}(\boldsymbol{G}'_W - S\boldsymbol{\rho}'_W)^2 + \lambda||\boldsymbol{\rho}'_W||_1 \equiv \chi^2(\boldsymbol{\rho}'_W) + \lambda||\boldsymbol{\rho}'_W||_1$$

7. Solve the optimization problem at various λ by using ADMM and find the most likely spectral function **ρ**.



An optimal value λ_{opt}:
The value of λ at the kink position of χ^2.
(Estimation of λ_{opt} is same as the previous study.)

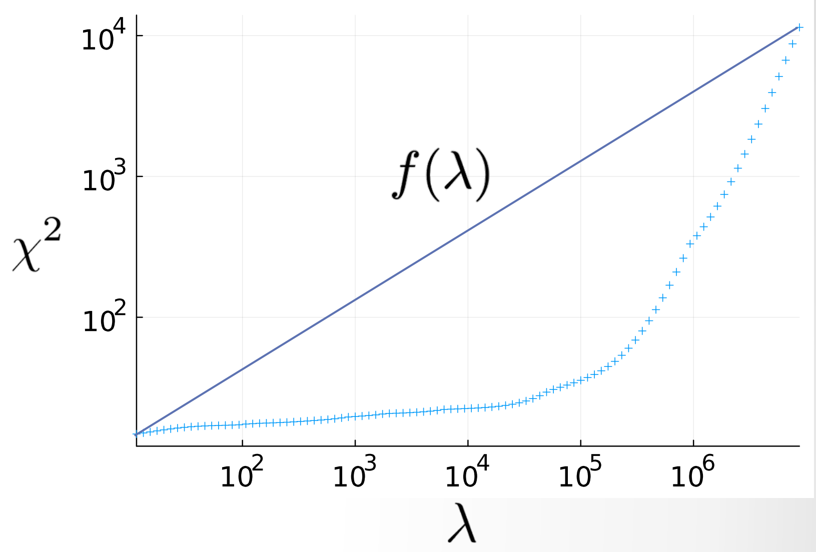
[E. Itou, Y. Nagai, J. High Energ. Phys. 2020, 7 (2020)]

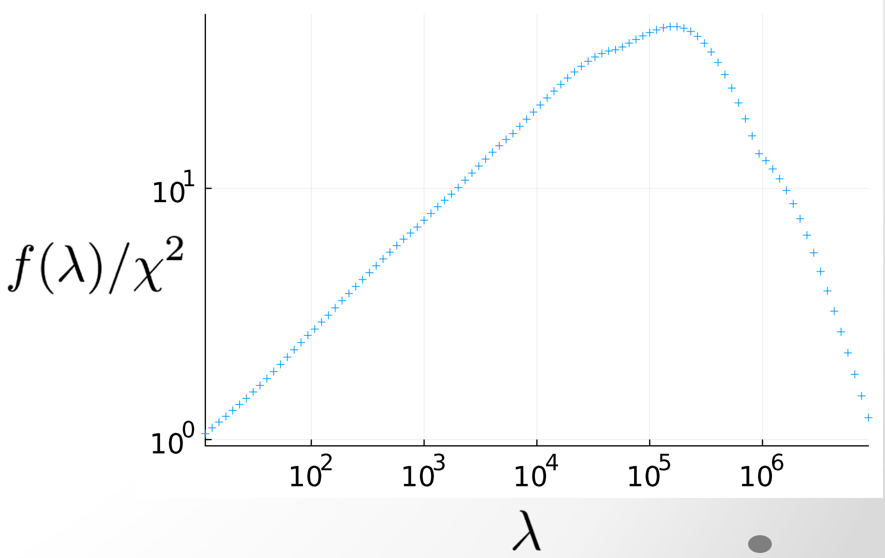


$f(\lambda)$: a line in log-log scale

Searching λ_{opt}

1. Fix a range of λ, $[\lambda_{min}, \lambda_{max}]$
2. Calculate $\chi^2(\rho')$ for each λ by using ADMM iterations
   - ✓ # of iterations: 10000
3. Obtain a function f(λ) in log-log scale by connecting $f(\lambda_{min})$ with $f(\lambda_{max})$
4. Calculate the ratio $f(\lambda)/\chi^2$
   - ✓ The λ located at the peak position of $f(\lambda)/\chi^2$ corresponds to λ_{opt}.





Mock spectral function

- $\mathrm{T < T_c}$: $\hat{\rho}_{T<T_c}(\hat{\omega}) = \tilde{\Theta}(\hat{\omega},\hat{\omega}_1,\Delta_1)(1-\tilde{\Theta}(\hat{\omega},\hat{\omega}_2,\Delta_2))\hat{\rho}_{res}(\hat{\omega}) + \tilde{\Theta}(\hat{\omega},\hat{\omega}_3,\Delta_3)\hat{\rho}_{Wilson}(\hat{\omega})$
- $\mathrm{T > T_c}$: $\hat{\rho}_{T>T_c}(\hat{\omega}) = \hat{\rho}_{trans}(\hat{\omega}) + \tilde{\Theta}(\hat{\omega},\hat{\omega}_4,\Delta_4)(1-\tilde{\Theta}(\hat{\omega},\hat{\omega}_5,\Delta_5))\hat{\rho}_{res}(\hat{\omega}) + \tilde{\Theta}(\hat{\omega},\hat{\omega}_6,\Delta_6)\hat{\rho}_{Wilson}(\hat{\omega})$

✓ Transport peak: $\hat{\rho}_{trans}(\hat{\omega}) = c_{trans}\dfrac{\hat{\omega}\eta}{\hat{\omega}^2+\eta^2}$ ✓ Resonance peak: $\hat{\rho}_{res}(\hat{\omega}) = c_{res}\dfrac{\Gamma M\hat{\omega}^2}{(\hat{\omega}^2-M^2)^2+M^2\Gamma^2}$

✓ Free Wilson spectral function:

$$\hat{\rho}_{Wilson}(\hat{\omega}) = c_{Wilson}\frac{4\pi N_c}{N_\sigma^3}\sum_{\vec{k}}\sinh\left(\frac{\hat{\omega}}{2\hat{T}}\right)\left[b^{(1)} - b^{(2)}\frac{\sum_{i=1}^3 \sin^2 k_i}{\sinh^2 E_{\vec{k}}(m)}\right]\frac{\delta(\hat{\omega}-2E_{\vec{k}}(m))}{2(1+M_{\vec{k}}(m))^2\cosh^2\left(\frac{E_{\vec{k}}(m)}{2\hat{T}}\right)}$$

$$\cosh E_{\vec{k}}(m) = 1 + \frac{K_{\vec{k}}^2 + M_{\vec{k}}^2(m)}{2(1+M_{\vec{k}}(m))}, \qquad K_{\vec{k}} = \sum_{i=1}^3 \gamma_i \sin k_i, \qquad M_{\vec{k}}(m) = \sum_{i=1}^3 (1-\cos k_i) + m$$

✓ Modified theta function: $\tilde{\Theta}(\hat{\omega},\hat{\omega}_i,\Delta_i) = \left(1+\exp\left(\dfrac{\hat{\omega}_i^2-\hat{\omega}^2}{\hat{\omega}\Delta_i}\right)\right)^{-1}$

$\hat{\omega} = \omega a,\ \hat{\rho} = \rho a^2$

Spectral function	Parameters
$\hat{\rho}_{res}$ for $T < T_c$	$c_{res} = 0.08/7,\ \Gamma = 0.05,\ M = 0.155$
$\hat{\rho}_{Wilson}$ for $T < T_c$	$c_{Wilson} = 0.5,\ b^{(1)} = 2,\ b^{(2)} = 1,\ m = 0.073,\ N_c = 3,\ N_\sigma = 4096$
$\hat{\rho}_{trans}$ for $T > T_c$	$c_{trans} = 5\times10^{-5},\ \eta = 0.006$
$\hat{\rho}_{res}$ for $T > T_c$	$c_{res} = 0.06,\ \Gamma = 0.15,\ M = 0.225$
$\hat{\rho}_{Wilson}$ for $T > T_c$	$c_{Wilson} = 1,\ b^{(1)} = 3,\ b^{(2)} = 1,\ m = 0.073,\ N_c = 3,\ N_\sigma = 4096$

$\hat{\omega}_1 = 0.145$	$\Delta_1 = 0.01$
$\hat{\omega}_2 = 0.155$	$\Delta_2 = 0.05$
$\hat{\omega}_3 = 0.225$	$\Delta_3 = 0.05$
$\hat{\omega}_4 = 0.225$	$\Delta_4 = 0.15$
$\hat{\omega}_5 = 0.225$	$\Delta_5 = 0.15$
$\hat{\omega}_6 = 0.350$	$\Delta_6 = 0.20$

Normalization

- Kernel

$$K(\omega,\tau) \equiv \frac{\cosh\left[\omega\left(\tau - \frac{1}{2T}\right)\right]}{\sinh\left(\frac{\omega}{2T}\right)}$$

← Diverges at $\omega = 0$.

- Correlation function has lattice cutoff effects at small distances.

⇨ $$\tilde{K}(\omega,\tau;\tau_0) \equiv \frac{K(\omega,\tau)}{K(\omega,\tau_0)} = \frac{\cosh\left[\omega\left(\tau - \frac{1}{2T}\right)\right]}{\cosh\left[\omega\left(\tau_0 - \frac{1}{2T}\right)\right]}$$

$$\tilde{\rho}(\omega;\tau_0) = \rho(\omega)K(\omega,\tau_0)$$

τ_0 : reference imaginary time

We used the correlation function data from τ_0/a to $N_\tau/2$ in our analysis.
We chose $\tau_0/a = 1$ in mock data tests and $\tau_0/a = 4$ for LQCD data.