Sparse modeling approach to obtaining the QCD shear viscosity from smeared correlation functions

Etsuko Itou
(RIKEN, Keio U., RCNP Osaka U.)

references:
E.I. and Y. Nagai, JHEP 07 (2020) 007
E.I. and S.Aoki , PoS **INPC2016** (2017) 342
Asakawa, Hatsuda, E.I., Kitazawa, Suzuki (FlowQCD coll.) Phys.Rev. D90 (2014) 1, 011501

Tackling The Real-Time Challenge In Strongly Correlated Systems: Spectral Properties From Euclidean Path Integrals, ECT* online workshop (14 Sep. 2021)

QCD shear viscosity

Around T=Tc (hadron/quark-gluon plasma phase transition), a small η/s has been suggested in RHIC experiment.

But in the theoretical side, it is hard to determine η/s.

Shear viscosity is given by the spectral function

$$\eta(T) = \pi \frac{d\rho(\omega)}{d\omega}|_{\omega=0}$$

$\rho(\omega)$ is defined from Euclidean correlation function of the renormalized spatial EMT

$$C(\tau) = \frac{1}{T^5}\int d\overrightarrow{x}\langle T^R_{12}(0,\overrightarrow{0})T^R_{12}(\tau,\overrightarrow{x})\rangle = \int_{-\infty}^{+\infty} d\omega K(\tau,\omega)\rho(\omega)$$

In lattice calculation,
(1) $C(\tau)$ is measured by generated configurations
(2) $\rho(\omega)$ is estimated.

Here, $K(\tau,\omega) = \frac{\cosh\left(\omega(\frac{1}{2T}-\tau)\right)}{\sinh(\frac{\omega}{2T})}$. It is independent of the Monte Carlo data.

Three difficulties to obtain the shear viscosity

(i) How to define the renormalized EMT on the lattice
(ii) How to improve a bad signal-to-noise ratio of the correlation function of EMT
(iii) How to estimate $\rho(\omega)$ from the limited number of the data $C(\tau)$

(i)EMT is generator of general covariance, but the covariance is explicitly broken on the lattice.

(ii) EMT has the same quantum number with QCD vacuum.

(iii) In finite T QCD, we have a few number of data, $C(\tau)$ (~$\mathcal{O}(10)$).
Serious inverse problem appears.

As for (i) and (ii), the gradient flow method looks promising.

Asakawa, Hatsuda, E.I., Kitazawa, Suzuki, Phys.Rev. D90 (2014) 1, 011501

As for (iii), the sparse modeling method is powerful.

H. Shinaoka, J. Otsuki, M. Ohzeki, K. Yoshimi, Phys. Rev. B 96 (2017) 035147 [arXiv:1702.03054].
J. Otsuki, M. Ohzeki, H. Shinaoka, K. Yoshimi, Phys. Rev. E 95 (2017) 061302 [arXiv:1702.03056].

Gradient flow method

Renormalized EMT

(i) How to define the renormalized EMT

Renormalized EMT is given by

Luescher and Weisz, JHEP 1102, 051(2011)
Suzuki, PTEP 2013, no8, 083B03

$$T^R_{\mu\nu}(x) = \lim_{t\to 0}\left\{\frac{1}{\alpha_U(t)}U_{\mu\nu}(t,x) + \frac{\delta_{\mu\nu}}{4\alpha_E(t)}\left[E(t,x) - \langle E(t,x)\rangle_0\right]\right\}$$

t : flow-time

$x = (\tau, \vec{x})$

Quantum field theory
(UV divergence)

perturbation with dim. reg.
+YM gradient flow
(general covariance OK!)

lattice reg.
+Wilson flow
(with a->0 limit)

At finite flow time, UV finite!

Gradient flow as a smearing

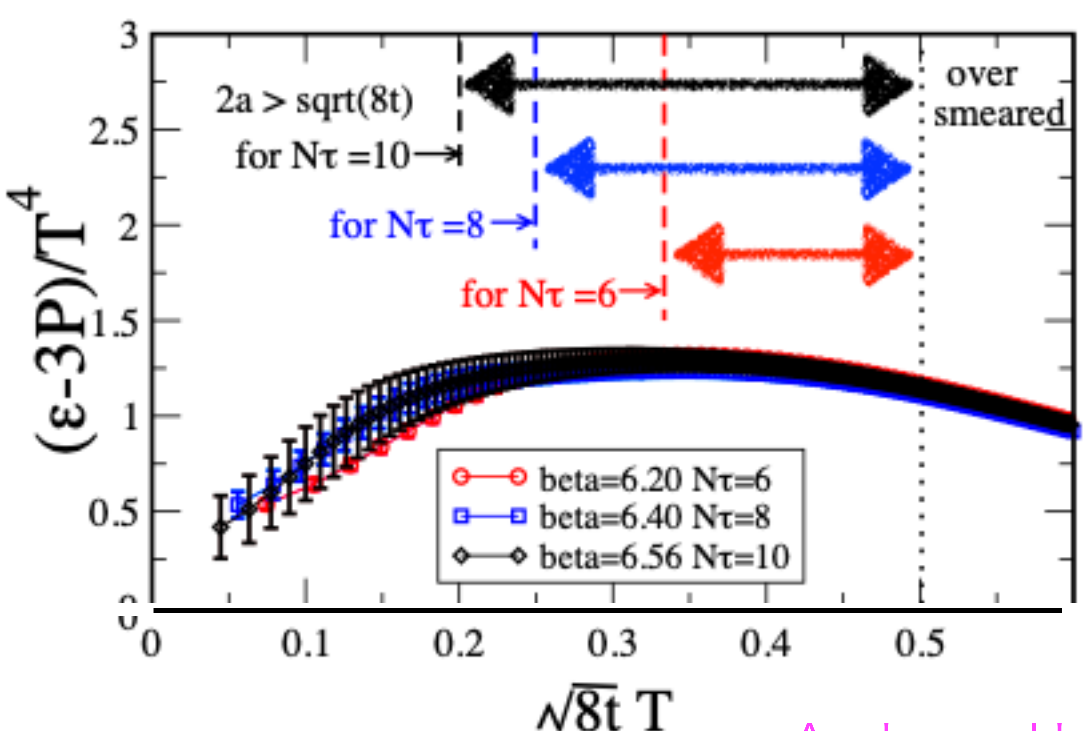
(ii) How to improve a bad signal-to-noise ratio of the correlation function of EMT

As for (ii), gradient flow is a continuous stout smearing (=block spin tranf.), so that it reduces the statistical errors.

The link op. around $|x| \lesssim \sqrt{8t}$ are smeared.

Note that there is a fiducial window of flow-time

One point fn. of EMT



t should be longer than lattice spacing

we want to avoid an over-smeared regime

theoretically, $2a < \sqrt{8t} < N_\tau a/2$

Actually, the data show a plateau.

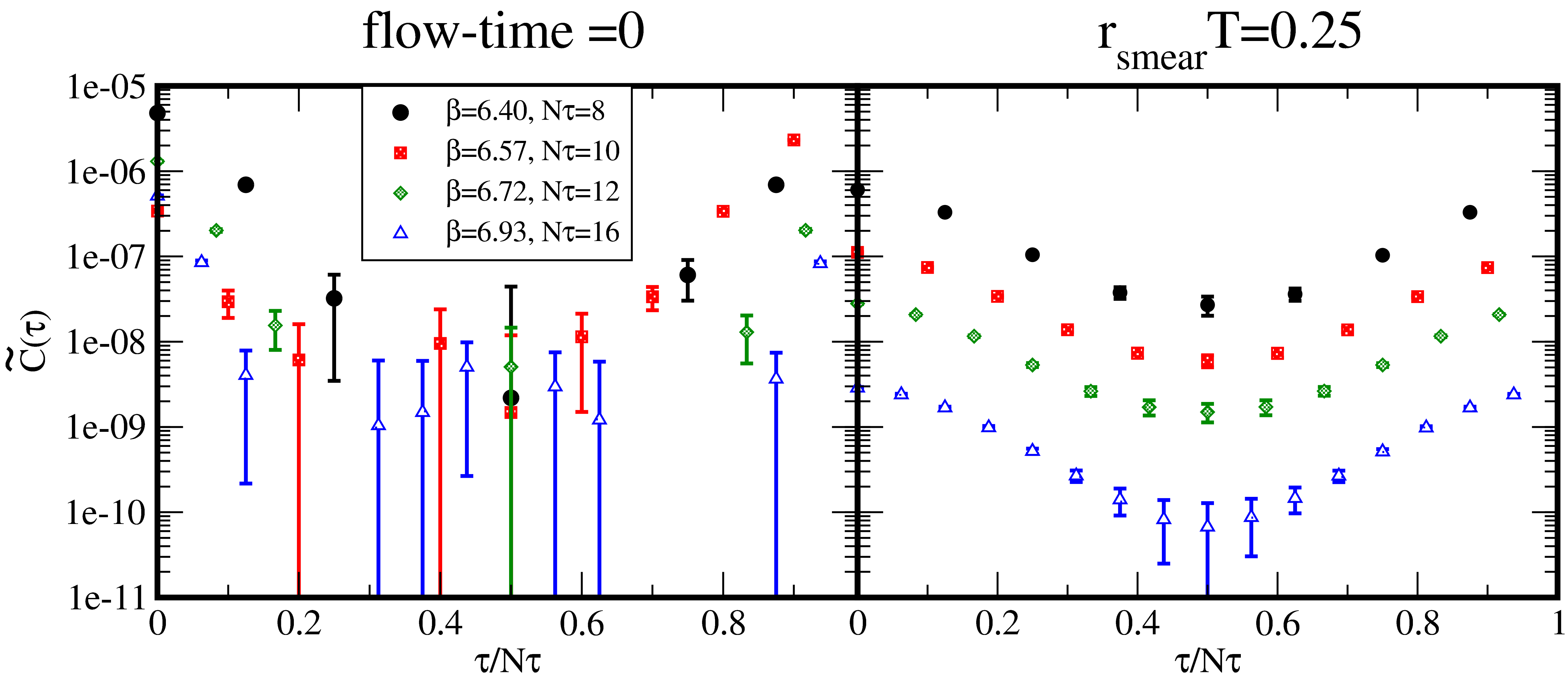
Asakawa, Hatsuda, E.I., Kitazawa, Suzuki Phys.Rev. D90 (2014) 1, 011501

Two-point fn. of EMT using the gradient flow method

$$C(\tau) = \frac{1}{T^5}\int d\overrightarrow{x}\langle T_{12}^R(0,\overrightarrow{0})T_{12}^R(\tau,\overrightarrow{x})\rangle$$

Improvement of signal-to-noise ratio

In these statistics, 2000-5000 configurations, $C(\tau)$ are highly fluctuated



Some non-smeared data take negative value because of the large fluctuation.

But smeared $C(t,\tau)$ correctly take positive value with small statistical errors.

Flow-time dependence of $C(t,\tau)$

Lattice size: $64^3 \times 16$, parameter: $\beta = 6/g_0^2 = 6.93$

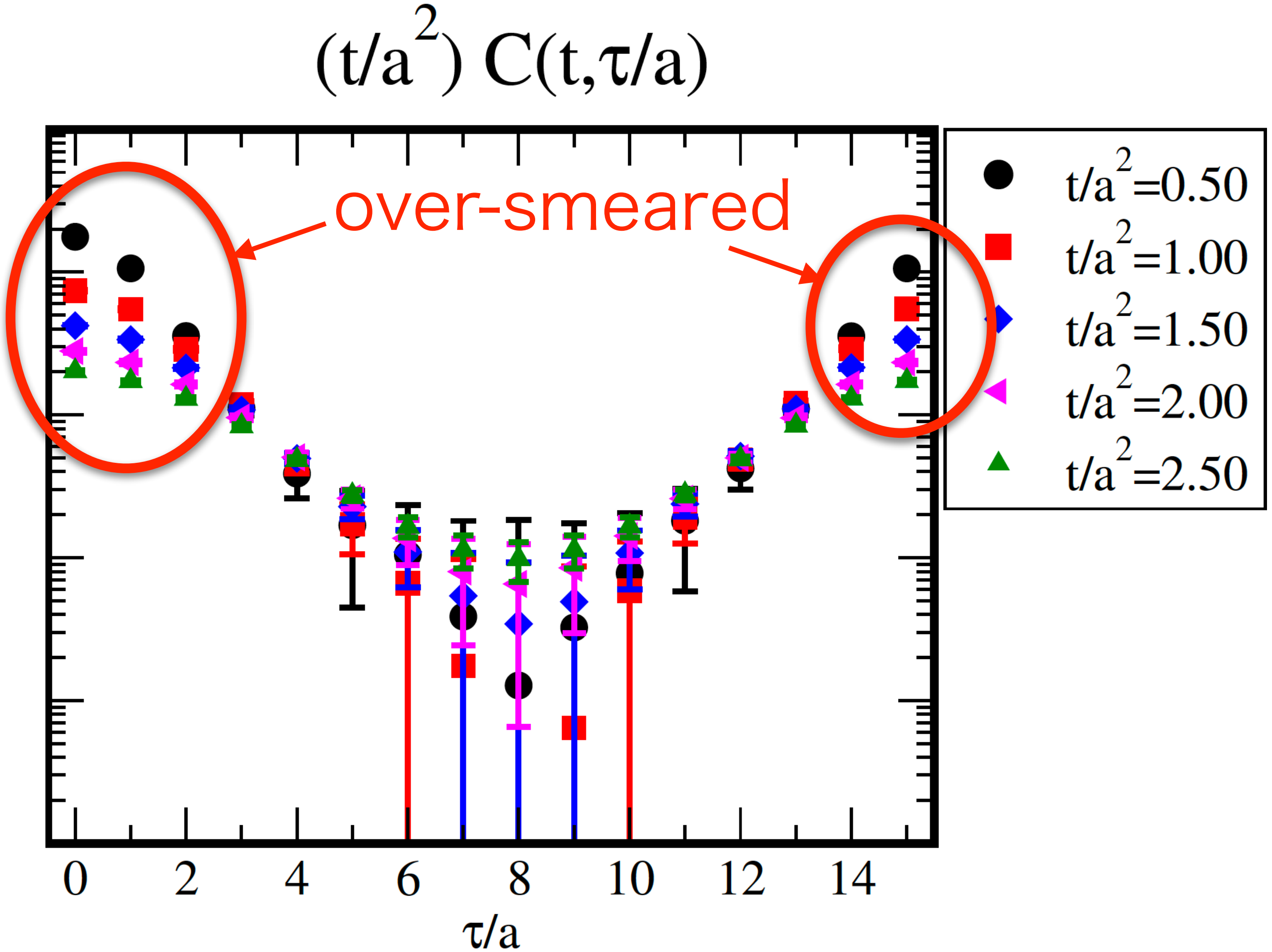
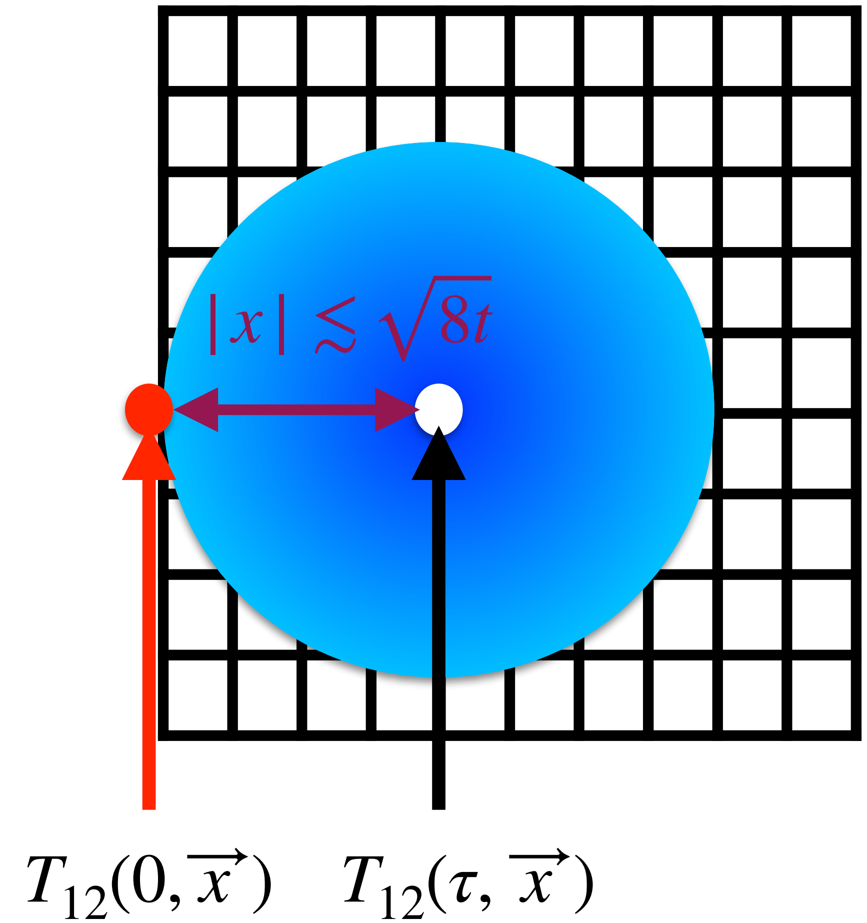


Image on the lattice



(1) Errors get smaller in whole τ-regime

(2) Slope is changed in shorter τ-regime.

- the data $\tau \lesssim \sqrt{8t}$ is over-smeared, since the smeared regime of $T_{12}(\tau,\vec{x})$ overlaps $T_{12}(0,\vec{0})$ in $\langle T^R_{12}(0,\vec{0})T^R_{12}(\tau,\vec{x})\rangle$ measurement.
- We would like to eliminate them from analysis to estimate $\rho(\omega)$.

Sparse modeling method

Gradient flow makes (iii) harder

(i) How to define the renormalized EMT
(ii) How to improve a bad signal-to-noise ratio of the correlation function of EMT
(iii) How to estimate $\rho(\omega)$ from the limited number of the data $C(\tau)$

As for (i) (ii), the gradient flow method looks promising.

On the other hand, the gradient flow method makes (iii) harder.

Essential difficulty (iii) comes from the smallness of the # of data points $C(\tau)$.

integral equation

$$C(\tau) = \int_{-\omega_{cut}}^{+\omega_{cut}} d\omega K(\tau,\omega)\rho(\omega)$$

↓

a set of linear eqs.

$$\begin{pmatrix} C(\tau_1) \\ C(\tau_2) \\ \cdots \\ C(\tau_{N_\tau}) \end{pmatrix} = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1N_\omega} \\ K_{21} & K_{22} & \cdots & K_{2N_\omega} \\ \cdots & \cdots & \cdots & \cdots \\ K_{N_\tau 1} & K_{N_\tau 2} & \cdots & K_{N_\tau N_\omega} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdots \\ \cdots \\ \cdots \\ \cdots \\ \rho_{N_\omega} \end{pmatrix}$$

If $N_\tau < N_\omega$, then several possible solutions are allowed.

As explained, the flowed $C(\tau)$ in short range are deformed by over-smearing.

If we eliminate the over-smeared data, then the situation gets worse.

Solve the inverse problem

We need a powerful tool to find $\rho(\omega)$ from very limited data of $C(\tau)$

Until now, several estimation methods, $\rho(\omega)$ from $C(\tau)$, have been proposed.

- fitting the data using some functional form of $\rho(\omega)$

  e.g., Breit-Wigner ansatz $\frac{\rho(\omega)}{\omega} = \frac{F}{1+b^2(\omega-\omega_0)^2} + \frac{F}{1+b^2(\omega+\omega_0)^2}$

- find a likely function based on Bayesian statistics

  e.g., maximum entropy method

Sparse-modeling (SpM) method

(1) perform the SVD of kernel matrix

It is independent of Monte Carlo data

$$K_{ij} = U_{ik} S_{kl} V^{\dagger}_{lj} \quad \begin{cases} S : N_\tau \times N_\omega \text{ diagonal matrix} \\ U : N_\tau \times N_\tau \text{ unitary matrix} \\ V : N_\omega \times N_\omega \text{ unitary matrix} \end{cases}$$

(2) transform vectors $\overrightarrow{C}$ and $\overrightarrow{\rho}$ into the IR (SVD) basis using unitary matrices,

the rank of $\overrightarrow{\rho}'$ becomes the same with $\overrightarrow{C}'(=N_\tau)$

(3) add L_1 regularization term to the optimization problem to be consistent with a reduction of modes

$C(\tau)$ and $\rho(\omega)$ in the IR basis

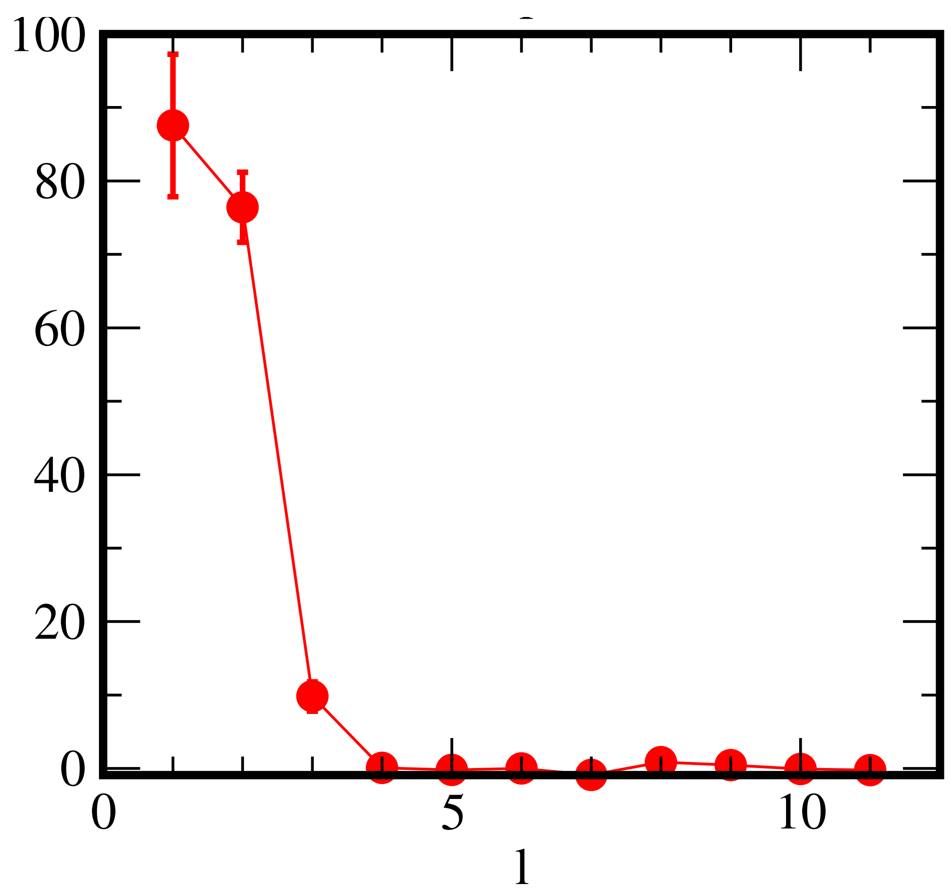
Original optimization problem $F(\vec{\rho}) \equiv \frac{1}{2}\|\vec{C} - K\vec{\rho}\|_2^2$. Here K is not diagonal.

The errors of $C(t,\tau)$ and $\rho(\omega)$ are not linearly related each other.

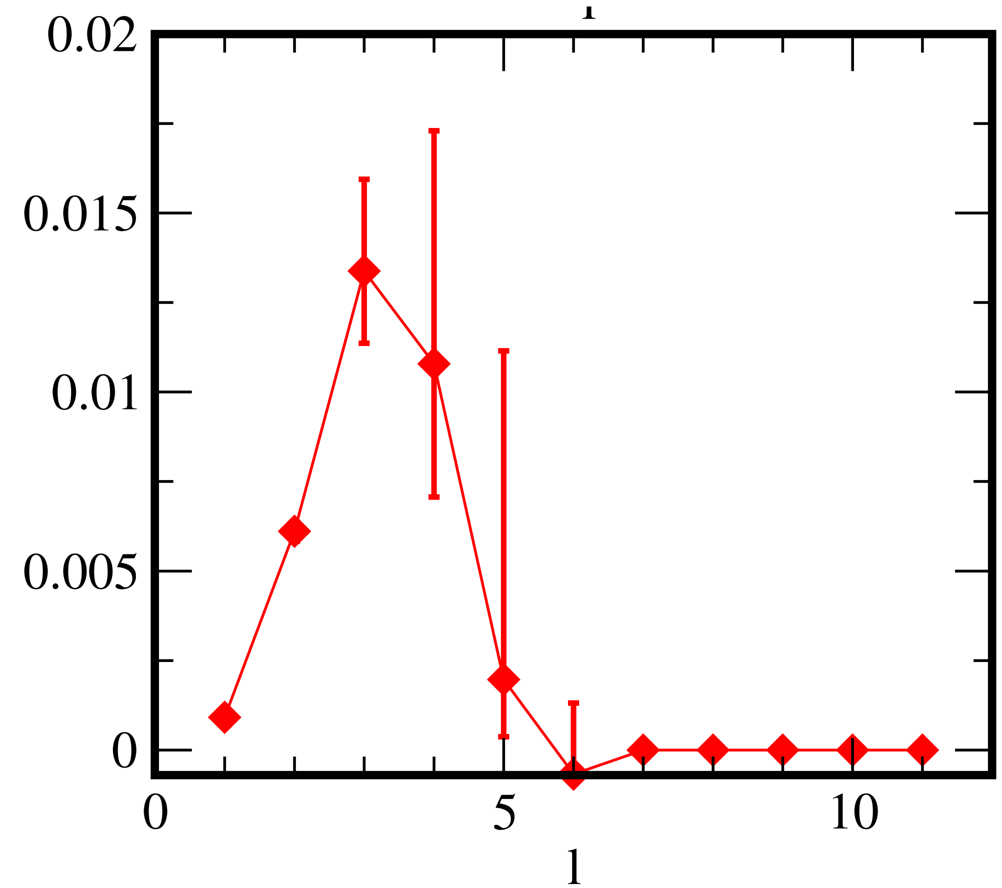
But in the IR basis, $F(\vec{\rho}') \equiv \frac{1}{2}\|\vec{C}' - S\vec{\rho}'\|_2^2$. Here, S is diagonal.

Then a linear relationship $[\vec{C}']_l = s_l[\vec{\rho}']_l$ for each l-th component exists.

$C'_l = [U^t\vec{C}]_l$ (input)



$\rho'_l = [V^t\vec{\rho}]_l$ (output)



The statistical error indicates an expected relationship:

$$\Delta C'_l \sim s_l \Delta\rho'_l$$

SpM

We introduce L_1 regularization into the optimization problem to obtain a stable solution.

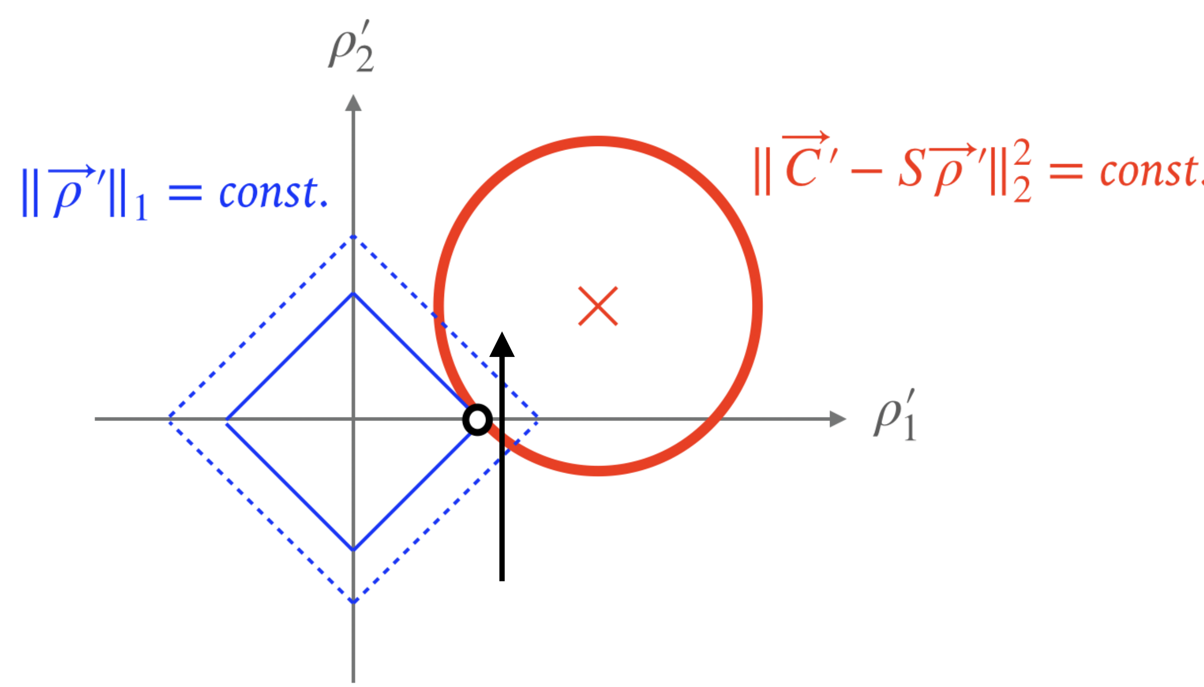
cost function: $F(\overrightarrow{\rho}') \equiv \frac{1}{2}\|\overrightarrow{C}' - S\overrightarrow{\rho}'\|_2^2 + \lambda\|\overrightarrow{\rho}'\|_1$

L_1 term : $\|\overrightarrow{\rho}'\|_1 \equiv \sum_l |\rho_l'|$

Find minimum of the sum

$\frac{1}{2}\|\overrightarrow{C}' - S\overrightarrow{\rho}'\|_2^2 = const.$ and $\|\overrightarrow{\rho}'\|_1 = const.$

by tuning λ (Lagrange multiplier)



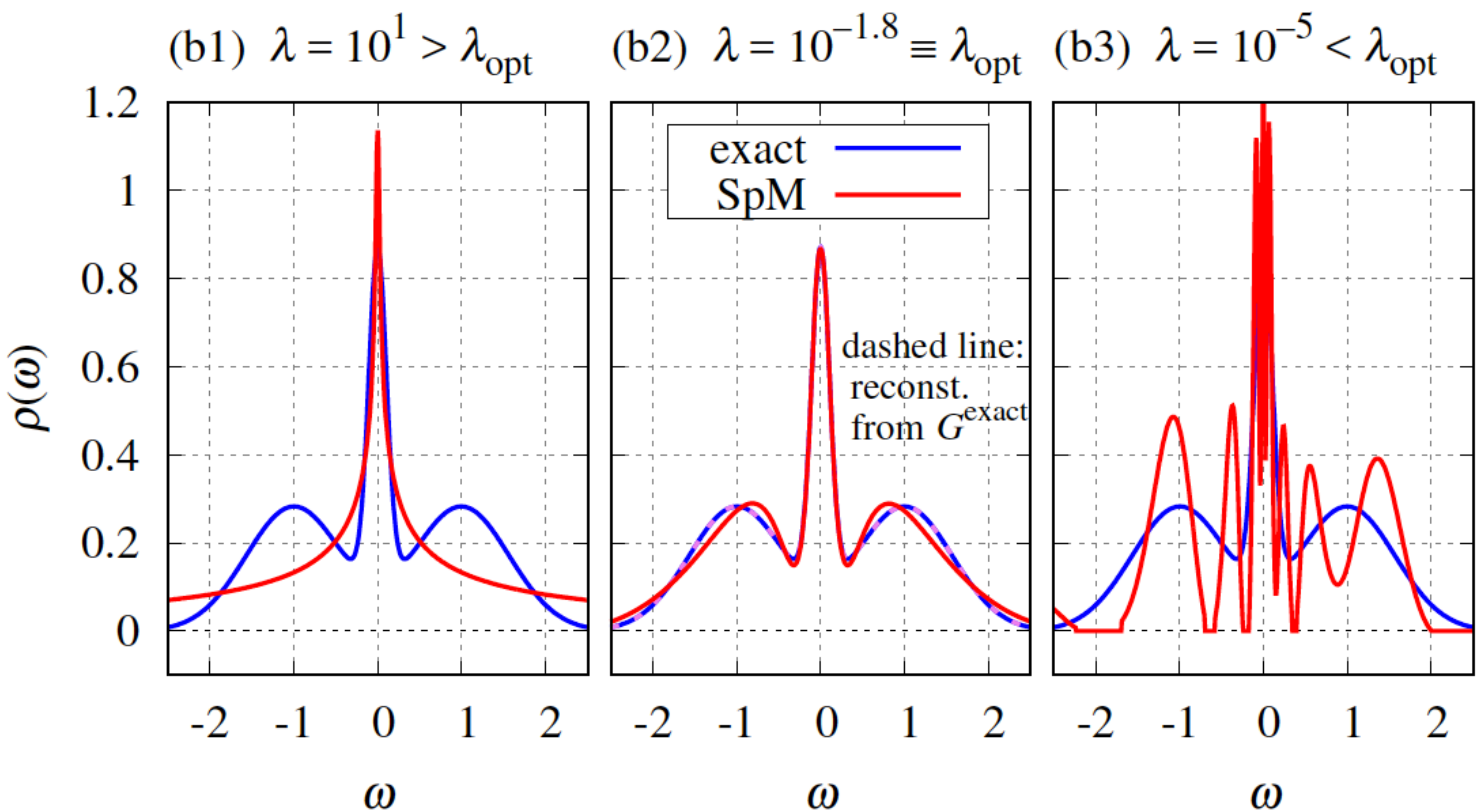
The point ($\rho_2' = 0$) gives the minimum.

L1 term favors the solution with a small number of components
The tendency will give a consistent solution with a cutoff s_l

Role of L_1 term

Test in a statistical model

J. Otsuki et al, Phys. Rev. E 95 (2017) 061302



spectral function becomes featureless

the under-fitting, where the L_1 regularization term is too strong and the number of components $\vec{\rho}'$ is too reduced.

artificial spikes appear

the over-fitting, where the L_1 term is too weak and the vector $\vec{\rho}'$ has redundancy.

Standard cost function

The standard cost function in our code on arXiv

https://arxiv.org/src/2004.02426v2/anc

$$\tilde{F}(\overrightarrow{\rho}', \vec{z}', \vec{z}) = \frac{1}{2\lambda}\|\overrightarrow{C}' - S\overrightarrow{\rho}'\|_2^2 - \nu(\langle V\overrightarrow{\rho}'\rangle - 1) + \|\vec{z}'\|_1 + \lim_{\gamma\to\infty}\sum_j \Theta(-z_j)$$

L1 regularization

sum rule $\sum_j \rho_j = 1$

non-negativity $\rho(\omega) \geq 0$

In actual analysis, we drop the sum rule (take $\nu = 0$)

since we eliminate some data $C(\tau)$.

The ADMM algorithm:

$\vec{z}, \vec{z}'$ are auxiliary vectors, and minimize the cost fn. to be $\vec{z}' = \overrightarrow{\rho}', \quad \vec{z} = V\overrightarrow{x}'$

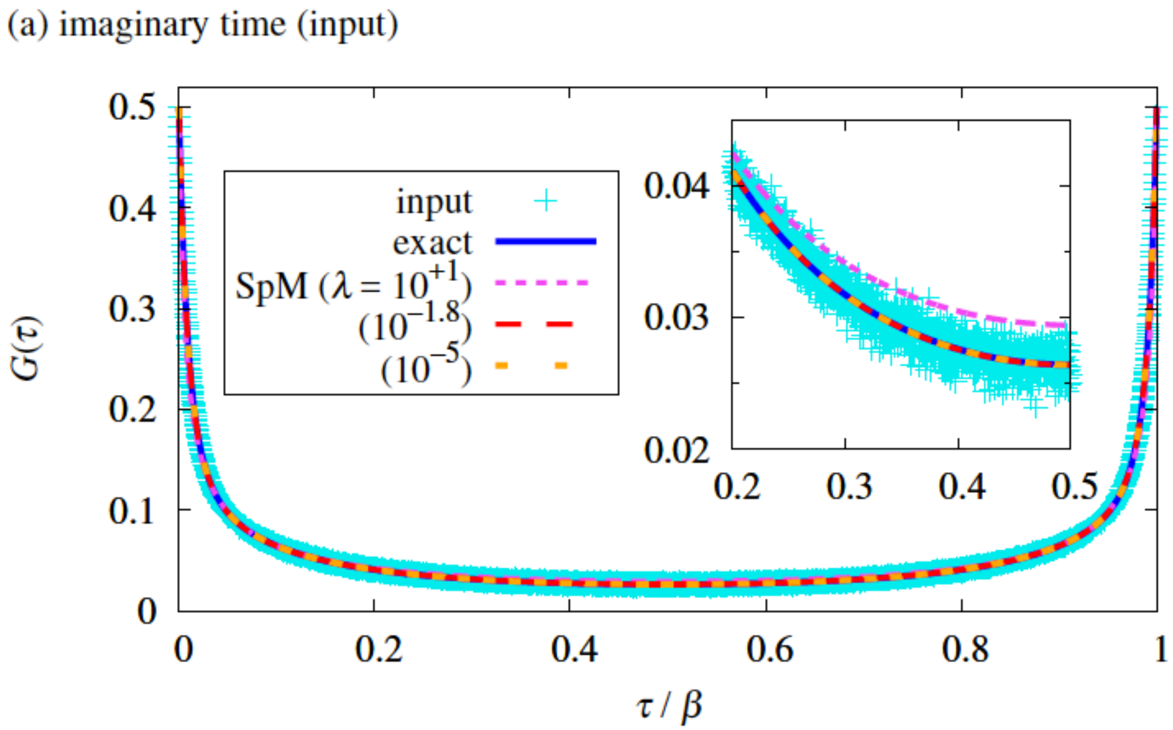
The algorithm for the optimization problem where some auxiliary vectors are introduced to satisfy the conditions.

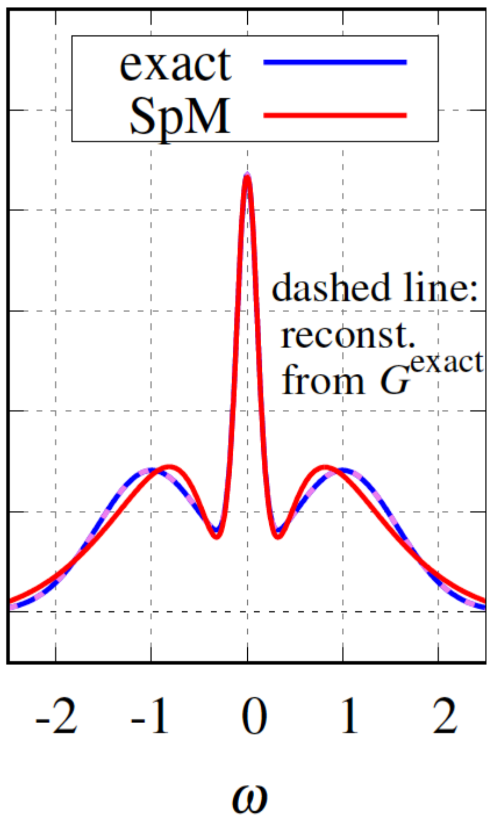
S. Boyd et al., Foundations and Trends R in Machine Learning 3, 1 (2011).
See also Appendix A in our paper

Results of test calculation

J. Otsuki, M. Ohzeki, H. Shinaoka, K. Yoshimi, Phys. Rev. E 95 (2017) 061302 [arXiv:1702.03056].

SpM can be applied for non-smeared $C(\tau)$ without the statistical error.





Preparations: $\rho(\omega)^{\mathrm{exact}}$ is given by hand (it has 3 Gaussian peaks)

$G(\tau)^{\mathrm{exact}}$ is constructed by $\rho(\omega)^{\mathrm{exact}}$

Test(1): Estimate $\rho(\omega)^{\mathrm{SpM}}$ from $G(\tau)^{\mathrm{exact}}$

Test(2): Make $G(\tau)^{\mathrm{input}} = G(\tau)^{\mathrm{exact}} + \eta$ and estimate $\rho(\omega)^{\mathrm{SpM}}$

The result by SpM looks very stable around $\omega \approx 0$.

(shear viscosity: $\eta \propto d\rho(\omega)/d\omega|_{\omega=0}$)

Simulation results for quenched QCD

simulation setup

- Lattice action: Wilson plaquette gauge action
- Lattice size: $64^3 \times 16$
- parameter: $\beta = 6/g_0^2 = 6.93$
- # of configurations: 2,000

  cf.) Nakamura-Sakai(2005): 800,000 conf.
  Borsanyi et al.(2018): 6 million conf.

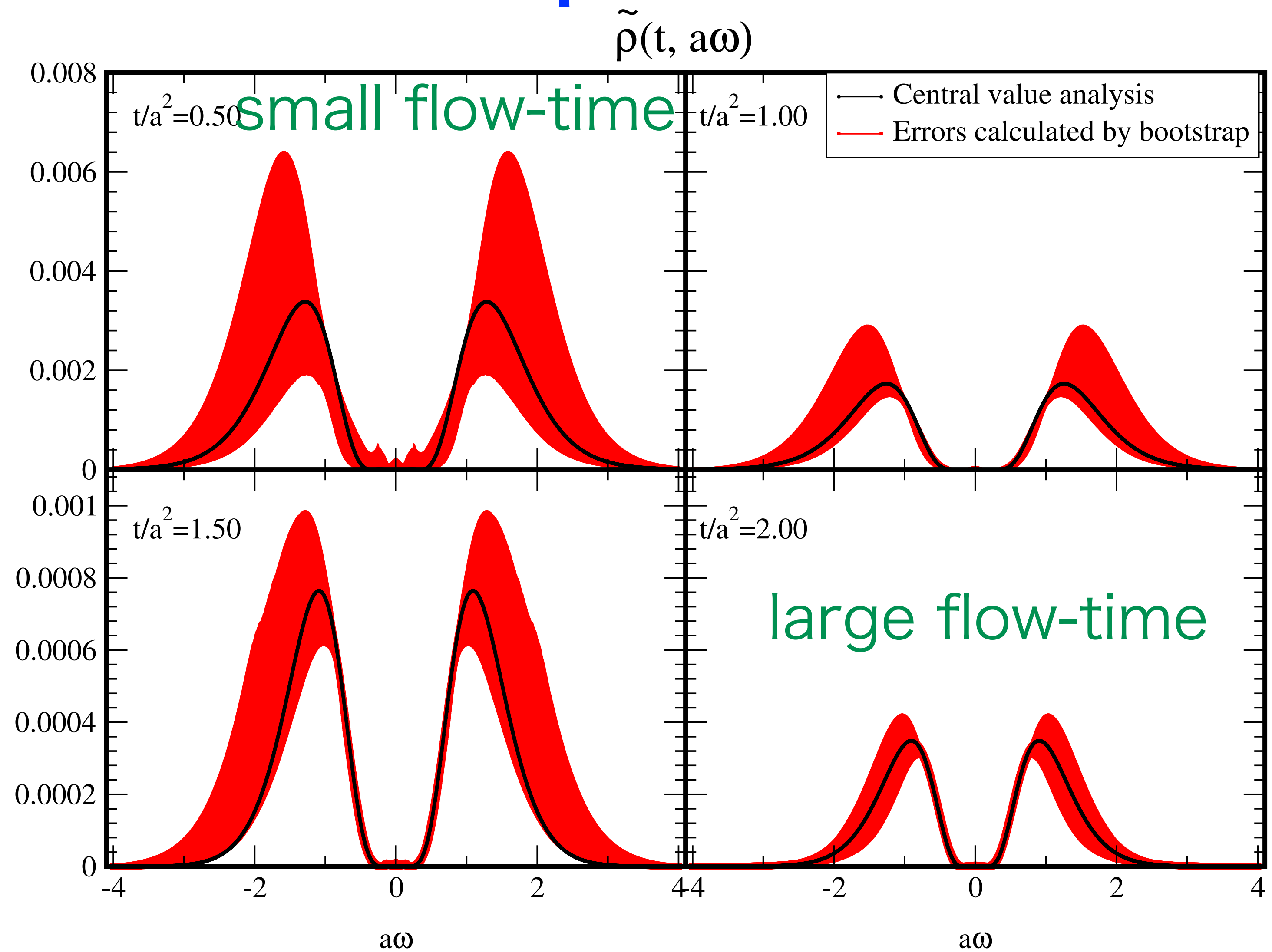
- Temperature, $T = 1.65T_c$

  ALPHA collaboration NPB535 (1998)389,
  G.Boyd et al., NPB469(1996)419

- the thermal entropy @ $T = 1.65T_c$: $s/T^3 = 4.98(24)$ in continuum limit

  Asakawa, Hatsuda, E.I., Kitazawa, Suzuki, PRD90 (2014) 1, 011501

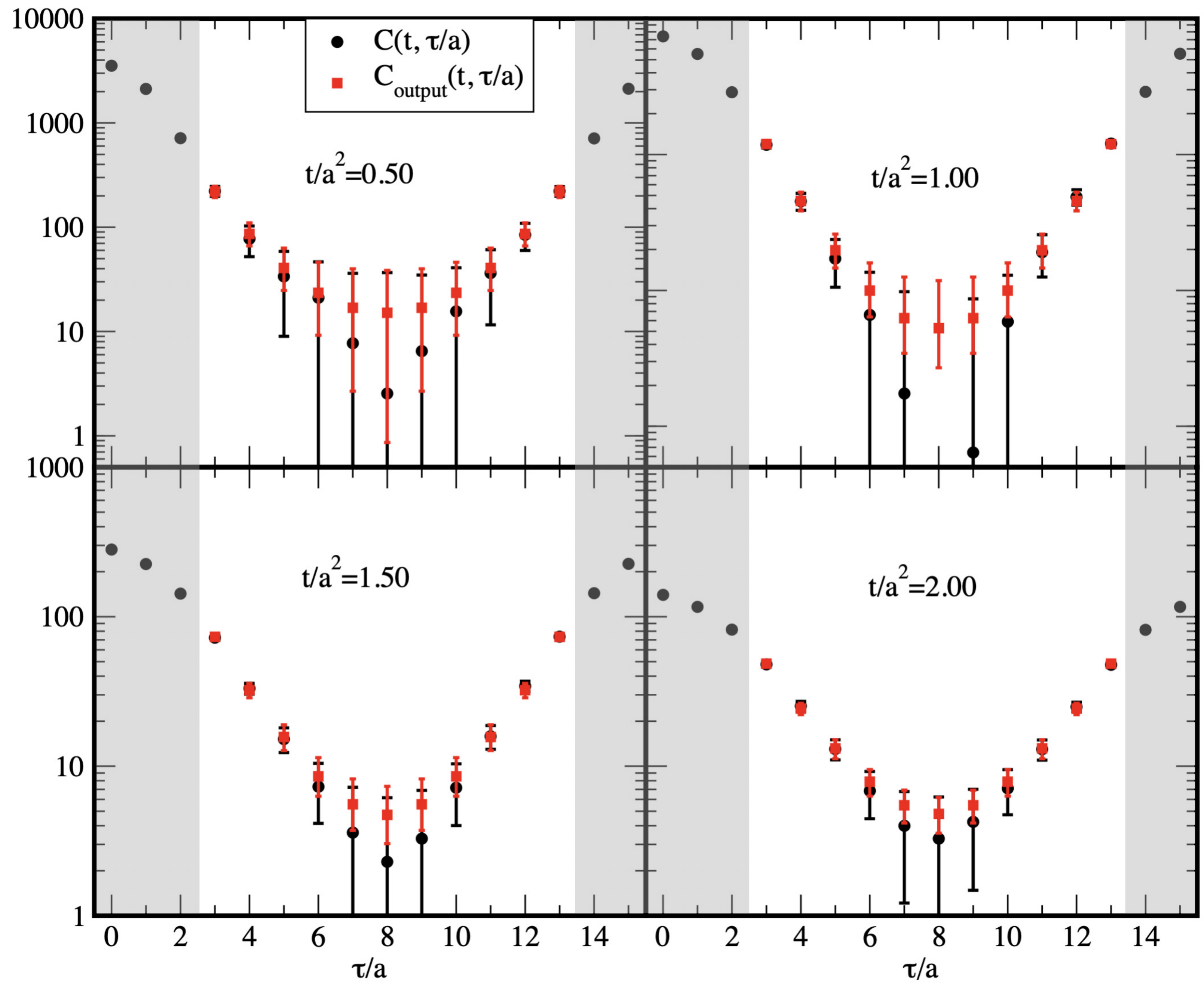
Results of spectral function



- sum of d.o.f. $\mathcal{N} = \int_{-a\omega_{cut}}^{a\omega_{cut}} \tilde{\rho}(t, a\omega) d(a\omega)$ gets smaller in large flow-time.
- The gradient flow can be interpreted as a renormalization group flow.
  Renormalization group decreases the d.o.f. of the system.
  The higher frequency modes are gradually suppressed by the gradient flow.
- The results support this intuitive property of the gradient flow.
- The statistical error in long flow time is small as expected.

Comparison with input and output $C(\tau)$

$C_{output}(t, \tau/a)$ is constructed by the obtained $\tilde{\rho}(\omega)$



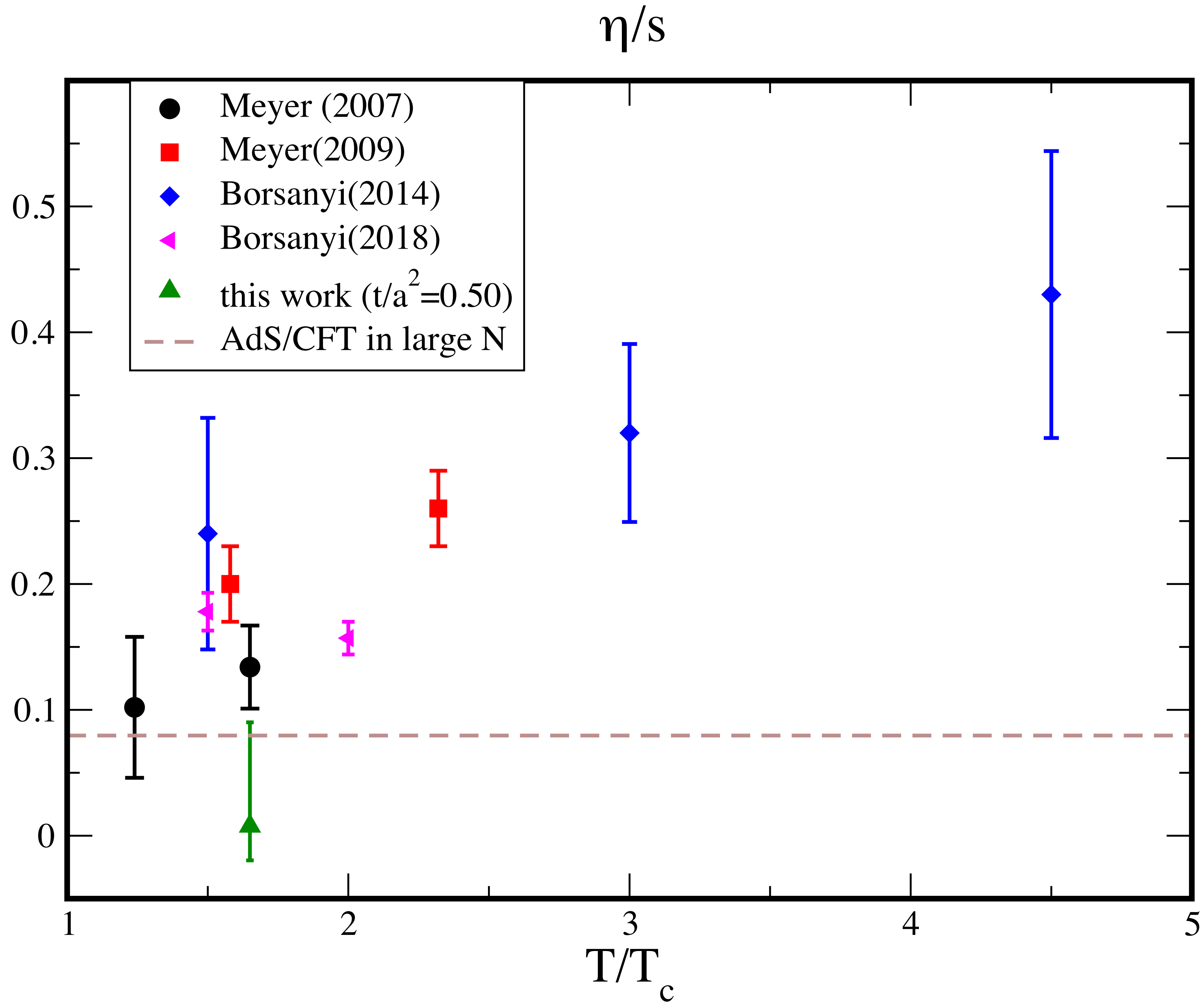
The output $C(\tau)$ almost reproduce the input data.

The SpM analysis works well.

Current status on shear viscosity

We have not taken $a \to 0$ limit and then $t \to 0$ limit yet,

Therefore, it may not be fair to compare our data with the other works….



A poor statistics of our data (only 2,000 conf.) can give a result of the viscosity by combination of the gradient flow and SpM methods.

Toward precise determination of viscosity

(i) How to define the renormalized EMT
(ii) How to improve a bad signal-to-noise ratio of the correlation function of EMT
(iii) How to estimate $\rho(\omega)$ from the limited number of the data $C(\tau)$

$$C(\tau) = \frac{1}{T^5} \int d\overrightarrow{x} \langle T^R_{12}(0,\overrightarrow{0}) T^R_{12}(\tau, \overrightarrow{x}) \rangle = \int_{-\infty}^{+\infty} d\omega K(\tau,\omega)\rho(\omega)$$

$C(\tau)$ is measured using the gradient flow for (i) and (ii)

$\rho(\omega)$ is estimated using the sparse modeling method for (iii)

Both theoretically and technically, these methods look promising.
It means that these methods reduce the noise and give a stable results.

The sparse modeling is a general framework to estimate $\rho(\omega)$ from $C(\tau)$.
I hope that it will be a powerful tool for various subjects in lattice QCD calculations.