Machine Learning and Many-Body Physics

中国科学院大学卡弗里理论科学研究所

Kavli Institute for Theoretical Sciences at UCAS

2017-06-28--2017-07-07 Beijing

Sparse modeling approach to *analytical continuation* and *compression* of imaginary-time quantum Monte Carlo data

Junya Otsuki

Tohoku University

科研費 KAKENHI

J-Physics：多極子伝導系の物理

J-Physics：Physics of Conductive Multipole Systems

Collaborators

M. Ohzeki
Tohoku University

Statistical information
Sparse modeling

H. Shinaoka
Saitama University

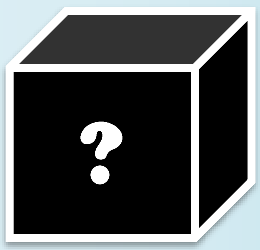
Comput. material sci.
QMC method

K. Yoshimi
ISSP, University of Tokyo

High Performance Comp.

application of Sparse modeling

＝ Technique: denoising, basis selection



New analytical continuation method

<u>JO</u>, Ohzeki, Shinaoka, Yoshimi, PRE 95, 061302(R) (2017)

extract Essense

Finding an efficient basis

Shinaoka, <u>JO</u>, Ohzeki, Yoshimi, arXiv:1702.03054

INTRODUCTION I:

Two Problems
in Quantum Many-body Computations

Imaginary time

Statistical mechanics

classical

$$Z = \sum_j e^{-\beta E_j}$$

$$\beta = 1/T$$

quantum

$$Z = \mathrm{Tr} e^{-\beta \mathcal{H}}$$

$\mathcal{H}$: Hamiltonian matrix

dim = $\mathcal{O}(e^N)$

cannot be diagonalized

c.f. Time evolution operator in quantum mechanics

$$U(t) = e^{-it\mathcal{H}}$$

imaginary time $it \to \tau$ $\int_0^\beta d\tau$

→ quantum Monte Carlo diagrammatic expansion

Why analytical continuation is necessary?

Quantity we want to know

$$\rho(\omega) = -\frac{1}{\pi}\mathrm{Im}G^{\mathrm{R}}(\omega)$$

Ex.: ARPES spectra
Spin excitations

Fourier transform

Analytical continuation

Retarded Green function

$$G^{\mathrm{R}}(t) = -i\theta(t)\langle [A(t), B]_{\pm}\rangle$$

$$A(t) = e^{iHt}Ae^{-iHt}$$

difficult to handle

"imaginary-time"

$$it \to \tau$$

Imaginary-time G

$$G(\tau) = -\langle T_\tau A(\tau) B\rangle$$

$$A(\tau) = e^{\tau H}Ae^{-\tau H}$$

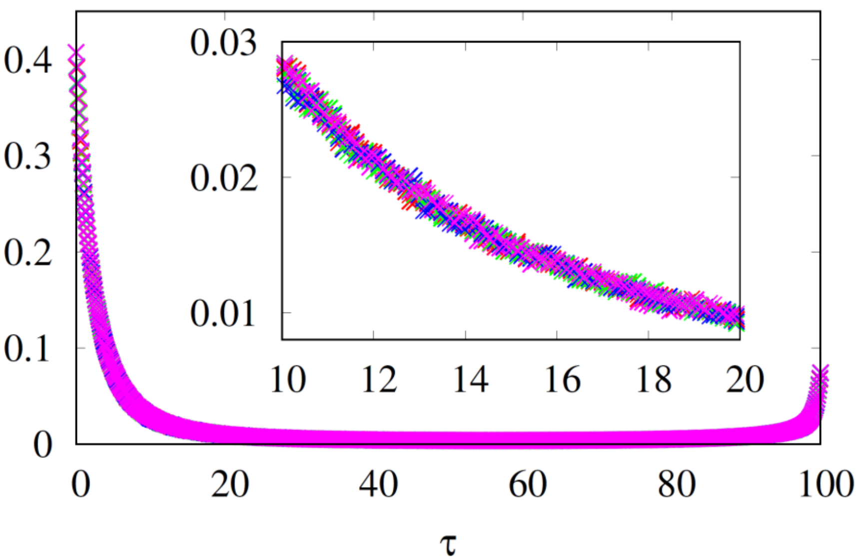
better for calculations
(diagrammatic expansion
quantum Monte Carlo)

Analytical continuation is noise sensitive

The standard method: Pade approximation

Vidberg, Serene, 1977

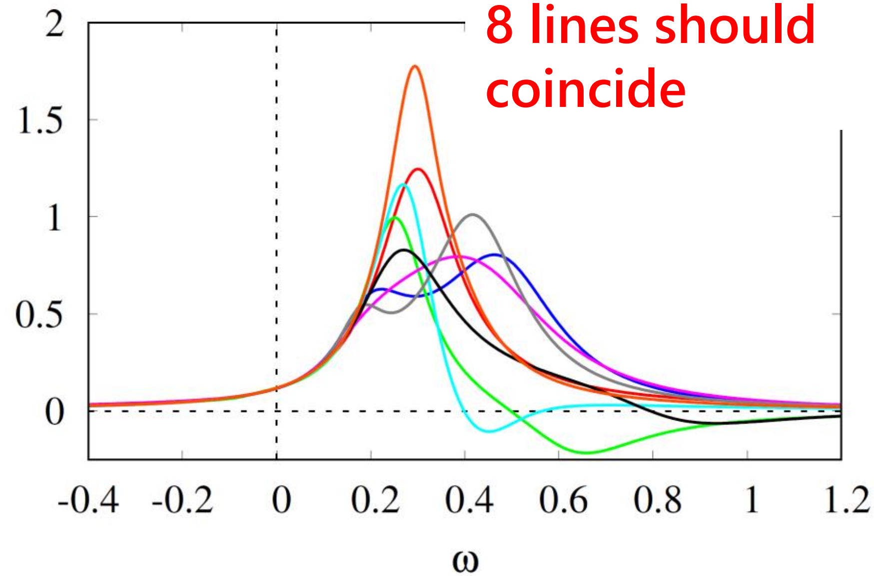
$G(\tau)$ **8-fold degenerate**
impurity Anderson model



CT-QMC data

$\rho(\omega)$

8 lines should coincide



Inverse problem

$$\boldsymbol{G} = K\boldsymbol{\rho}$$

discretize

Lehmann rep

$$G(\tau) = \int_{-\infty}^{\infty} d\omega K_{\pm}(\tau,\omega)\rho(\omega)$$

$$K_{\pm}(\tau,\omega) = \frac{e^{-\tau\omega}}{1 \pm e^{-\beta\omega}}$$ fermion / boson

Evaluate ρ for a given $\boldsymbol{G}$

difficulty: K is an ill-conditioned matrix

~~Least square solution~~

~~$$\boldsymbol{\rho} = (K^{\mathrm{t}}K)^{-1}K^{\mathrm{t}}\boldsymbol{G}$$~~

unstable (NaN)

~~taking errors into account~~

~~$$\chi^2(\boldsymbol{\rho}) \equiv \frac{1}{2}\|\boldsymbol{G} - K\boldsymbol{\rho}\|_2^2 < \eta$$~~

infinite number of solutions
(almost all are unphysical)

Related investigations

Maximum entropy method

M. Jarrell, J. E. Gubernatis, Phys. Rep. 269, 133 (1996)

$$F(\boldsymbol{\rho}) = \frac{1}{2}\|\boldsymbol{G} - K\boldsymbol{\rho}\|_2^2 + \boxed{\alpha \sum_i [\rho_i - m_i - \rho_i \log(\rho_i/m_i)]}$$

m : "default model" = priorknowledge

penalty against deviation from m

Stochastic method

A. W. Sandvik, PRB 57, 10287 (1998)
A. S. Mishchenko et al. PRB 62, 6317 (2000)
S. Fuchs, T. Pruschke, and M. Jarrell, PRE 81, 056701 (2010)
K. S. D. Beach, arXiv:cond-mat/0403055
A. W. Sandvik, PRE 94, 063308 (2016)

Growing attempts

K. S. D. Beach, R. J. Gooding, and F. Marsiglio, PRB 61, 5147 (2000)
A. Dirks *et al.*, Phys. Rev. E 87, 023305 (2013).
F. Bao *et al.*, PRB 94, 125149 (2016)
O. Goulko *et al.*, PRB 95, 014102 (2017).
G. Bertaina, D. Galli, and E. Vitali, arXiv:1611.08502.
L.-F. Arsenault *et al.*, arXiv:1612.04895.

PROBLEM I

analytical continuation

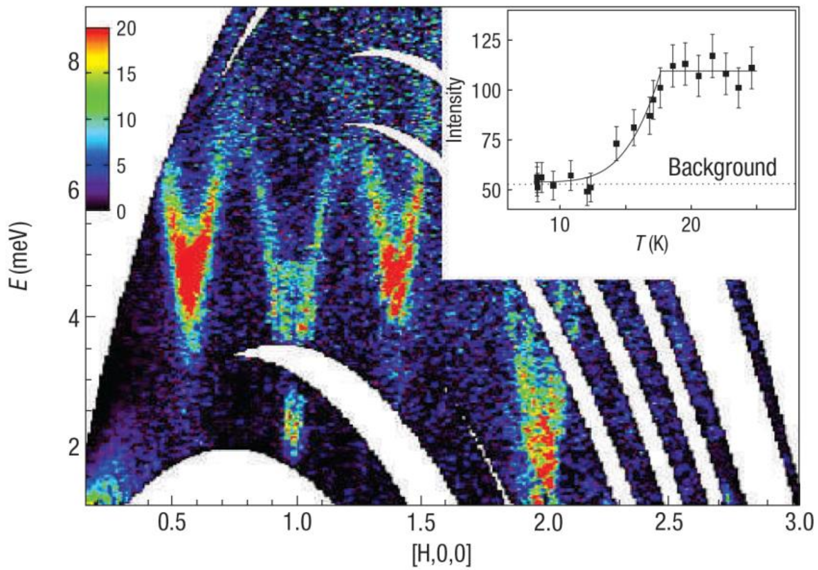
$$G(\tau) \rightarrow \rho(\omega)$$

imaginary time real frequency

Dynamical susceptibility

$$\chi(\boldsymbol{q}, \omega)$$

Ex.:
spin excitation

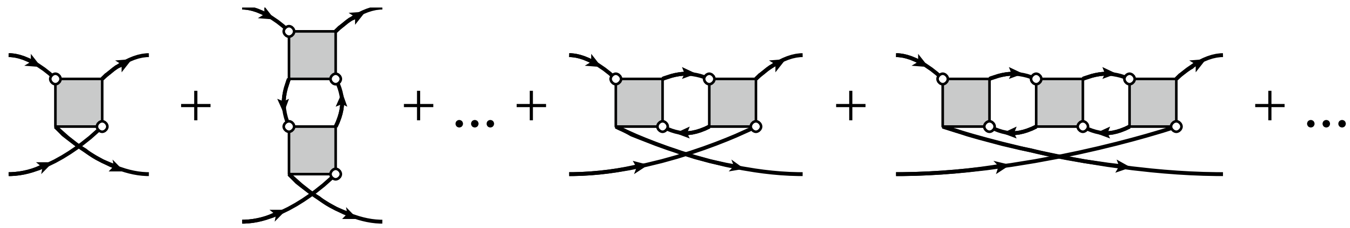


inelastic neutron scat.

URu_2Si_2
Wiebe et al. 2007

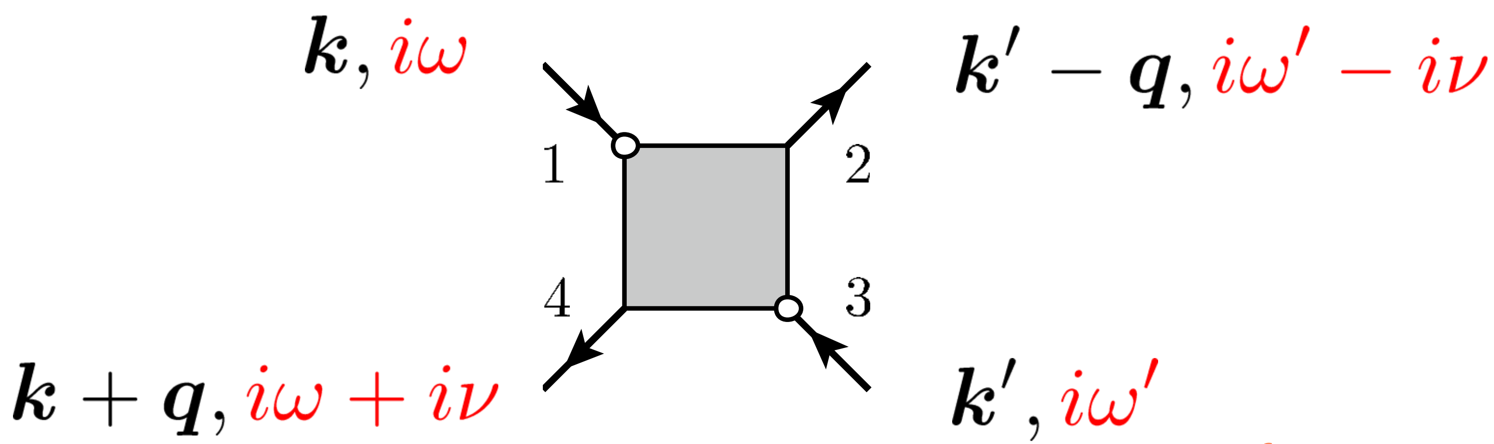
To compute χ(*q*, iω)

Tensor product



Effective interactions

$$\Gamma_{1234}(\boldsymbol{k}, i\omega, \boldsymbol{k}', i\omega'; \boldsymbol{q}, i\nu)$$

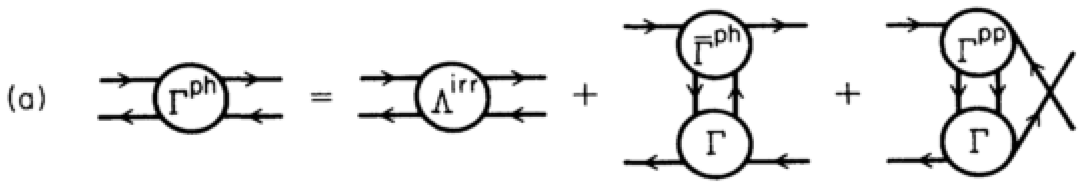


frequency dep.
is crucially important
in SCES

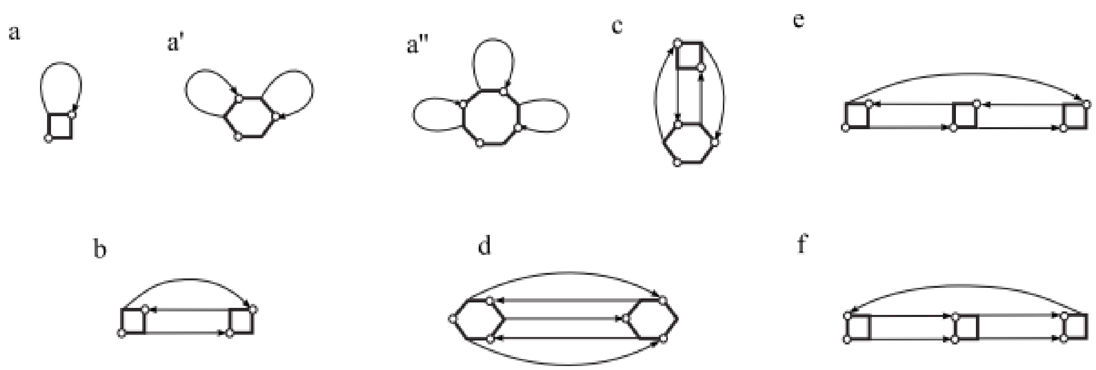
$1, 2, 3, 4$	internal DOF (spin, orbital)
ω, ω'	Fermionic Matsubara freq.
ν	Bosonic Matsubara freq.

Increasing demands

Parquet equation Bickers, White 1991
Functional RG (fRG) Metzner et al. 2012



Beyond dynamical mean-field theory (DMFT) Georges et al. 1996
- DΓA Kusunose 2006, Toschi et al 2006
- dual fermion approach Rubtsov et al. 2008, Hafermann et al. 2009
- fRG extensions Tranto et al. 2014



PROBLEM II

How to handle efficiently

$$\Gamma(i\omega_n, i\omega_{n'}; i\nu_m)$$

A *sparse-modeling* solution

- Solution to
  - Problem I (analytical continuation)

    extract relevant information, basis selection denoising

  - Problem II (two-particle objects)

    Compact representation of correlation functions

SOLUTION to problem I

Sparse-Modeling (SpM) Analytical Continuation

$$G(\tau) \to \rho(\omega)$$

JO, Ohzeki, Shinaoka, Yoshimi, PRE 95, 061302(R) (2017)

Analytical continuation $$G = K\rho$$

Sparseness

||

"There is only a little information that is not disturbed by noise"

Sparseness is basis-dependent

Procedure 1: Basis transformation

Q. Which basis makes ρ sparse ?

A. $\boldsymbol{\rho}' = V^{\mathrm{t}}\boldsymbol{\rho}$ $\quad$ $\boldsymbol{G}' = U^{\mathrm{t}}\boldsymbol{G}$

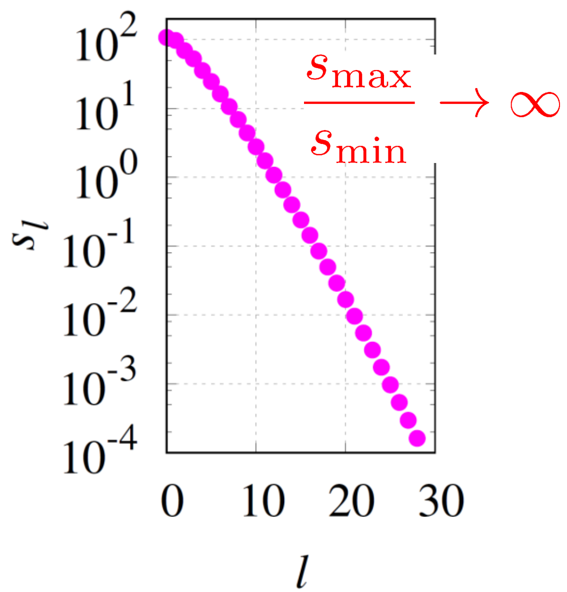
SVD

$$K = USV^{\mathrm{t}}$$

∵ K is an ill-conditioned matrix

$$\chi^2(\boldsymbol{\rho}) = \frac{1}{2}\|\boldsymbol{G} - K\boldsymbol{\rho}\|_2^2 = \frac{1}{2}\|\boldsymbol{G}' - S\boldsymbol{\rho}'\|_2^2$$

Components of ρ' that has small s_l is indefinite



Procedure 2: L1 regularization

Sparseness

$$F(\boldsymbol{\rho}') \equiv \frac{1}{2}\|\boldsymbol{G}' - S\boldsymbol{\rho}'\|_2^2 + \lambda\|\boldsymbol{\rho}'\|_1$$

L1 norm

$$\|\boldsymbol{\rho}'\|_1 \equiv \sum_i |\rho'_i|$$

LASSO-type optimization
(Least Absolute Shrinkage of Selection Operators)
R. Tibshirani, J. R. Stat. Soc. B 58, 267 (1996)

Minimize F under constraints

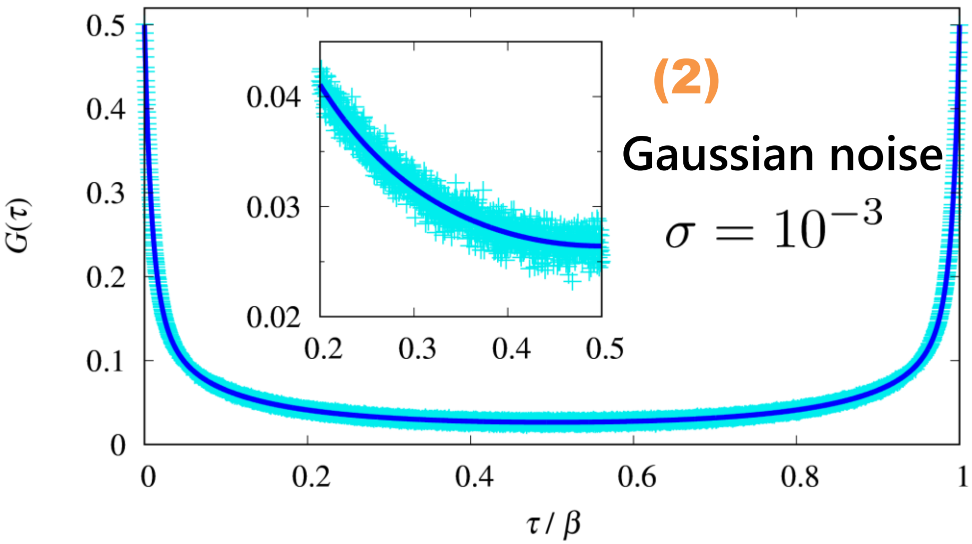
non-negative sum-rule

$$\rho_i \geq 0, \quad \sum_i \rho_i = 1$$

ADMM algorithm (alternating direction method of multipliers)
Boyd et al., Foundations and Trends in Machine Learning 3, 1 (2011)

Test data

$$G(\tau)$$



(2)

Gaussian noise

$$\sigma = 10^{-3}$$

M=4000点

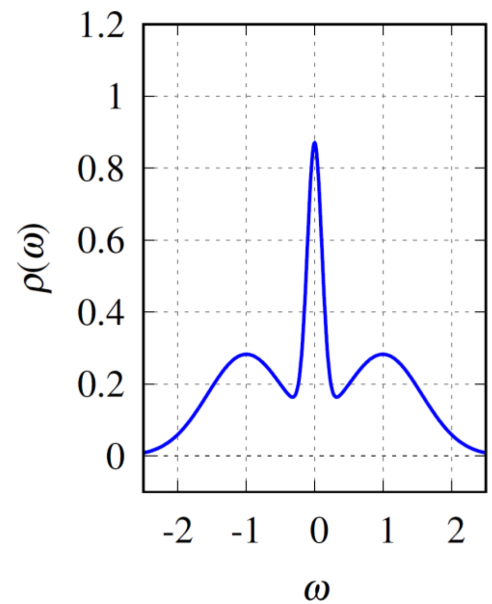
(1)

$$G = K\boldsymbol{\rho}$$

SpM

?

$$\boldsymbol{\rho}$$

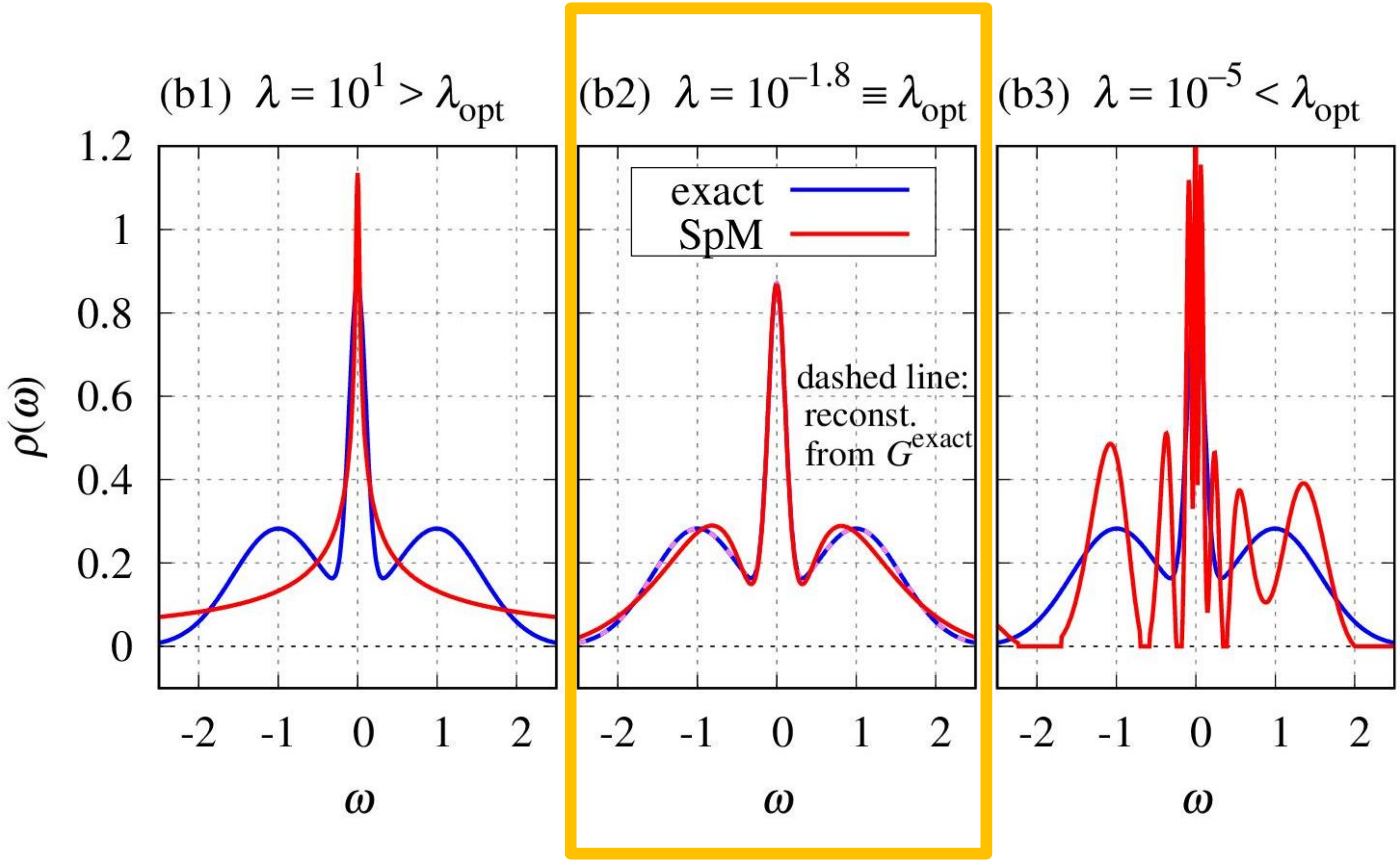


In ordinary situation,
this is directly computed e.g. by QMC

Results

regularization parameter
λ

too strong | optimal | too weak



How it works ?

input with noise

imaginary time

$$G' = U^{\mathrm{t}} G$$

Sparse modeling

ordinary way

real frequency

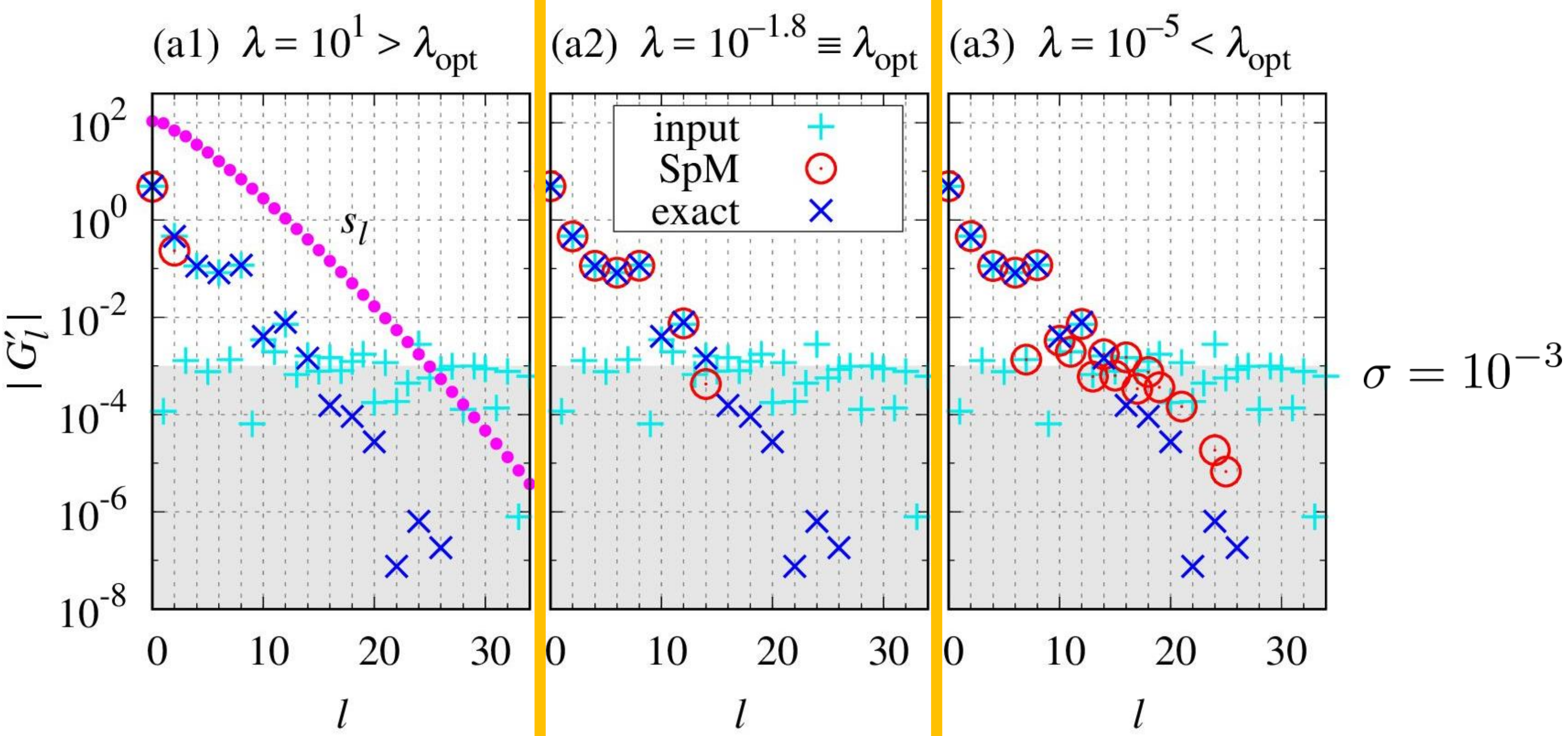
$$\rho' = V^{\mathrm{t}} \rho$$

output

input data $\boldsymbol{G}' = U^{t}\boldsymbol{G}$

regularization parameter λ

too strong | optimal | too weak



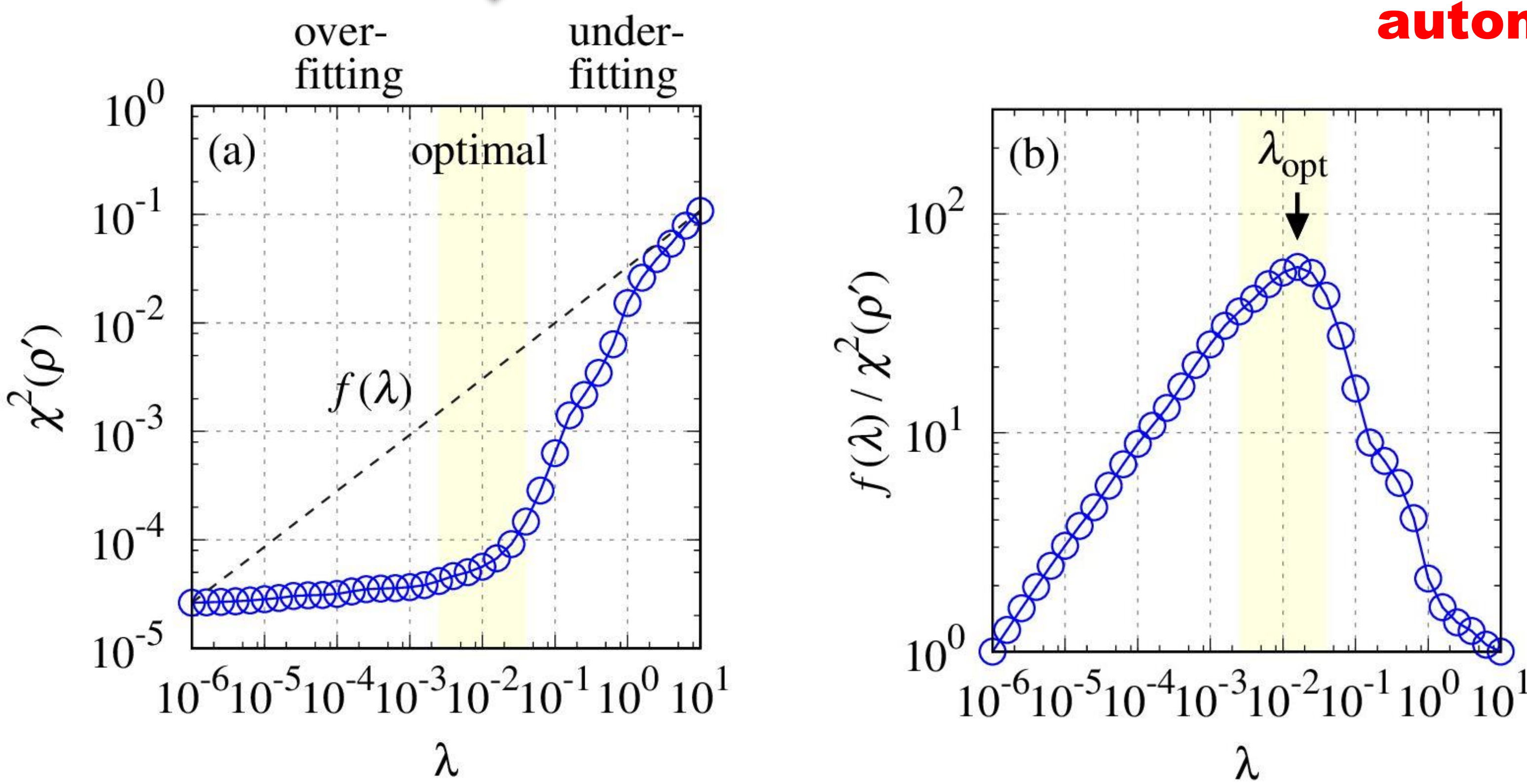
input (w/ noise)
exact (w/o noise)
result

How to fix λ

For a given λ

$$\min_{\boldsymbol{\rho}'} \left\{ \frac{1}{2}\|\boldsymbol{G}' - S\boldsymbol{\rho}'\|_2^2 + \lambda\|\boldsymbol{\rho}'\|_1 \right\} \quad \text{subj. to} \quad \text{constraints}$$

automatic!



SOLUTION to problem II

Intermediate Representation (IR)

$$\Gamma(i\omega_n, i\omega_{n'}; i\nu_m)$$

Shinaoka, JO, Ohzeki, Yoshimi, arXiv:1702.03054

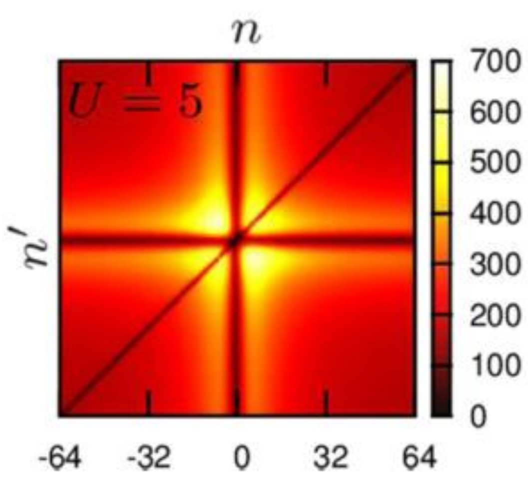
Effective interactions

Dynamical mean-field approximation

$$\Sigma(\boldsymbol{k}, i\omega) \to \Sigma_{\rm loc}(i\omega)$$

$$\Gamma(\boldsymbol{k}, i\omega, \boldsymbol{k}', i\omega'; \boldsymbol{q}, i\nu) \to \Gamma_{\rm loc}(i\omega, i\omega'; i\nu)$$

$\Gamma(i\omega_n, i\omega_{n'}; 0)$



Hafermann PRB 2014

c.f. Hubbard interactions

$$\Gamma(i\omega, i\omega'; i\nu) = U$$

Effective interactions in SCES
are strongly frequency dependent

$$\Gamma(i\omega, i\omega'; i\nu)$$

Can we parameterize?

c.f. Landau parameter in Fermi liquid theory

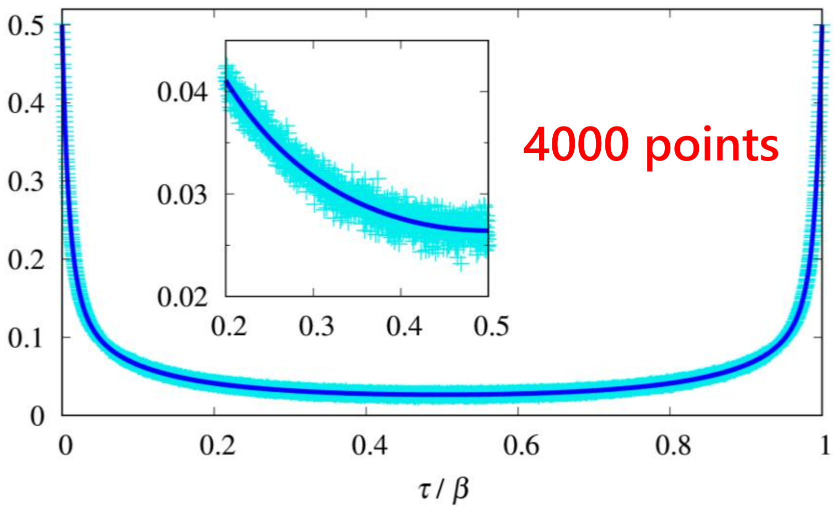
$$\Gamma(\boldsymbol{k}, \boldsymbol{k}'; \boldsymbol{q} = 0) = \sum_{l=0}^{\infty} F_l P_l(\cos\theta_{\boldsymbol{k},\boldsymbol{k}'})$$

Which basis best describes frequency dependences?

Look at the input data again

Original imaginary time data

$$G(\tau)$$

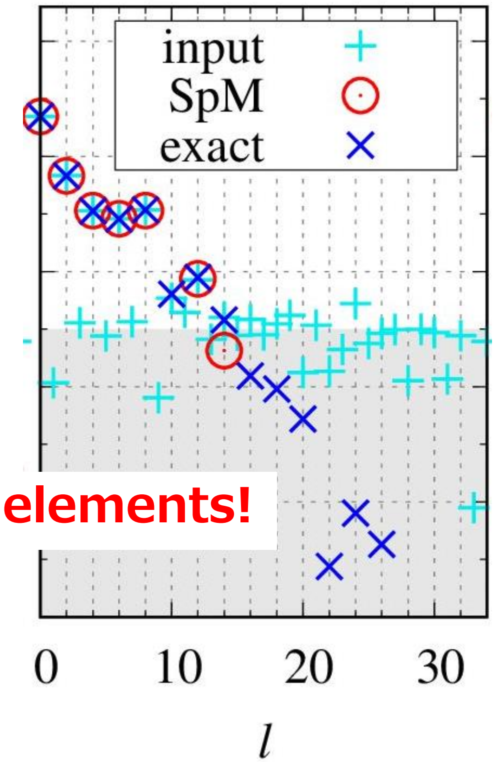


4000 points

unitary trans

After transformation

$$G' = U^{\mathrm{t}} G$$



Only 7 elements!

$$\boldsymbol{G}' = U^{\mathrm{t}} \boldsymbol{G}$$

can be used for data compression

A new orthogonal basis set

Lehmann rep

$$G(\tau) = \int_{-\omega_{\max}}^{\omega_{\max}} d\omega K_{\pm}(\tau,\omega)\rho(\omega)$$

$$K_{\pm}(\tau,\omega) = \frac{e^{-\tau\omega}}{1 \pm e^{-\beta\omega}}$$

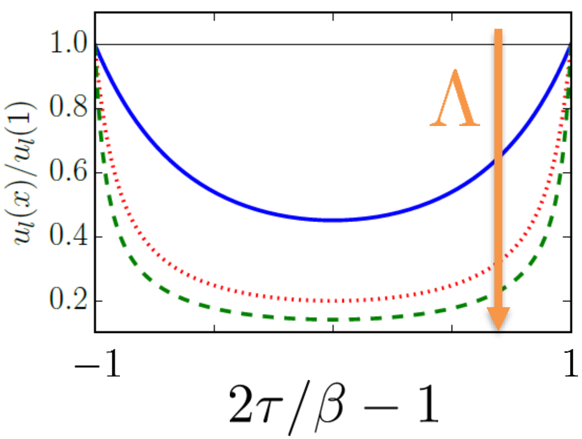
SVD

$$K(\tau,\omega) = \sum_{l=0}^{\infty} s_l u_l(\tau) v_l(\omega)$$

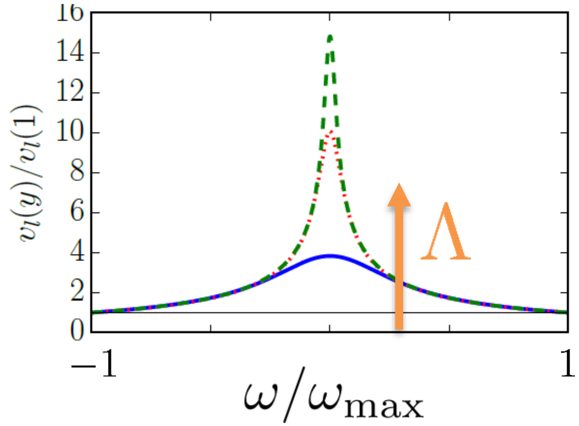
$$\Lambda \equiv \beta\omega_{\max}$$

dimensionnless

$u_{l=0}(\tau)$



$v_{l=0}(\omega)$



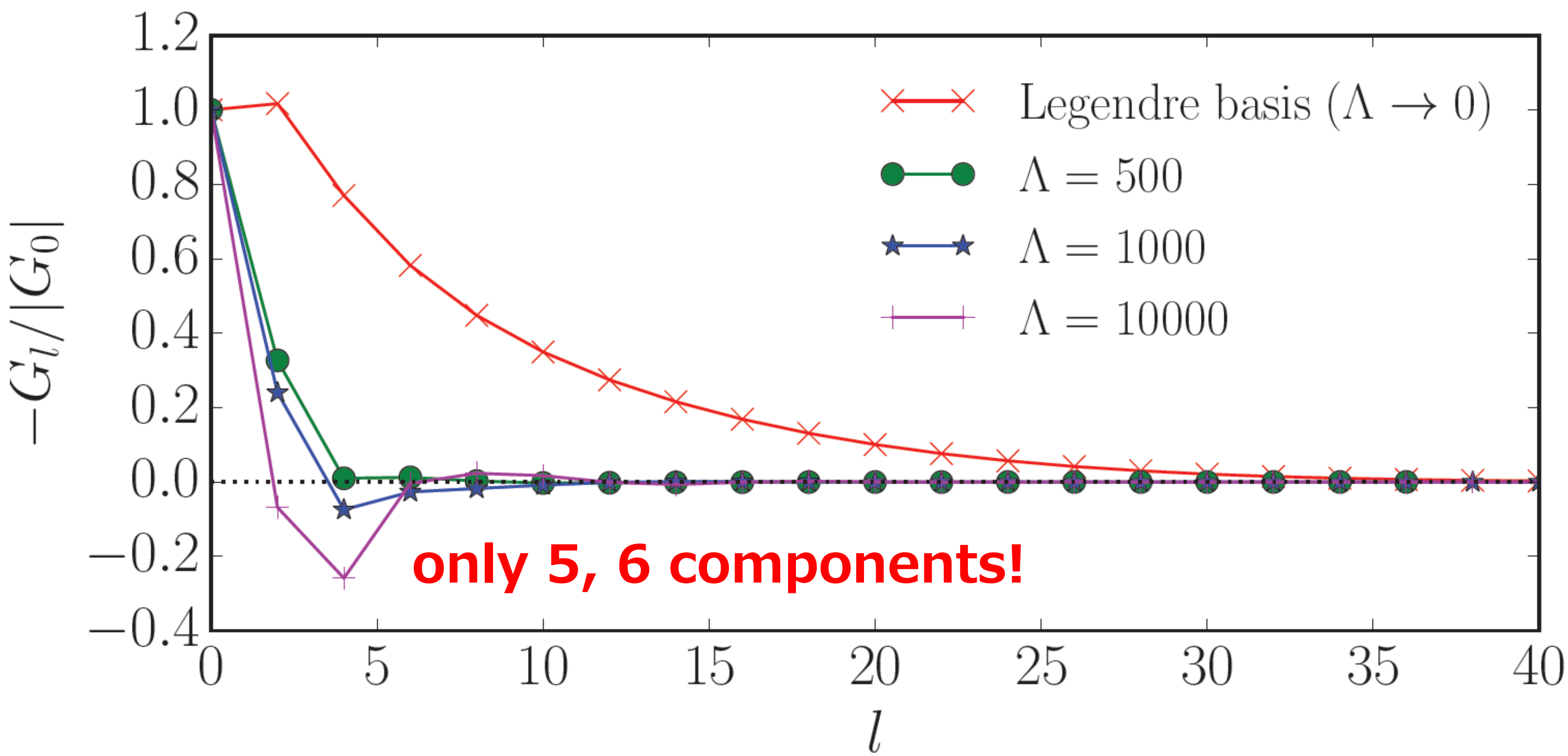
Legendre polynomial in the limit Λ→0

(high-T)

single-particle Green function

$$G(\tau) = \sum_{l=0}^{\infty} G_l u_l(\tau)$$

can be measured directly by CT-QMC

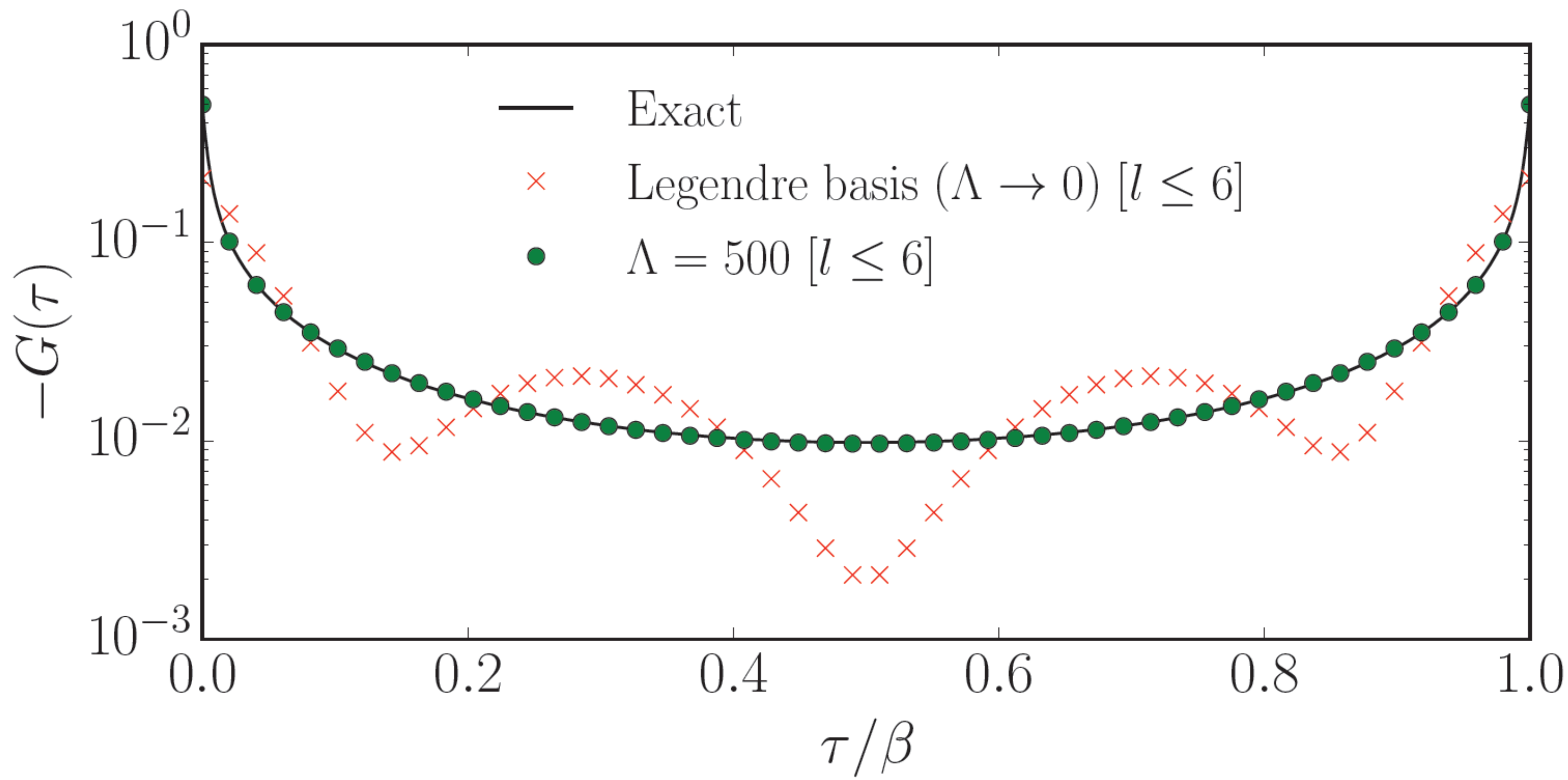


c.f. power decay in Fourier rep

Legendre expansion Boehnke et al. 2011

Check if the original function can be reproduced

$$G(\tau) \overset{?}{=} \sum_{l=0}^{l_{\mathrm{cutoff}}} G_l u_l(x)$$

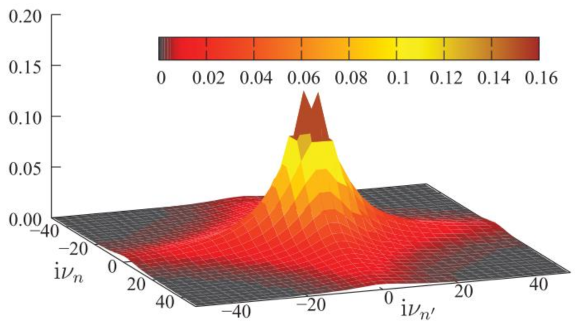


Two-particle Green function

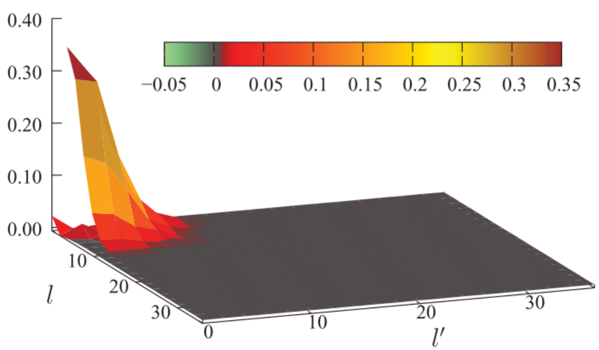
$$\chi(\tau, \tau'; 0) = \sum_{l,l'=0}^{\infty} \chi_{ll'} u_l(\tau) u_{l'}(\tau')$$

Former study Boehnke et al. 2011

Matsubara frequency



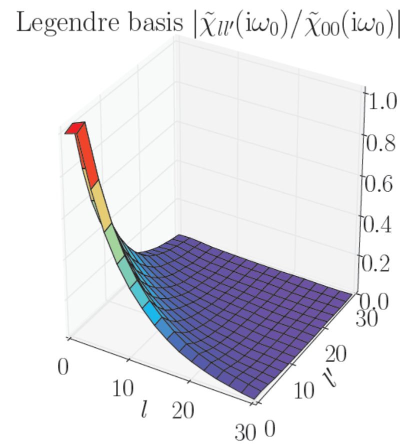
Legendre



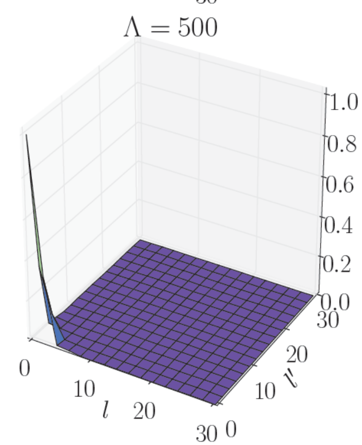
SVD basis

$\Lambda = 0$

(Legendre)



$\Lambda = 500$



Parameterizing effective interactions

$$\Gamma_{1234}(\boldsymbol{k}, i\omega_n, \boldsymbol{k}', i\omega_{n'})$$

$$\boldsymbol{q} = 0, \nu_m = 0$$

Spin, Charge, Orbital
(irreducible rep)

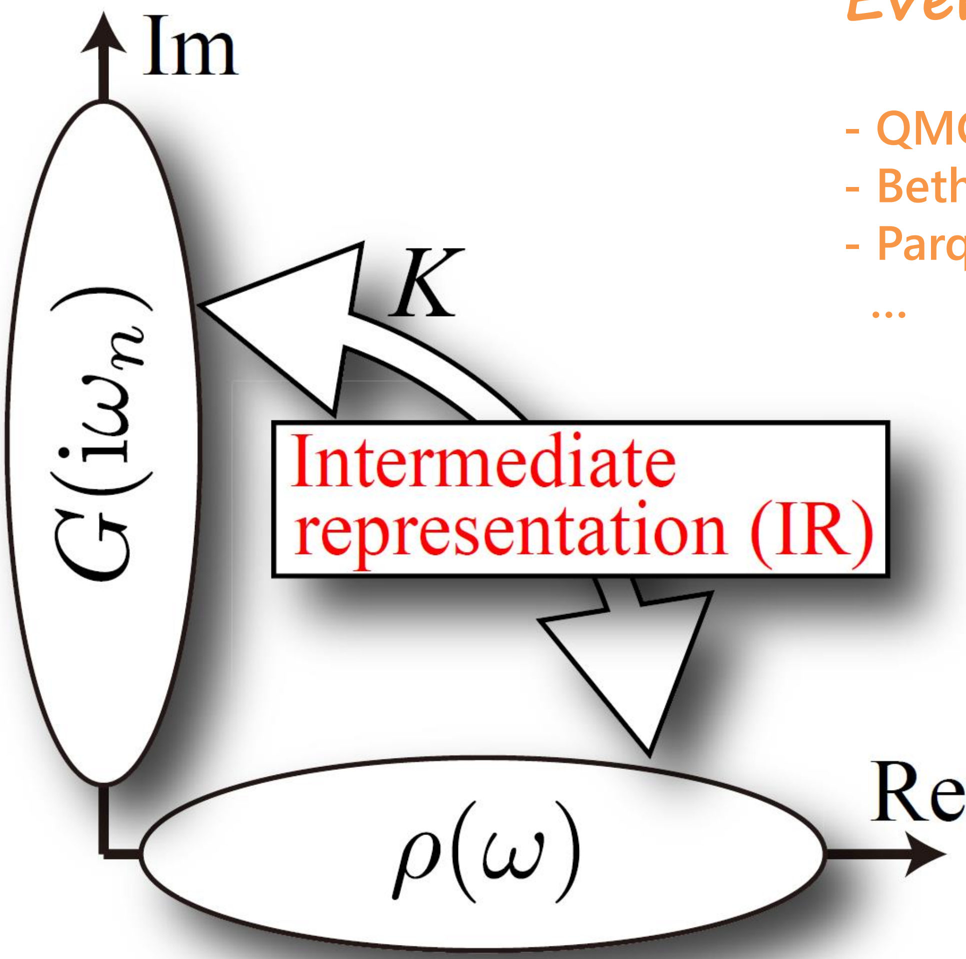
$$\sum_{l=0}^{\infty} F_l P_l(\cos\theta_{\boldsymbol{k},\boldsymbol{k}'})$$

Landau parameter

$$\sum_{l,l'}^{\infty} \gamma_{ll'} u_l(i\omega_n) u_{l'}(i\omega_{n'})$$

Dynamical version of
Landau parameters?

A "new" rep



Everything in IR basis!

- QMC measurement
- Bethe-Salpeter equation
- Parquet equation

...

Summary

- *Analytical continuation*
  I. SVD of the kernel
  II. Basis selection by L1 regularization

https://github.com/j-otsuki/SpM

- *Efficient basis set*
  - Extremely compact representation
  - QMC measurement, diagrammatic calculations, etc

extract essence by sparse modeling