Controlled analytic continuation of Matsubara correlation functions using minimal information principle

Lei Zhang
University of Michigan
lzphy@umich.edu

Phys. Rev. B 110, 035154 (2024), by **L. Z.** and Emanuel Gull
Phys. Rev. B 110, 235131 (2024), by **L. Z.**, Yang Yu and Emanuel Gull

Outline

- Introduction
  - Green’s functions, NAC problem, Prony-like methods
- Minimal pole method
  - Complex pole representation, Causal projection, Minimal information principle, Holomorphic mapping, Generalization
- Numerical results
  - Convergence, Robustness, Versatility, Imaginary time, Matrix-valued, Realistic simulations
- Other applications
  - Compact representation, bath fitting
- Code
- Conclusion

Introduction

Green's functions

$$H = H^\dagger = \sum_{ij}(t_{ij} - \mu\delta_{ij})d_i^\dagger d_j + \sum_{ijkl} U_{ijkl} d_i^\dagger d_j^\dagger d_l d_k \text{ with } [d_i, d_j^\dagger]_\pm = \delta_{ij}$$

Imaginary axis (simulation):

$$G_{ij}(\tau) = -\langle T_\tau d_i(\tau) d_j^\dagger \rangle_{\text{th}} \text{ with } \tau \in [-\beta, \beta],\ \ G_{ij}(\tau < 0) = \mp G_{ij}(\tau + \beta)$$

$$G_{ij}^{\text{Mat}}(i\omega_n) = \int_0^\beta d\tau e^{i\omega_n \tau} G_{ij}(\tau) \text{ with } i\omega_n = i\frac{(2n+1)\pi}{\beta}\text{(fermion) or } i\frac{2n\pi}{\beta}\text{(boson)}$$

Real axis (interpretation):

$$G_{ij}^{\text{Ret}}(t) = -i\Theta(t)(\langle d_i(t) d_j^\dagger \rangle_{\text{th}} \pm \langle d_j^\dagger d_i(t) \rangle_{\text{th}}) \text{ with } t \in (-\infty, +\infty)$$

$$G_{ij}^{\text{Ret}}(\omega) = \int_{-\infty}^{+\infty} dt e^{i\omega t} G_{ij}^{\text{Ret}}(t) \text{ with } \omega \in (-\infty, +\infty)$$

$$A_{ij}(\omega) = -\frac{1}{\pi}\text{Im}[G_{ij}^{\text{Ret}}(\omega)] \text{ (measured by ARPES) and } G_{ij}^{\text{Mat}}(i\omega_n) = \int_{-\infty}^{+\infty} d\omega \frac{A_{ij}(\omega)}{i\omega_n - \omega}$$

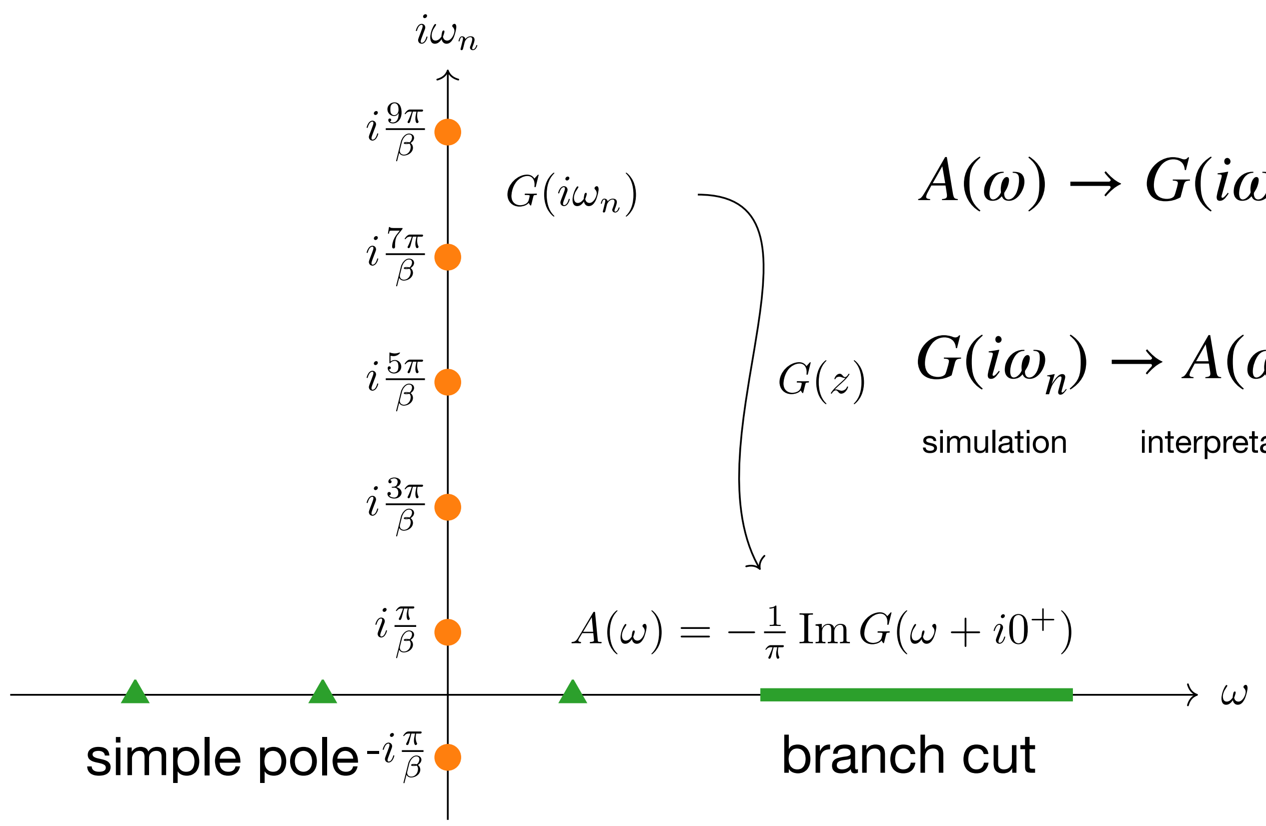
Lehmann representation:

$$H|n\rangle = E_n|n\rangle,\ \ \sum_n |n\rangle\langle n| = I,\ \ \boldsymbol{d}^\dagger = (d_1^\dagger, \cdots, d_n^\dagger)$$

$$\boldsymbol{G}(z) = \frac{1}{\mathcal{Z}}\sum_{m,n}\frac{e^{-\beta E_m} \pm e^{-\beta E_n}}{z - (E_n - E_m)}\langle m|\boldsymbol{d}|n\rangle\langle n|\boldsymbol{d}^\dagger|m\rangle = \sum_l \frac{\boldsymbol{A}_l^{(\text{real})}}{z - \xi_l^{(\text{real})}}$$

with $\boldsymbol{G}(z = i\omega_n) = \boldsymbol{G}^{\text{Mat}}(i\omega_n)$ and $\boldsymbol{G}(z = \omega + i0^+) = \boldsymbol{G}^{\text{Ret}}(\omega)$

$$\boldsymbol{A}(\omega) = \sum_l \boldsymbol{A}_l^{(\text{real})}\delta(\omega - \xi_l^{(\text{real})}) \text{ with } \boldsymbol{A}_l \text{ (fermion) or } \text{sgn}(\omega)\boldsymbol{A}_l \text{ (boson) being PSD}$$



$A(\omega) \rightarrow G(i\omega_n)$: by numerical integration (controlled)

$G(i\omega_n) \rightarrow A(\omega)$: numerical analytic continuation (NAC)

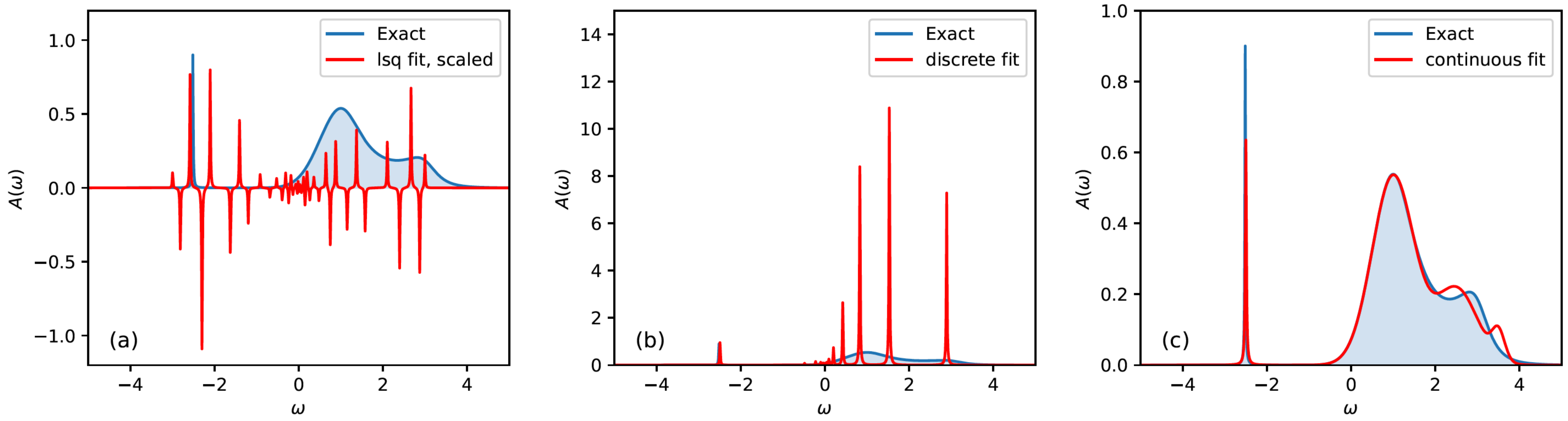
simulation interpretation

ill-conditioned!

NAC problem

Input: $G(i\omega_n)$ obtained from an exact $A(\omega)$ (blue)

Output: recovered solutions $\{A^{(i)}(\omega)\}$ (red)



NAC is ill-conditioned. In any finite precision:

- infinite causal solutions $\{A^{(i)}(\omega)\}$
- $A^{(i)}(\omega)$ and $A^{(j)}(\omega)$ could differ A LOT:

  $\sup_{ij}|A^{(i)}(\omega) - A^{(j)}(\omega)| = +\infty$ for any $\omega \in (-\infty, +\infty)$
- Impossible for *any* method to completely eliminate bias

Without prior knowledge, choose which one?
How reliable?

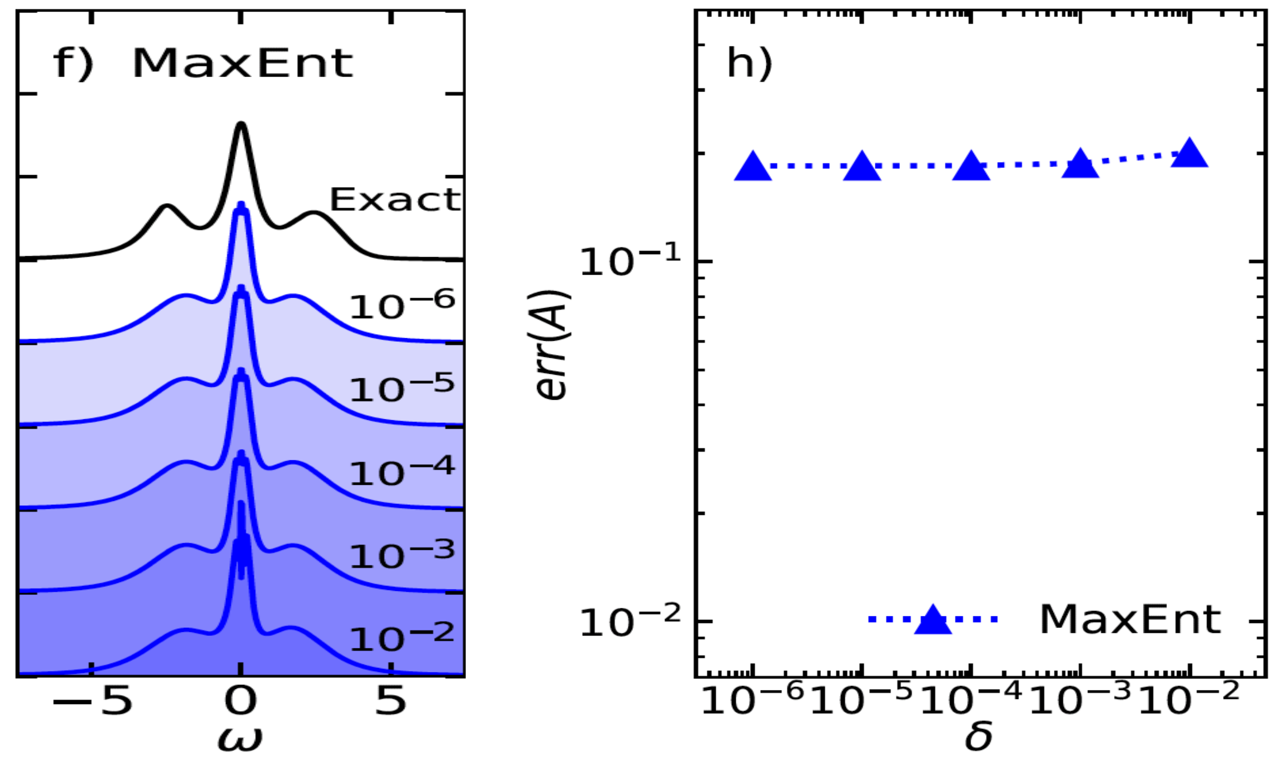
It was commonly believed that there just wasn't enough information in finite-temperature field theories to reliably extract spectra.

NAC methods

Fitting methods:

- Maximum entropy method (MaxEnt) Eur. Biophys. J. 18, 165 (1990) Phys. Rep. 269, 133 (1996)
- Stochastic optimization method (SOM) Phys. Rev. B 62, 6317 (2000)
- …

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

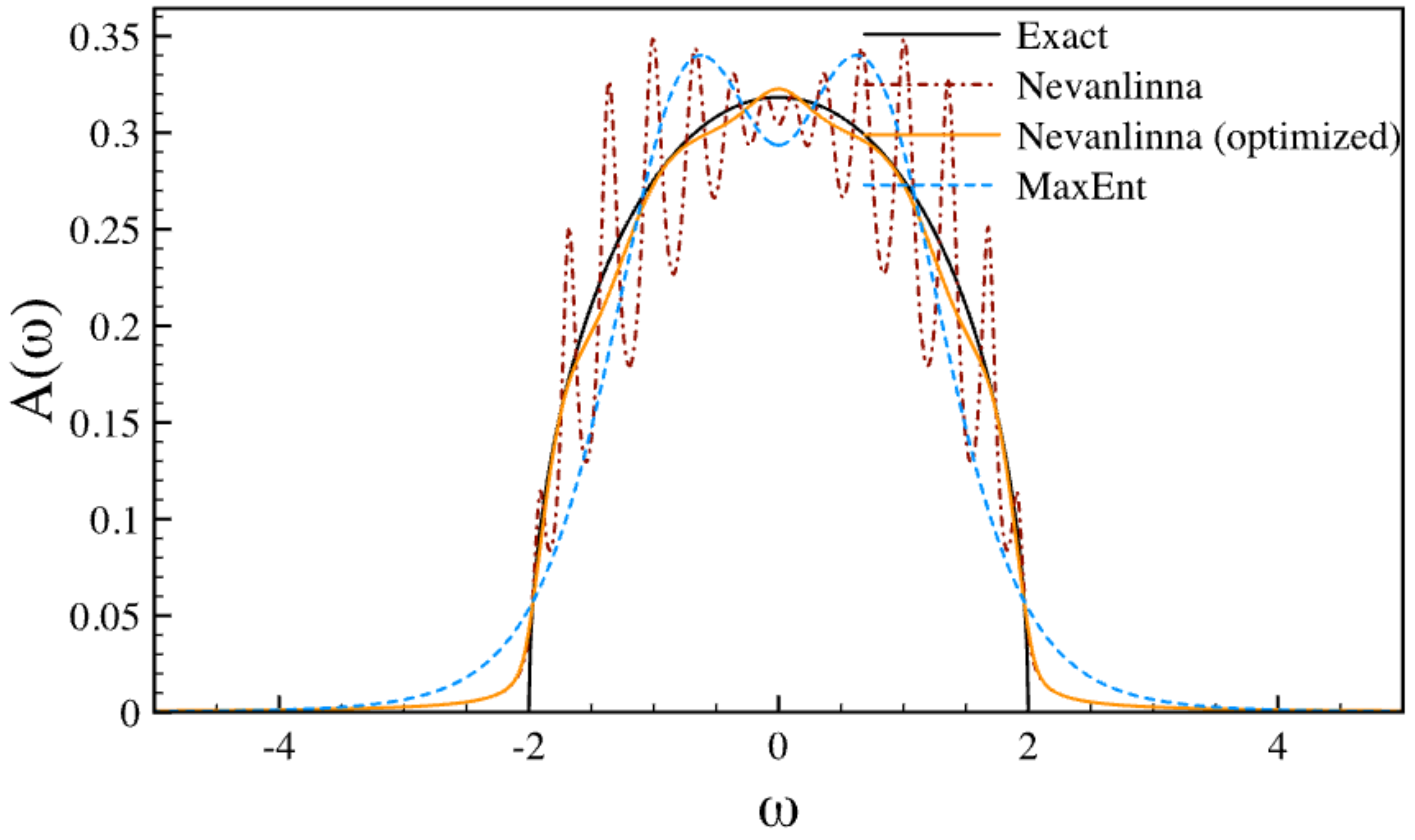


Robust to noise
Powerful for large noise
Not systematically improvable

Interpolation methods:

- Padé interpolation method J. Low Temp. Phys. 29, 179 (1977) Phys. Rev. B 61, 5147 (2000)
- Nevanlinna / Carathéodory method Phys. Rev. Lett. 126, 056402 (2021) Phys. Rev. B 104, 165111 (2021)
- …

S. Iskakov et al., Comput. Phys. Commun. 304, 109299 (2024)



Not robust to noise
Powerful for precise data
Ambiguity still exists

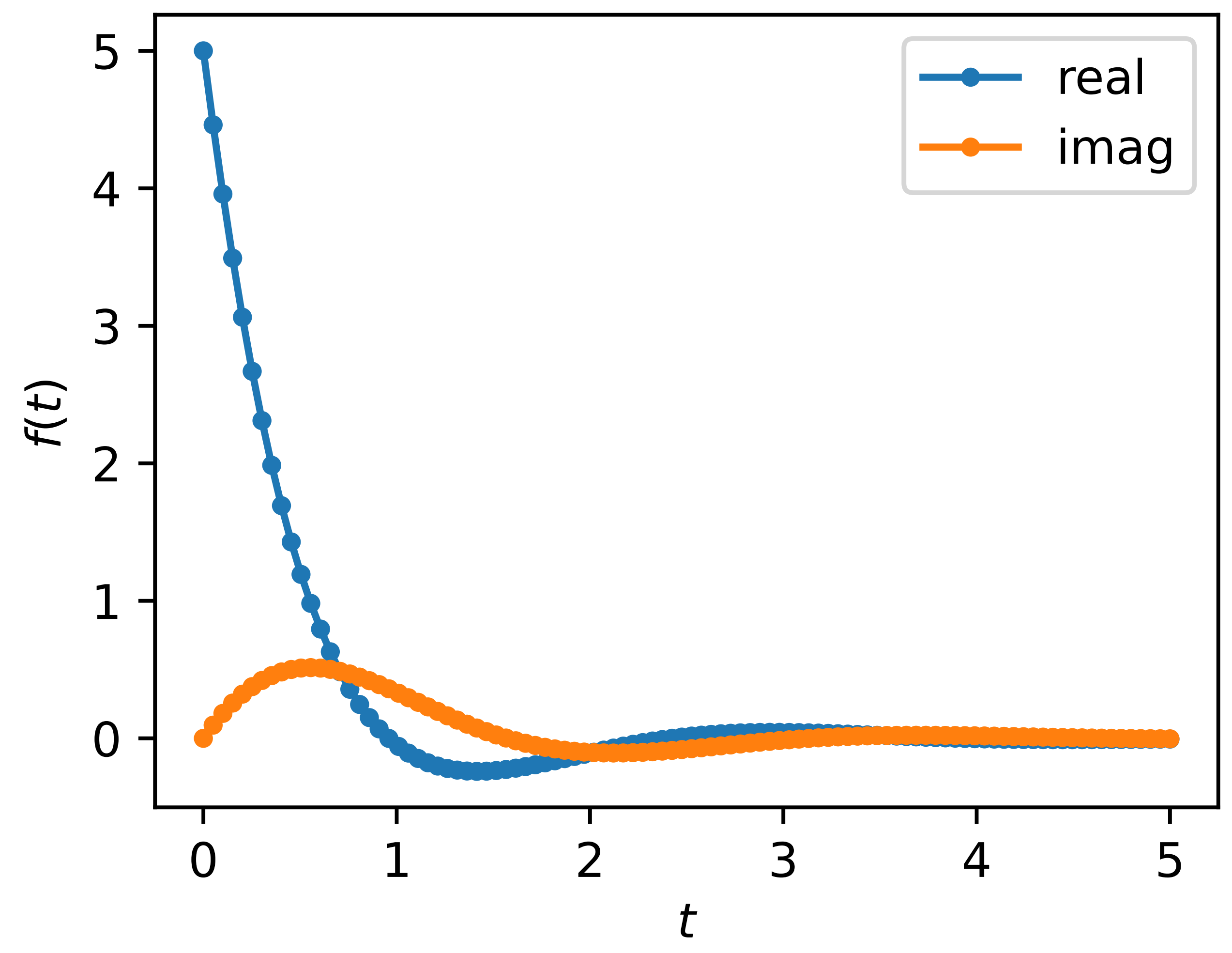
Our claim: Even without prior knowledge, it is still possible to design a systematically improvable NAC method.

Prony-like methods

Input: $\{f(t_k), t_k\}$ on a uniform grid

Ansatz: $f(t) \approx \sum_{i=1}^{M} R_i e^{s_i t} \xrightarrow{\text{discretize}} f(t_k) \approx \sum_{i=1}^{M} R_i z_i^k$

Goal: estimate $M \in \mathbb{Z}$, $R_i \in \mathbb{C}$ and $z_i \in \mathbb{C}$



Solution: ESPRIT, Matrix Pencil, Prony approximation…

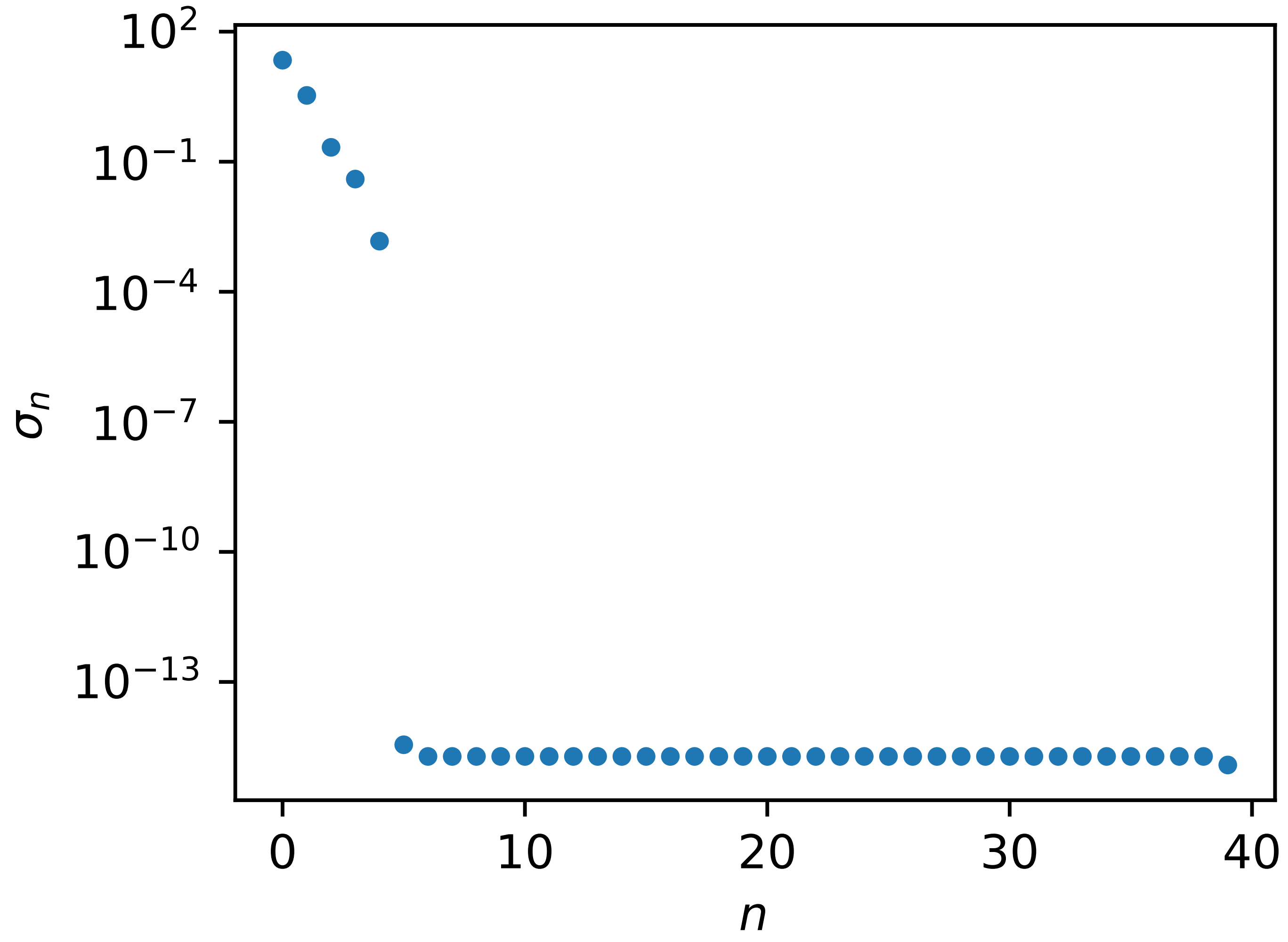
IEEE Trans. Acoust. Speech, Signal Process. 37, 984 (1989)

IEEE Trans. Acoustics, Speech, Signal Process. 38, 814 (1990)

Appl. Comput. Harmonic Anal. 19, 17 (2005)

ESPRIT:

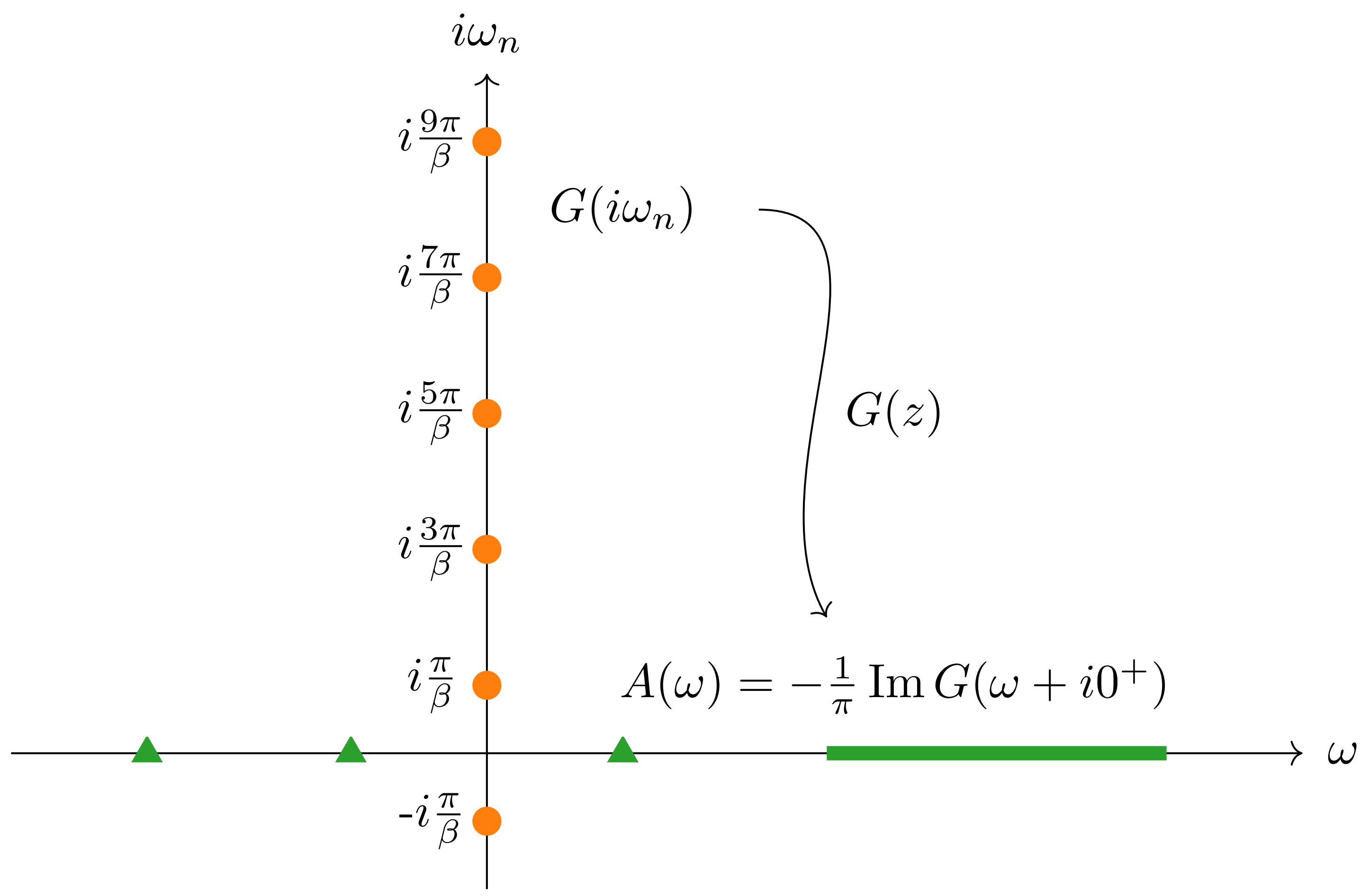
1. Construct $(N-L)\times(L+1)$ Hankel matrix $H_{ij} = f(t_{i+j})$, with $\frac{N}{3} \le L \le \frac{N}{2}$
2. Perform SVD: $H = U\Sigma W$
3. Obtain M from $\sigma_{M+1} < \varepsilon$
4. Obtain z_i from eigenvalues of $(W(1:M, 1:L)^T)^{+} W(1:M, 2:L+1)^T$
5. Obtain R_i from least-square fits



Robust to noise; use a minimal number of complex exponentials to fit the data to the given precision ε

Minimal Pole Method (MPM)

Ingredient 1: spectral representation

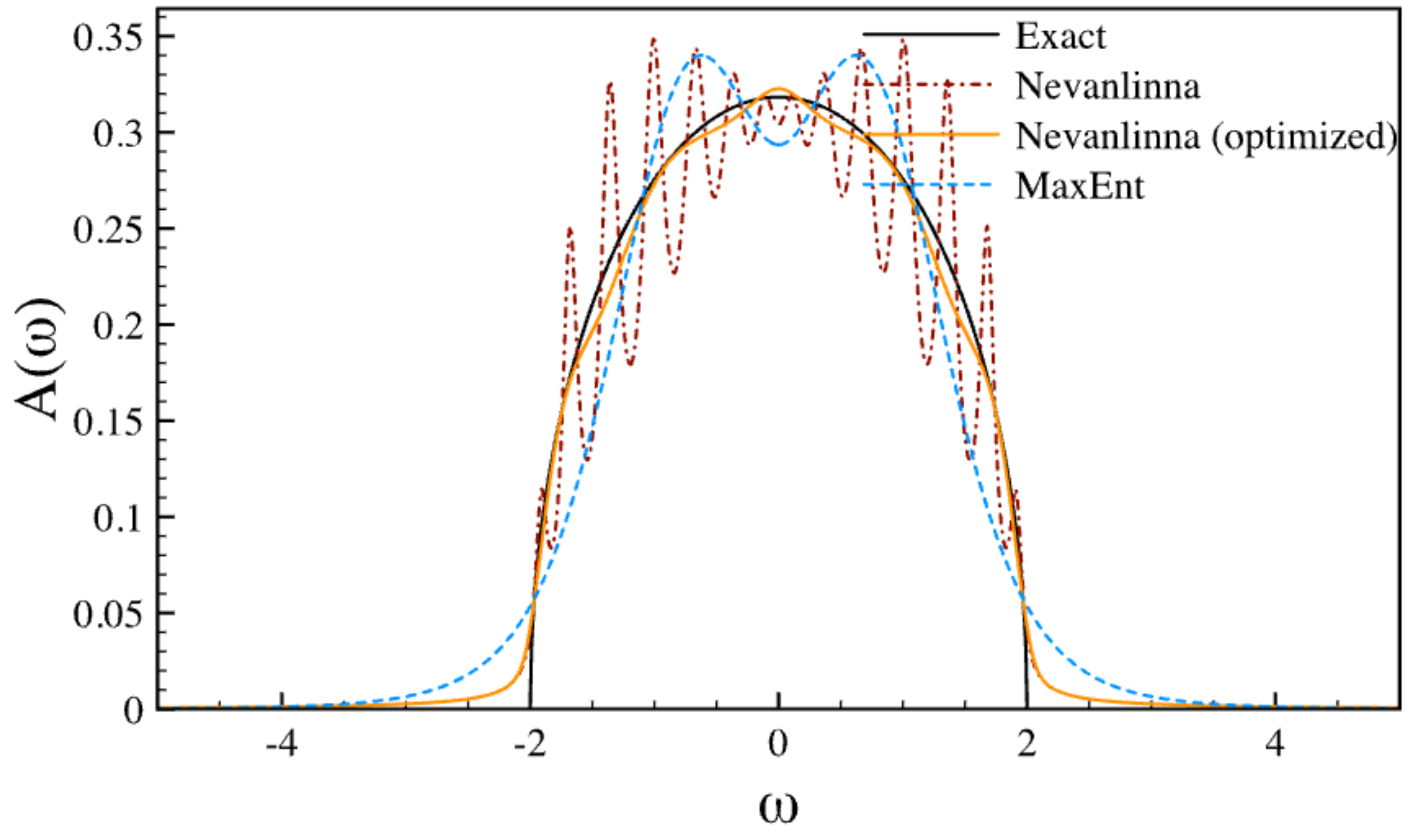


Options:

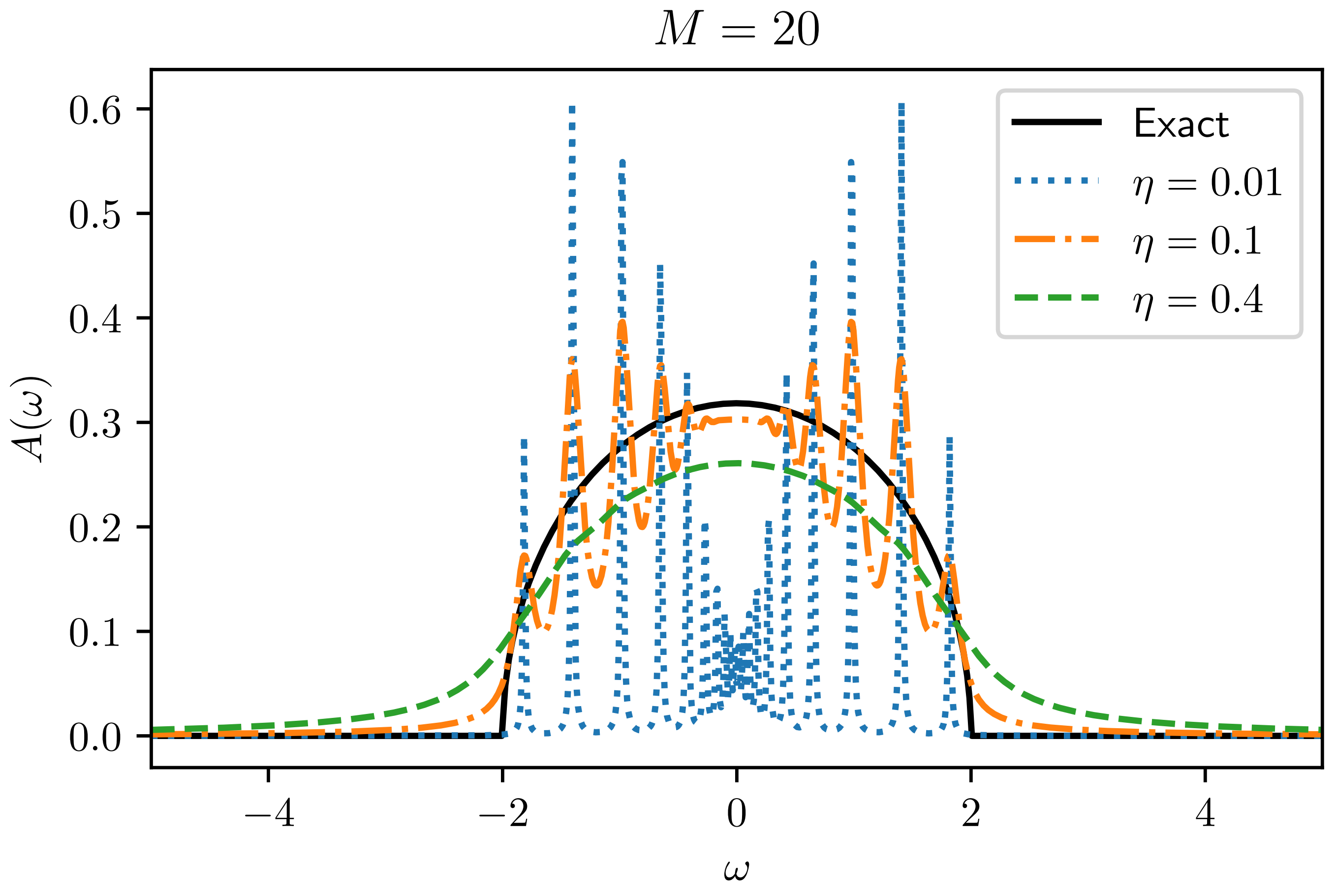
1. Predetermined frequency grid + smoothness (sharp features ✕)
2. Real pole + artificial broadening (smooth features ✕)
3. Complex conjugates of poles on the entire plane (incorrect ✕)
4. Complex poles in the lower-half plane (sharp & smooth features ✓)

Finite grid + smoothness criterion

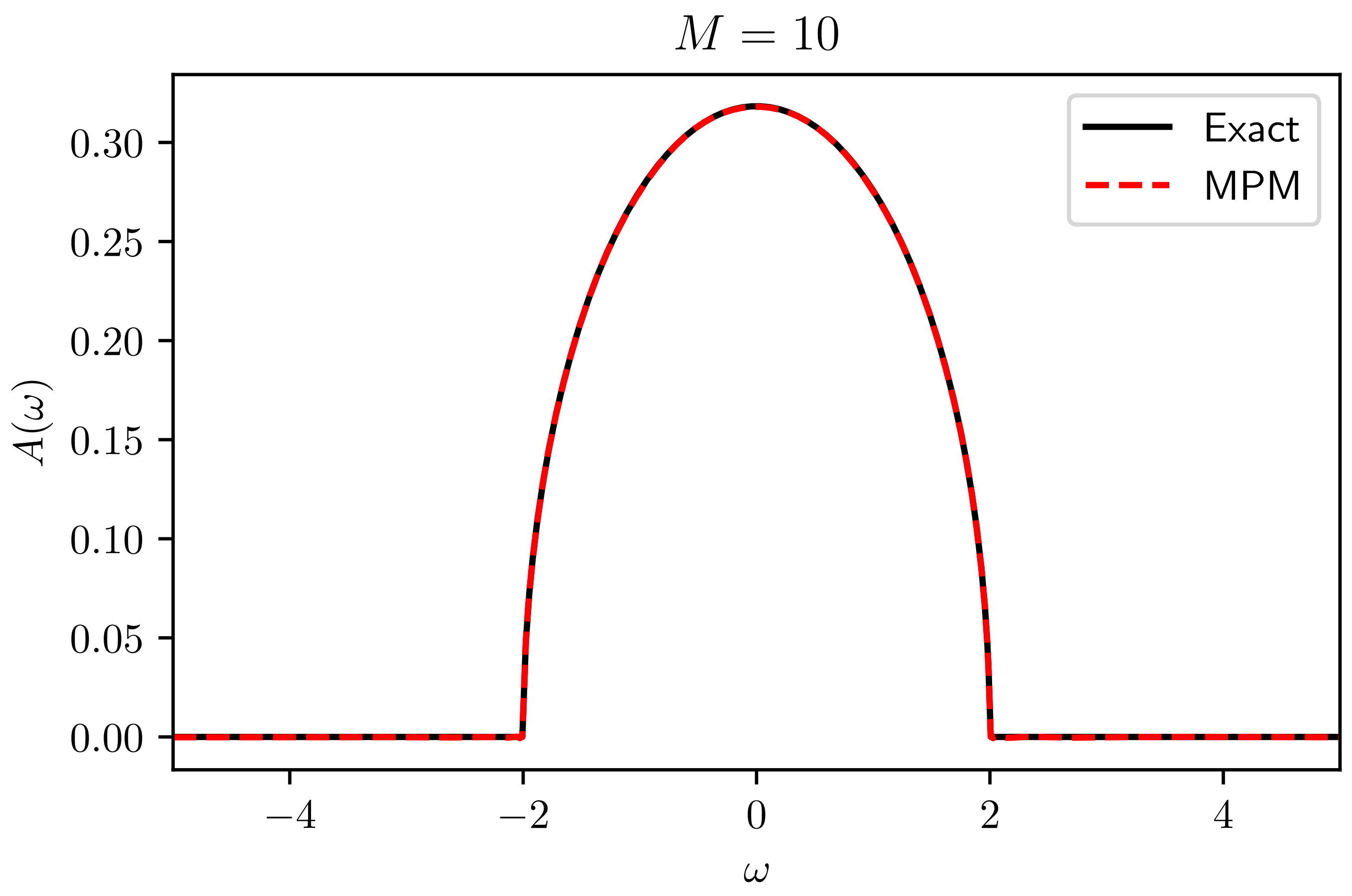
S. Iskakov et al., Comput. Phys. Commun. 304, 109299 (2024)



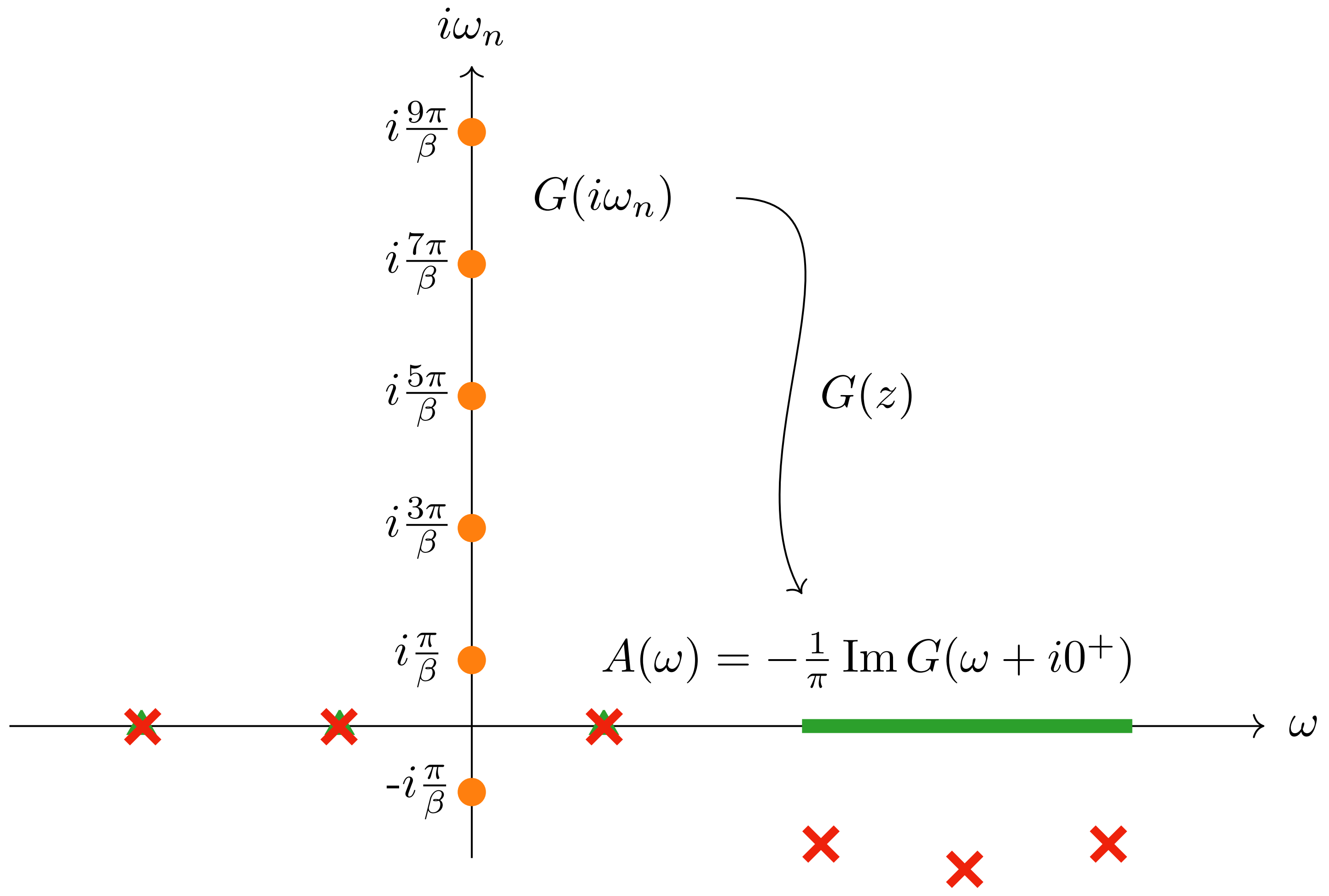
Real pole



Complex pole



Ingredient 1: complex pole representation



For $\mathrm{Im}z > 0$:

$$G(z) = \sum_{l=1}^{M} \frac{A_l}{z - \xi_l}, \ \mathrm{Im}\xi_l \leq 0$$

$A_l \in \mathbb{C}$ and $\xi_l \in \mathbb{C}$

For $\mathrm{Im}z < 0$:

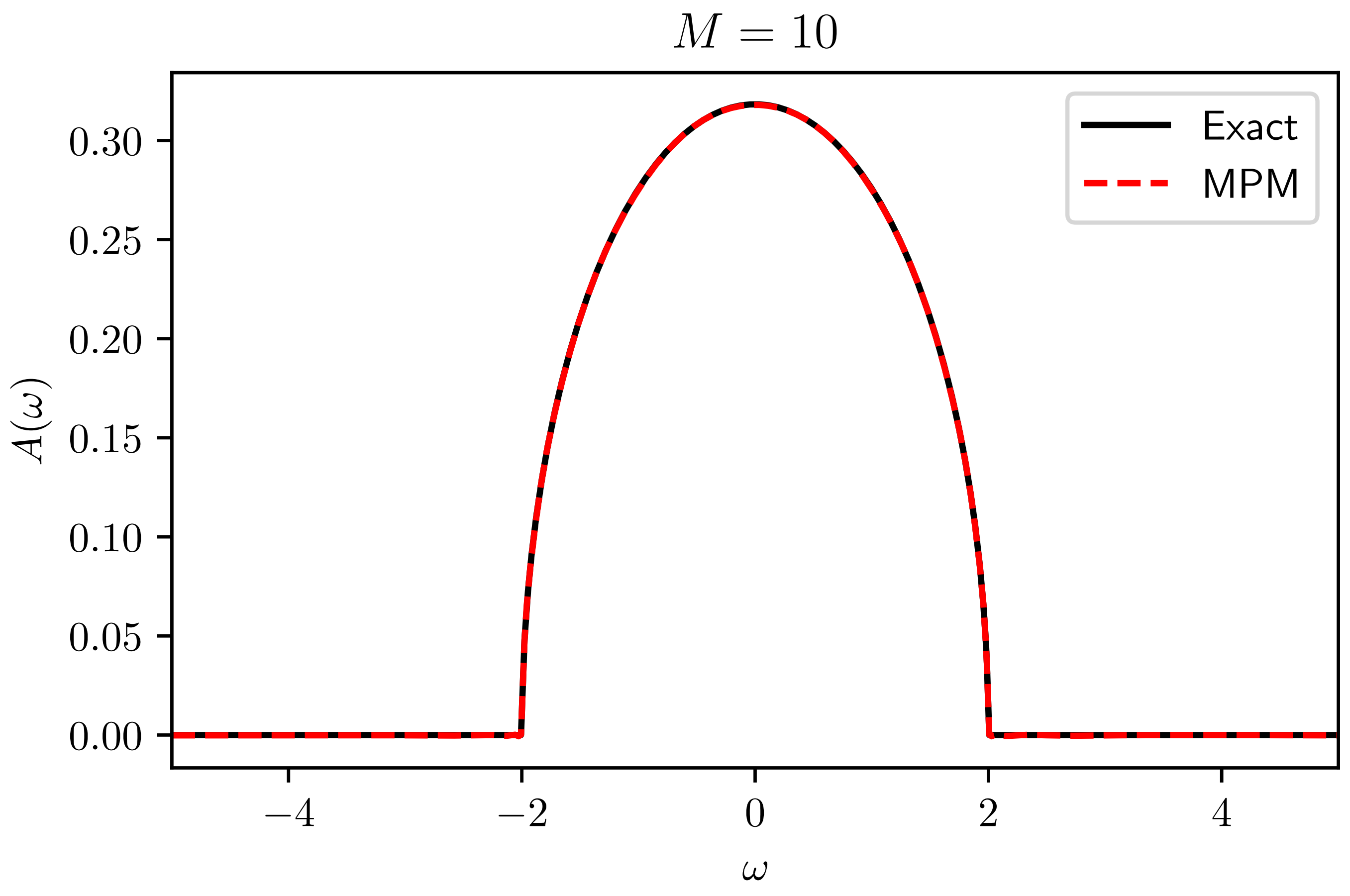
$$G(z) = \sum_{l=1}^{M} \frac{A_l^*}{z - \xi_l^*}, \ \mathrm{Im}\xi_l^* \geq 0$$

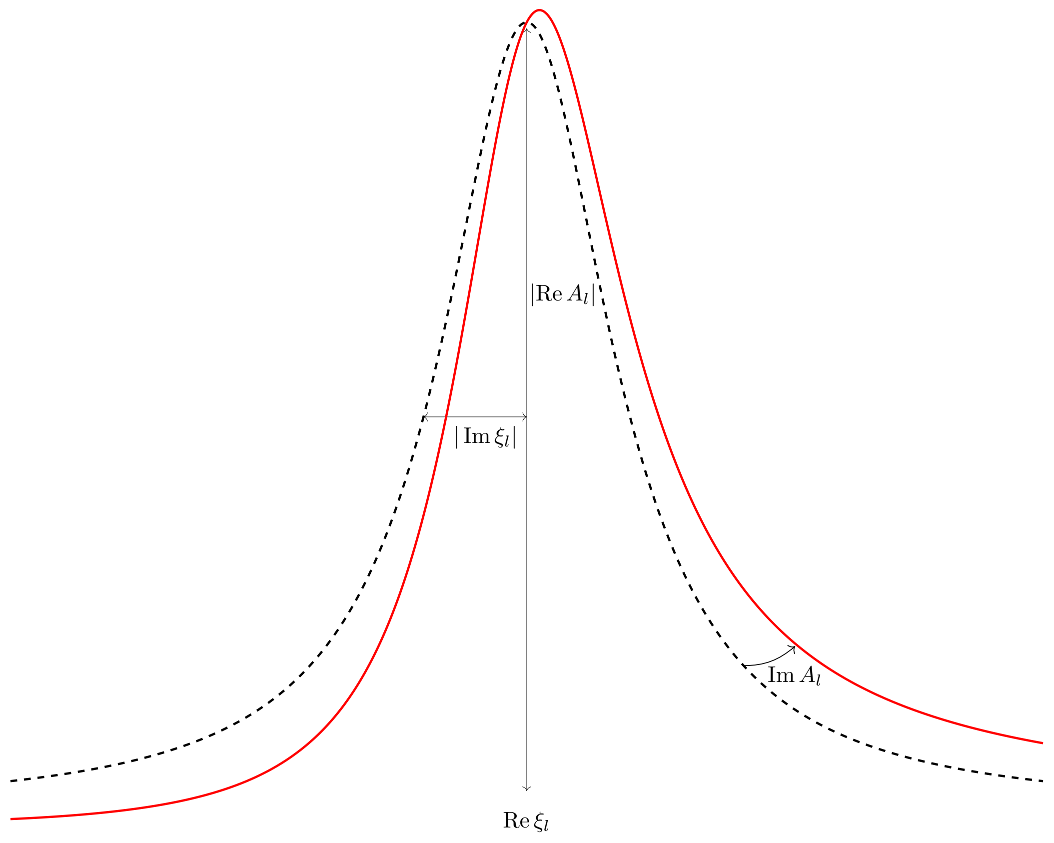
Also used in other NAC methods:

- L. Ying, J. Comput. Phys. 469, 111549 (2022)
- L. Ying, J. Sci. Comput. 92, 107 (2022)
- L. Huang et al., Phys. Rev. B 111, 125139 (2025)
- …

Also useful after applying the Fourier transform:

- HEOM — Y. Tanimura et al., J. Phys. Soc. Jpn. 58, 101 (1989)
- Quasi-Linblad — G. Park et al., Phys. Rev. B 110, 195148 (2024)
- Real-time QMC — A. Erpenbeck, Session MAR-N47
- …





- A sum of “generalized Lorentzians”
- $\mathrm{Im}A_l$ compensates for asymmetric parts
- Numerically sufficient to capture all spectral features and to perform calculations
  (J. Chem. Phys. 162, 214111 (2025), L. Zhang et al.)

Ingredient 2: controlled approximation for $G(iy)$

Maximum modulus principle:

$$|G_{\text{approx}}(\omega + i0^+) - G_{\text{exact}}(\omega + i0^+)| \le \varepsilon$$

$$\to |G_{\text{approx}}(iy) - G_{\text{exact}}(iy)| \le \varepsilon \text{ for } y \in [\omega_n, \omega_{n+1}]$$

Approximation in the continuous interval must be controlled!

Not possible for an arbitrary function

But … we only care about functions with physical meaning

Fermion: $-\operatorname{Im} \boldsymbol{G}(iy) = \sum_l \frac{y}{y^2 + \xi_l^{(r)2}} \boldsymbol{A}_l$ $(y > 0)$, $\boldsymbol{A}_l$ is positive semi-definite (PSD)

"least oscillation"

Boson: $-\operatorname{Re} \boldsymbol{G}(iy) = \sum_l \frac{|\xi_l^{(r)}|}{y^2 + \xi_l^{(r)2}} \operatorname{sign}(\xi_l^{(r)}) \boldsymbol{A}_l$, $\operatorname{sign}(\xi_l^{(r)}) \boldsymbol{A}_l$ is PSD

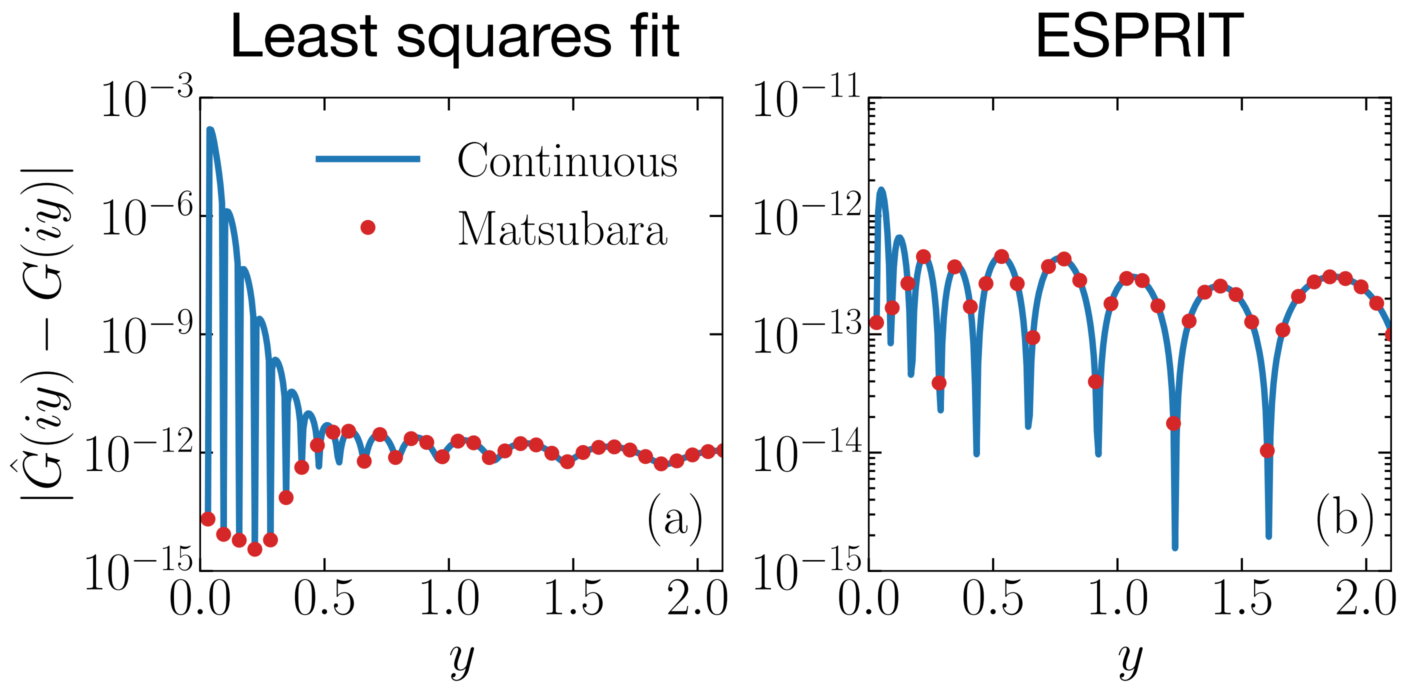
R.Roy et al., IEEE Trans. Acoust. Speech, Signal Process. 37, 984 (1989)

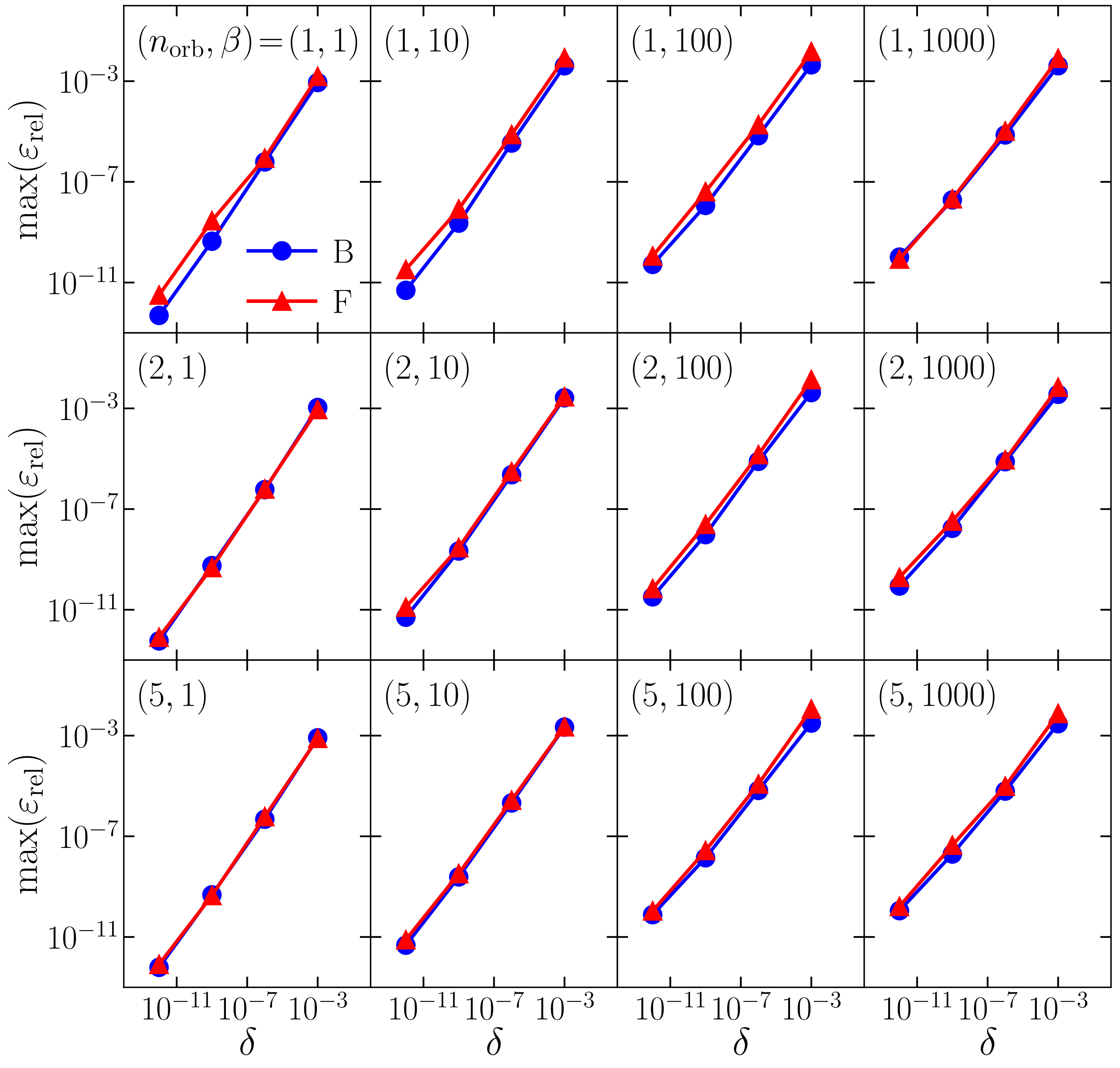
Prony-like problem + ESPRIT algorithm: $G_{\text{approx}}(i\omega_n) = \sum_i R_i z_i^n$

- minimal # of exponentials
- fit but not over-fit
- arbitrary precision as data allows
- suppress oscillations in between

1. $\left\| \boldsymbol{G}^{(L)}_{\text{approx}}(i\omega_n) - \boldsymbol{G}_{\text{input}}(i\omega_n) \right\| \leqslant \varepsilon$ for $0 \le n \le n_\omega - 1$

2. $\left\| \boldsymbol{G}^{(L')}_{\text{approx}}(iy) - \boldsymbol{G}^{(L)}_{\text{approx}}(iy) \right\| \leqslant \varepsilon$ for $n_0 \le n \le n_\omega - 1$





Controlled $|G_{\text{approx}}(iy) - G_{\text{exact}}(iy)| \le C\varepsilon$ for $y = \omega_0, \cdots, \omega_{n_0 - 1}$ and $\omega_{n_0} \le y \le \omega_{n_\omega - 1}$ (Typically, $n_0 \approx 0$)

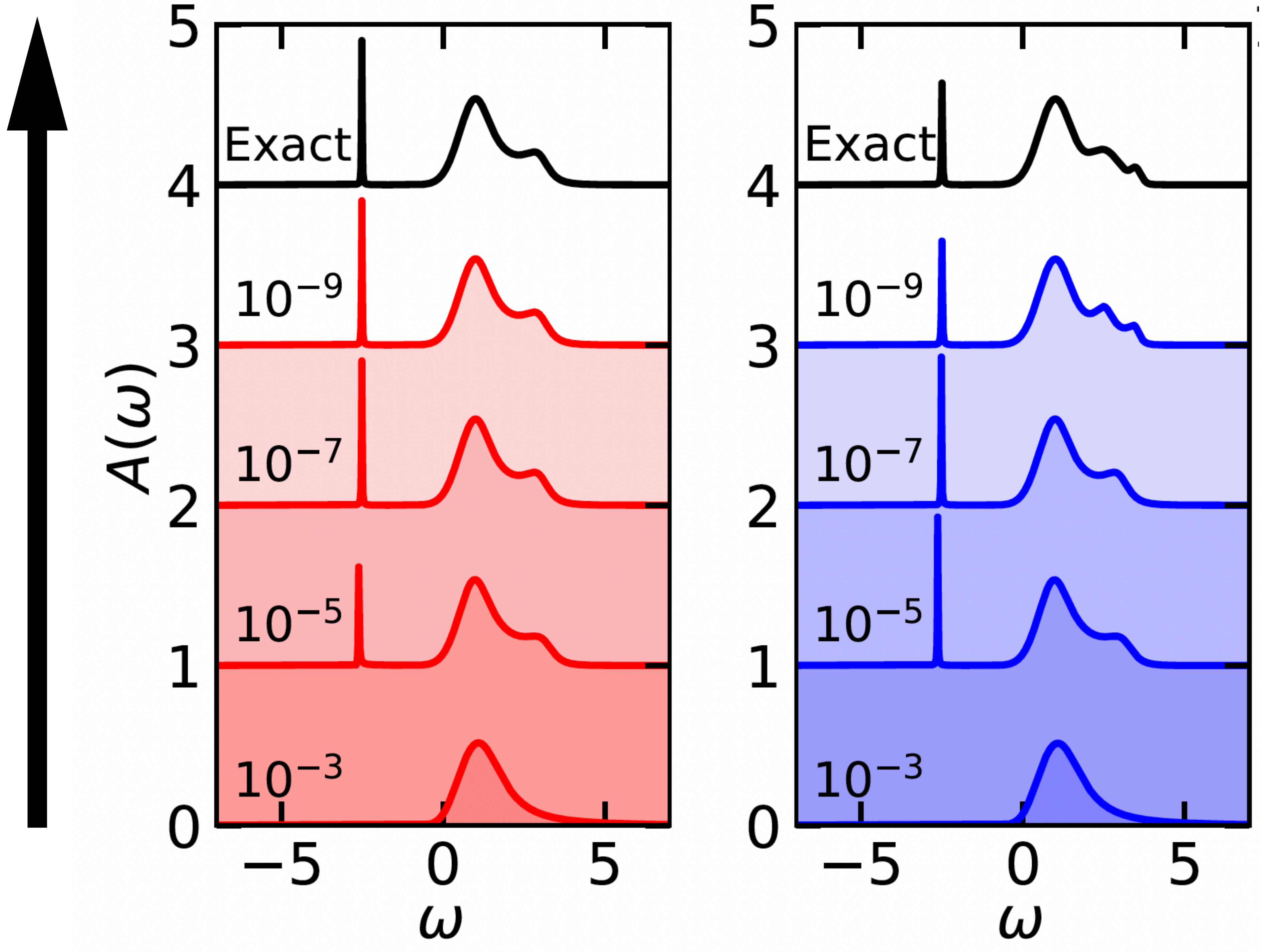
L. Zhang et al., Phys. Rev. B 110, 235131 (2024)

Ingredient 3: minimal information principle

In our method, minimal information = minimal number of complex poles

Any different $A^{(i)}(\omega)$ and $A^{(j)}(\omega)$ can always be distinguished at sufficiently high precision.

Required minimal information increases



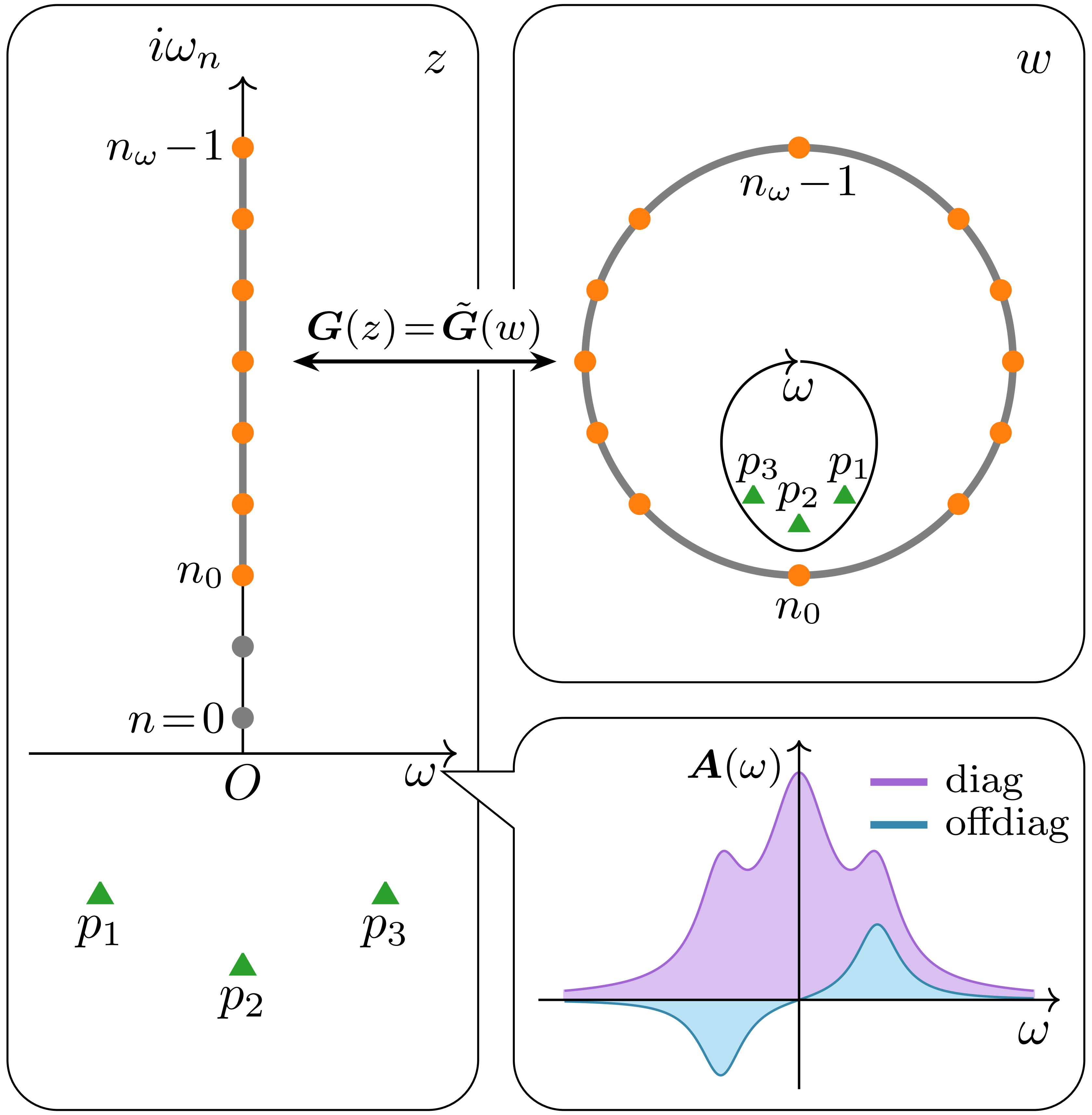
Distinguishable and results have converged

Undistinguishable

Spoiler: results obtained by L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

Ingredient 4: conformal mapping

L. Zhang et al., Phys. Rev. B 110, 235131 (2024)



$$G(z) = \sum_{l=1}^{M} \frac{A_l}{z - \xi_l}, \ \mathrm{Im}\, \xi_l \leq 0$$

- Controlled $|G_{\text{approx}}(iy) - G_{\text{exact}}(iy)| \leq \varepsilon$ for $y = \omega_0, \cdots, \omega_{n_0 - 1}$ and $\omega_{n_0} \leq y \leq \omega_{n_\omega - 1}$

  ε: data precision   Typically, $n_0 \approx 0$

- Shown by **L. Ying, J. Comput. Phys. 469, 111549 (2022)** and **L. Ying, J. Sci. Comput. 92, 107 (2022)**
  - Conformal mapping + residue theorem -> Prony's problem

    Contour integral $\quad h_k := \dfrac{1}{2\pi i} \displaystyle\int_{\partial \bar{D}} \tilde{G}(w) w^k dw$

    Residue theorem $\quad h_k = \displaystyle\sum_l \tilde{A}_l \tilde{\xi}_l^k, \ k \geq 0$. recovered by ESPRIT

  - Transform back

$$\xi_l = g^{-1}(\tilde{\xi}_l) = \frac{\Delta\omega_{\mathrm{h}}}{2}(\tilde{\xi}_l - \frac{1}{\tilde{\xi}_l}) + i\omega_{\mathrm{m}} ,$$
$$A_l = \mathrm{Res}[G(z), \xi_l] = \frac{\Delta\omega_{\mathrm{h}}}{2}(1 + \frac{1}{\tilde{\xi}_l^2})\tilde{A}_l .$$

optional restrictions to causality and sum rules

One major difference in our work: We account for the approximation on $y \in (\omega_n, \omega_{n+1})$, which is necessary for the method to be systematically improvable.

Minimal Pole Method:

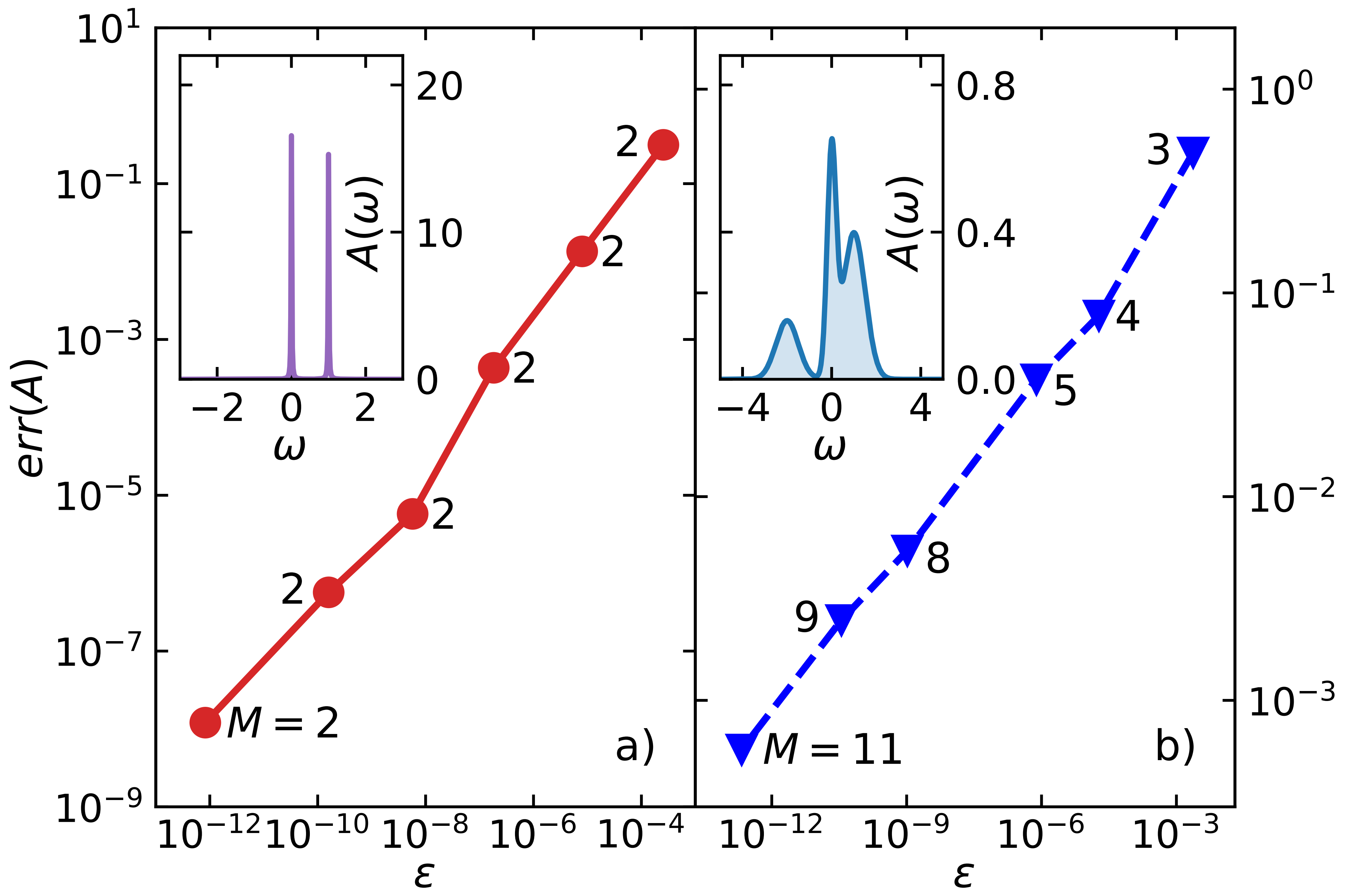
Recover a minimal number of complex poles to fit transformed spectral moments to within ε

Generalization

- Restriction to real poles
  - Modify the holomorphic mapping to map $(-\infty, -i\omega_{n_0}] \cup [i\omega_{n_0}, +\infty)$ to the unit circle
- Generalize to multiple-orbital cases
  - Shared poles with matrix-valued weights
  - Achieved by matrix-valued ESPRIT (L. Zhang et al., Phys. Rev. B 110, 235131 (2024))
- Combination with discrete Lehmann representation (J. Kaye et al., Phys. Rev. B 105, 235115 (2022))
  - h_k is estimated by $\sum_{i=1}^{r} \tilde{g}_i^{(\mathrm{dlr})} \tilde{\omega}_i^{(\mathrm{dlr})\,k}$
  - Larger n_0 (slower convergence) but way higher computational efficiency

Numerical results

Results — convergence



$$G(z) = \sum_{l=1}^{M} \frac{A_l}{z - \xi_l}, \ \mathrm{Im}\, \xi_l \leq 0$$

(a) $A(\omega) = 0.52\delta(\omega) + 0.48\delta(\omega - 1)$

(b) $A(\omega) = 0.2g(\omega, -2, 0.5) + 0.3g(\omega, 0, 0.2) + 0.5g(\omega, 1, 0.5)$

ε : precision of the approximation $err(A) = \int d\omega \, |A_{\mathrm{cont}}(\omega) - A_{\mathrm{exact}}(\omega)|$

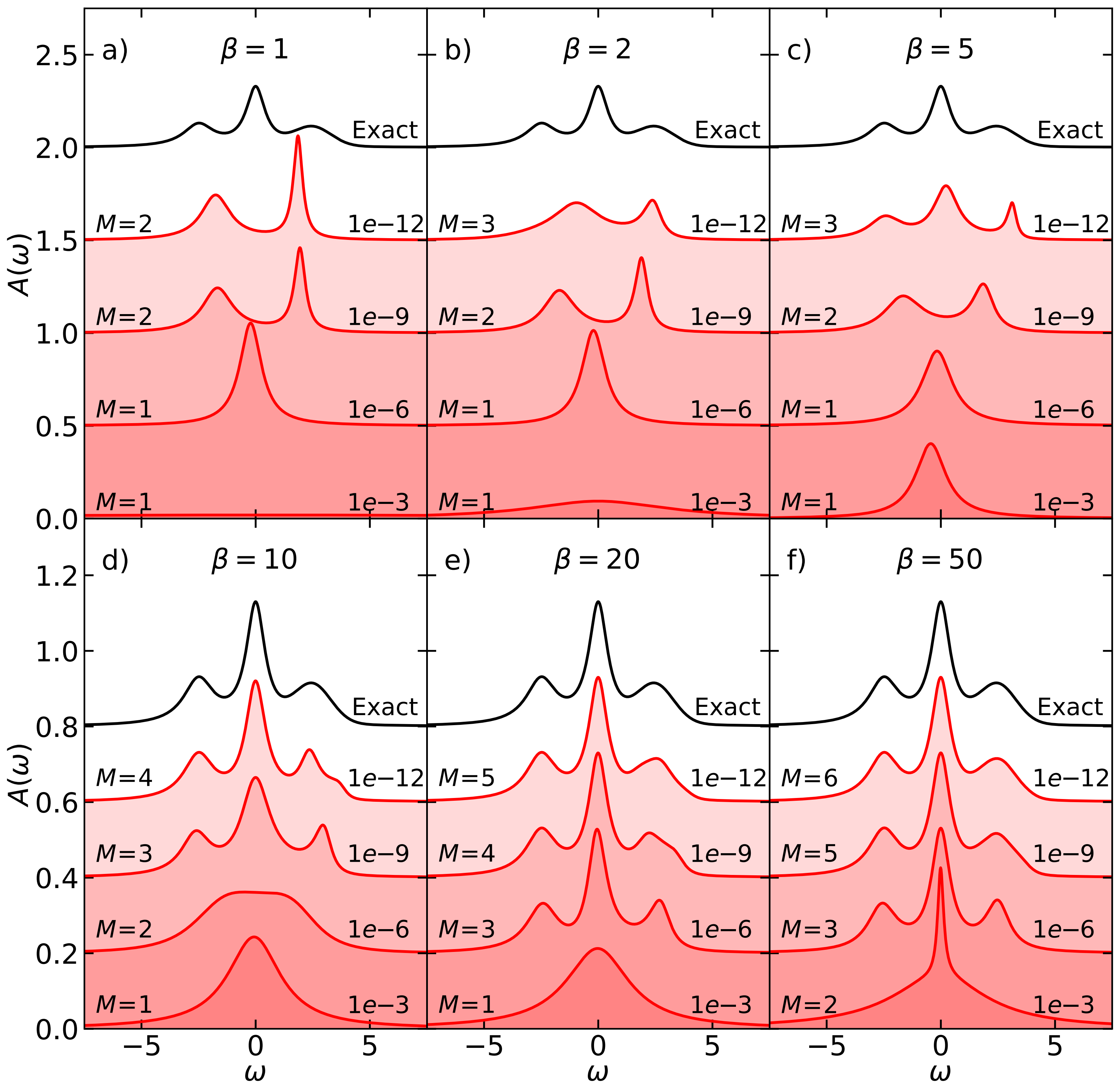
$\beta = 200, N_\omega = 2001$

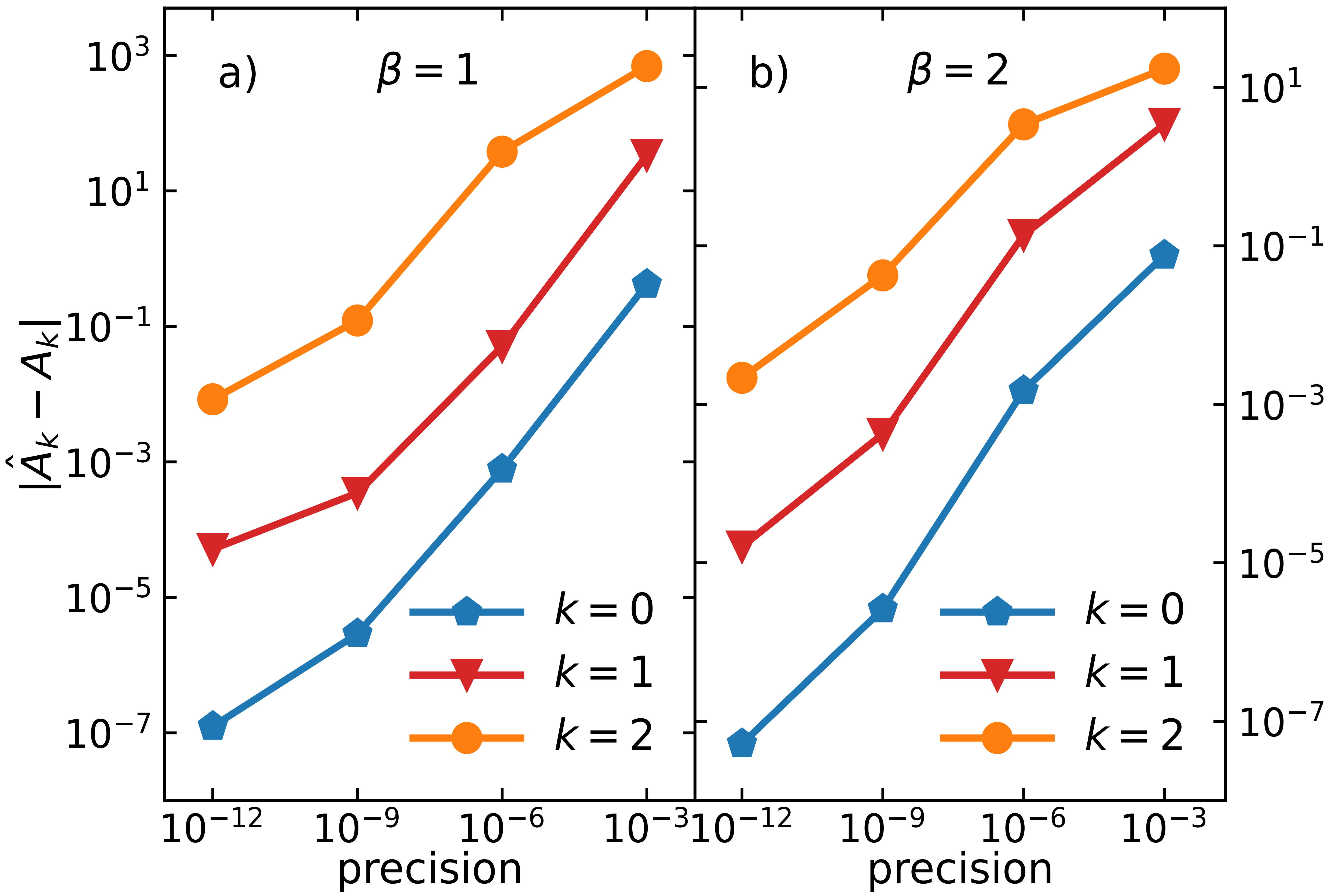
As data precision increases, results converges

- Regardless of temperature and systems
- As temperature increases, convergence becomes slower
- Robust to n_ω: $(n_\omega)_{\min} \gtrapprox 2n_{\mathrm{pole}} + 1$

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

Results — dependence on temperature

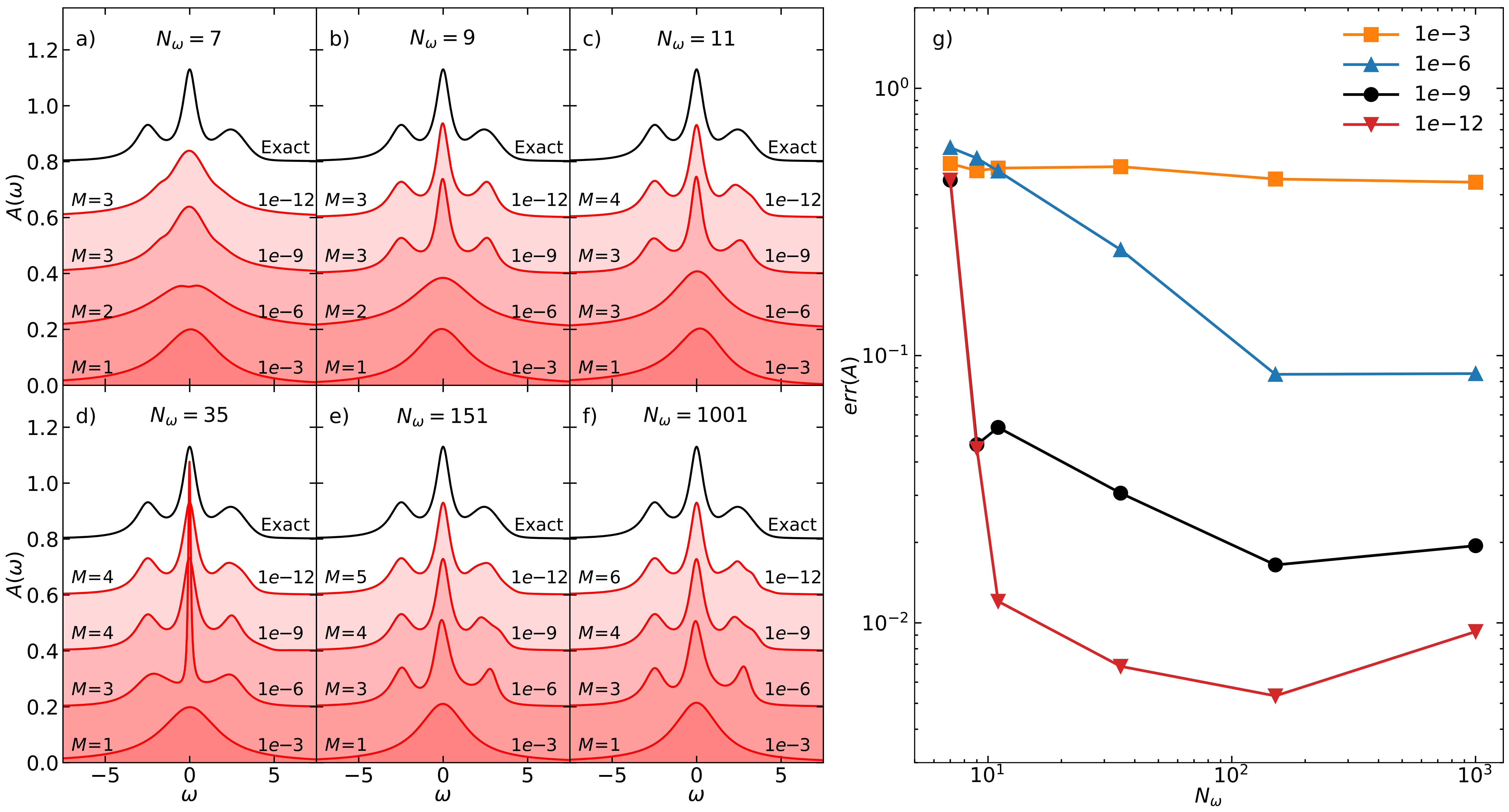




Data precision increases -> spectral moments more accurate -> recovered spectrum closer to the correct result

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

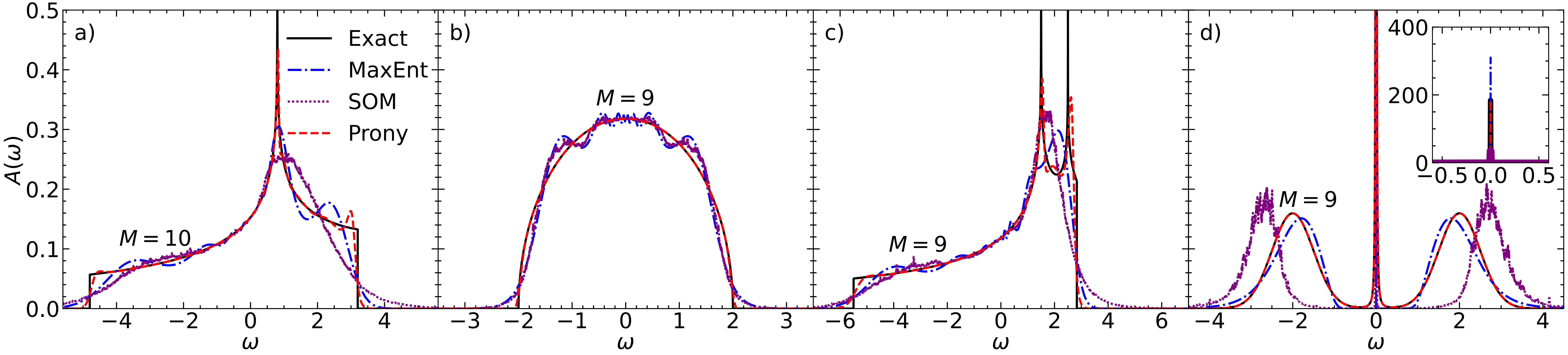
Results — dependence on # of data points



- Required number of points: $N_{\min} \gtrsim 2N_{\text{pole}} + 1$
- N_{pole} is the number of poles needed to resolve the fine structure of the spectrum
- When $N_\omega \geq N_{\min}$, convergence should always be observed at sufficient precision
- Larger N_ω doesn't necessarily leads to faster convergence

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

Results — challenging examples

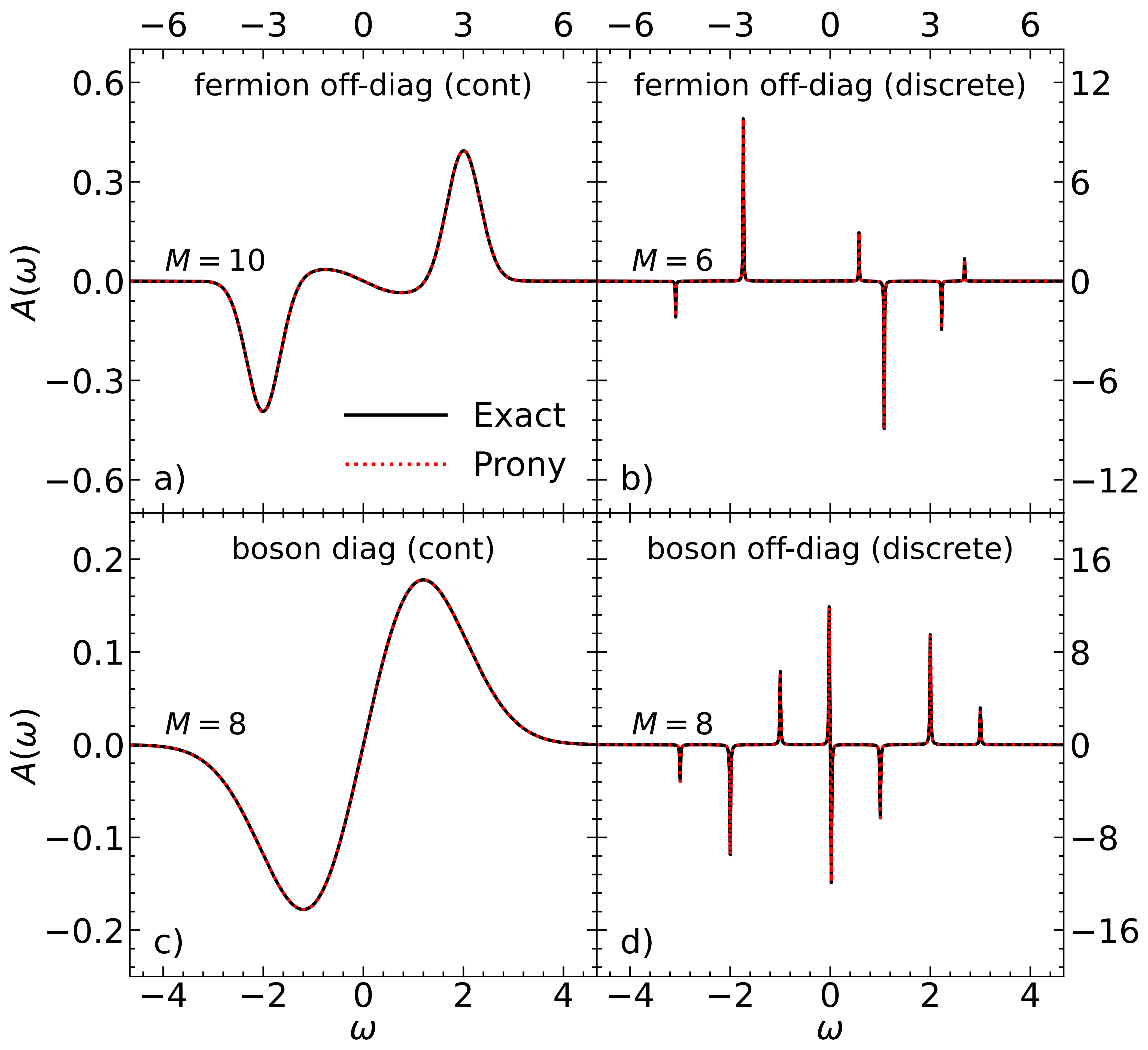


(a) dos of the tight binding model on square lattice
(b) dos of the tight binding model on Bethe lattice
(c) dos of the tight binding model on anisotropic triangular lattice
(d) “Kondo”-like spectral function

- Still systematically improvable
- Convergence slows down because of singularities

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

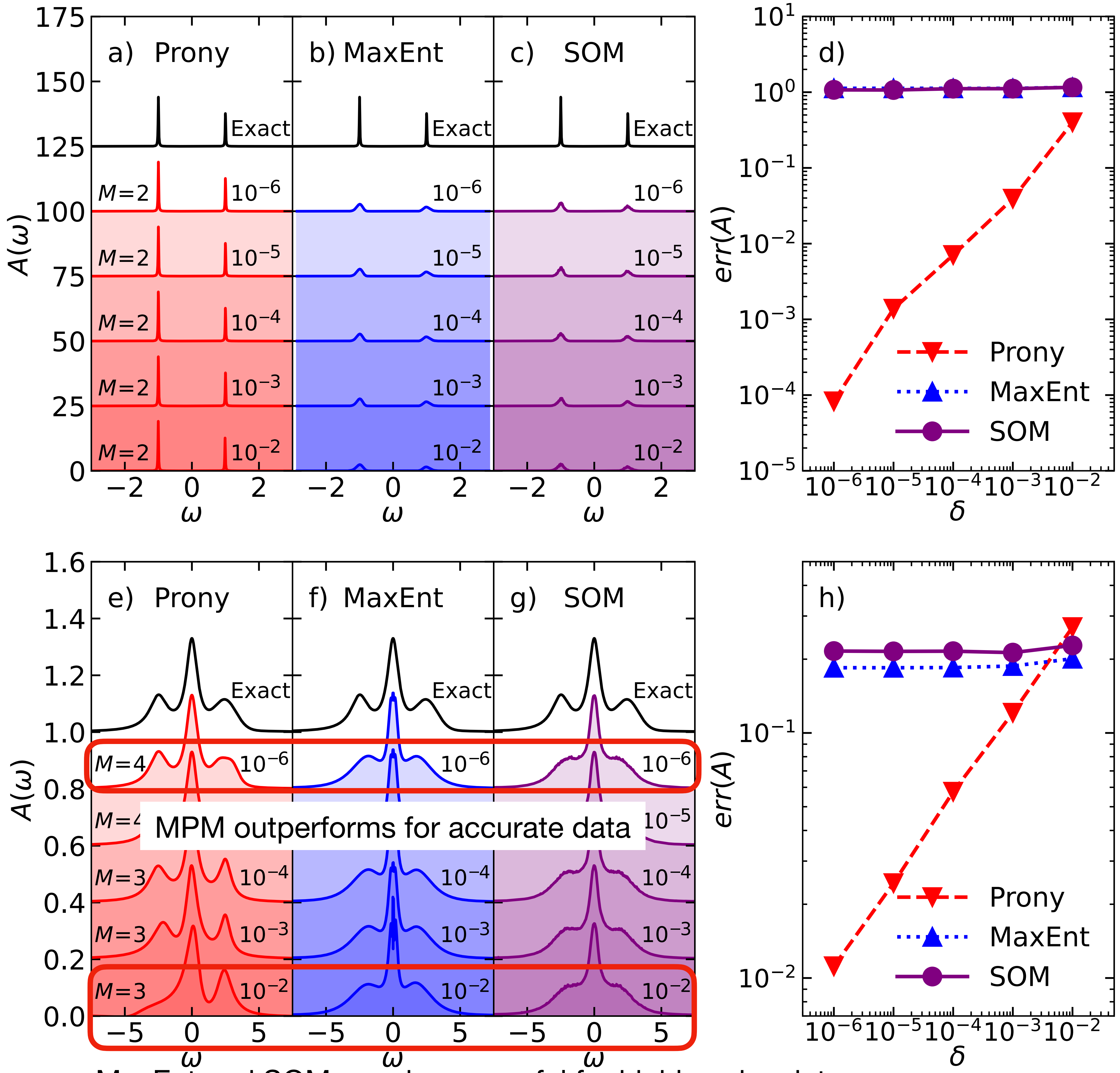
Results — versatility



- Applies to all systems
- Systematically improvable
- Recover the correct results for accurate data

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

Results — noisy cases



MaxEnt and SOM remains powerful for highly noisy data

MaxEnt: R. K. Bryan, Eur. Biophys. J. 18, 165 (1990)

M. Jarrell et al., Phys. Rep. 269, 133 (1996)

code: R. Levy et al., CPC 215, 149 (2017)

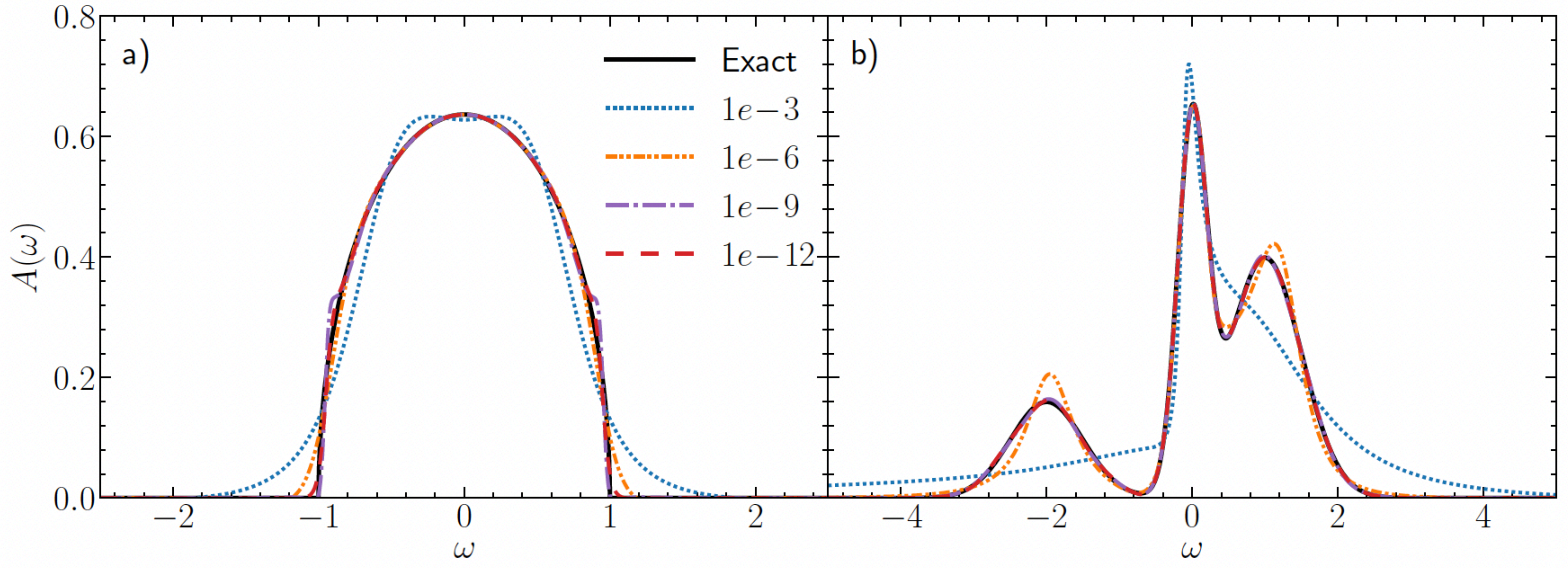
SOM: A. S. Mishchenko et al., Phys. Rev. B 62, 6317 (2000)

code: I. Krivenko et al., CPC 239, 166 (2019)

L. Zhang et al., Phys. Rev. B 110, 035154 (2024)

Results — imaginary-time data

convergence observed also for $G(\tau)$ (DLR transform for $G(\tau) \to G(i\omega_n)$) J. Kaye et al., Phys. Rev. B 105, 235115 (2022)



Results — matrix-valued cases

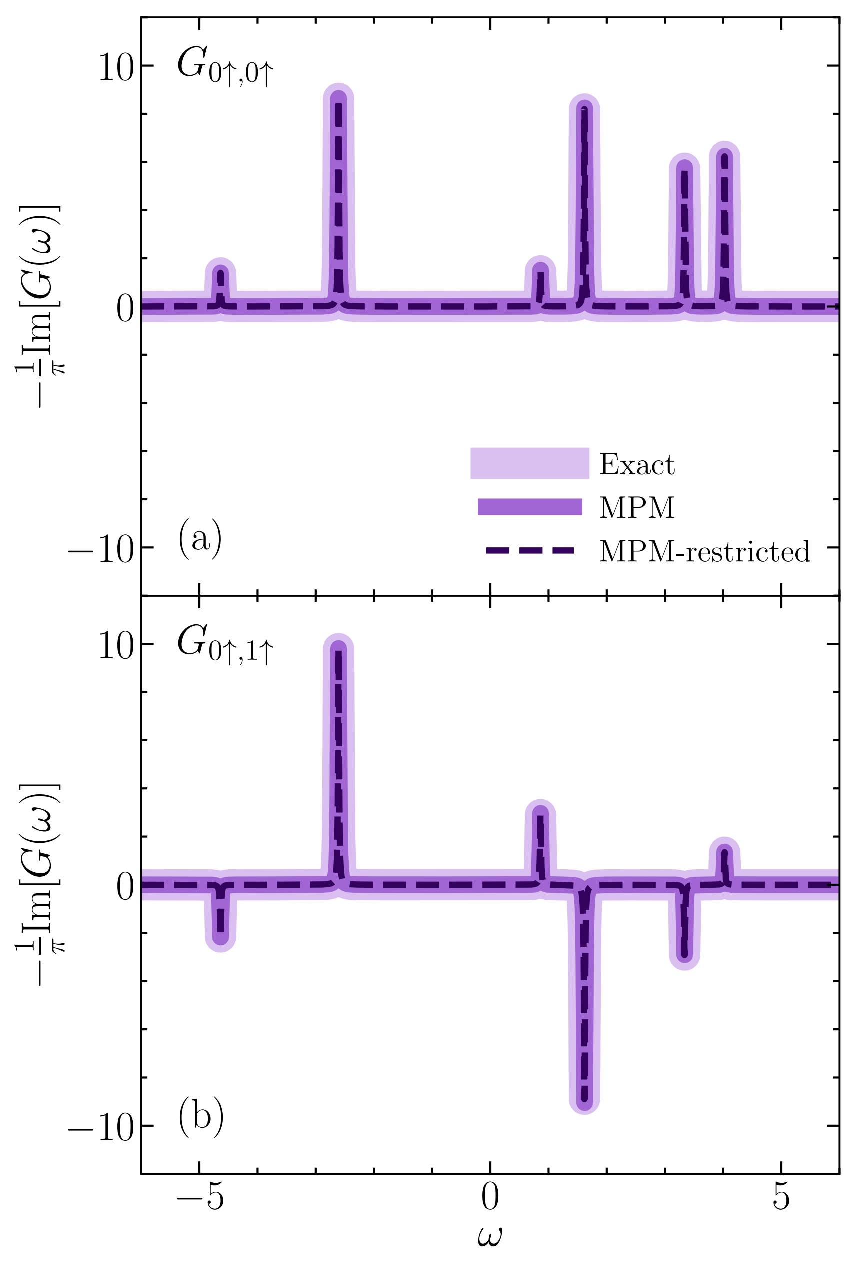
$$\boldsymbol{G}(z) = \sum_{l=1}^{M} \frac{\boldsymbol{A}_l}{z - \xi_l}, \ \mathrm{Im}\,\xi_l \leq 0$$

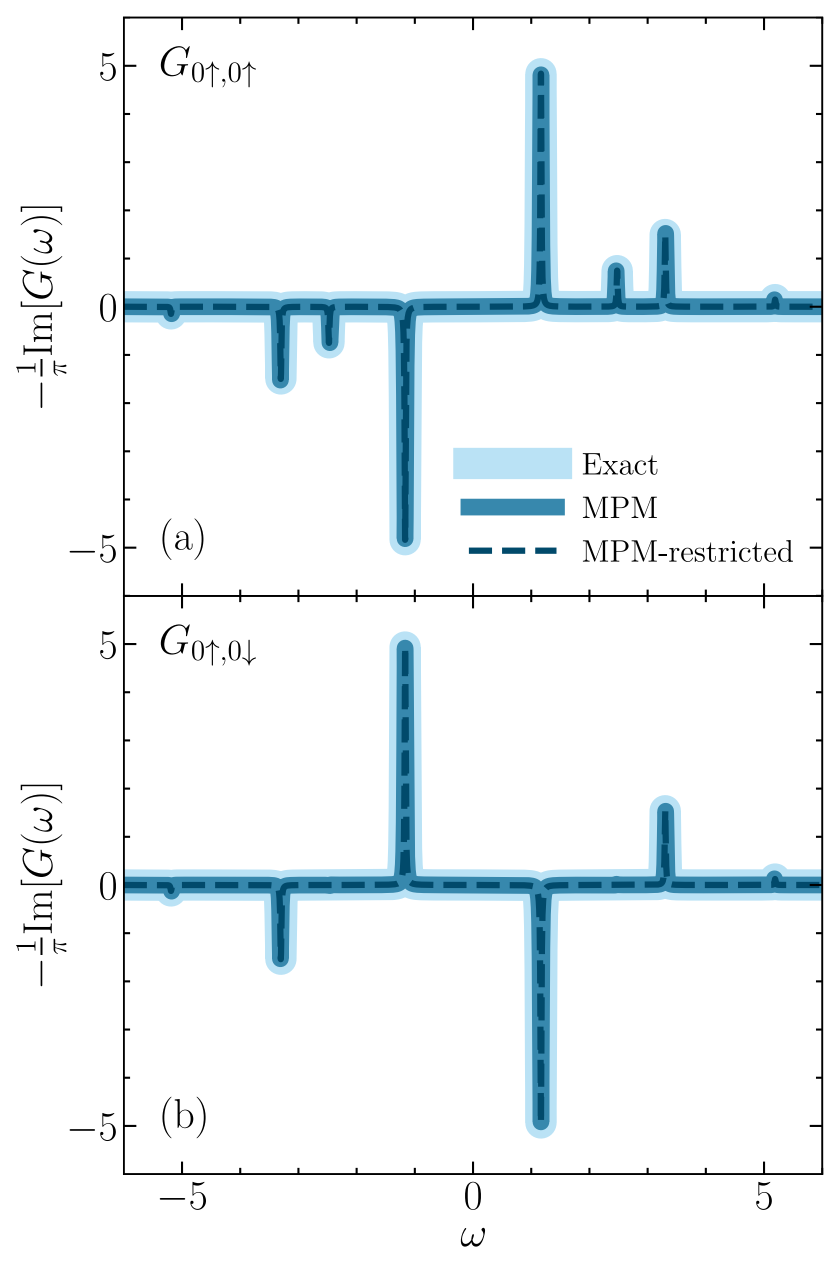
x_l shared poles

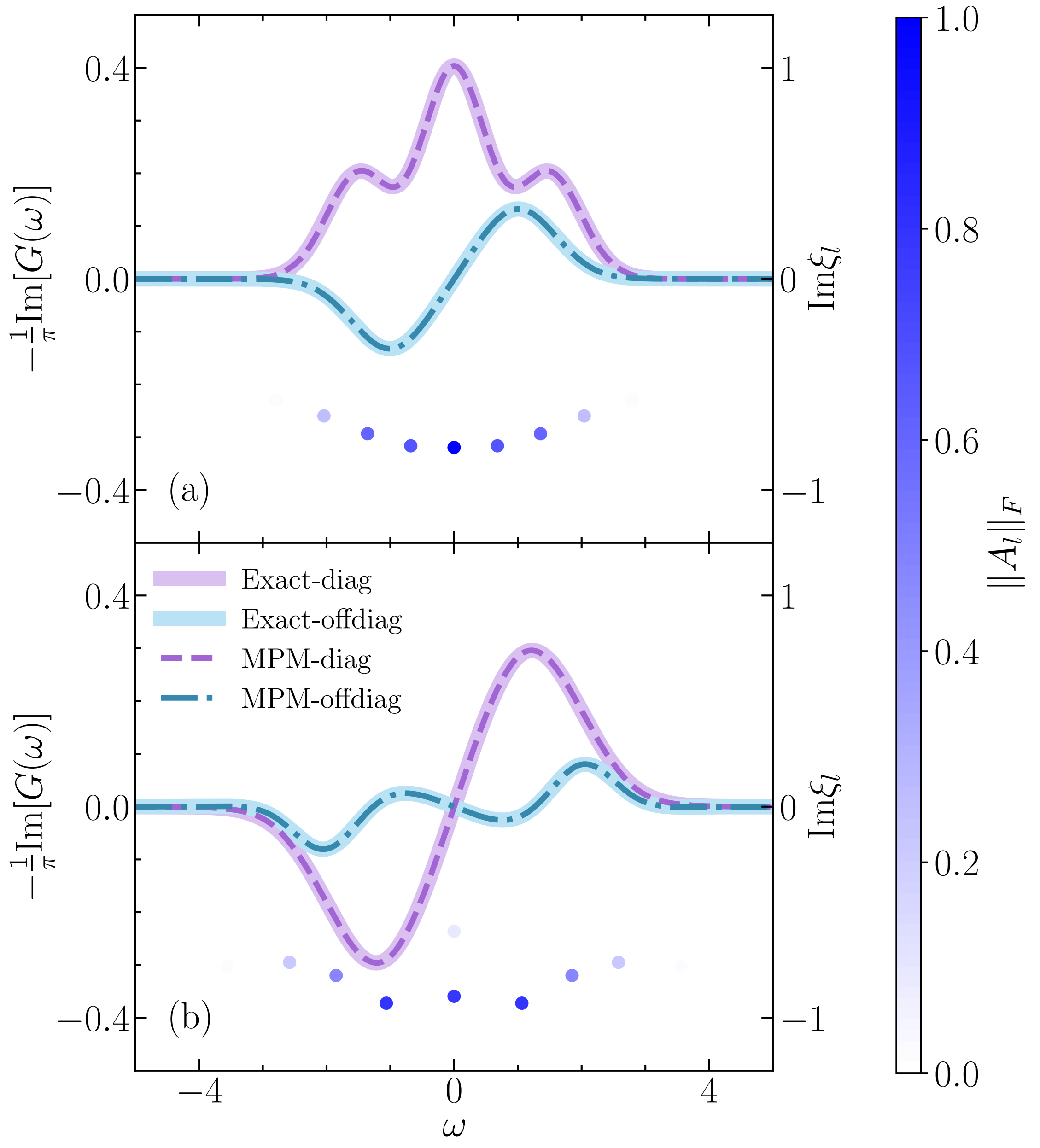
$\boldsymbol{A}_l$ matrix-valued weights

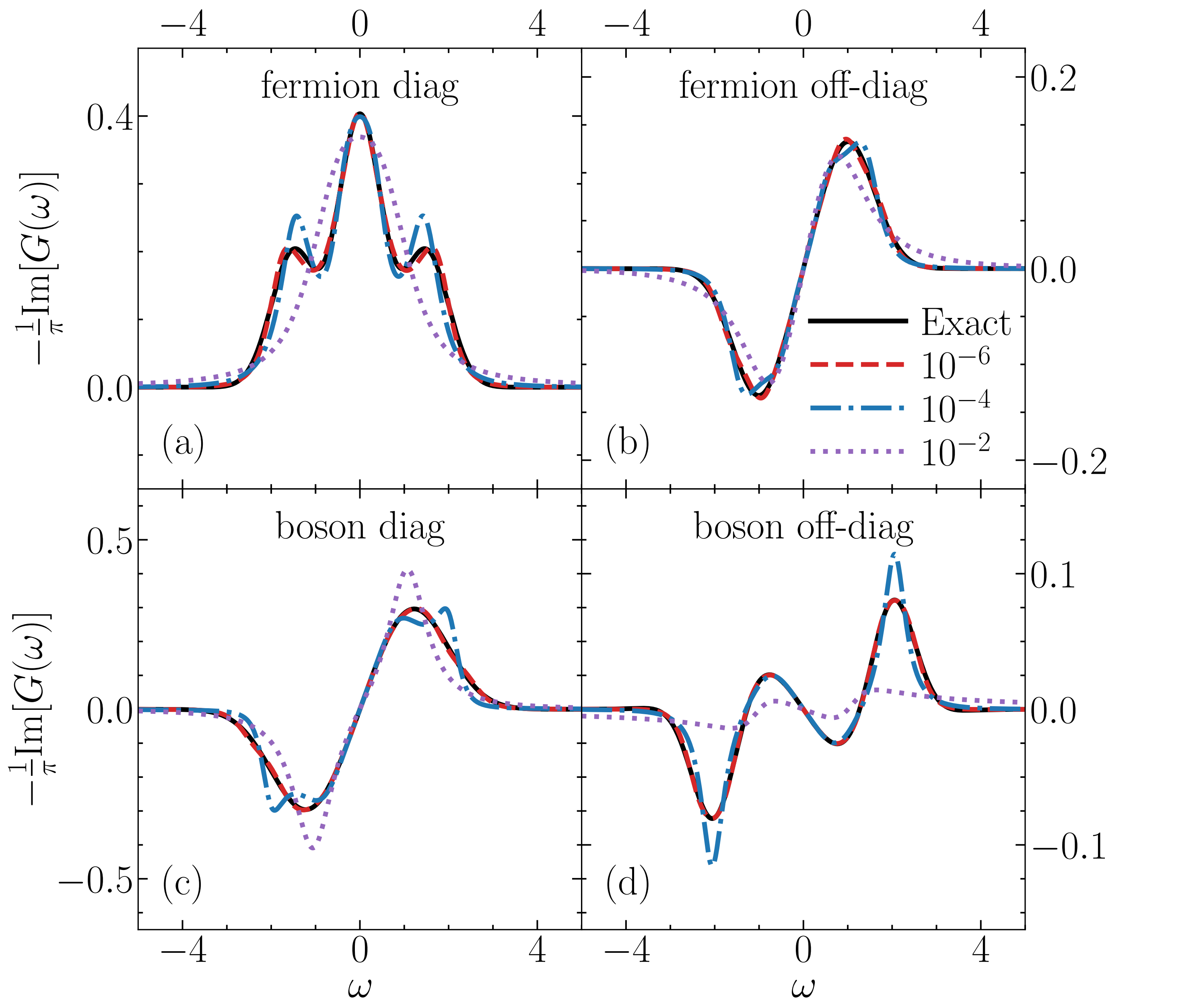
Matrix-valued ESPRIT,

followed from matrix-valued Prony L. Ying, J. Sci. Comput. 92, 107 (2022)





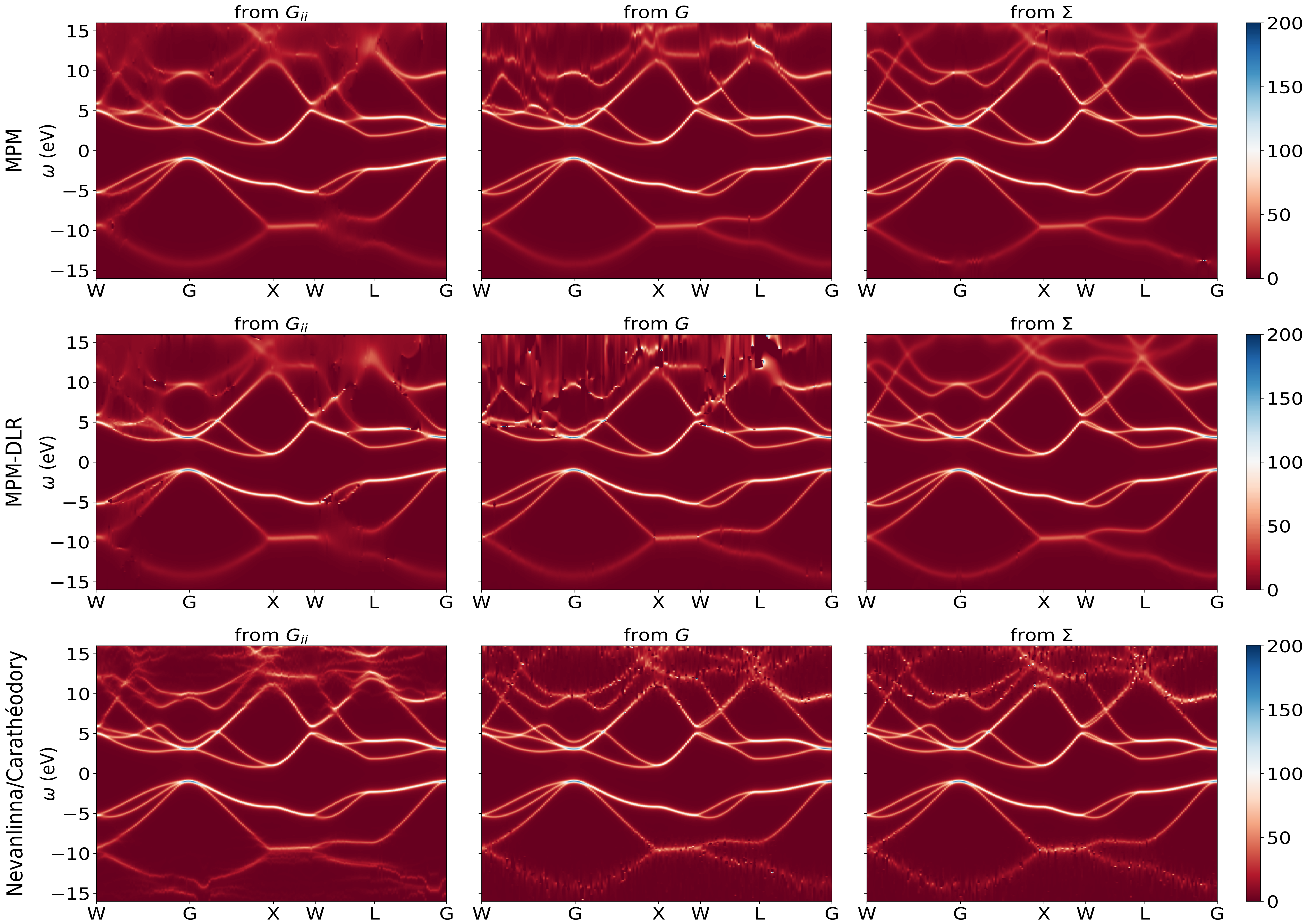




L. Zhang et al., Phys. Rev. B 110, 235131 (2024)

Results — band structure

Si, 6x6x6 lattice, self-consistent GW



$$\boldsymbol{G}(i\omega_n)^{-1} = \boldsymbol{G}_0(i\omega_n)^{-1} - \boldsymbol{\Sigma}(i\omega_n)$$

$$\boldsymbol{G}(\omega + i0^+)^{-1} = \boldsymbol{G}_0(\omega + i0^+)^{-1} - \boldsymbol{\Sigma}(\omega + i0^+)$$

Nevanlinna / Carathéodory:

- no substantial difference

MPM:

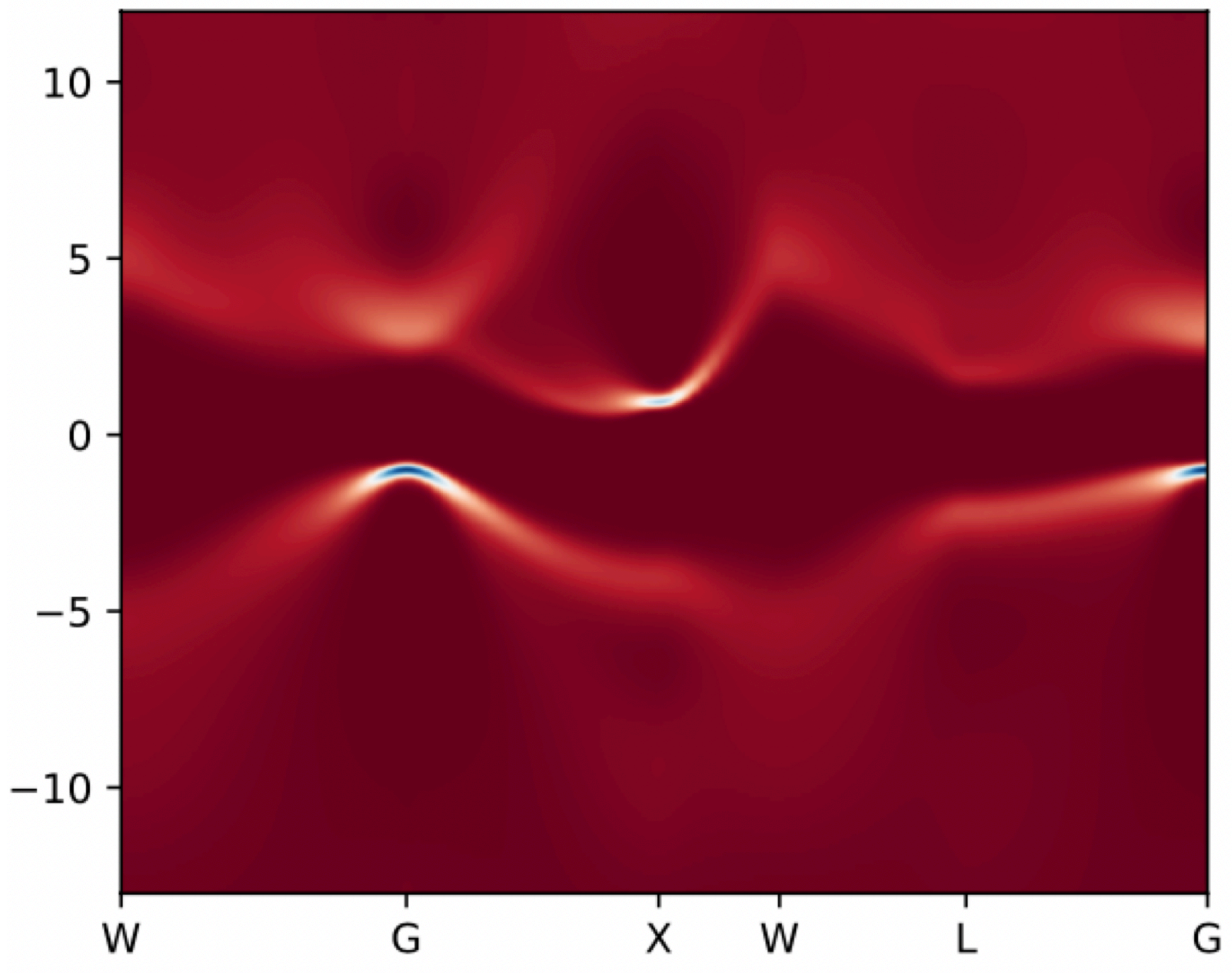
- No improvement from G_{ii} to $\boldsymbol{G}$
  (MPM better than MPM-DLR)
- Better resolution for $\boldsymbol{\Sigma}$
  (MPM ≈ MPM-DLR)

L. Zhang et al., Phys. Rev. B 110, 235131 (2024)

Results — band structure

Si, 6x6x6 lattice, self-consistent GW

Maxent

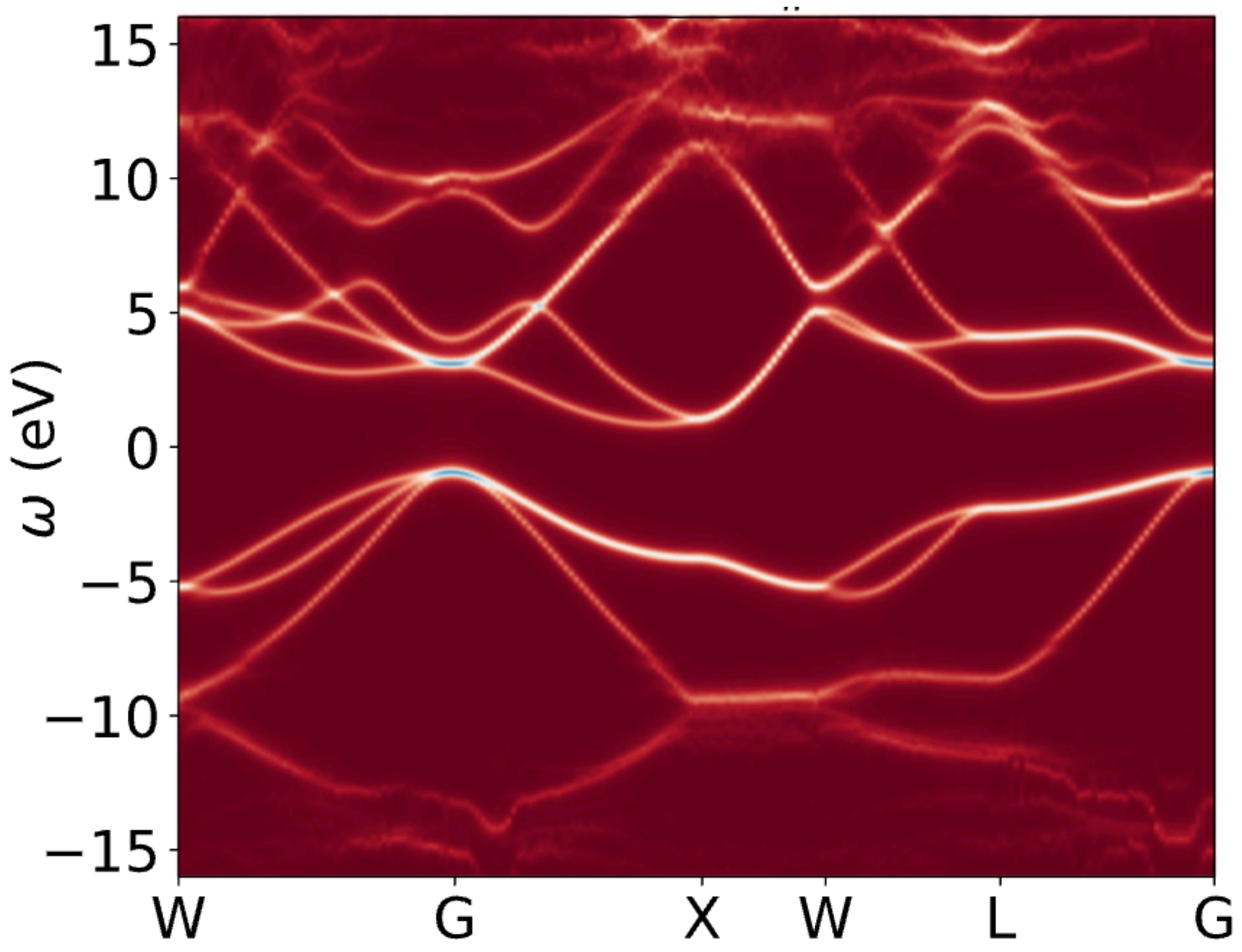


200 k points

26 orbitals per unit cell

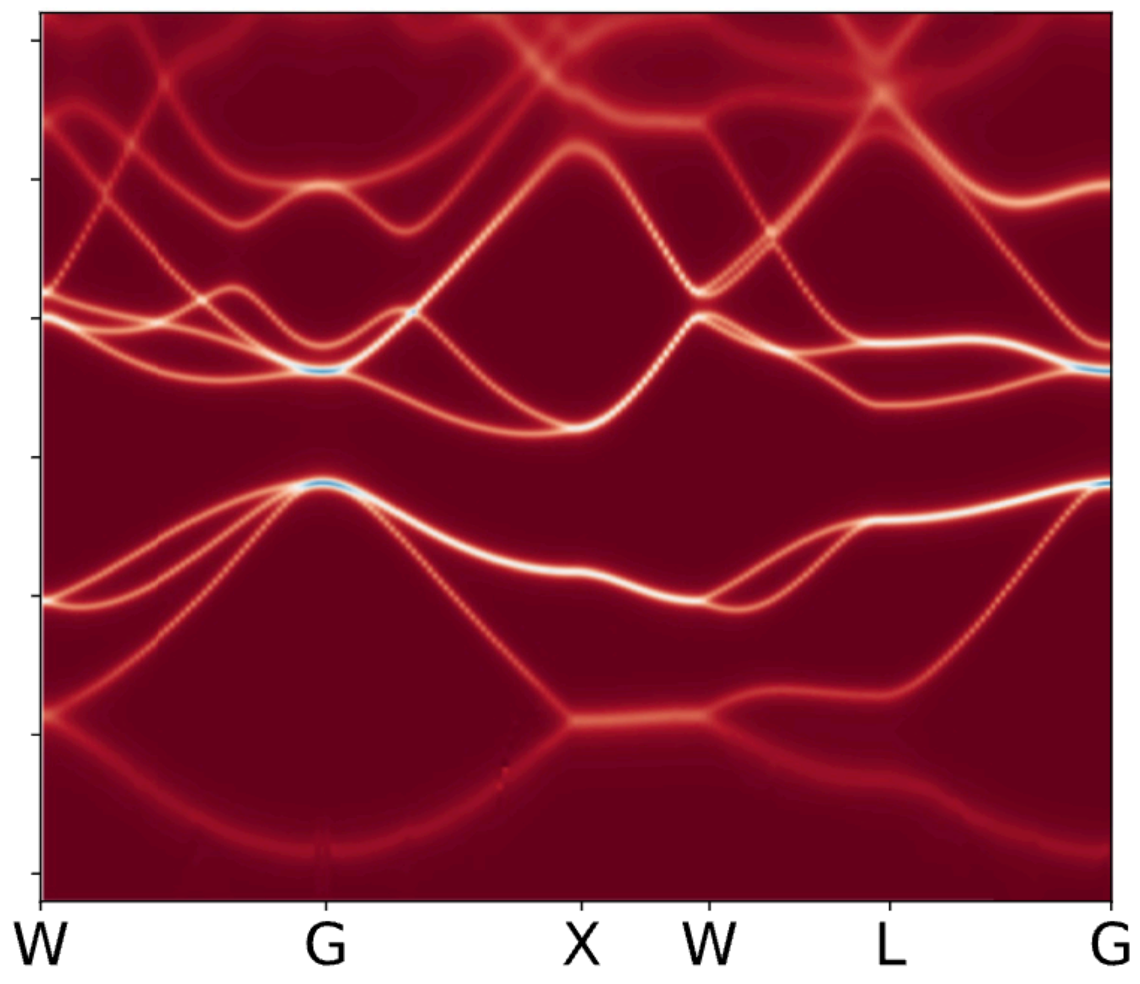
$\beta = 700\ \mathrm{Ha}^{-1}$

State-of-the-art
Nevanlinna / Carathéodory



- multiprecision arithmetic
- not robust to noise
- artificial broadening

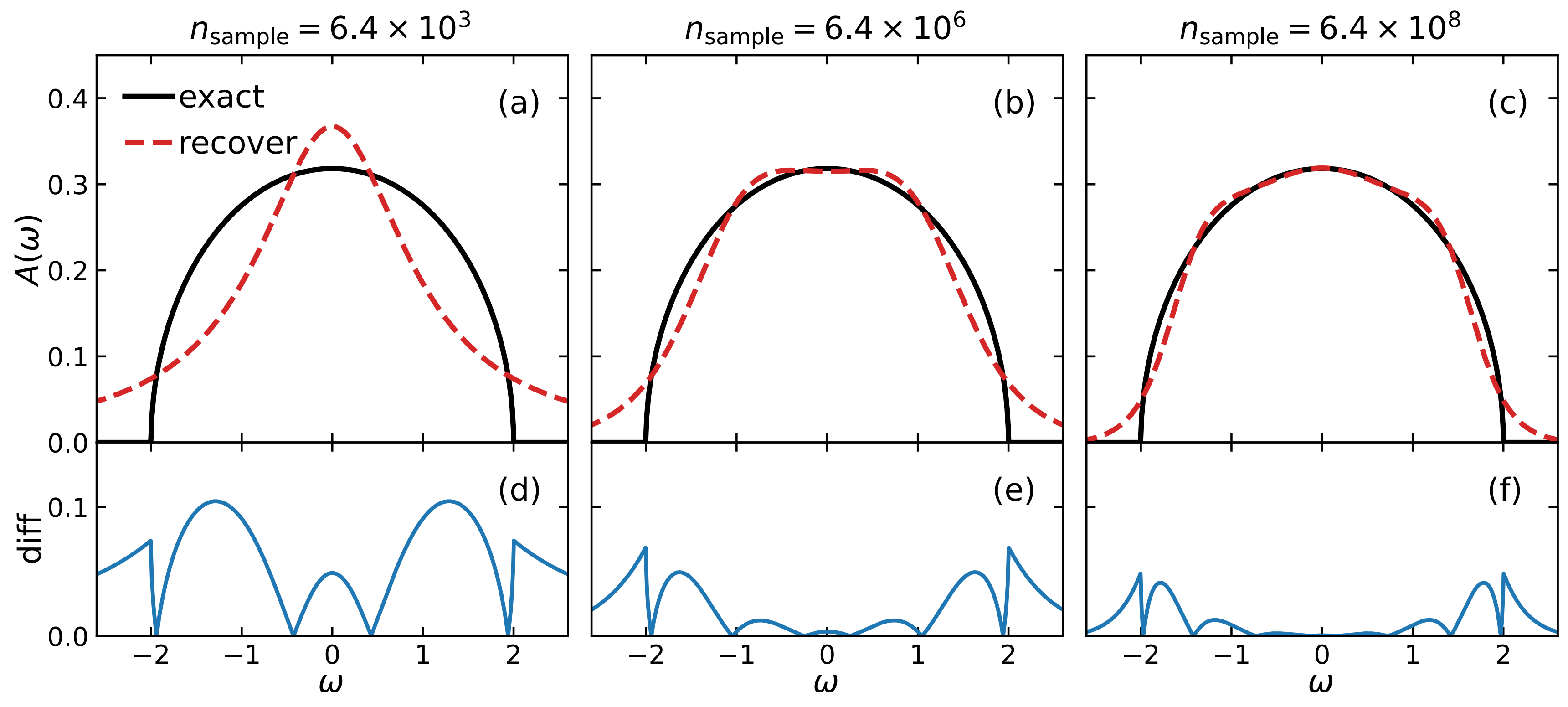
L. Zhang et al., Phys. Rev. B 110, 235131 (2024)
Minimal Pole Method



- double precision
- robust to noise
- no artificial broadening
- analytic expression
- extremely efficient:
  80 CPU hrs for MiniPole
  &**160 CPU secs** after utilizing DLR [Phys. Rev. B 105, 235115 (2022)]
  (800 CPU hrs for Carathéodory [CPC 304, 109299 (2024)])

Continuation for real-material data on a laptop is now possible!

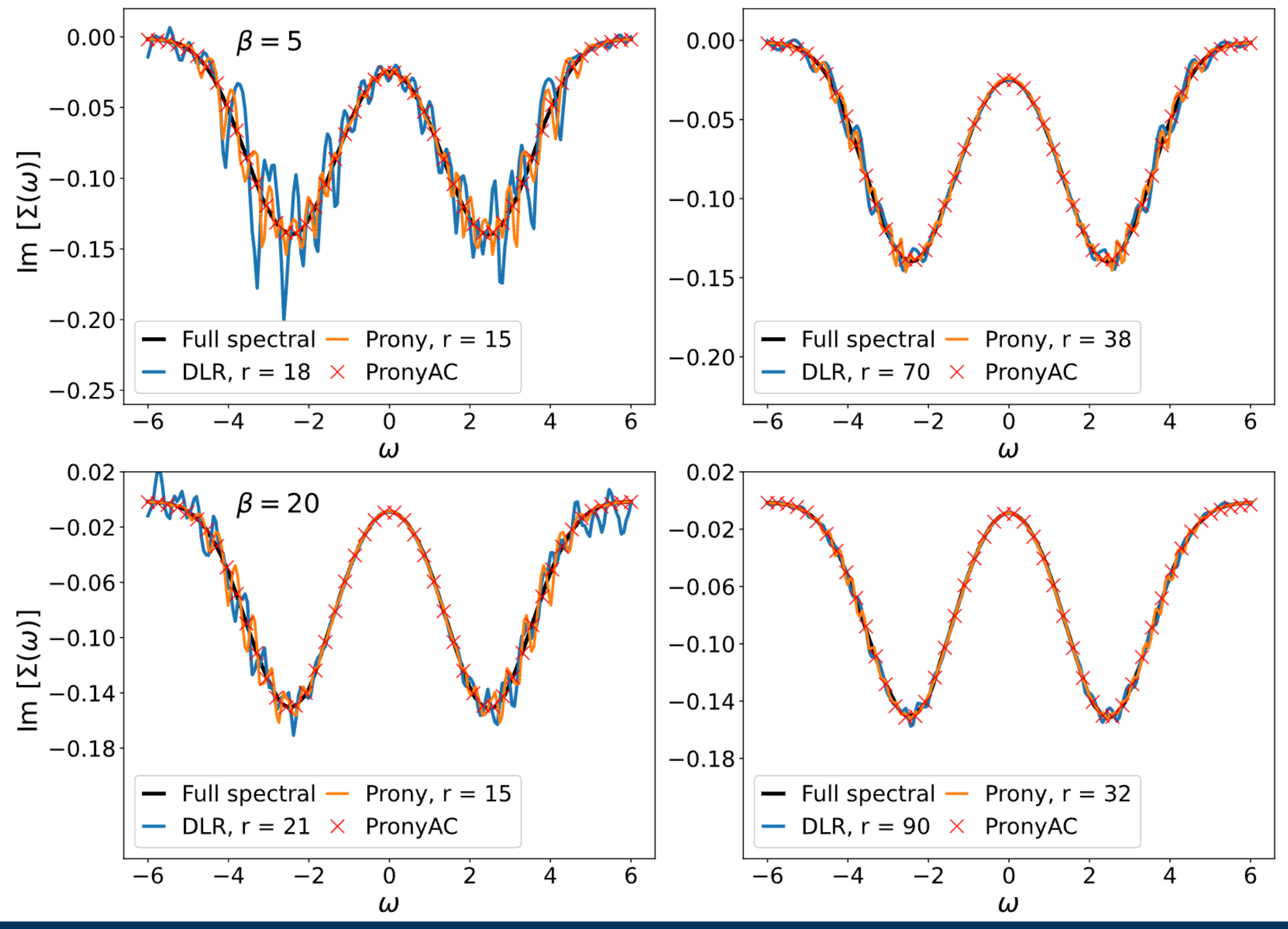
Results — other simulations



CT-HYB

Theory: P. Werner et al., Phys. Rev. Lett. 97, 076405 (2006)

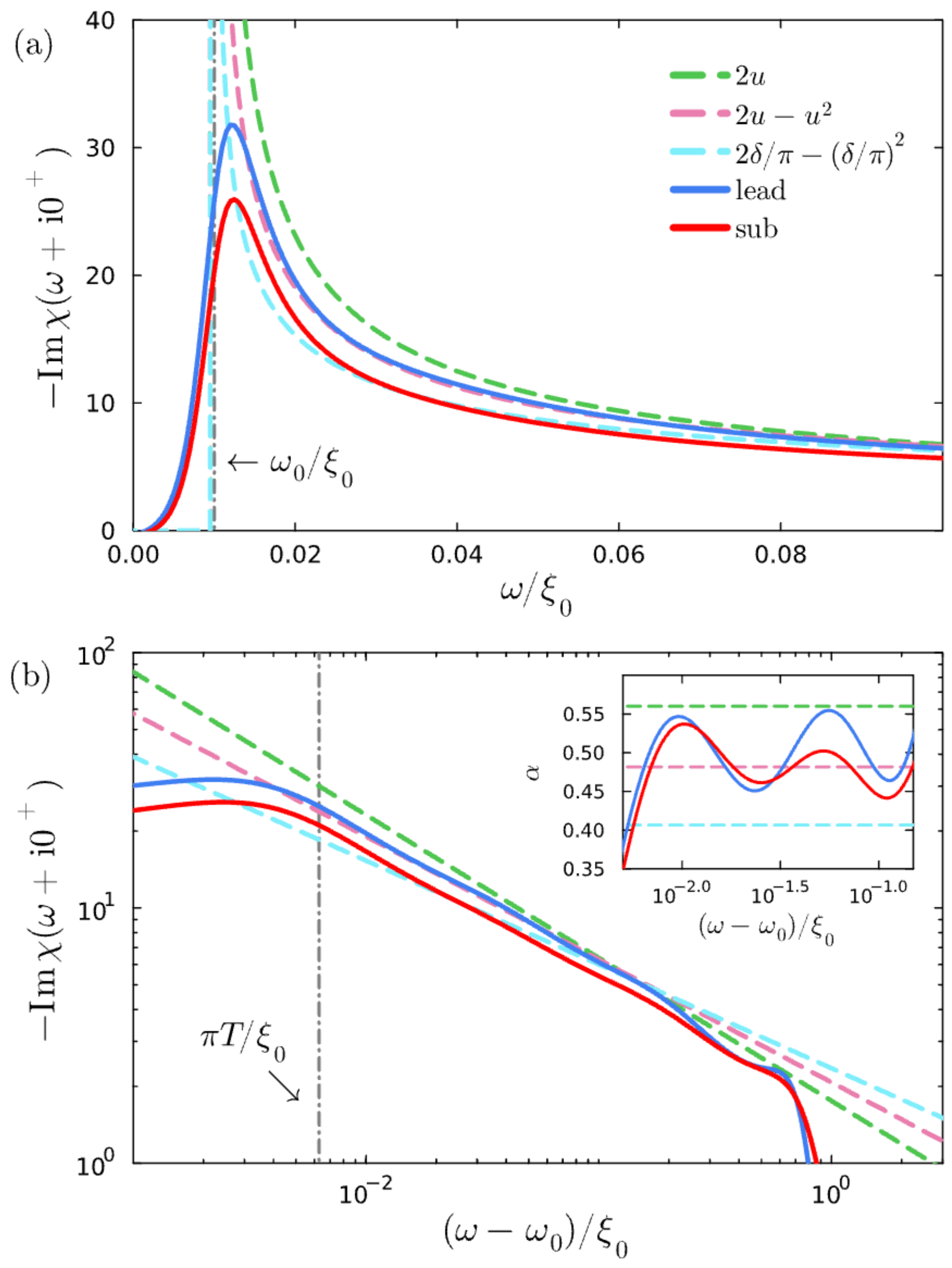
Continuously improves as runtime increases



Perturbation theory

D. Gazizova et al., Phys. Rev. B 110, 075158 (2024)

Continuation (red crosses) undistinguishable with exact results (solid black line)



Reliably capture the power-law decay

M. Gievers et al., Phys. Rev. B 111, 085151(2025)

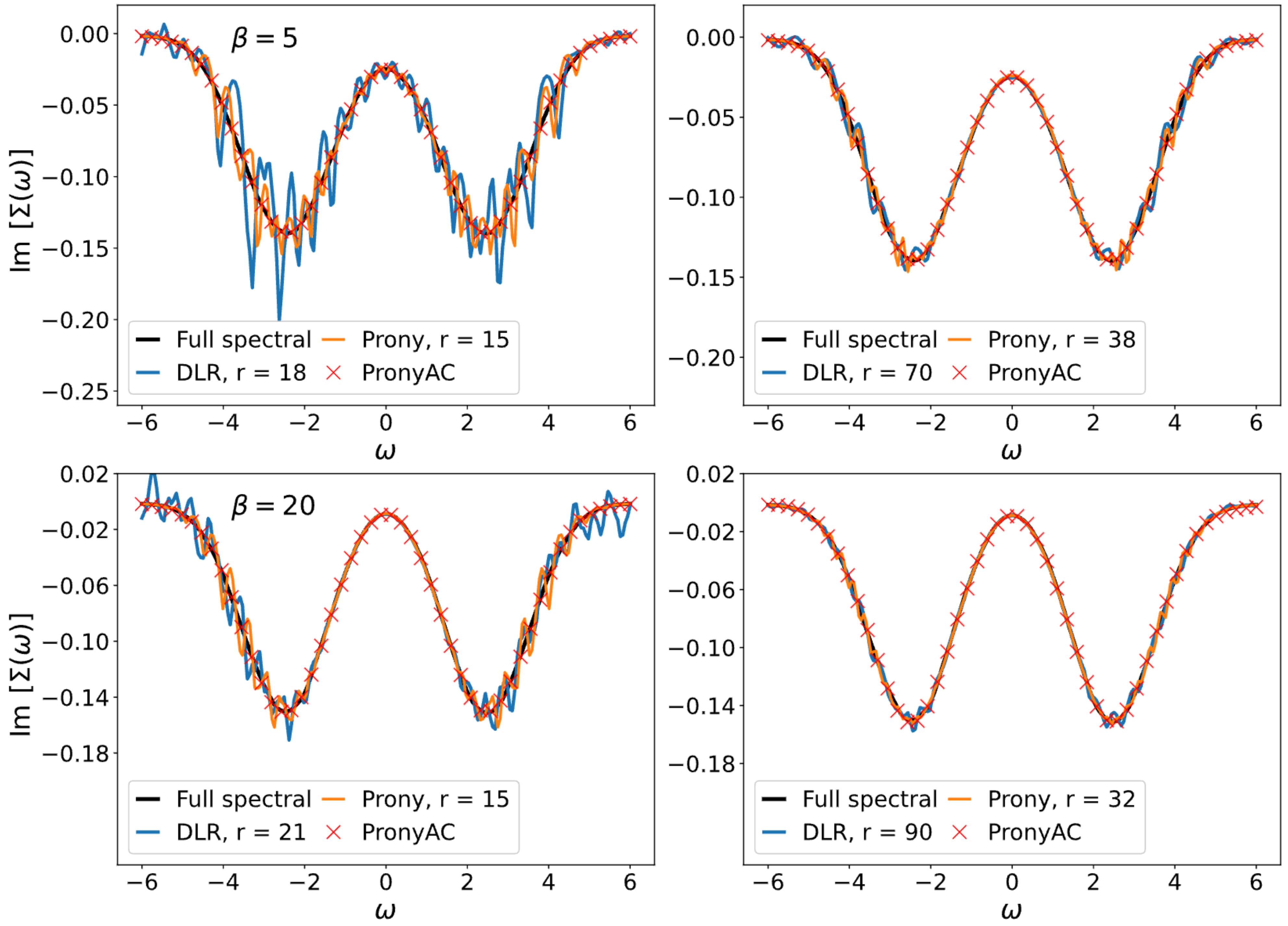
Other applications

Results - realistic

Compact representation

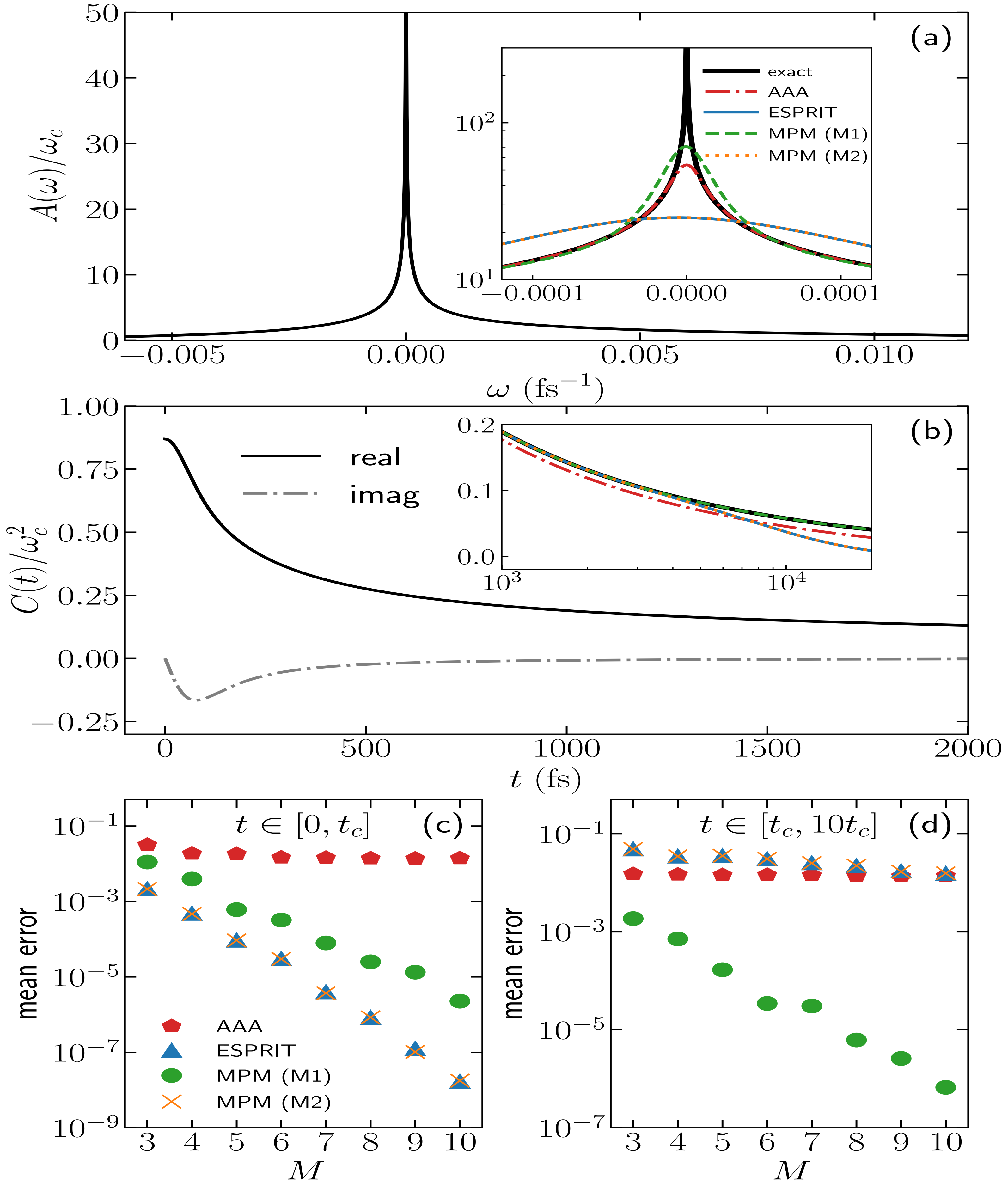
D. Gazizova et al., Phys. Rev. B 110, 075158 (2024)

- Usually more compact than DLR
- Advantageous to minimize oscillation
- Lack of a predetermined grid



Bath fitting

J. Chem. Phys. 162, 214111 (2025), L. Zhang et al.



Code

Try it!

Green Software Package

https://green-phys.org/docs/

> Components > Many-Body Framework > Analytic Continuation > Prony method

Analytic Continuation – Prony method

The Prony Analytic Continuation method, also called the Minimal Pole Method, is described in detail in the following papers:

1. https://doi.org/10.1103/PhysRevB.110.035154 (scalar-valued version)
2. https://doi.org/10.1103/PhysRevB.110.235131 (matrix-valued version)

Green provides a pedagogical implementation of the method as it was submitted to the journal. The implementation is generic; follow the instructions in the README to get started. A more elaborate version that seamlessly integrates with the GREEN code will be presented at a later time.

To get started, head to the repositories at:

1. https://github.com/Green-Phys/PronyAC (scalar-valued version, deprecated as it is a special case of MiniPole)
2. https://github.com/Green-Phys/MiniPole (matrix-valued version)

https://github.com/Green-Phys/MiniPole

(also available: pypi, zenodo)

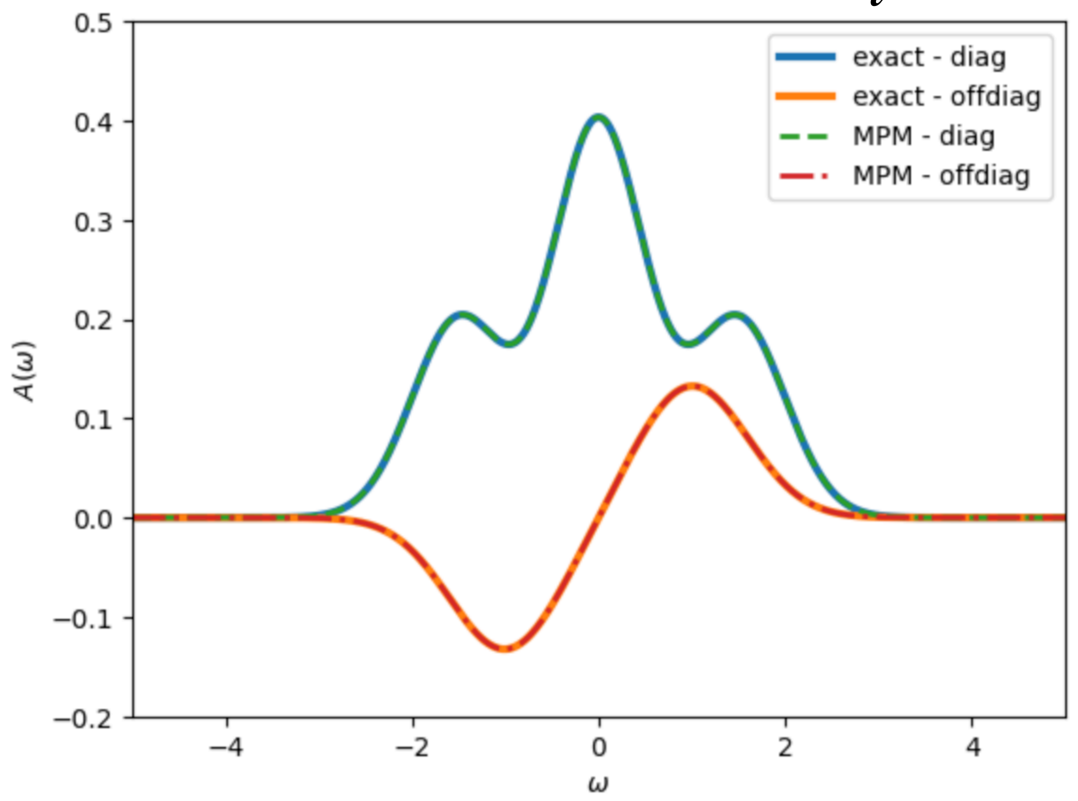
pip install mini_pole

As simple as possible!

```
p = MiniPole(G_w, w)
```

p.pole_weight -> A_l

p.pole_location -> x_l



Note: v0.4 was released recently!

- Can determine error tolerance automatically
- Can perform real-frequency fitting directly

A variety of parameters to satisfy most purposes

If get stuck, feel free to ask me for help!

Status

- Functionality (completed)
- Efficiency (can be improved)
- Tutorial (to be published)

Lei Zhang University of Michigan lzphy@umich.edu

Conclusion

Conclusion

- Despite of ill-conditioned nature, MPM are able to be systematically improvable
- Any different types of systems can be dealt in a unified framework in MPM
- MPM could possibly have better resolution even compared to the state-of-the-art method
- Expect to be more and more powerful as advanced simulation methods are developed
- MPM has broad applications: compact representation, bath fitting…
- Unresolved question: improve convergence speed for highly noisy data

Thanks!