

UC Berkeley

UC Berkeley Previously Published Works

Title

Robust analytic continuation of Green's functions via projection, pole estimation, and semidefinite relaxation

Permalink

<https://escholarship.org/uc/item/5m79k49n>

Journal

Physical Review B, 107(7)

ISSN

2469-9950

Authors

Huang, Zhen

Gull, Emanuel

Lin, Lin

Publication Date

2023-02-01

DOI

10.1103/physrevb.107.075151

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Robust analytic continuation of Green’s functions via projection, pole estimation, and semidefinite relaxation

Zhen Huang,¹ Emanuel Gull,² and Lin Lin^{1,3,*}

¹*Department of Mathematics, University of California, Berkeley, California 94720 USA*

²*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*

³*Applied Mathematics and Computational Research Division,
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

(Dated: October 11, 2022)

Green’s functions of fermions are described by matrix-valued Herglotz-Nevanlinna functions. Since analytic continuation is fundamentally an ill-posed problem, the causal space described by the matrix-valued Herglotz-Nevanlinna structure can be instrumental in improving the accuracy and in enhancing the robustness with respect to noise. We demonstrate a three-pronged procedure for robust analytic continuation called PES: (1) Projection of data to the causal space. (2) Estimation of pole locations. (3) Semidefinite relaxation within the causal space. We compare the performance of PES with the recently developed Nevanlinna and Carathéodory continuation methods and find that PES is more robust in the presence of noise and does not require the usage of extended precision arithmetics. We also demonstrate that a causal projection improves the performance of the Nevanlinna and Carathéodory methods. The PES method is generalized to bosonic response functions, for which the Nevanlinna and Carathéodory continuation methods have not yet been developed. It is particularly useful for studying spectra with sharp features, as they occur in the study of molecules and band structures in solids.

I. INTRODUCTION

Green’s functions of quantum many-body systems are central subjects in condensed matter physics, quantum chemistry, and quantum field theory. Due to the difficulty of performing correlated finite-temperature calculations on the real axis, numerical methods typically obtain results on the Matsubara axis and employ analytic continuation as a postprocessing tool to infer Green’s functions along the real axis. Analytic continuation is fundamentally an ill-posed problem, and many numerical methods have been developed, such as the Padé fit [1, 2], maximum entropy (MaxEnt) [3–7], stochastic analytic continuation and its variants [8–11], sparse modeling [12, 13], and machine learning approaches [14, 15]. All these methods encounter difficulties in satisfying causality, recovering sharp features, resolving multiple features, and/or capturing high frequency information in the spectra. Furthermore, most methods have been developed only for diagonal entries of the Green’s function. The analytic continuation of off-diagonal entries of the Green’s function [16–20] is often attempted by diagonalizing the Green’s function at a certain Matsubara frequency, conducting the analytic continuation only of the diagonal entries with respect to the transformed basis at other frequencies, and neglecting the remaining non-diagonal entries.

The fermionic Green’s functions is a special type of *matrix-valued Herglotz-Nevanlinna functions*. These functions have a wide range of scientific and engineering applications, and have been (somewhat incon-

sistently) named after renowned mathematicians including Carathéodory, Herglotz, Nevanlinna, Pick, and Riesz. They are also sometimes called *R-functions* [21]. The name matrix-valued Herglotz-Nevanlinna functions adopted in this paper follows the suggestion in [22]. This crucial analytic structure has only been taken into account recently by the Nevanlinna continuation method [23] for diagonal entries of Green’s functions, and by the Carathéodory method [24] for both diagonal and off-diagonal entries. In the absence of noise, these methods have reached unprecedented accuracy in analytic continuation by resolving complicated spectral functions with multiple features. However, these methods also have notable drawbacks. (1) They are not numerically stable, and extended precision arithmetic operations (128 bits of precision or higher) are needed to carry out such calculations (even in the absence of noise). Therefore the computational cost of the analytic continuation (especially for the Carathéodory continuation) can be large. (2) The Nevanlinna/Carathéodory interpolants exist if and only if Pick’s criterion [23, 25, 26] is satisfied by the Matsubara data. In practice, noisy data often violates Pick’s criterion, meaning that the computational result is not guaranteed to be causal. (3) In their standard form, matrix-valued Herglotz-Nevanlinna structure is only applicable to fermions, and hence these methods are not directly applicable to bosonic systems.

The Herglotz-Nevanlinna function and its interpolation have been thoroughly studied in areas such as the control theory (see e.g., [27] and [28, Chapter 9.2]). The problem of analytic continuation is also intimately related to problems in signal processing, and in particular fitting with exponential functions (see e.g., [29–31]). One advantage of Nevanlinna and Carathéodory methods is that they require little prior knowledge of the structure

* linlin@math.berkeley.edu

of Green's functions. With further prior knowledge of the lower and upper bound of the absolute value of the pole locations, Ying recently showed that a modified Prony's method can be used for performing analytic continuation with noisy data with multiple features using only double precision arithmetic operations [32, 33].

In this article, we propose a three-pronged Projection-Estimation-Semidefinite relaxation (PES) method to perform analytic continuation within the causal space, while avoiding the aforementioned drawbacks. The meaning of the three steps are as follows: (1) Projection of the noisy data into the causal space; (2) Estimation of pole locations using the adaptive Antoulas–Anderson (AAA) algorithm [34]; (3) Semidefinite relaxation (SDR) fitting of Green's functions (diagonal and off-diagonal elements) using a bi-level optimization approach [35]. Each step of the approach aims at resolving certain aspects of the difficulties in the analytic continuation of noisy Matsubara Green's functions. The prior knowledge needed for the PES method is comparable to that in Nevanlinna and Carathéodory methods. We demonstrate that the PES method can robustly perform analytical continuation of noisy data using standard double precision arithmetic operations, and is applicable to both fermionic and bosonic systems. Table I compares the PES method with the Nevanlinna, Carathéodory, MaxEnt, and Padé method. We emphasize that each of the three steps of PES can also be useful in improving the robustness of other analytic continuation methods. For instance, the projection step can be employed as a data pre-processing step that significantly improves the robustness of the Nevanlinna and Carathéodory continuation for noisy data. The semidefinite relaxation can be combined with other continuation methods as a post-processing refinement.

This article is organized as follows. In Section II, we introduce the theory of the PES method. After giving a brief discussion on the setup of the analytic continuation problem in Section II.1, we explain in detail each of the three steps of the PES method: the pre-processing step using semidefinite projection in Section II.2, the pole-estimation step using AAA algorithm in Section II.3 and the SDR fitting step using SDR and bi-level optimization in Section II.4. We summarize the PES method in Section II.5. The numerical results are presented in Section III. With the Hubbard dimer example, we compare the results of our methods with the Nevanlinna, Carathéodory and MaxEnt methods, and show that the PES method is both noise-robust and efficient for sharp features, and recovers both diagonal and off-diagonal entries. We also demonstrate the necessity of the pole estimation step using this model. In Section III.2, we show that the PES method is also applicable to bosonic response functions and has a much better performance compared to MaxEnt. In Section III.3, in order to demonstrate the strength of the PES method, we conduct analytic continuations for the band structure of solids with hundreds of orbitals.

II. THEORY

II.1. Analytic structure of Green's functions

Let $\hat{c}_i, \hat{c}_i^\dagger$ be the annihilation and creation operator for the i -th orbital, $i = 1, \dots, N_{\text{orb}}$, where N_{orb} is the number of orbitals. In the Lehmann representation, a Green's function can be expressed as follows (see e.g., [36, Chapter 5.2]):

$$\mathbb{G}_{ij}(z) = \frac{1}{Z} \sum_{r,s} \frac{\langle \Psi_s | \hat{c}_i | \Psi_r \rangle \langle \Psi_r | \hat{c}_j^\dagger | \Psi_s \rangle}{z + E_s - E_r} (e^{-\beta E_s} \mp e^{-\beta E_r}). \quad (1)$$

Here β is the inverse temperature. $|\Psi_s\rangle$ is the s -th eigenfunction of the Hamiltonian \hat{H} with energy E_s , i.e.,

$$\hat{H} |\Psi_s\rangle = E_s |\Psi_s\rangle, \quad s = 1, 2, \dots, N_S,$$

where $N_S = 2^{N_{\text{orb}}}$, \hat{H} is the many-body Hamiltonian and $Z = \sum_{s=1}^{N_S} e^{-\beta E_s}$ is the partition function. The negative sign corresponds to bosons, while the positive sign corresponds to fermions.

Since $\mathbb{G}_{ij}(z)$ is defined for $i, j \in \{1, \dots, N_{\text{orb}}\}$, the Green's function \mathbb{G} can be viewed as a $N_{\text{orb}} \times N_{\text{orb}}$ matrix. In other words, we have $\mathbb{G}(z) = \sum_{r,s} \frac{\mathbb{X}_{r,s}}{z - \lambda_{r,s}}$ where $\lambda_{r,s} = E_r - E_s \in \mathbb{R}$ and $\mathbb{X}_{r,s} = c_{r,s} v_{r,s} v_{r,s}^\dagger$ is a rank-1 $N_{\text{orb}} \times N_{\text{orb}}$ matrix, in which

$$c_{r,s} = \frac{e^{-\beta E_s} \mp e^{-\beta E_r}}{Z}, \quad v_{r,s}(i) = \langle \Psi_r | \hat{c}_i^\dagger | \Psi_s \rangle. \quad (2)$$

Here $v_{r,s}(i)$ is the i -th component of $v_{r,s}$. Note that for fermionic systems, $c_{r,s}$ is always nonnegative. For bosonic systems, we have $c_{r,s} > 0$ if $E_r > E_s$, or equivalently $\lambda_{r,s} > 0$; while $c_{r,s} < 0$ if $E_r < E_s$, or equivalently $\lambda_{r,s} < 0$.

For orthogonal orbitals, the creation and annihilation operators satisfy the canonical relation $[\hat{c}_i, \hat{c}_j^\dagger]_{\mp} = \hat{c}_i \hat{c}_j^\dagger \mp \hat{c}_j^\dagger \hat{c}_i = \delta_{ij}$. As a result, the matrices $\mathbb{X}_{r,s}$ have the following property:

$$\begin{aligned} & \sum_{r,s} (\mathbb{X}_{r,s})_{ij} \\ &= \frac{1}{Z} \sum_{r,s} \langle \Psi_s | \hat{c}_i | \Psi_r \rangle \langle \Psi_r | \hat{c}_j^\dagger | \Psi_s \rangle (e^{-\beta E_s} \mp e^{-\beta E_r}) \\ &= \frac{1}{Z} \sum_s \left\langle \Psi_s \left| \hat{c}_i \left(\sum_r |\Psi_r\rangle \langle \Psi_r| \right) \hat{c}_j^\dagger \right| \Psi_s \right\rangle e^{-\beta E_s} \\ & \mp \frac{1}{Z} \sum_r \left\langle \Psi_r \left| \hat{c}_j^\dagger \left(\sum_s |\Psi_s\rangle \langle \Psi_s| \right) \hat{c}_i \right| \Psi_r \right\rangle e^{-\beta E_r} \\ &= \frac{1}{Z} \sum_s \left\langle \Psi_s \left| \hat{c}_i \hat{c}_j^\dagger \mp \hat{c}_j^\dagger \hat{c}_i \right| \Psi_s \right\rangle e^{-\beta E_s} = \delta_{ij}. \end{aligned} \quad (3)$$

This indicates the sum rule $\sum_{r,s} \mathbb{X}_{r,s} = \mathbb{I}_{N_{\text{orb}}}$, where $\mathbb{I}_{N_{\text{orb}}}$ is the $N_{\text{orb}} \times N_{\text{orb}}$ identity matrix. For the rest of this

Method	Noise Robustness	Calculation Precision Requirement	Sharp Features	Causality	Applicability
This work	✓	Double	✓	✓	Fermion Boson
Nevanlinna and Carathéodory	✗	Extended	✓	✓ if clean ✗ if noisy	Fermion
MaxEnt	✓	Double	✗	✓	Fermion Boson
Padé	✗	Extended	✓	✗	Fermion Boson

Table I. Comparison of approaches for analytic continuation. Double means double precision arithmetic operations, and extended means extended precision arithmetic operations (requiring 128 bits or larger).

article, let us compress the indices r, s into l , and let N_p be the number of matrices $\mathbb{X}_{r,s}$ that are nonzero, i.e., the number of poles $\lambda_{r,s}$ that actually contribute to the Green's function. In this way, the Green's function $\mathbb{G}(z)$ can be written as

$$\mathbb{G}(z) = \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{z - \lambda_l}, \quad (4)$$

where \mathbb{X}_l satisfies the rank-1 semidefinite condition:

$$\begin{aligned} \text{For bosons:} \quad & \text{sign}(\lambda_l) \cdot \mathbb{X}_l \text{ is an } N_{\text{orb}} \times N_{\text{orb}} \\ & \text{rank-1 positive semidefinite matrix;} \end{aligned} \quad (5)$$

$$\begin{aligned} \text{For fermions:} \quad & \mathbb{X}_l \text{ is an } N_{\text{orb}} \times N_{\text{orb}} \\ & \text{rank-1 positive semidefinite matrix;} \end{aligned} \quad (6)$$

and the sum rule

$$\sum_{l=1}^{N_p} \mathbb{X}_l = \mathbb{I}_{N_{\text{orb}}}. \quad (7)$$

The function $\mathbb{G}(z)$ is defined on $\{0\} \cup \mathbb{C} \setminus \mathbb{R}$, which excludes the real axis. The Matsubara Green's function $\mathbb{G}^M(\omega)$ and the retarded Green's function $\mathbb{G}^R(\omega)$ share the same formula $\mathbb{G}(z)$ (Eq. (1) and Eq. (4)) in the following way:

- When $z = i\omega_n \in i\mathbb{R}$ is on the imaginary axis, $\mathbb{G}^M(\omega_n) = \mathbb{G}(i\omega_n)$ is the Matsubara (or imaginary frequency) Green's function, ω_n is called the Matsubara frequency. For fermionic systems, $i\omega_n \in i\frac{2\mathbb{Z}+1}{\beta}$; for bosonic systems, $i\omega_n \in i\frac{2\mathbb{Z}}{\beta}$.
- Since $\mathbb{G}(z)$ has poles on the real axis, the real-time (or real-frequency) Green's function could only be evaluated at positions infinitesimally close to the real axis. Let η be a positive infinitesimal number. When $z = \omega + i\eta$, $\mathbb{G}^R(\omega) = \mathbb{G}(\omega + i\eta)$ is the retarded Green's function.

The spectral function $A(\omega)$ is defined from $\mathbb{G}^R(\omega)$ as follows:

$$A(\omega) = -\frac{1}{\pi} \text{Im} (\text{Tr} (\mathbb{G}^R(\omega))) \quad (8)$$

The spectral function contains information of the excitation spectra in quantum systems. In practice, the retarded Green's function and the spectral function are evaluated with η a very small positive number.

Green's functions of fermionic systems are closely related to the matrix-valued Herglotz-Nevanlinna functions (see e.g., [22]). A matrix-valued function $\mathbb{N}(z)$ is said to be Herglotz-Nevanlinna if $\mathbb{N} : \mathbb{C}_+ \rightarrow \mathbb{C}^{N_{\text{orb}} \times N_{\text{orb}}}$ is analytic and $\text{Im}(\mathbb{N}(z))$ is a positive semidefinite matrix for $z \in \mathbb{C}_+$. Here \mathbb{C}_+ is the open complex upper half-plane, $\mathbb{C}^{N_{\text{orb}} \times N_{\text{orb}}}$ is the set of $N_{\text{orb}} \times N_{\text{orb}}$ matrices with entries in \mathbb{C} , and $\text{Im}(\mathbb{N}(z))$ is the imaginary part of $\mathbb{N}(z)$. The matrix-valued Herglotz-Nevanlinna functions admit the following integral representation, (see e.g., [37, Theorem 5.4] for proof):

$$\mathbb{N}(z) = Mz + B + \int_{\mathbb{R}} \left(\frac{1}{t-z} - \frac{t}{1+t^2} \right) d\Sigma(t), \quad z \in \mathbb{C}_+, \quad (9)$$

where $M, B \in \mathbb{C}^{N_{\text{orb}} \times N_{\text{orb}}}$, M is positive semidefinite, B is Hermitian, and $\Sigma(t)$ is a nondecreasing matrix-valued function on \mathbb{R} such that $\int_{\mathbb{R}} \frac{u^\dagger(d\Sigma(t))u}{1+t^2} < \infty$ for any vector $u \in \mathbb{C}^{N_{\text{orb}}}$. This condition on $\Sigma(t)$ ensures that the above integration in Eq. (9) is well-defined. For the detailed mathematical theory of Herglotz-Nevanlinna functions, we refer the readers to [38, Chapter 2] and [39, Chapter 3].

From Eq. (9), we may set

$$d\Sigma(t) = \sum_{l=1}^{N_p} \mathbb{X}_l \delta(t - \lambda_l), \quad M = 0, \quad B = \int_{\mathbb{R}} \frac{t}{1+t^2} d\Sigma(t), \quad (10)$$

then

$$\mathbb{N}(z) = \int_{\mathbb{R}} \frac{d\Sigma(t)}{t-z} = \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{\lambda_l - z} = -\mathbb{G}(z), \quad z \in \mathbb{C}_+. \quad (11)$$

In other words, $-\mathbb{G}(z)$ is a matrix-valued Herglotz-Nevanlinna function. This is the mathematical foundation of Nevanlinna [23] and Carathéodory continuation [24] methods.

The causal space for Green's functions of fermionic systems is a subset of the Herglotz-Nevanlinna function class denoted by \mathcal{S}_F :

$$\mathcal{S}_F = \left\{ \mathbb{G} \left| \begin{array}{l} \mathbb{G}(z) = \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{z - \lambda_l}, \text{ for some } N_p \in \mathbb{Z}, \\ \text{for some } \lambda_l \in \mathbb{R}, \text{ and for } \mathbb{X}_l \text{ satisfying} \\ \text{fermionic rank-1 semidefinite conditions} \\ \text{Eq. (6) and the sum rule Eq. (7).} \end{array} \right. \right\}. \quad (12)$$

Similarly, the causal space of Green's functions of bosonic systems is

$$\mathcal{S}_B = \left\{ \mathbb{G} \left| \begin{array}{l} \mathbb{G}(z) = \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{z - \lambda_l}, \text{ for some } N_p \in \mathbb{Z}, \\ \text{for some } \lambda_l \in \mathbb{R}, \text{ and for } \mathbb{X}_l \text{ satisfying} \\ \text{fermionic rank-1 semidefinite conditions} \\ \text{Eq. (5) and the sum rule Eq. (7).} \end{array} \right. \right\}. \quad (13)$$

In the definition of \mathcal{S}_{\pm} , the rank-1 condition in \mathbb{X}_l can be dropped, because any semi-definite matrix can be represented as the sum of several rank-1 semi-definite matrices. Degenerate excitations can be treated similarly by allowing the values of some λ_l 's to be the same. This is the mathematical foundation of the semidefinite relaxation fitting [35].

Now we are ready to introduce the setup for analytic continuation problems of Matsubara data. Given several (possibly noisy) Matsubara data $\mathbb{G}_n \approx \mathbb{G}(i\omega_n)$ for $n = 1, 2, \dots, N_w$, our goal is to obtain the fitting of these data into the following analytic form $\mathbb{G}(z) = \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{z - \lambda_l}$, where $\lambda_l \in \mathbb{R}$ and \mathbb{X}_l is a $N_{\text{orb}} \times N_{\text{orb}}$ rank-1 semidefinite matrix, and the matrices $\{\mathbb{X}_l\}_{l=1}^{N_p}$ satisfies the sum rule $\sum_l \mathbb{X}_l = \mathbb{I}_{N_{\text{orb}}}$. For fermionic systems, \mathbb{X}_l is positive semidefinite, while for bosonic cases, $\text{sign}(\lambda_l) \cdot \mathbb{X}_l$ is positive semidefinite.

In summary, we would like to obtain the poles λ_l , the semidefinite matrices \mathbb{X}_l , and the number of poles N_p which we may not have *a priori* knowledge. From such information we can infer the spectral function, as well as other diagonal and off-diagonal entries of the Green's function.

Finally, we would like to remark that the Green's functions are only one type of response functions. Other response functions commonly considered, such as the charge, magnetic, or superconducting susceptibilities (see e.g., [36, Chapter 5.5]) admit the same formula Eq. (4), in which \mathbb{X}_l also satisfies the bosonic/fermionic semidefinite conditions (see Eqs. (5) and (6)). The only difference is that \mathbb{X}_l may satisfy a different set of sum rules. In other words, the analytic structure of Green's function and other response functions are similar, therefore the same methodology of analytic continuation applies.

II.2. Step 1: Projection of the noisy data to the causal space

The Matsubara data often contains unphysical noise, i.e., $\mathbb{G}^M(i\omega_n)$ cannot be expressed as a matrix-value func-

tion in the set $\mathcal{S}_{F/B}$. Therefore the first step of our algorithm is to project the noisy data into the causal space. This can be achieved efficiently using semidefinite programming. For simplicity, let us choose a fine uniform grid on the real-axis

$$x_m = -\Lambda_x + mh_x, \quad h_x = \frac{2C}{M}, \quad m = 0, \dots, M, \quad (14)$$

where Λ_x is the real-axis cutoff, and h_x is the grid size. Our goal is to fit the Matsubara data \mathbb{G}_n via the following form:

$$\mathbb{G}_n \approx \sum_{m=0}^M \frac{\mathbb{P}_m}{i\omega_n - x_m}, \quad n = 1, \dots, N_w, \quad (15)$$

where $\{\mathbb{P}_m\}$ are semidefinite matrices. The objective function $\mathcal{E}_{\text{proj}}(\{\mathbb{P}_m\}_{m=0}^M)$ is defined as

$$\mathcal{E}_{\text{proj}}(\{\mathbb{P}_m\}_{m=0}^M) = \left(\sum_{n=1}^{N_w} \left\| \mathbb{G}_n - \sum_{m=0}^M \frac{\mathbb{P}_m}{i\omega_n - x_m} \right\|_F^2 \right)^{1/2}. \quad (16)$$

Here $\|\cdot\|_F$ is the Frobenius norm. The solutions of the projection $\mathbb{P}_m^{\text{proj}}$ is obtained by solving the following optimization problem:

$$\begin{aligned} \{\mathbb{P}_m^{\text{proj}}\} &= \arg \min_{\{\mathbb{P}_m\}} \mathcal{E}_{\text{proj}}(\{\mathbb{P}_m\}_{m=1}^M) \\ \text{subject to} &\begin{cases} \text{i) semidefinite constraint:} \\ \quad \begin{cases} \mathbb{P}_m \geq 0 & (\text{Fermions}), \\ \text{sign}(x_m) \cdot \mathbb{P}_m \geq 0 & (\text{Bosons case}); \end{cases} \\ \text{ii) sum rules:} \\ \quad \sum_m \mathbb{P}_m = \mathbb{I}_{N_{\text{orb}}}. \end{cases} \end{aligned} \quad (17)$$

This is a convex optimization problem, and can be solved efficiently using software packages such as CVX [40]. If we are performing analytic continuation of other correlation functions, $\{\mathbb{P}_m\}_{m=1}^M$ may be subject to a different sum rule.

For scalar-valued fermionic Green's function, the positive semidefinite condition $\mathbb{P}_m \geq 0$ becomes the non-negativity condition $\mathbb{P}_m \geq 0$. Therefore the sum rule $\sum_m \mathbb{P}_m = 1$ can also be written as $\sum_m |\mathbb{P}_m| = 1$, i.e., an ℓ^1 -norm constraint on the vector $(\mathbb{P}_1, \dots, \mathbb{P}_M)$. The least squares problem with a ℓ^1 -norm constraint is similar to the well-known Least Absolute Selection and Shrinkage Operator (LASSO) problem [41], which favors solutions with a sparse structure (see also [42]). This agrees with our numerical observation that the solution \mathbb{P}_m often has relatively few nonzero entries. For matrix-valued Green's function, there is a similar mechanism that induces the sparsity of solution. The natural generalization of the ℓ^1 norm is the nuclear norm $\|\cdot\|_*$, defined as $\|A\|_* = \sqrt{\text{Tr}(A^\dagger A)} = \sum_k \sigma_k(A)$, where $\sigma_k(A)$ is the k -th largest singular value for A . Note that each \mathbb{P}_m is a positive semidefinite matrix, therefore the k -th largest singular value $\sigma_k(\mathbb{P}_m)$ is equal to the k -th largest eigenvalue $\lambda_k(\mathbb{P}_m)$. Therefore, we have $\sum_m \|\mathbb{P}_m\|_* = \sum_m \sum_k \sigma_k(\mathbb{P}_m) = \sum_m \sum_k \lambda_k(\mathbb{P}_m) =$

$\sum_m \text{Tr}(\mathbb{P}_m) = \text{Tr}(\sum_m \mathbb{P}_m)$. Combined with the sum rule $\sum_m \mathbb{P}_m = \mathbb{I}_{N_{\text{orb}}}$, we can see that the matrices $\{\mathbb{P}_m\}$'s are enforced to satisfy the nuclear norm constraint $\sum_m \|\mathbb{P}_m\|_* = N_{\text{orb}}$. The nuclear norm is also a sparsity-inducing property (see e.g., [43]).

After obtaining $\mathbb{P}_m^{\text{proj}}$, we can construct the projected Matsubara data $\mathbb{G}_n^{\text{proj}}$:

$$\mathbb{G}_n^{\text{proj}} = \sum_{m=0}^M \frac{\mathbb{P}_m^{\text{proj}}}{i\omega_n - x_m}, \quad n = 1, \dots, N_w, \quad (18)$$

This projection step can be used to improve the robustness of other analytic continuation methods, and therefore is of independent interest. For instance, the Nevanlinna analytic continuation method requires the Pick matrix to be positive semidefinite (see [23] for details). Numerical results indicate that this criterion can be violated when very small perturbations are applied to the Matsubara data. Given that noise is inevitable in many Green's function solvers (notably, quantum Monte Carlo solvers), the application range of the Nevanlinna analytic continuation is thus significantly limited by the nature of the noise. Our numerical results suggest that when the Nevanlinna method is applied to the projected noisy Matsubara data (for diagonal entries), the quality of the analytic continuation is significantly improved. (See Section III.1 and Section III.3 for details.) Similar improvements are observed on both diagonal and off-diagonal entries of the Green's functions while applying Carathéodory continuations on the projected data (also see Section III.1 for details).

II.3. Step 2: Estimation of pole locations using the AAA algorithm

While the projection step can be viewed as an analytic continuation algorithm by itself, the quality of the continuation is constrained by the resolution of the grid on the real axis. In practice, the number of grid points (i.e., M) in the uniform grid is often too small to accurately resolve the pole locations, but is too large to be directly used in the subsequent semidefinite relaxation (SDR) step to be detailed in Section II.4. Furthermore, the loss function in the SDR step is highly non-convex, and the optimization with respect to this loss function requires a proper initial guess. These considerations lead to the second step of the algorithm for estimating the locations of a relatively small number of poles.

We use the adaptive Antoulas–Anderson (AAA) algorithm [34] for the pole estimation, which is available as a subroutine in the Chebfun package [44]. The AAA algorithm is able to obtain the poles of a scalar complex-valued function $g(z)$. In the context of analytic continuation, the scalar function $g(z)$ could either be each diagonal entry $\mathbb{G}_{ii}(z)$ ($i = 1, \dots, N_{\text{orb}}$) of the Green's function, or the trace $\text{Tr}(\mathbb{G}(z))$. In practice, we find that performing the pole estimation on each diagonal entry separately often gives better results.

The AAA algorithm performs a rational approximation to a scalar function $g(z)$ using barycentric representations [45]:

$$g(z) \approx \frac{n(z)}{d(z)} = \frac{\sum_{j=1}^n \frac{w_j g_j}{z - z_j}}{\sum_{j=1}^n \frac{w_j}{z - z_j}}. \quad (19)$$

Here $n(z) = \sum_{j=1}^n \frac{w_j g_j}{z - z_j}$ and $d(z) = \sum_{j=1}^n \frac{w_j}{z - z_j}$ are the numerator and denominator part of the barycentric representation, respectively. Note that the above formulation Eq. (19) automatically satisfies that $\frac{n(z_j)}{d(z_j)} = g_j$.

We briefly describe the AAA algorithm below, and refer readers to Ref. [34] for more details. Assume that we are given the values of $g(z)$ on a set $\mathcal{Z} \subset \mathbb{C}$ of N points (in analytic continuation, $\mathcal{Z} = \{i\omega_n, n = 1, \dots, N_w\}$), the AAA algorithm aims at finding a set of support points $\{z_j\}_{j=1}^n$, which allows us to solve for the weights $\{w_j\}_{j=1}^n$ in Eq. (19). Here $g_j = g(z_j)$ ($j = 1, \dots, n$). Starting from $n = 1$, the AAA algorithm gradually expands the set $\{z_j\}_{j=1}^n$ following a greedy algorithm.

At the $(n - 1)$ -th step ($n = 1, 2, \dots$), if the residual $\text{res}(z) = g(z) - \frac{n(z)}{d(z)}$ is sufficiently small for all $z \in \mathcal{Z}$, then the algorithm terminates. Otherwise, we pick the next support point z_n , by choosing $z \in \mathcal{Z} \setminus \{z_1, \dots, z_{n-1}\}$ where the residual $\text{res}(z)$ takes its maximum absolute value. Let us write $\mathcal{Z} \setminus \{z_1, \dots, z_n\}$ as a column vector $\mathbf{Z}^{(n)} := (Z_1^{(n)}, \dots, Z_{N-n}^{(n)})^T$ and define $\mathbf{g}^{(n)} = (g_1^{(n)}, \dots, g_{N-n}^{(n)})^T := (g(Z_1^{(n)}), \dots, g(Z_{N-n}^{(n)}))^T$. Since $g(z)$ is supposed to be approximated by $g(z) \approx n(z)/d(z)$, i.e. $g(z)d(z) \approx n(z)$, we choose the normalized weight $w = (w_1, \dots, w_n)$ by solving the least square problem:

$$\begin{aligned} \min_{\|w\|_2=1} \sum_{i=1}^{N-n} \left| g_i^{(n)} d(Z_i^{(n)}) - n(Z_i^{(n)}) \right|^2 \\ = \sum_{i=1}^{N-n} \left| \sum_{j=1}^n \frac{w_j (g_i^{(n)} - g_j)}{Z_i^{(n)} - z_j} \right|^2. \end{aligned} \quad (20)$$

The minimization problem of Eq. (20) could be viewed as the following linear algebra problem:

$$\min_{\|w\|_2=1} \|A^{(n)} w\|_2, \quad (21)$$

where $A^{(n)}$ is

$$A^{(n)} = \begin{pmatrix} \frac{g_1^{(n)} - g_1}{Z_1^{(n)} - z_1} & \dots & \frac{g_1^{(n)} - g_n}{Z_1^{(n)} - z_n} \\ \vdots & \ddots & \vdots \\ \frac{g_{N-n}^{(n)} - g_1}{Z_{N-n}^{(n)} - z_1} & \dots & \frac{g_{N-n}^{(n)} - g_n}{Z_{N-n}^{(n)} - z_n} \end{pmatrix}. \quad (22)$$

This optimization problem Eq. (21) could be solved by performing a singular value decomposition (SVD) on the

matrix $A^{(n)} = U\Sigma V^\dagger$, where U is a $(N - n) \times r$ orthogonal matrix, V is a $n \times r$ orthogonal matrix, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$ is a diagonal matrix ($\sigma_1 \geq \dots \geq \sigma_r$), and $r \leq \min(N - n, n)$ is the rank of $A^{(n)}$. The solution w of Eq. (21) should be taken as the last column of V .

After the iteration terminates at the K -th step, we have $\{z_1, \dots, z_K\}$ and its associated weights $\{w_1, \dots, w_K\}$. We can calculate the zeros of $d(z)$, which serve as the estimated poles of $g(z)$ that we want to obtain, using the following generalized eigenvalue problem [34]:

$$\begin{pmatrix} 0 & w_1 & w_2 & \cdots & w_K \\ 1 & z_1 & & & \\ 1 & & z_2 & & \\ \vdots & & & \ddots & \\ 1 & & & & z_K \end{pmatrix} = \xi \begin{pmatrix} 0 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}. \quad (23)$$

At least two of the eigenvalues in this problem are infinite, and the remaining $(K - 1)$ eigenvalues $\{\xi_a\}_{a=1}^{K-1}$ are the zeros of d , i.e. the complex-valued poles of g .

Since all poles of the Green's function in Eq. (4) are real-valued, after obtaining $\{\xi_a\}_{a=1}^{K-1}$, we discard the poles far away from the real axis. In other words, we only keep poles with $|\text{Im} \xi_a| \leq \varepsilon_p$ for some $\varepsilon_p > 0$, and define $\{\lambda_l^{\text{in}}\}_{l=1}^{N_p}$ to be real parts of the remaining poles. These N_p poles will be used as the initial guess of the poles in the semidefinite relaxation step below.

II.4. Step 3: Semidefinite relaxation

In the final step, we use numerical optimization to obtain an effective fitting of the Matsubara data in the form of Eq. (4). The fitting error is defined as

$$\begin{aligned} & \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l\}_{l=1}^{N_p} \right) \\ &= \left(\sum_{n=1}^N \left\| \mathbb{G}_n - \mathbb{G} \left(i\omega_n; \{\lambda_l, \mathbb{X}_l\}_{l=1}^{N_p} \right) \right\|_F^2 \right)^{1/2} \\ &= \left(\sum_{n=1}^N \left\| \mathbb{G}_n - \sum_{l=1}^{N_p} \frac{\mathbb{X}_l}{i\omega_n - \lambda_l} \right\|_F^2 \right)^{1/2}. \end{aligned} \quad (24)$$

The fitting error expressed in Eq. (24) is a highly nonconvex function, and the minimization of this function can frequently be trapped in local minima, which strongly depend on the choice of the initial guess.

Recall that in Eq. (4) $\{\mathbb{X}_l\}$ are required to be rank-1 semidefinite matrix. The semidefinite relaxation drops this rank-1 constraint. In other words, $\{\mathbb{X}_l\}$ (or $\text{sign}(\lambda_l) \cdot \mathbb{X}_l$) are only required to be positive semidefinite matrices for the analytic continuation of fermionic (bosonic) Green's function.

For simplicity, the following discussion focuses on fermionic Green's functions, while the same numerical

treatment is applicable to bosonic Green's functions. The minimization of $\text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l \geq 0\}_{l=1}^{N_p} \right)$ can be formulated as a bi-level optimization problem:

$$\begin{aligned} & \min_{\lambda_l, \mathbb{X}_l \geq 0} \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l\}_{l=1}^{N_p} \right) \\ &= \min_{\lambda_l \in \mathbb{R}} \min_{\mathbb{X}_l \geq 0} \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l\}_{l=1}^{N_p} \right), \\ &= \min_{\lambda_l \in \mathbb{R}} \mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right) \end{aligned} \quad (25)$$

where

$$\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right) = \min_{\mathbb{X}_l \geq 0} \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l \geq 0\}_{l=1}^{N_p} \right). \quad (26)$$

When the poles $\{\lambda_l\}_{l=1}^{N_p}$ are fixed, the optimization with respect to semidefinite matrices $\{\mathbb{X}_l \geq 0\}$ is a convex optimization problem, and could be efficiently solved using software packages such as CVX [40]. The optimization of $\text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l \geq 0\}_{l=1}^{N_p} \right)$ is transformed into the optimization of $\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right)$, but note that such an optimization over the pole positions is still a nonconvex problem.

Therefore, we have the following SDR fitting procedure through the bi-level optimization framework: starting from an initial value for poles $\{\lambda_l^{\text{in}}\}_{l=1}^{N_p}$ obtained by the previous pole estimation step, we use numerical optimization technique to minimize $\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right)$, where

- The value of $\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right)$ and the corresponding optimal $\{\mathbb{X}_l^{\text{opt}} \geq 0\}$ for fixed poles $\{\lambda_l\}_{l=1}^{N_p}$ are found by a convex optimization solver.
- Since $\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right) = \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l^{\text{opt}}\}_{l=1}^{N_p} \right)$ and the optimal matrices $\{\mathbb{X}_l^{\text{opt}}\}$ should be regarded as functions of the poles $\{\lambda_l\}$, the partial derivative of $\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right)$ could be written as

$$\begin{aligned} \partial_m \mathcal{E} &= \frac{\partial}{\partial \lambda_m} \text{Err} \\ &+ \sum_{l=1}^{N_p} \sum_{i,j=1}^{N_{\text{orb}}} \frac{\partial \text{Err}}{\partial (\mathbb{X}_l)_{ij}} \Big|_{\mathbb{X}_l^{\text{opt}}} \cdot \frac{\partial (\mathbb{X}_l^{\text{opt}})_{ij}}{\partial \lambda_m}, \end{aligned} \quad (27)$$

The first order optimality condition for solving Eq. (26) implies $\frac{\partial \text{Err}}{\partial (\mathbb{X}_l)_{ij}} \Big|_{\mathbb{X}_l^{\text{opt}}} = 0$. Such an evaluation of the derivative quantities is similar to the treatment in the Hellmann-Feynman's theorem [46]. Therefore we have

$$\begin{aligned} \partial_m \mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right) &= \frac{\partial}{\partial \lambda_m} \text{Err} \left(\{\lambda_l\}_{l=1}^{N_p}, \{\mathbb{X}_l\}_{l=1}^{N_p} \right) \Big|_{\mathbb{X}_l = \mathbb{X}_l^{\text{opt}}} \\ &= - \frac{\sum_{n=1}^N \sum_{i,j=1}^{N_{\text{orb}}} \text{Re} \left(\left((\mathbb{G}_n)_{ij} - \sum_l \frac{(\mathbb{X}_l^{\text{opt}})_{ij}}{i\omega_n - \lambda_l} \right) \frac{(\mathbb{X}_m^{\text{opt}})_{ij}^*}{(i\omega_n + \lambda_m)^2} \right)}{\sqrt{\mathcal{E} \left(\{\lambda_l\}_{l=1}^{N_p} \right)}}. \end{aligned} \quad (28)$$

Therefore, both the function value and gradient of \mathcal{E} can be efficiently evaluated. The optimization of $\mathcal{E}(\{\lambda_l\}_{l=1}^{N_p})$ could be conducted with standard gradient-based optimization solvers, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method (see e.g. [47]).

Note that the SDR fitting step enforces that the Green's functions must be in the causal space, and is thus robust to unphysical noises. Since the projection step may introduce biases especially when the uniform mesh in Eq. (14) is relatively coarse, we find that the performance of SDR fitting is often better when applied to the original noisy data, instead of the data obtained after the projection.

II.5. Summary of the PES method

We summarize the PES method in Algorithm 1 below.

Algorithm 1 Projection-Estimation-Semidefinite relaxation (PES) method

- 1: **Parameters:** real axis cutoff Λ_x and mesh size h_x in the projection step; imaginary axis cutoff ε_p in the pole estimation step.
 - 2: **Input:** Matsubara data $\{\mathbb{G}_n\}_{n=1}^{N_w}$, where \mathbb{G}_n is an $N_{\text{orb}} \times N_{\text{orb}}$ matrix.
 - 3: **Output:** Poles $\{\lambda_l\}_{l=1}^{N_p}$, weights $\{\mathbb{X}_l\}_{l=1}^{N_p}$.
 - 4: **Projection** of the noisy data $\{\mathbb{G}_n\}_{n=1}^{N_w}$: form a uniform mesh in Eq. (14), minimize $\mathcal{E}_{\text{proj}}(\{\mathbb{P}_m\}_{m=1}^M)$ in Eq. (17), and then obtain the projected data $\{\mathbb{G}_n^{\text{proj}}\}_{n=1}^{N_w}$ using Eq. (18).
 - 5: **Estimation** of the pole locations using the projected data $\{\mathbb{G}_n^{\text{proj}}\}_{n=1}^{N_w}$ via the AAA algorithm. The output is $\{\lambda_l^{\text{in}}\}_{l=1}^{N_p}$.
 - 6: **Semidefinite relaxation fitting** through bi-level optimization: with the initial poles $\{\lambda_l^{\text{in}}\}_{l=1}^{N_p}$, we can conduct SDR fitting by minimizing $\mathcal{E}(\{\lambda_l\}_{l=1}^{N_p})$ in Eq. (26) to obtain the final poles $\{\lambda_l\}_{l=1}^{N_p}$ and weights $\{\mathbb{X}_l\}_{l=1}^{N_p}$.
-

Computational cost: Both the projection and the semidefinite relaxation steps require the solution of convex optimization problems that can be reformulated into semidefinite programming (SDP) problems. The cost for solving the convex optimization problems can depend on the SDP reformulation, and the choice of algorithms. The default solver in the CVX software package is SDTP3 [48], which implements the interior point method with a primal-dual path-following strategy [49] to solve SDP problems. In the primal-dual interior point method, the cost is dominated by solving the Newton equation to obtain the primal-dual search direction. The unknown variables are N_v positive semidefinite matrices each of size $N_d \times N_d$, and the total number of variables is $N_v N_d^2$. Therefore the cost of solving Newton's equation can be as large as $\mathcal{O}(N_v^3 N_d^6)$, and there is an additional cost in forming the matrices in Newton's equation (see e.g.,

[50]). While the cost of solving Newton's equation may be reduced using preconditioned and iterative solvers in certain cases [51], in general, the solution the SDP problem can become expensive when both N_v and N_d are large.

In the projection step, $N_v = M$ is usually large. Note that the pole estimation step only uses the diagonal entries of the Green's function, and it is sufficient to only conduct projection of these diagonal entries. If M is very large, we may further reduce the cost by performing projection for each diagonal entry of the Green's functions separately. The semidefinite relaxation fitting requires the solution of $N_v = N_p$ matrices each of size $N_d = N_{\text{orb}}$. Therefore it is important for the pole estimation step to obtain a relatively small number of poles. Furthermore, if the non-diagonal information of the Green's function is not needed (for example if the goal is to calculate the spectral function), the cost can be significantly reduced by conducting the semidefinite relaxation step for the diagonal entries only. Compared to solving the SDP problems, the cost of the AAA algorithm for the pole estimation is usually negligible.

Noiseless data: If the data is noiseless or if the Green's functions are causal by construction (e.g., Green's functions are approximated via causal diagrammatic methods), we can skip the projection step, and conduct estimation and SDR fitting directly. Furthermore, since the projection is only performed on a finite grid along the real axis, the pre-processing step may become detrimental to the accuracy when the grid size is relatively small.

Parameters in AAA: Numerical results suggest that the final result is often insensitive to the choice of the parameter ε_p . For noisy data, the Lawson refinement step [52] (also implemented in Chebfun), which is based on an iteratively reweighted least-squares (IRLS) iteration can also be used to improve the fitting result. Also, the AAA algorithm in Chebfun [44] could be implemented with or without specifying the number of poles. Prior knowledge on the number of peaks in the physical system can be used to control the maximal number of poles obtained by the AAA algorithm, and to enhance the robustness with respect to noise.

III. RESULTS

In order to demonstrate the strengths of the PES method and give a fair comparison with other methods, we present various results using the Hubbard dimer system and the band structure calculation of different materials. When testing the performances of analytic continuation methods in the presence of noise, we will manually add multiplicative Gaussian noise to the clean Matsubara data:

$$\mathbb{G}_{\text{noisy}} = \mathbb{G}_{\text{exact}} \cdot (1 + \sigma N_{\mathbb{C}}(0, 1))$$

Here $N_{\mathbb{C}}(0, 1)$ is the complex-valued normal Gaussian noise and σ is referred to as the (multiplicative) noise level of the data.

III.1. Hubbard dimer

III.1.1. Setup of the Hubbard Dimer

The Hubbard dimer example used in [24] is a simple but nontrivial test case, since it has multiple sharp features while the true value of the Green's function could be obtained by the exact diagonalization. It is also a prototypical system for studying the excitation spectra of molecules. There are only four orbitals in the Hubbard dimer, namely $|0 \uparrow\rangle, |0 \downarrow\rangle, |1 \uparrow\rangle, |1 \downarrow\rangle$. The Hamiltonian is

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \quad (29)$$

where

$$\begin{aligned} \hat{H}_0 = & - \sum_{\sigma \in \{\uparrow, \downarrow\}} t \left(\hat{c}_{0\sigma}^\dagger \hat{c}_{1\sigma} + \hat{c}_{1\sigma}^\dagger \hat{c}_{0\sigma} \right) \\ & - \mu (\hat{n}_{0,\uparrow} + \hat{n}_{0,\downarrow} + \hat{n}_{1,\uparrow} + \hat{n}_{1,\downarrow}), \end{aligned} \quad (30)$$

$$\begin{aligned} \hat{H}_1 = & (U + U_a) \hat{n}_{0,\uparrow} \hat{n}_{0,\downarrow} + (U - U_a) \hat{n}_{1,\uparrow} \hat{n}_{1,\downarrow} \\ & - \frac{U}{2} (\hat{n}_{0,\uparrow} + \hat{n}_{0,\downarrow} + \hat{n}_{1,\uparrow} + \hat{n}_{1,\downarrow}) \\ & + h (\hat{n}_{0\uparrow} - \hat{n}_{0\downarrow} + \hat{n}_{1\uparrow} - \hat{n}_{1\downarrow}) \\ & + \mu_a (\hat{n}_{0\uparrow} + \hat{n}_{0\downarrow} - \hat{n}_{1\uparrow} - \hat{n}_{1\downarrow}) \\ & + h_a (\hat{n}_{0\uparrow} - \hat{n}_{0\downarrow} - \hat{n}_{1\uparrow} + \hat{n}_{1\downarrow}), \end{aligned} \quad (31)$$

and

$$\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}, \quad i = 0, 1, \quad \sigma = \uparrow, \downarrow. \quad (32)$$

The parameters are chosen the same as in [24], namely $t = 1, \mu = 0.7, h = 0.3, U_a = 0.5, \mu_a = 0.2, h_a = 0.03, \beta = 31.1$ and $U = 5$.

III.1.2. Fitting results, with comparison to other methods

Let us first demonstrate the efficiency of the PES method. We conduct analytic continuation for four sets of Matsubara data with different noise levels: $\sigma = 0, \sigma = 4 \times 10^{-6}, \sigma = 2.56 \times 10^{-4}, \sigma = 1.6384 \times 10^{-2}$. In Fig. 1, we plot the spectral function (evaluated at $\mathbb{R} + i\eta, \eta = 0.01$) obtained from analytic continuation with comparison to its true values. The methods that we are comparing with PES are Nevanlinna [23], Carathéodory [24] and MaxEnt [5]. We also implement the ES method, in which we skip the projection step and conduct the estimation and the SDR fitting directly. To make it a fair comparison, we also perform Nevanlinna and Carathéodory continuations on the projected data. Since the amplitude of MaxEnt

spectral functions are very small compared to the delta peaks in the true spectral functions, we rescale the MaxEnt results for comparison with true values and results from other methods.

Our observation is summarized as follows:

- When the Matsubara data is clean (noise-level $\sigma = 0$), all methods except the MaxEnt continuation could retrieve all the peaks perfectly. The MaxEnt method can not deal with sharp features. In all cases, MaxEnt can only result in two (and up to three) significantly broadened peaks.
- Both the PES method and the ES method are robust to noise. It can retrieve almost all features of the Green's function when the noise level is small ($\sigma = 4 \times 10^{-6}$ and 2.56×10^{-4}), and can retrieve quite a few features when the noise level is relatively large ($\sigma = 1.6384 \times 10^{-2}$). Particularly, the band gap could be accurately resolved even in the large noise scenario.
- For noisy data, PES behaves better than ES (see noise level $\sigma = 4 \times 10^{-6}$ and 2.56×10^{-4}), which implies that the projection of noisy data is indeed helpful. For noiseless data ($\sigma = 0$), ES behaves better than PES.
- Even when the data is only slightly noisy (noise-level $\sigma = 4 \times 10^{-6}$), the Nevanlinna and Carathéodory continuation breaks down. This is due to the violation of the Pick criterion [23]. The spectral function is no longer nonnegative and multiple features are missing.
- If combined with the projection of data into the causal space, the results of the Nevanlinna and Carathéodory continuation are greatly improved. However, the quality of Projection+Nevanlinna and Carathéodory is still not as good as PES. This could be seen in all noisy levels. What's more, the calculation results are still not guaranteed to be causal, as shown in the results of Projection+Carathéodory methods for relatively large noise (i.e. $\sigma = 2.56 \times 10^{-4}$ and $\sigma = 1.6384 \times 10^{-2}$).

Since both PES method and Carathéodory continuation are able to obtain non-diagonal entries of the Green's functions, let us show the calculation result of the real part of G_{13} (also evaluated at $\mathbb{R} + i\eta, \eta = 0.01$) in Fig. 2, in which we compare the performances of PES, ES, Carathéodory and Projection+Carathéodory.

Similar to what we have learned from the results of spectral functions, we have the following observations:

- The Projection+Carathéodory does not perform as good as the PES and ES method. This could be seen at all noise levels. The two high peaks are perfectly retrieved by the PES and ES method but not by Projection+Carathéodory.

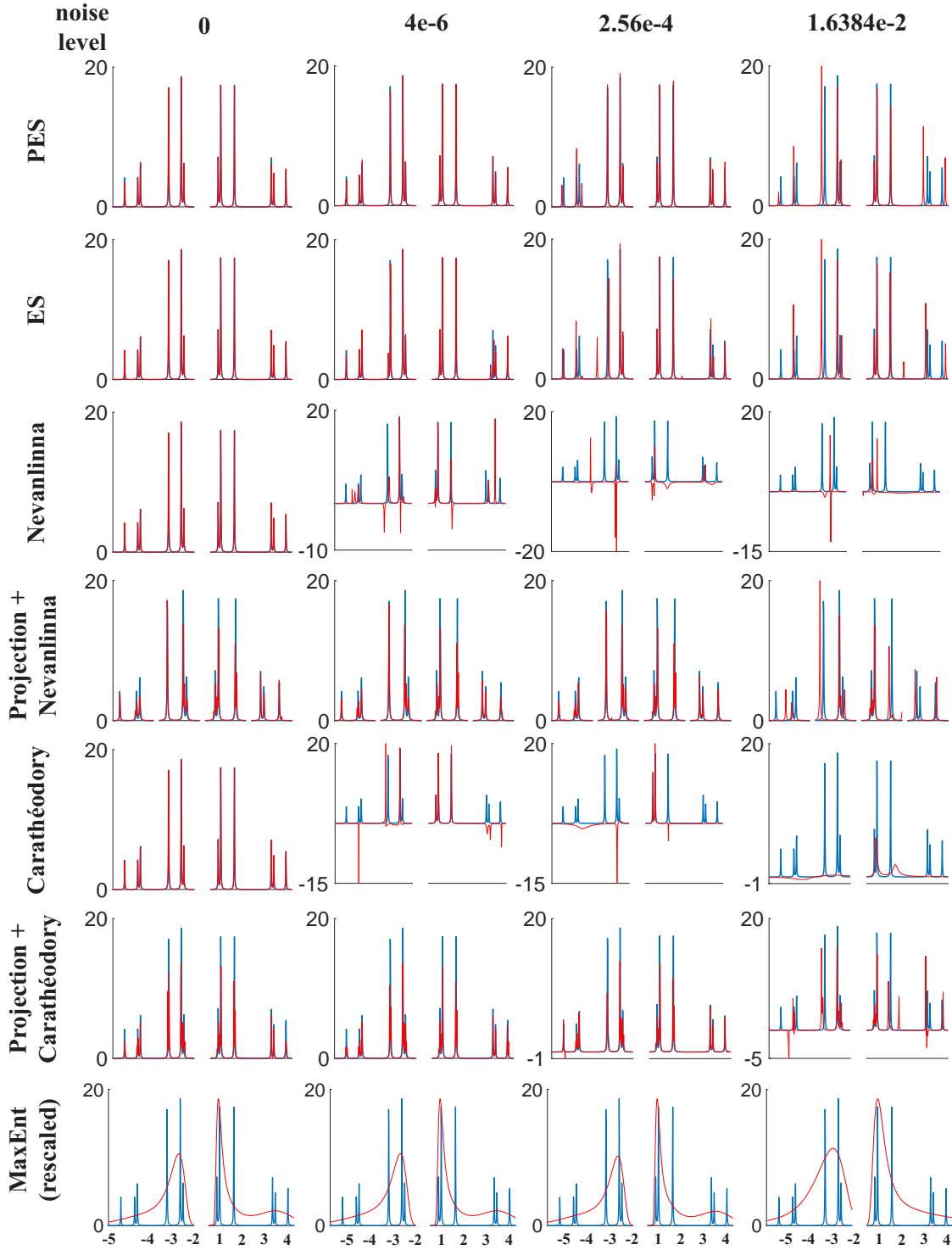


Figure 1. Spectral functions of the Hubbard dimer system obtained from PES, ES, Nevanlinna, Projection+Nevanlinna, Carathéodory, Projection+Carathéodory and MaxEnt methods. Blue line: true value. Red line: calculation results from analytic continuation.

- At noise-level $\sigma = 1.6384 \times 10^{-2}$, the noise is too large and no method could recover all peaks perfectly.

With these observations, we conclude that the PES

method has the most robust performance, especially for analytic continuation of the matrix-valued noisy Matsubara data. The causal projection could help cure the noise-sensitivity issue of Nevanlinna and Carathéodory contin-

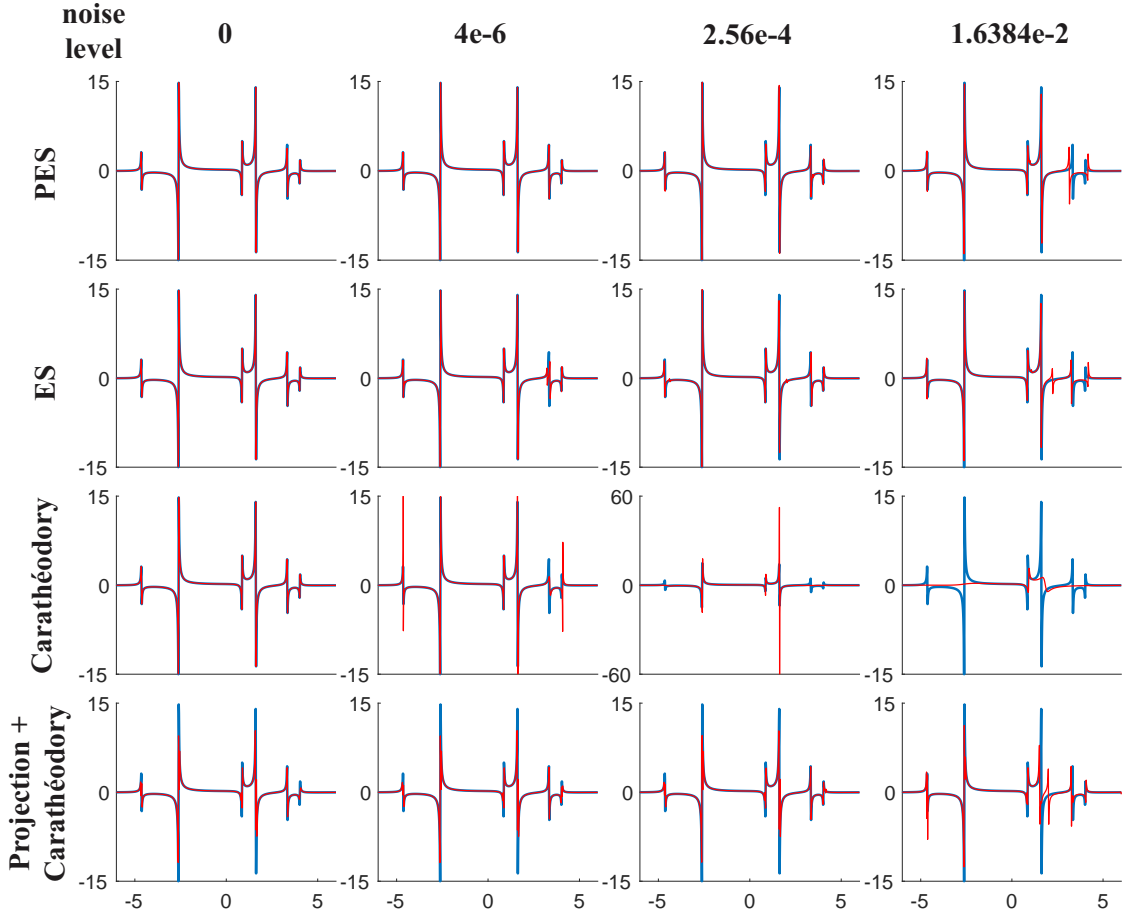


Figure 2. The non-diagonal element G_{13} obtained by different analytic continuation methods. Blue line: true value. Red line: calculation results from analytic continuation.

uation, but does not always guarantee a causal result, and not necessarily perform as well as the PES method.

III.1.3. Importance of the pole estimation

Now we would like to demonstrate the importance of the pole estimation step. In [35], the SDR step is implemented in the context of hybridization fitting without the pole estimation. A random initialization of poles could fail to converge towards the global minimum of the loss function (see [35, Figure 2]). In the context of analytic continuation, we also find that a random selection of initial poles will not converge to its true positions. Particularly, the high frequency features, i.e. the poles away from zero may not be identified accurately.

To illustrate the difficulty in the SDR step, we fix the poles to be its true values except for λ_1 . The true value of λ_1 is -4.6361 . We choose its initial value to be $\lambda_1^{\text{in}} = -5.25, -5.00, -4.75, -4.50, -4.25, -4.00$, and then conduct the SDR fitting step. The corresponding result of the spectral function at the interval $[-5, -4]$ is shown in Fig. 3 (top panel). We see that even such a

small deviation from the true value of λ_1 can result in large errors in analytic continuation.

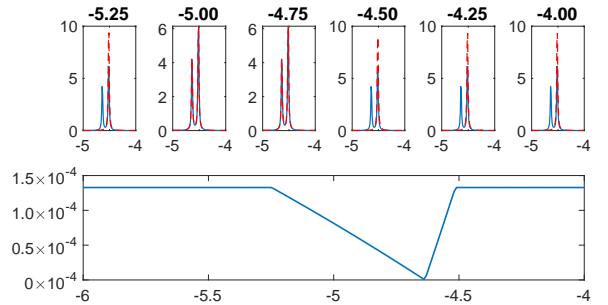


Figure 3. Top panel: calculation results of spectral functions (red dashed lines) with comparison to true value (blue lines) on the interval $[-5, -4]$. The initial value of λ_1 are chosen as $\lambda_1^{\text{in}} = -5.25, -5.00, -4.75, -4.50, -4.25$ and -4.00 . Bottom panel: landscape of $\mathcal{E}(\{\lambda_l\}_{l=1}^{N_p})$ with respect to λ_1 on the interval $[-6, -4]$. Other poles are kept fixed at their true positions: $-5.2479, -4.5143, -3.1957, -2.6075, -2.4857, 0.8619, 0.9837, 1.6181, 3.4599, 3.3381$ and 4.0255 .

The reason behind this is as follows. Let us keep other poles fixed at their true values and let λ_1 vary from -6 to -4 , and plot the landscape of the loss function $\mathcal{E}(\{\lambda_l\})$ in Fig. 3 (bottom panel). We can see that the loss function is highly nonconvex. When λ_1^{in} is outside the interval $[-5.25, -4.5]$, the loss function is essentially flat. In fact, if $\lambda_1^{\text{in}} \notin [-5.25, -4.5]$, the optimized semidefinite matrix $\mathbb{X}_1^{\text{opt}}$ is a zero matrix. This will result in a vanishing gradient for λ_1^{in} (see Eq. (28)), and therefore the optimization could not continue. In such an adversarial scenario, a random initialization of poles cannot recover the true pole locations.

III.2. Bosonic response functions

In order to demonstrate that the PES method is also applicable to bosonic response functions, let us consider the quantity $b(\tau) = \langle \hat{n}_{0\uparrow}(\tau) \hat{n}_{0\uparrow}(0) \rangle$ of the Hubbard dimer model, which indicates the same-spin susceptibility for the 0-th site of the system. This function is also used in [2] to test the performance of analytic continuation methods for bosonic functions. We use the Hubbard dimer example with the same parameters as in Section III.1 other than taking $U = 0.9$.

Let us use $b(w)$ to denote the time-response function in the frequency domain. We are trying to fit the Matsubara data of $b(\tau)$ into the following form:

$$b(w) = \sum_{l=1}^{N_p} \frac{x_l}{w - \lambda_l}, \quad (33)$$

where the quantities x_l satisfies the semidefinite constraints Eq. (5) and the constraint

$$\sum_l x_l = \frac{1}{Z} \sum_{r,s} |\langle \Psi_s | \hat{n}_{0\uparrow} | \Psi_r \rangle|^2 (e^{-\beta E_s} - e^{-\beta E_r}) = 0. \quad (34)$$

Here $N_{\text{orb}} = 1$ since we are only considering the time response function with respect to the 0-th site.

In the current physical system, the spectral function of $b(w)$ has four peaks, with two on the positive half axis, and two on the negative half axis. In fact, $b(w)$ is an odd function. Since the susceptibility is mirrored for negative frequencies, we only choose to plot the positive half axis. Note that since the Nevanlinna and Carathéodory continuation are not applicable to bosonic functions, we can only compare the PES method to the MaxEnt continuation method, for Matsubara data with different noise levels $\sigma = 0, 10^{-6}, 10^{-4}$ and 10^{-2} . The result is shown in Fig. 4.

The MaxEnt method could not resolve the peaks of the bosonic response function under any noise level. We can see that when the noise level is small ($\sigma = 0, 10^{-6}, 10^{-4}$), the PES method could recover both peaks on the positive half axis. When the noise level is large ($\sigma = 10^{-2}$), the PES method could only recover one peak. This means

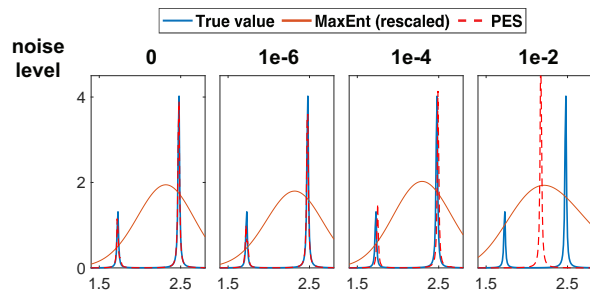


Figure 4. Analytic continuation of bosonic response functions.

that the noise-robustness of PES method for bosonic functions are comparable to that for fermionic functions.

III.3. Band structure

We now applied the PES method to *ab initio* band structure calculations of solids. We use two data sets obtained from GW calculations: the matrix-valued Green's functions for crystalline Si in the equilibrium geometry, previously analyzed in Ref. [24], and for AgI obtained with a relativistic exact two-component formulation in the one-electron approximation (x2c1e) [53]. The data have been obtained by fully self-consistent GW formulated in Gaussian Bloch orbitals, as described in detail in Ref. [54]. Both AgI and Si employ the gth-dzvp-moloptsr basis [55] with gth-pbe pseudopotential [56], using integrals generated by the PySCF package [57] and a code based on ALPS [58]. Calculations are performed on a $6 \times 6 \times 6$ grid [59] and interpolated using a Wannier interpolation of Matsubara data [54] on 52 positive IR frequencies [60, 61] on a grid with 200 momentum points along a high-symmetry path. The broadening parameter of $\eta = 0.01\text{Ha} \approx 3157\text{K}$ is larger than the temperature of $T = 0.001\text{Ha} \approx 316\text{K}$ and correlation effects.

The analytic continuation is conducted as follows. Since the number of orbitals is much larger than that in the Hubbard dimer example ($N_{\text{orb}} \approx 70$ for silicon and $N_{\text{orb}} \approx 230$ for AgI), the computational cost will be too high if we conduct the analytic continuation for the entire matrix. Therefore we only implement the analytic continuation for the diagonal elements. The projection step is conducted on each diagonal entry separately with a fine grid of $M = 12000$ points. After the projection step, we estimate the poles of each diagonal entry using AAA algorithm, and then conduct the SDR fitting of each diagonal entry separately.

We conduct analytic continuation for Matsubara data with different noise levels $\sigma = 0, 10^{-6}, 10^{-4}$ and 10^{-2} . The results of the spectral functions along a high symmetry path reveal the band structure, and are plotted in Fig. 5. We place a white marker in the plot if the corresponding component of the spectral function becomes negative to indicate that the result is unphysical.

Once again, we can see that Nevanlinna continua-

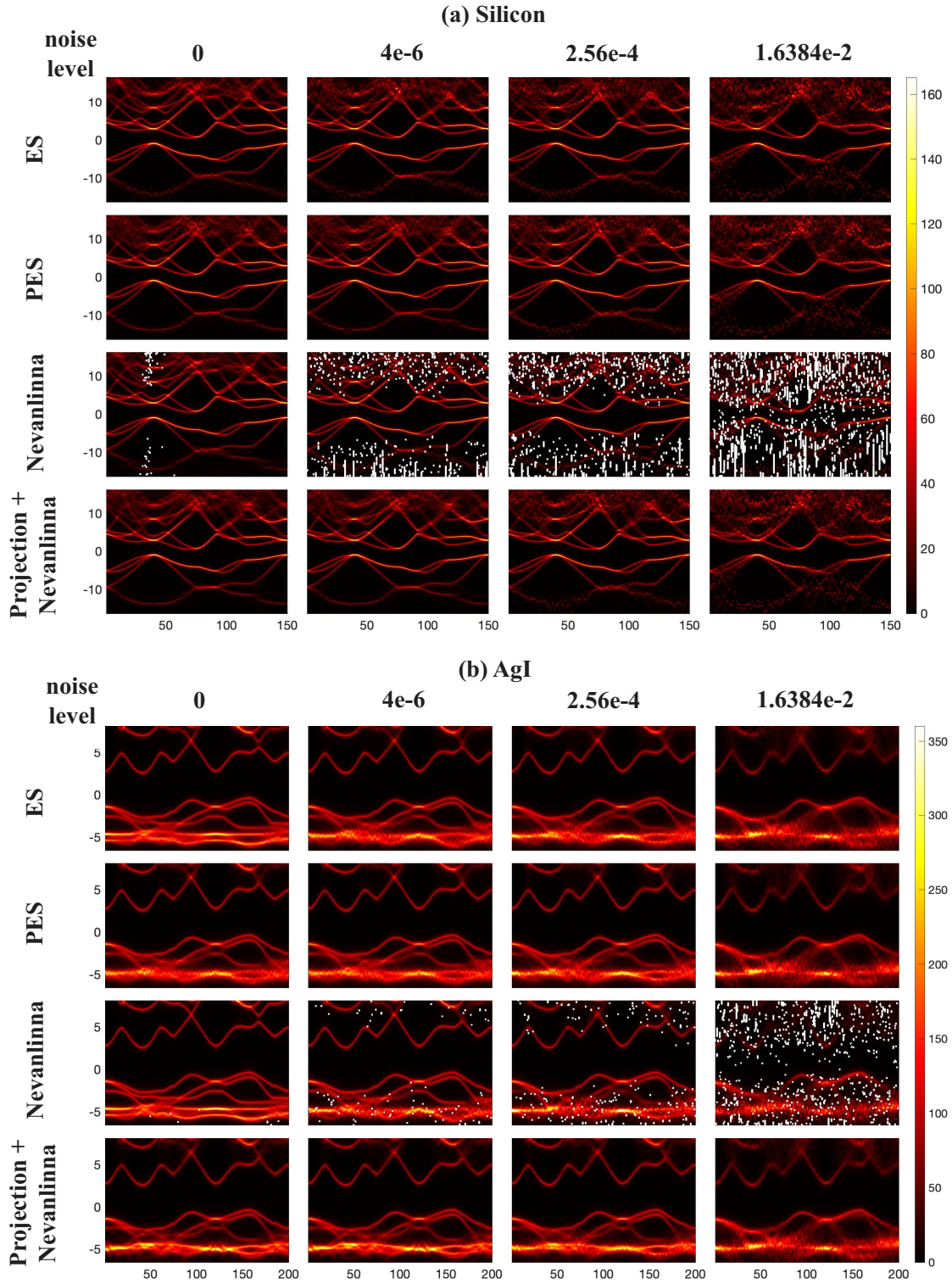


Figure 5. Band structure of Silicon and AgI, revealed by the analytic continuation. Data from Ref.[54]. White entries correspond to negative (non-causal) spectral functions.

tions could result in negative spectral functions even for the clean data, while Nevanlinna continuations combined with the causal projection can largely resolve the issue of noise-sensitivity. PES method, ES method and the

Nevanlinna continuation on projected data routine are all noise-robust. Even at a relatively large noise-level $\sigma = 10^{-2}$, the structure of the valence and conduction bands can be recovered. This means that our noise-

robust analytic continuation methods could successfully retrieve information such as the band gap from Matsubara data even under a significant noise level.

IV. DISCUSSION

In this article, we have proposed a robust analytic continuation method called PES, consisting of the following three steps: causal projection, pole estimation and SDR fitting through bi-level optimization. The PES method is noise-robust, could deal with matrix-valued Green's functions and is applicable to both fermionic and bosonic systems. As a by product, we also find that the causal projection could significantly improve the performance of noise-sensitive methods such as Nevanlinna and Carathéodory continuation. We have also demonstrated these properties with extensive numerical experiments.

Our current strategy of pole estimation is to perform AAA algorithm on the projected data. Note that the AAA algorithm is still an interpolation procedure can be susceptible to noise. The causal projection alleviates this drawback, and in practice it often provides a good enough initial guess for SDR fitting even in the noisy scenario. A better strategy to estimate the pole locations (e.g., by combining with the Prony method [33]) may still be helpful in improving the overall performance of the PES method. Moreover, the current SDP solver (SDPT3) can be computationally expensive when the number of variables is large. In such a case, more efficient solvers based on first-order methods may reduce the computa-

tional time in both the projection and the SDR fitting steps.

As we have shown, the PES method is useful for retrieving the excitation spectra of molecules and band structures in solids. These data feature multiple sharp peaks modeled by poles on the real axis. It could also be used to perform bath fitting in the Dynamical Mean Field Theory (DMFT) [35], and it is applicable in lattice gauge theory (for example, see [62, 63]). Broadened spectral functions that cannot be resolved into discrete peaks may be modeled by a relatively small number of poles away from the real axis. This would require modification of the definition of the causal space. Note that in the noisy scenario, it may be very difficult to distinguish the spectra described by a pole in the complex plane and the spectra due to many poles on or near the real axis. Therefore prior information such as the maximum number of complex poles and approximate pole locations may become necessary. This scenario will be investigated in future work.

Acknowledgment

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of Basic Energy Sciences, Scientific Discovery through Advanced Computing (SciDAC) program under Award Number DE-SC0022198 (Z.H.) and DE-SC0022088 (E.G.). The work is partially supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0095 (L.L.). L.L. is a Simons Investigator. We thank Jiahao Yao, Lexing Ying, Yang Yu and Dominika Zgid for helpful discussions.

-
- [1] G. A. Baker, G. A. Baker Jr, P. Graves-Morris, G. Baker, and S. S. Baker, *Pade Approximants: Encyclopedia of Mathematics and It's Applications, Vol. 59 George A. Baker, Jr., Peter Graves-Morris*, vol. 59 (Cambridge University Press, 1996).
- [2] J. Schött, I. L. Locht, E. Lundin, O. Grånäs, O. Eriksson, and I. Di Marco, *Physical Review B* **93**, 075104 (2016).
- [3] M. Jarrell and J. E. Gubernatis, *Physics Reports* **269**, 133 (1996).
- [4] M. Asakawa, Y. Nakahara, and T. Hatsuda, *Progress in Particle and Nuclear Physics* **46**, 459 (2001).
- [5] R. Levy, J. LeBlanc, and E. Gull, *Computer Physics Communications* **215**, 149 (2017).
- [6] R. Bryan, *European Biophysics Journal* **18**, 165 (1990).
- [7] A. Rothkopf, *Data* **5**, 85 (2020).
- [8] A. W. Sandvik, *Physical Review B* **57**, 10287 (1998).
- [9] K. Vafayi and O. Gunnarsson, *Physical Review B* **76**, 035115 (2007).
- [10] S. Fuchs, T. Pruschke, and M. Jarrell, *Physical Review E* **81**, 056701 (2010).
- [11] H. Shao and A. W. Sandvik, *Progress on stochastic analytic continuation of quantum monte carlo data* (2022), 2202.09870.
- [12] K. Yoshimi, J. Otsuki, Y. Motoyama, M. Ohzeki, and H. Shinaoka, *Computer Physics Communications* **244**, 319 (2019).
- [13] Y. Motoyama, K. Yoshimi, and J. Otsuki, *Physical Review B* **105**, 035139 (2022).
- [14] R. Fournier, L. Wang, O. V. Yazyev, and Q. Wu, *Physical Review Letters* **124**, 056401 (2020).
- [15] H. Yoon, J.-H. Sim, and M. J. Han, *Physical Review B* **98**, 245101 (2018).
- [16] J. M. Tomczak and S. Biermann, *Journal of Physics: Condensed Matter* **19**, 365206 (2007).
- [17] H. T. Dang, X. Ai, A. J. Millis, and C. A. Marianetti, *Physical Review B* **90**, 125114 (2014).
- [18] E. Gull and A. J. Millis, *Physical Review B* **90**, 041110 (2014).
- [19] G. J. Kraberger, R. Triebl, M. Zingl, and M. Aichhorn, *Physical Review B* **96**, 155128 (2017).
- [20] A. Reymbaut, A.-M. Gagnon, D. Bergeron, and A.-M. S. Tremblay, *Phys. Rev. B* **95**, 121104 (2017).
- [21] I. S. Kac and M. G. Krein, *Amer. Math. Soc. Transl* **103**, 1–18 (1974).
- [22] S. Belyi, S. Hassi, H. de Snoo, and E. Tsekanovskii, *Linear algebra and its applications* **419**, 331 (2006).
- [23] J. Fei, C.-N. Yeh, and E. Gull, *Physical Review Letters* **126**, 056402 (2021).

- [24] J. Fei, C.-N. Yeh, D. Zgid, and E. Gull, *Physical Review B* **104**, 165111 (2021).
- [25] G. Pick, *Mathematische annalen* **78**, 270 (1917).
- [26] P. Khargonekar and A. Tannenbaum, *IEEE Transactions on Automatic Control* **30**, 1005 (1985).
- [27] A. Tannenbaum, in *26th IEEE Conference on Decision and Control* (IEEE, 1987), vol. 26, pp. 1635–1636.
- [28] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback control theory* (Courier Corporation, 2013).
- [29] M. R. Osborne and G. K. Smyth, *SIAM J. Sci. Comput.* **16**, 119 (1995).
- [30] G. Golub and V. Pereyra, *Inverse problems* **19**, R1 (2003).
- [31] G. Beylkin and L. Monzón, *Applied and Computational Harmonic Analysis* **19**, 17 (2005).
- [32] L. Ying, arXiv preprint arXiv:2202.09719 (2022).
- [33] L. Ying, arXiv preprint arXiv:2202.02670 (2022).
- [34] Y. Nakatsukasa, O. Sète, and L. N. Trefethen, *SIAM Journal on Scientific Computing* **40**, A1494 (2018).
- [35] C. Mejuto-Zaera, L. Zepeda-Núñez, M. Lindsey, N. Tubman, B. Whaley, and L. Lin, *Physical Review B* **101**, 035143 (2020).
- [36] J. W. Negele and H. Orland, *Quantum many-particle systems* (Westview, 1988).
- [37] F. Gesztesy and E. Tsekanovskii, *Mathematische Nachrichten* **218**, 61 (2000).
- [38] J. A. Shohat and J. D. Tamarkin, *The problem of moments*, vol. 1 (American Mathematical Society (RI), 1950).
- [39] N. I. Akhiezer, *The classical moment problem and some related questions in analysis* (SIAM, 2020).
- [40] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1* (2014).
- [41] R. Tibshirani, *J. Royal Stat. Soc. B* **58**, 267 (1996).
- [42] E. J. Candes, J. K. Romberg, and T. Tao, *Communications on Pure and Applied Mathematics* **59**, 1207 (2006).
- [43] Bach, Francis and Jenatton, Rodolph and Mairal, Julien and Obozinski, Guillaume (2012).
- [44] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun guide* (2014).
- [45] J.-P. Berrut and L. N. Trefethen, *SIAM Rev.* **46**, 501 (2004).
- [46] R. P. Feynman, *Physical review* **56**, 340 (1939).
- [47] J. Nocedal and S. J. Wright, *Numerical optimization* (Springer Verlag, 1999).
- [48] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, in *Handbook on semidefinite, conic and polynomial optimization* (Springer, 2012), pp. 715–754.
- [49] S. Mehrotra, *SIAM Journal on optimization* **2**, 575 (1992).
- [50] L. Vandenberghe and S. Boyd, *SIAM Rev.* **38**, 49 (1996).
- [51] Jiang, Kaifeng and Sun, Defeng and Toh, Kim-Chuan, *SIAM J. Optim.* **22**, 1042 (2012).
- [52] Y. Nakatsukasa and L. N. Trefethen, *SIAM Journal on Scientific Computing* **42**, A3157 (2020).
- [53] C.-N. Yeh, A. Shee, Q. Sun, E. Gull, and D. Zgid (2022).
- [54] C.-N. Yeh, S. Isakov, D. Zgid, and E. Gull, *Fully self-consistent finite-temperature gw in gaussian bloch orbitals for solids* (2022).
- [55] J. VandeVondele and J. Hutter, *The Journal of Chemical Physics* **127**, 114105 (2007), <https://doi.org/10.1063/1.2770708>.
- [56] S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- [57] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, et al., *The Journal of Chemical Physics* **153**, 024109 (2020), <https://doi.org/10.1063/5.0006074>.
- [58] A. Gaenko, A. Antipov, G. Carcassi, T. Chen, X. Chen, Q. Dong, L. Gamper, J. Gukelberger, R. Igarashi, S. Isakov, et al., *Computer Physics Communications* **213**, 235 (2017), ISSN 0010-4655.
- [59] S. Isakov, C.-N. Yeh, E. Gull, and D. Zgid, *Phys. Rev. B* **102**, 085105 (2020).
- [60] H. Shinaoka, J. Otsuki, M. Ohzeki, and K. Yoshimi, *Phys. Rev. B* **96**, 035147 (2017).
- [61] J. Li, M. Wallerberger, N. Chikano, C.-N. Yeh, E. Gull, and H. Shinaoka, *Phys. Rev. B* **101**, 035144 (2020).
- [62] Y. Burnier and A. Rothkopf, *Phys. Rev. Lett.* **111**, 182003 (2013).
- [63] R.-A. Tripolt, P. Gubler, M. Ulybyshev, and L. von Smekal, *Computer Physics Communications* **237**, 129 (2019), ISSN 0010-4655.