Reconstructing lattice QCD spectral functions with stochastic pole expansion and Nevanlinna analytic continuation

Li Huang[1,*] and Shuang Liang[2,3]

[1]*Science and Technology on Surface Physics and Chemistry Laboratory,*
P.O. Box 9-35, Jiangyou 621908, China
[2]*Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[3]*Yangtze River Delta Physics Research Center, Liyang, Jiangsu 213300, China*

The reconstruction of spectral functions from Euclidean correlation functions is a well-known, yet ill-conditioned inverse problem in the fields of many-body and high-energy physics. In this paper, we present a comprehensive investigation of two recently developed analytic continuation methods, namely stochastic pole expansion and Nevanlinna analytic continuation, for extracting spectral functions from mock lattice QCD data. We examine a range of Euclidean correlation functions generated by representative models, including the Breit-Wigner model, the Gaussian mixture model, the resonance-continuum model, and the bottomonium model. Then, we apply the two methods to reconstruct spectral functions from charmonium correlation functions computed using lattice QCD simulations at a finite temperature. Our findings demonstrate that the stochastic pole expansion method, when combined with the constrained sampling algorithm and the self-adaptive sampling algorithm, successfully recovers the essential features of the spectral functions and exhibits excellent resilience to noise of input data. In contrast, the Nevanlinna analytic continuation method suffers from numerical instability, often resulting in the emergence of spurious peaks and significant oscillations in the high-energy regions of the spectral functions, even with the application of the Hardy basis function optimization algorithm.

I. INTRODUCTION

Lattice QCD (LQCD) is a well-established first-principles and nonperturbative approach for studying strong interactions [1–3]. It serves as a valuable tool in understanding the genesis and evolution of the quark-gluon plasma (QGP) [4] and mapping out the phase diagram of strong-interaction matter [5–7]. In LQCD, spectral functions play a vital role in scrutinizing and elucidating high-energy physical phenomena that involve quarks and gluons, such as the melting of heavy quarkonium [8–15] and the transport properties [16–18] of the QGP formed through relativistic heavy-ion collisions. However, accessing the spectral functions and other dynamical properties of the QCD medium from lattice simulations remains challenging due to LQCD's typical formulation on a discrete Euclidean space-time grid [1–3]. Therefore, researchers must reconstruct the spectral functions from numerically

computed Euclidean correlation functions on the lattice to understand the relevant physics and compare the theoretical results with corresponding experimental data obtained from the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC).

Mathematically, the Euclidean correlators $G(t)$ and the correspondingly spectral functions $\rho(s)$ are connected through a Fredholm equation of the first kind: $G(t) = K(t, s) \circledast \rho(s)$, where $K(t, s)$ represents a continuous kernel function and $\circledast$ signifies convolution. Mapping the spectral functions to the Euclidean correlators is a straightforward process that can be easily accomplished using numerical integration. Nevertheless, extracting spectral functions from Euclidean correlators through numerical analytic continuation poses a formidable challenge [19]. We observe that similar inverse problems are quite common in many-body and high-energy physics [1,20]. They are considered ill conditioned. There are two main reasons for this statement. First, the Euclidean correlators are evaluated at a finite number of points due to the space-time discretization in LQCD [1–3]. Secondly, as LQCD simulations rely on stochastic Monte Carlo sampling, the resulting Euclidean correlators are inherently noisy [21]. Even small deviations or fluctuations in the Euclidean correlators result in significant uncertainties in the spectral functions. As a result, the

majority of resulting spectral functions exhibit high oscillations and lack physical significance, thereby inhibiting a reliable comparison between the theoretical spectra and experimental data. To address these issues, researchers have developed a plethora of analytic continuation approaches in the past decades. Next, we will briefly introduce some of the most commonly employed methods in LQCD simulations.

Maximum entropy method. The maximum entropy method (MaxEnt) is perhaps the most popular analytic continuation tool, and it has dominated this field for a long time. In this method, the spectral density is interpreted as a probability distribution [22–25]. The primary objective is to extract the most probable spectral density ρ from the correlation function G and maximize the posterior probability $\Pr[\rho|G]$. According to Bayes's theorem, $\Pr[\rho|G] \propto \Pr[G|\rho] \Pr[\rho]$, where $\Pr[G|\rho]$ is the likelihood function and $\Pr[\rho]$ is the prior probability. It is important to incorporate analytical knowledge related to spectral properties in LQCD, such as positive definiteness or even the presence of pole structures within the spectra, into the probability distribution. A significant portion of the prior information could be encoded within the prior probability $\Pr[\rho]$, which is proportional to $\exp(\alpha S)$, where α is a regulation parameter and S denotes entropy. It is worth emphasizing that the entropic term αS is not unique and typically takes the form of the generalized Shannon-Jaynes entropy [22,23]. Another popular alternative is the Bayesian reconstruction entropy [25,26]. While the MaxEnt method is widely recognized for its efficiency and noise-tolerance, it sometimes struggles to faithfully recover sharp, subtle, and high-frequency features within the spectral functions.

Stochastic analytic continuation. In the past decade, the stochastic analytic continuation method (SAC) and its variants have emerged as formidable contenders to surpass the MaxEnt method [27–37]. Unlike the MaxEnt method, the SAC method treats all spectral functions equally instead of selecting the most probable one. Initially, the spectral functions are parametrized with hundreds or thousands of δ-like functions. And then these parameters, such as the amplitudes and locations of the δ functions, are stochastically sampled at a fictitious temperature Θ using a Boltzmann-like weight function, which essentially serves as a likelihood function [27,28]. Finally, the gathered spectral functions are filtered and averaged. We note that Shao and Sandvik have proven the equivalence in a generalized thermodynamic limit (large number of degrees of freedom) of the average spectrum and the maximum entropy solution [38]. Although the SAC method supplements the MaxEnt method by enabling the resolution of subtle structures in spectra, it requires significant computational resources [32].

Machine learning approaches. In recent years, several machine learning aided methods have been developed to address the analytic continuation challenges in LQCD simulations. These methods include the deep neural networks (DNN) [39,40], radial basis functions network (RBFN) [41], entropy variational autoencoder (SVAE) [42], kernel ridge regression (KRR) [43], automatic differentiation (AD) [44], Gaussian processes regression (GPR) [45], and many others. Although these methods may offer improved performance in certain situations, their universality is not guaranteed. Furthermore, some studies have adopted supervised approaches to train machine learning network models, which incorporate prior knowledge from specific physics insights into the training sets. However, caution must be exercised as there is a risk of introducing biases in the training data.

Recently, two new analytic continuation methods, namely the stochastic pole expansion (SPX) [46] and the Nevanlinna analytic continuation (NAC) [47,48], have been proposed. The SPX method inherits the spirit of the SAC method, where the Matsubara Green's function is initially parametrized with hundreds or thousands of poles. Subsequently, the amplitudes and positions of these poles are optimized using a stochastic algorithm that based on stimulated annealing [49]. The SPX method is applicable to fermionic and bosonic systems. It has been extended to support analytic continuation of matrix-valued Green's functions [46]. On the other hand, similar to the Padé approximation (PA) [50–53], the NAC method aims to interpolate the Matsubara data in the complex plane using some form of continued fraction expansion [47,48]. It takes the "Nevanlinna" analytic structure of the Matsubara Green's function into consideration, ensuring that the calculated spectral functions are inherently positive and normalized. However, this method is highly sensitive to the noise level in the raw Matsubara data. With noiseless data as input, it can successfully resolve complex spectral functions across a wide energy range with unprecedented accuracy. Unfortunately, when the Matsubara data contains noise, the Nevanlinna interpolants may not exist, and the resulting spectral functions are not guaranteed to be causal.

Both the SPX and NAC methods have not yet been employed in addressing the issue of analytic continuation in the LQCD simulations, and it remains uncertain whether they are applicable for the analytic continuation of Euclidean correlation functions. Therefore, the purpose of this study is to fill this knowledge gap and expand the potential applications of these methods. We at first generate noisy Euclidean data using four representative models: the Breit-Wigner model, the Gaussian mixture model, the resonance-continuum model, and the bottomonium model. These synthetic datasets are then processed using the SPX, NAC, and MaxEnt methods. We also test these methods using the realistic LQCD data, namely the charmonium correlation functions. Finally, we conduct a comprehensive comparison between the calculated spectra and the exact solutions if available. The results suggest that the SPX method manifests comparable or even superior performance in comparison to the commonly used MaxEnt

method. On the other hand, the NAC method tends to suffer from numerical instability, even in the absence of noise in the input data.

The structure of the remaining sections of this paper is as follows: Sec. II provides an introduction to the basic formulations of Euclidean correlation functions and offers a brief overview of the SPX and NAC methods. In Sec. III, we present the computational setups, demonstrate, and discuss five representative examples. Section IV explores the robustness of the SPX method and numerical instability of the NAC method in the presence of noisy Euclidean data. Additionally, we analyze the effects of the constrained sampling algorithm and the self-adaptive sampling algorithm on the SPX method, along with the impact of the Hardy basis function optimization algorithm on the NAC method. Finally, our findings are summarized in Sec. V.

II. METHOD

A. Euclidean correlation function

At finite temperature, the Euclidean correlation function $G(\tau)$ is related to the spectral function $\rho(\omega)$ through [19]

$$G(\tau) = \int_0^\infty d\omega\, \rho(\omega) \frac{\cosh[\omega(\tau - \beta/2)]}{\sinh(\beta\omega/2)}. \quad (1)$$

Here $\beta = 1/T$ represents the inverse system temperature, and τ represents the Euclidean (imaginary) time interval ($\tau \in [0, \beta]$). In momentum space, the expression for the Euclidean correlation function is [54]

$$G(p) = \int_0^\infty d\omega\, \rho(\omega) \frac{\omega}{\omega^2 + p^2}, \quad (2)$$

where p represents the Euclidean (Matsubara) frequency, the value of $G(p)$ can be derived by performing a discrete Fourier transform on $G(\tau)$. In the references, Eq. (2) is also known as the Källén-Lehmann (KL) spectral representation [55]. By applying the process of analytic continuation, the retarded propagator $G^R(\omega)$ can be attained, enabling the extraction of the spectral function through the following expression:

$$\rho(\omega) = -\frac{1}{\pi} \mathrm{Im} G^R(\omega), \quad (3)$$

where $\omega = -ip$. It is important to note that $\rho(\omega)$ must be an odd function for bosonic system, i.e., $\rho(-\omega) = -\rho(\omega)$.

B. Stochastic pole expansion

According to the textbooks of many-body physics, the Lehmann representation of the finite temperature many-body Green's functions is given by the following formula [20,56]:

$$G(z) = \frac{1}{Z} \sum_{m,n} \frac{\langle n|d|m\rangle \langle m|d^\dagger|n\rangle}{z + E_n - E_m} \left(e^{-\beta E_n} \pm e^{-\beta E_m}\right). \quad (4)$$

In this expression, d and $d^\dagger$ represent the annihilation and creation operators, respectively. $|n\rangle$ and $|m\rangle$ are the eigenstates of the Hamiltonian $\hat{H}$, and E_n and E_m are the corresponding eigenvalues. $Z = \sum_n \exp(-\beta E_n)$ is the partition function. $z \in \mathbb{C} \backslash \mathbb{R}$. The positive sign corresponds to fermions, while the negative sign corresponds to bosons. By introducing $A_{mn} = \langle n|d|m\rangle \langle m|d^\dagger|n\rangle (e^{-\beta E_n} \pm e^{-\beta E_m})/Z$ and $P_{mn} = E_m - E_n$, Eq. (4) can be simplified as:

$$G(z) = \sum_{m,n} \frac{A_{mn}}{z - P_{mn}}. \quad (5)$$

It is evident that only terms where $A_{mn} \neq 0$ can be present. The indices m and n can also be compressed as γ, resulting in the following expression:

$$G(z) = \sum_{\gamma=1}^{N_p} \frac{A_\gamma}{z - P_\gamma}. \quad (6)$$

Equation (6) is referred to as the *pole representation* of the many-body Green's functions [20]. In this representation, N_p denotes the number of poles, and A_γ and P_γ mean the amplitude and position of the γth pole. For the Euclidean correlation in momentum space, its pole representation can be reformulated as follows:

$$G(p) = \sum_{\gamma=1}^{N_p} \Xi(p, P_\gamma) \tilde{A}_\gamma. \quad (7)$$

Here, Ξ represents the kernel matrix, which is calculated using the following equation:

$$\Xi(p, \omega) = -\frac{G(0)\omega}{p - \omega}. \quad (8)$$

$\tilde{A}_\gamma$ is the renormalized amplitude of the γth pole, given by:

$$\tilde{A}_\gamma = -\frac{A_\gamma}{G(0)P_\gamma}. \quad (9)$$

It can be easily proven that $\tilde{A}_\gamma$ and P_γ must satisfy the following constraints:

$$\forall\, \gamma,\ \ 0 \leq \tilde{A}_\gamma \leq 1,\ \ \sum_\gamma \tilde{A}_\gamma = 1,\ \ \text{and}\ \ P_\gamma \in \mathbb{R}. \quad (10)$$

We assume that the input Euclidean correlation function is denoted as $\mathcal{G}(p_n)$, and the input data consists of N frequency points. We then utilize Eq. (7) to approximate the Euclidean data. To assess the discrepancy between $\mathcal{G}(p_n)$

and $G(p_n)$, we introduce the so-called goodness-of-fit function χ^2. Its definition is as follows:

$$\chi^2\left[\{\tilde{A}_\gamma, P_\gamma\}_{\gamma=1}^{N_p}\right] = \frac{1}{N}\sum_{n=1}^{N}\left\|\mathcal{G}(p_n) - \sum_{\gamma=1}^{N_p}\Xi(p_n, P_\gamma)\tilde{A}_\gamma\right\|_F^2,$$
(11)

where $\|\cdot\|_F$ represents the Frobenius norm. Hence, the objective of the analytic continuation is to solve the subsequent multivariate optimization problem:

$$\arg\min_{\{\tilde{A}_\gamma, P_\gamma\}_{\gamma=1}^{N_p}} \chi^2\left[\{\tilde{A}_\gamma, P_\gamma\}_{\gamma=1}^{N_p}\right].$$
(12)

Once the optimized parameters N_p, $\tilde{A}_\gamma$, and P_γ are determined, evaluating the retarded Green's function is straightforward by substituting p with $\omega + i0^+$ in Eq. (7). Additionally, the spectral function $\rho(\omega)$ is computed using Eq. (3). It is important to note that this optimization problem [i.e., Eq. (12)] is highly nonconvex. Traditional gradient-based optimization methods typically fail to identify the global minimum unless the initial solution is of high quality [57]. Therefore, in the SPX method, we employ the simulated annealing algorithm [49] to optimize the $\tilde{A}_\gamma$ and P_γ parameters subject to the constraints defined by Eq. (10). For technical details regarding the possible Monte Carlo random walking rules in the configuration space $\mathcal{C} = \{\tilde{A}_\gamma, P_\gamma\}$, please refer to Ref. [46]. The advantages of the SPX method include its ability to derive approximate expressions for correlation functions and its ease of extension to support the analytic continuation of bosonic systems, two-particle Green's functions, matrix-valued Green's functions, and so on. Applications to noisy Matsubara data suggest that the SPX method can accurately resolve both continuum spectra for condensed matter cases and multiple δ-like peaks for molecule cases. Notably, it performs well in reproducing sharp high-frequency features.

C. Nevanlinna analytic continuation

It is well known that the retarded Green's function, denoted as $G^R(\omega + i0^+)$, and the Matsubara Green's function, denoted as $G(i\omega_n)$, can both be consistently represented as $G(z)$, where $z \in \mathbb{C}\backslash\mathbb{R}$. The NAC method utilizes the fact that the negative fermionic Green's function, denoted as $f(z) = -G(z)$, belongs to the class of Nevanlinna functions. By applying the invertible Möbius transform $h(z) = (z-i)/(z+i)$ to the function value of $f(z)$, the Nevanlinna function is mapped in a one-to-one fashion to a contractive function $\theta(z) = h[f(z)]$. This contractive function $\theta(z)$ can be expressed in the form of a continued fraction expansion, and an iterative algorithm can be constructed accordingly [47]. The recursion relation between two steps $\theta_j(z)$ and $\theta_{j+1}(z)$ is given by:

$$\theta_j(z) = \frac{\theta_{j+1}(z) + \gamma_j}{\gamma_j^* h_j(z)\theta_{j+1}(z) + 1}.$$
(13)

In this equation, $h_j(z) = (z - Y_j)/(z + Y_j)$, $Y_j = i\omega_j$ represents the jth Matsubara frequency used, and $\gamma_j = \theta_j(Y_j)$ represents the function value of the jth contractive function at the point Y_j. The final expression of the recursive function $\theta(z)$ can be written as [58]:

$$\theta(z)[z; \theta_{N_s+1}(z)] = \frac{a(z)\theta_{N_s+1}(z) + b(z)}{c(z)\theta_{N_s+1}(z) + d(z)},$$
(14)

where

$$\begin{pmatrix} a(z) & b(z) \\ c(z) & d(z) \end{pmatrix} = \prod_{j=1}^{N_s}\begin{pmatrix} h_j(z) & \gamma_j \\ \gamma_j^* h_j(z) & 1 \end{pmatrix},$$
(15)

with j increasing from left to right. Here N_s is the overall iteration step, which is equivalent to the number of data points. After obtaining $\theta(z)$, one can immediately get the Green's function by an inverse Möbius transform as $G(z) = -h^{-1}[\theta(z)]$. Note that the Pick criterion [59] should be fulfilled for the existence of the Nevanlinna interpolation.

Additionally, it is worth noting that there is flexibility in choosing $\theta_{N_s+1}(z)$, which can be used to select the most desirable spectral function. In Ref. [47], $\theta_{N_s+1}(z)$ is expanded in the Hardy basis and chosen in such a way that it achieves the smoothest possible spectral function [60]. The loss function employed in this selection process is given by:

$$\mathcal{L} = \left|1 - \int\frac{d\omega}{2\pi}\rho_{\theta_{N_s+1}}(\omega)\right|^2 + \lambda\left\|\frac{d^2\rho_{\theta_{N_s+1}}(\omega)}{d\omega^2}\right\|_F^2.$$
(16)

This loss function consists of two terms. The first term enforces the proper sum rule, while the second term incorporates the smoothness condition. λ is an adjustable parameter. By preserving the "Nevanlinna" analytic structure of Green's functions, the NAC method automatically generates positive and normalized spectral functions [47]. However, it is important to emphasize that the method is sensitive to noise, and either a large number of data points N or a high Hardy order H can potentially lead to numerical instabilities.

Although the NAC method has been extended to support the analytic continuation of matrix-valued Green's functions [48], it cannot be directly applied to bosonic systems in its original formalism [61]. Quite recently, Nogaki *et al.* suggest an ingenious trick to work around this limitation. Their basic idea is to introduce an auxiliary fermionic function [62]. Let us start with a bosonic Green's function $G(\tau)$ that satisfies the periodic condition $G(\tau + \beta) = G(\tau)$.

One can construct an artificial antiperiodic fermionic Green's function $\tilde{G}(\tau)$ as follows:

$$\tilde{G}(\tau) = \begin{cases} G(\tau) & (0 < \tau < \beta) \\ -G(\tau + \beta) & (-\beta < \tau < 0) \end{cases} \quad (17)$$

Clearly, this auxiliary fermionic Green's function exhibits the same value as the bosonic Green's function in the range $0 < \tau < \beta$. It is easy to prove that the relation between the bosonic spectral function $\rho(\omega)$ and the auxiliary fermionic spectral function $\tilde{\rho}(\omega)$ is as follows:

$$\rho(\omega) = \tilde{\rho}(\omega) \tanh(\beta\omega/2). \quad (18)$$

Furthermore, the sum rule for $\tilde{\rho}(\omega)$ is given by:

$$\int_{-\infty}^{\infty} d\omega\, \tilde{\rho}(\omega) = G(\tau = 0^+) + G(\tau = \beta - 0^-). \quad (19)$$

Given $\tilde{G}(\tau)$, it is easy to construct $\tilde{G}(i\nu_n)$ via direct Fourier transformation, where $\nu_n = (2n + 1)\pi/\beta$ are the fermionic Matsubara frequencies. Since

$$\tilde{G}(i\nu_n) = \int_{-\infty}^{\infty} d\omega \frac{\tilde{\rho}(\omega)}{i\nu_n - \omega}, \quad (20)$$

one can perform analytic continuation for $\tilde{G}(i\nu_n)$ via the standard NAC method to get $\tilde{\rho}(\omega)$. And then the bosonic spectral function $\rho(\omega)$ can be derived according to Eq. (18). This procedure has been outlined in Fig. 1 in Ref. [62].

III. BENCHMARKS

A. Computational setups

To benchmark the SPX and NAC methods, we consider five typical models, namely the Breit-Wigner model, the Gaussian mixture model, the resonance-continuum model, the bottomonium model, and the charmonium correlation function in the present investigation. The first three models provide analytic formulas for generating the exact spectral functions, denoted as $\rho(\omega)$. For the bottomonium model, we take the "exact" spectral function from Ref. [63] as input. Using these spectral functions, one can create clean Euclidean data, denoted as $\mathcal{G}_{\text{clean}}$, using Eq. (2) [64]. The charmonium correlation function is extracted from Ref. [15]. Note that the raw data is in time domain. Thus, they have to be interpolated into a denser grid and then converted into frequency domain using Fourier transformation. To mimic the noise present in LQCD simulations [21], we manually add multiplicative Gaussian noises to the clean Euclidean data. The formula we use is as follows:

$$\mathcal{G}_{\text{noisy}} = \mathcal{G}_{\text{clean}}[1 + \delta N_{\mathbb{C}}(0, 1)], \quad (21)$$

where δ measures the noise level of the input data and $N_{\mathbb{C}}(0, 1)$ represents complex-valued normal Gaussian noise [65]. In our subsequent analysis, unless explicitly stated, we set $\delta = 10^{-4}$ for the SPX method and $\delta = 0.0$ for the NAC method. We then supply the noisy Euclidean data, denoted as $\mathcal{G}_{\text{noisy}}$, into the SPX and NAC codes to extract the spectral functions. Finally, we compare the calculated spectral functions with the corresponding exact solutions.

The SPX method has been implemented within the ACFlow package [66]. In this study, the number of poles (N_p) is fixed to 2000. For each test, we perform a total of 2×10^3 individual SPX runs. Each SPX run consists of 2×10^5 Monte Carlo sampling steps. The spectral functions generated in all SPX runs are gathered and the corresponding χ^2 values are recorded. Assuming the mean value of the collected χ^2 is denoted as $\langle\chi^2\rangle$, we only retain the solutions whose χ^2 values are smaller than $\langle\chi^2\rangle/\alpha_{\text{good}}$, and collect them to calculate the averaged spectrum. Here, α_{good} is an adjustable parameter ($\alpha_{\text{good}} \geq 1.0$). Its optimal value is about 1.2. In order to obtain smoother spectra, one can further increase the number of Monte Carlo sampling steps and number of indvidual SPX runs, or decrease α_{good}.

Regarding the NAC method, we utilize another open-source toolkit, namely the Nevanlinna.jl package [67]. In the present calculations, the lowest 100 Matsubara frequencies are kept as input. In order to avoid breaking the Pick criterion [59], the optimal number of data points is automatically determined, which is denoted as N_{opt}, through a "Pick selection" procedure in the algorithm. Usually a typical value for N_{opt} is 10. During the simulated process, the Hardy basis function optimization algorithm is always enabled. The highest Hardy order, denoted as H_{max}, is set to be 50. To ensure numerical stability, the cutoff value of Hardy order, denoted as H_{cut}, should be determined automatically on a case-by-case basis. The λ parameter, as seen in Eq. (16), is set to be 10^{-6}.

In addition to the SPX and NAC methods, the classic MaxEnt method [24] is also employed to yield analytic continuation results for comparison. We again utilize the ACFlow package, which provides a state-of-the-art implementation of the MaxEnt method [66]. For the MaxEnt simulations, we usually choose a flat default model and calibrate the regularization parameter α by using the χ^2-kink algorithm [68]. The starting and ending values of α are 10^{16} and 10^1, respectively. The ratio between two successive α parameters, i.e., α_i/α_{i+1}, is 10. For the prior probability (the entropic term), we adopt the generalized Shannon-Jaynes entropy [22,23], but the Bayesian reconstruction entropy [25,26] is also examined. The simulated results are in good agreement with each other. Thus, we only present results obtained with the generalized Shannon-Jaynes entropy in the following discussions.
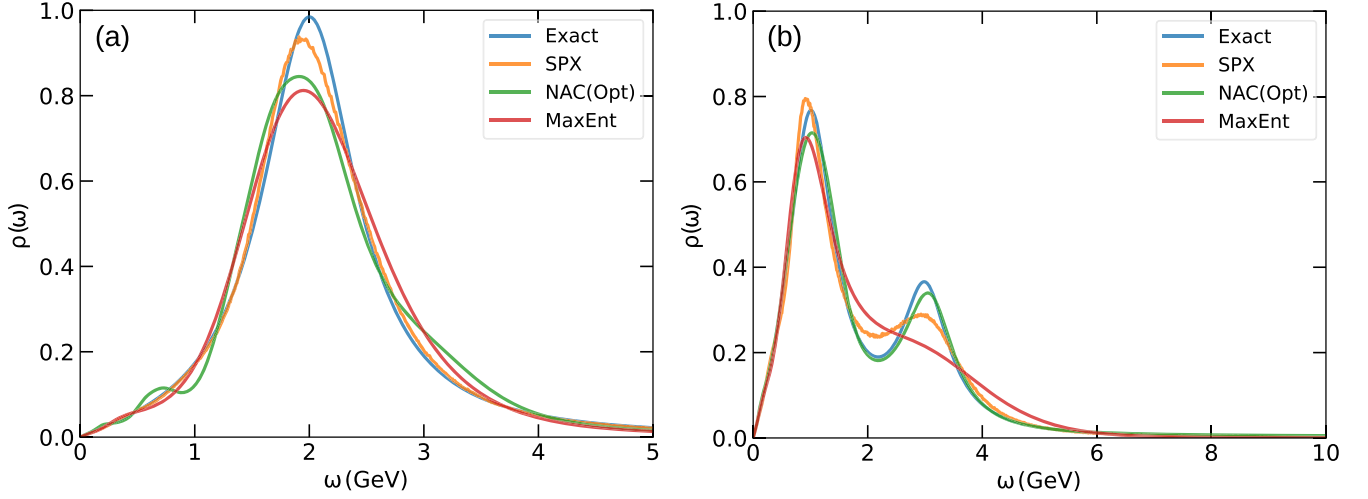
FIG. 1. Analytic continuations of the Breit-Wigner models. For the NAC method, the Hardy basis function optimization algorithm is enabled. (a) Single Breit-Wigner peak. $N_{\text{opt}} = 13$. (b) Two Breit-Wigner peaks. $N_{\text{opt}} = 14$. In panel (b) the spectra are scaled by a factor of 0.5 for a better view.

B. Breit-Wigner model

The Breit-Wigner spectral function, obtained from a parametrization derived directly from one-loop perturbative quantum field theory [19,39], is expressed as follows:

$$\rho(\omega) = \frac{4A\Gamma\omega}{(M^2 + \Gamma^2 - \omega^2)^2 + 4\Gamma^2\omega^2}, \quad (22)$$

where M represents the mass of the corresponding state, Γ is the width, and A is a positive constant. We start with a superposed collection of Breit-Wigner peaks. Specially, two typical scenarios are investigated in the present work: (1) Single Breit-Wigner peak (dubbed 1BW model) with $M = 2.0$ GeV, $\Gamma = 0.5$ GeV, and $A = 1.0$ GeV. (2) Two Breit-Wigner peaks (dubbed 2BW model) with $M_1 = 1.0$ GeV, $M_2 = 3.0$ GeV, $\Gamma_1 = \Gamma_2 = 0.5$ GeV, $A_1 = 0.8$ GeV, and $A_2 = 1.0$ GeV. The system temperature T is fixed to be 0.02 GeV. The synthetic Euclidean data ($\mathcal{G}_{\text{noisy}}$) consists of 50 frequency points for both the SPX method and the MaxEnt method, and 100 frequency points for the NAC method. The analytical continuation results obtained by the SPX, NAC, and MaxEnt methods are illustrated in Fig. 1.

Overall, the SPX method demonstrates better performance for the 1BW model. It captures precisely not only the height, but also the position of the Breit-Wigner peak. In comparison, the NAC and MaxEnt methods tend to overestimate the width and underestimate the height of the Breit-Wigner peak. Furthermore, the NAC method leads to an obvious oscillation phenomenon around 0.5 GeV. For the 2BW model, all three methods are able to reproduce the low-energy peak around 1.0 GeV, but encounter difficulty in resolving the high-energy peak near 3.0 GeV. Although the SPX method successfully identifies the position of the high-energy peak, its weight is not accurately resolved. The spectrum obtained by the NAC method is quite sensitive to

the λ parameter in this case. If $\lambda = 10^{-4}$ (it is the default choice of the code), the high-energy peak is shifted toward a higher energy (~ 4.5 GeV) and broadened significantly. Only when $\lambda = 10^{-6}$, the high-energy peak is well reproduced. The spectrum obtained by the MaxEnt method is also not ideal. The high-energy peak is smeared out and replaced with a shoulderlike feature.

C. Gaussian mixture model

Just as its name implies, the spectral function of the Gaussian mixture model [24] is a superposition of some Gaussian peaks. It can be expressed by the following equation:

$$\rho(\omega) = \sum_i A_i \exp\left[-\frac{(\omega - M_i)^2}{\Gamma_i}\right], \quad (23)$$

where A_i, M_i, and Γ_i represent the amplitude, position, and broadening of the ith Gaussian peak. In this example, we consider a three Gaussian peaks model. The specific values for the model parameters are as follows: $A_1 = 1.0$ GeV, $A_2 = 0.4$ GeV, $A_3 = 0.2$ GeV, $M_1 = 0.5$ GeV, $M_2 = 2.5$ GeV, $M_3 = 6.5$ GeV, $\Gamma_1 = 0.01$ GeV, $\Gamma_2 = 0.2$ GeV, and $\Gamma_3 = 1.5$ GeV. The mock Euclidean data consists of 100 points, and $T = 0.02$ GeV. In the simulations, we have enhanced the SPX method by utilizing the self-adaptive sampling algorithm, which we refer to as SA-SPX [46]. The results of the analytical continuation are presented in Fig. 2.

It is expected that the true spectrum will exhibit three well-defined peaks. We find that all three methods are successful in recovering the low-energy sharp peak at M_1. However, resolving the two high-energy peaks presents some challenges. Specifically, for the SA-SPX method, it is
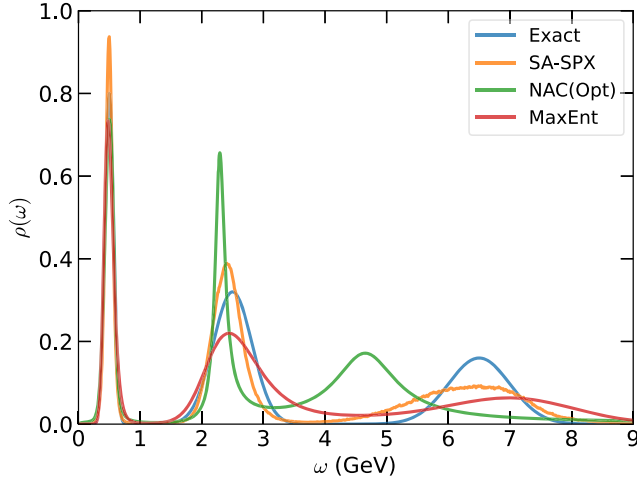
FIG. 2. Analytic continuations of the Gaussian mixture model. The spectra are scaled by a factor of 0.8 for a better view. As for the SPX method, the self-adaptive sampling algorithm is enabled to obtain a more reasonable spectrum. Once the self-adaptive sampling algorithm is turned off, the obtained spectrum by the SPX method will resemble the one by the MaxEnt method. The spectrum should become smoother and almost featureless in the high energy region ($\omega > 4.0$ GeV). The NAC method is enhanced by the Hardy basis function optimization algorithm and $N_{\text{opt}} = 11$. Please see the main text for more details.

able to roughly resolve the peak at M_2, but it tends to overestimate the width of the peak at M_3. In order to save the computational resources, the iteration number of the self-adaptive sampling algorithm is fixed to 10. Nevertheless, it should be emphasized that increasing the iteration number could further reduce the peak's width at M_3. In Sec. IV B, we will delve into the combination of the self-adaptive sampling algorithm with the SPX method and discuss the usefulness of the SA-SPX method in resolving complex LQCD spectral functions. As for the NAC method, it produces a sharp peak around 2.1 GeV and a broad peak around 4.5 GeV, which are both in incorrect positions. More specifically, this method underestimates the energies of the two high-energy peaks. We also adjust the λ parameter, but it does not help. Regarding the MaxEnt method, it can recover the peak at M_2, although with a larger width. However, it fails to resolve the peak at M_3, instead exhibiting a broad hump around 7.0 ± 3.0 GeV.

D. Resonance-continuum model

The resonance-continuum model is a physics-motivated model borrowed from Refs. [37,44]. The spectral function of the resonance-continuum model can be regarded as a nonlinear combination of the resonance part (ρ_r) and the continuum part (ρ_c):

$$\rho(\omega) = \xi_1(\omega)\rho_r(\omega) + \xi_2(\omega)\rho_c(\omega). \quad (24)$$

Here, ξ_1 and ξ_2 are the mixing coefficients. Their definitions are as follows:

$$\xi_1(\omega) = \xi(\omega, M_r, \Gamma)[1 - \xi(\omega, M_r + \Gamma, \Gamma)], \quad (25)$$

and

$$\xi_2(\omega) = \xi(\omega, M_c + \Gamma, \Gamma), \quad (26)$$

where ξ is a cutoff function:

$$\xi(\omega, M, \Delta) = \left[1 + \exp\left(\frac{M^2 - \omega^2}{\omega\Delta}\right)\right]^{-1}. \quad (27)$$

It is used to smooth out the constructed spectral function. The resonance part of the spectral function is given by:

$$\rho_r(\omega) = C_r\omega^2 \left[\frac{(M_r^2 - \omega^2)^2}{M_r^2\Gamma^2} + 1\right]^{-1}, \quad (28)$$

which follows a relativistic Breit-Wigner form. The continuum part of the spectral function is expressed as:

$$\rho_c(\omega) = \frac{3C_c}{8\pi\omega}\tanh\left(\frac{\omega}{4T}\right)\sqrt{\omega^2 - 4M_c^2}(2\omega^2 + 4M_c^2). \quad (29)$$

In this example, the model parameters are $M_r = 0.10$ GeV, $M_c = 0.05$ GeV, $C_r = 2.0$ GeV, $C_c = 2.10$ GeV, and $\Gamma = 0.06$ GeV. The synthetic Euclidean data consist of 100 frequency points, and $T = 0.02$ GeV. In this study, we consider three different cases: (i) the resonance-continuum model, (ii) the resonance model, and (iii) the continuum model. The analytic continuation results are shown in Fig. 3.

For the resonance-continuum model [see Fig. 3(a)], it is evident that the resonance peak at M_r is approximately reproduced. However, both the SPX and MaxEnt methods exhibit moderate oscillations in the continuum part ($\omega > 0.4$ GeV). These oscillations will decay when ω is increased. The NAC method results in huge oscillations, especially when ω is large. This unphysical feature cannot be eliminated or suppressed by the Hardy basis function optimization algorithm. This fact suggests that the three analytic continuation methods do not accurately describe the continuum part. In the case of the resonance model [see Fig. 3(b)], all three methods successfully recover the location, width, and height of the resonance peak. As for the continuum model [see Fig. 3(c)], all three methods produce oscillating spectra. In particular, the NAC method leads to more pronounced oscillations. Even worse, these oscillations are enhanced with the increment of ω. The only useful information that can be extracted from the calculated spectra is the location of the band edge.
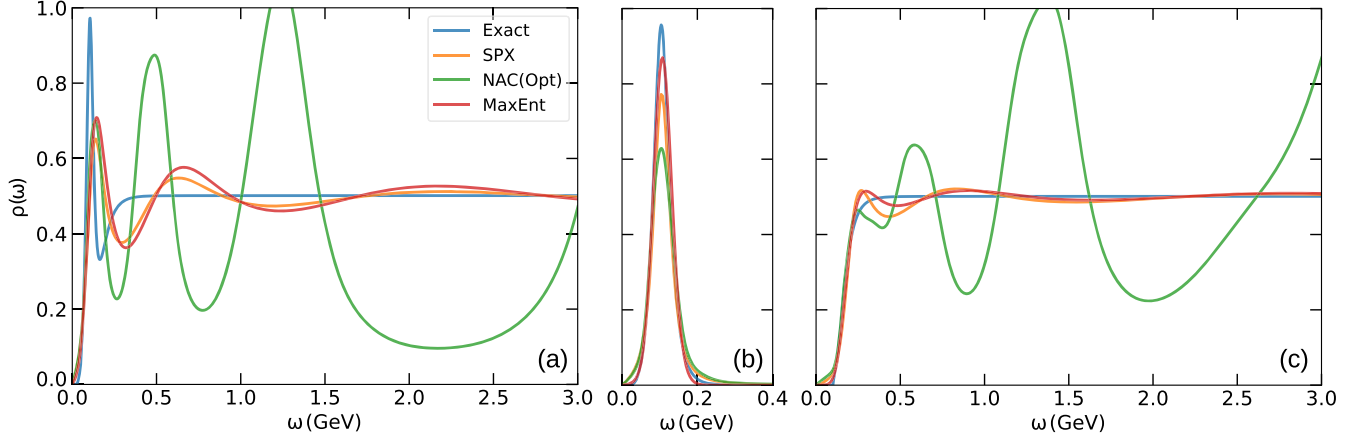
FIG. 3. Analytic continuations of the resonance-continuum model, the resonance model, and the continuum model. The NAC method is enhanced with the Hardy basis function optimization algorithm. (a) Spectra of the resonance-continuum model. $N_{opt} = 12$. (b) Spectra of the resonance model. $N_{opt} = 10$. (c) Spectra of the continuum model. $N_{opt} = 13$.

E. Bottomonium spectral functions

The "exact" bottomonium spectrum utilized in this study is taken directly from Refs. [63,69]. It is generated by a $N_f = 2 + 1$ LQCD calculation. The temperature employed in the LQCD simulation is 201 MeV, which exceeds the deconfinement crossover temperature (T_c). The spectrum is specifically for the Υ channel. Initially, we synthesize the Euclidean data by Eq. (2) for the first 100 Matsubara frequencies. Then, random Gaussian noises are added by Eq. (21). In this example, we adopt the combination of the SPX method with the constrained sampling algorithm (dubbed C-SPX) [46] to improve the performance. Specifically, locations for the randomly generated poles (P_γ) are restricted to the energy range: $\omega \in$ [9.5 GeV, 16.0 GeV] [70]. For the NAC method, the Hardy optimization trick is applied [47]. Regarding the MaxEnt method, its default model is a shifted Gaussian function [71]. The analytic continuation solutions, together with the "exact" bottomonium spectrum, are illustrated in Fig. 4.

As is seen in Fig. 4, the ideal bottomonium spectrum consists of a single resonance peak at approximately 9.6 GeV and a "rise-and-decay" feature with two sizable bumps around 10.8 GeV and 12.0 GeV. By employing the constrained algorithm, the SPX method successfully resolves the left boundary of the resonance peak and captures the long tail of the "rise-and-decay" feature. However, it falls short in resolving the resonance peak and the two bumps. In Sec. IV B, we will demonstrate the application of the self-adaptive sampling algorithm to cure this problem partly [46]. For the NAC method, it accurately reproduces the resonance peak. But it fails to recover the desired "rise-and-decay" feature, which is instead replaced by two distinct peaks located at approximately 11.3 GeV and 14.0 GeV. According to our experience, though the spectrum obtained by the NAC method is not accurate enough, it can provide some hints about the energy range of

the true spectrum (such as the position of the resonance peak in this case). We can use this information to refine subsequent simulations by imposing more reasonable constraints for the C-SPX method and more appropriate default models for the MaxEnt method. By employing the MaxEnt method, the resonance peak and the "rise-and-decay" pattern are smoothed out. Such that the calculated spectrum only features a Gaussian-like peak centered around 11.5 GeV. In addition to the shifted Gaussian model, we also benchmark alternative default models, such as the flat model, the shifted Lorentzian model, and the two Lorentzians model, etc. However, they do not contribute to improving the results.
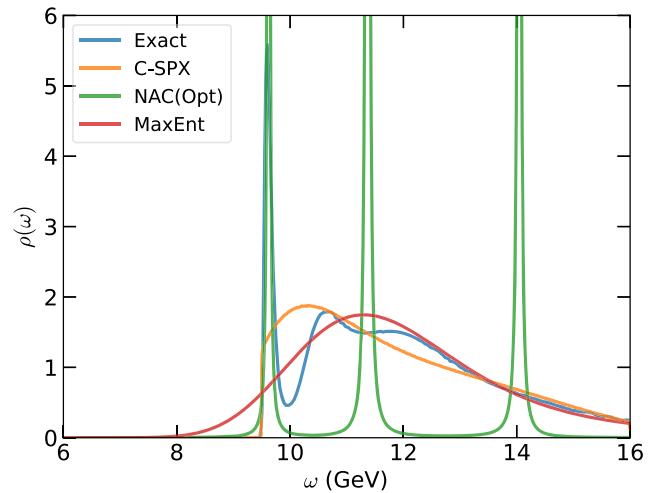

FIG. 4. Analytic continuations of the bottomonium correlation function. Here, the terminology "C-SPX" implies that the positions of the poles are restricted in the SPX simulations [46,70]. The Hardy basis function optimization algorithm is enabled for the NAC method and $N_{opt} = 8$. A shifted Gaussian function, instead of a constant, is used as the default model for the MaxEnt method [71]. See main text for more details.

F. Charmonium spectral functions

Next let us turn to a realistic example: extracting spectral functions from charmonium correlation functions.

The Euclidean data is extracted from Ref. [15]. The original correlation functions are obtained by LQCD calculations using clover-improved Wilson fermions on large, fine, isotropic lattices (i.e. $128^3 \times 96$) at $T = 0.73$ T_c, where T_c being the deconfinement transition temperature of a pure gluon plasma. Only the charmonium correlation function in the pseudoscalar (η_c) channel is considered. It is converted into the frequency domain at first. Then the random Gaussian noises are supplemented. The final input data contains 100 points. In this example, the standard SPX method without any constraints is used. For the MaxEnt method, the default model is flat. The analytic continuation results are shown in Fig. 5.

Note that the NAC method fails in this example. Therefore, only the spectra obtained by the SPX method and the MaxEnt method are shown in Fig. 5. In addition, the original paper provides a reference spectral function, which is calculated by the MaxEnt method with a physics-motivated default model [15]. It is also extracted and shown in this figure for comparison. The reference spectrum covers a very wide range of frequency ($0.0 < \omega < 75.0$). Besides the resonance peak (at 3.31 GeV), the reference spectrum features three peaks (around 7.0 GeV, 17.0 GeV, and 43.0 GeV) and a long tail until $\omega > 70.0$ GeV. The SPX method can roughly reproduce the resonance peak, the fourth peak at 43.0 GeV, and the long tail. But it fails to distinguish the two peaks at intermediate frequency. The
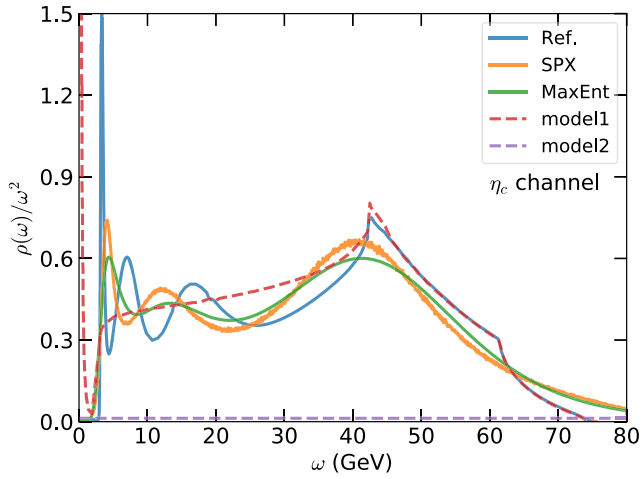

FIG. 5. Analytic continuations of the charmonium correlation functions (pseudoscalar channel, $T = 0.73\,T_c$) from LQCD simulations. There is no exact solution. The reference spectrum is taken from Ref. [15]. Here, the dashed line "model1" denotes the default model (a physics-motivated model) used by the MaxEnt method to generate the reference spectrum, while the dashed line "model2" denotes the default model (a flat model) used in the MaxEnt simulation in the present work. See main text for more details.

MaxEnt output just resembles the characteristics of the spectrum obtained by the SPX method. We find that the resonance peak (the first peak) remains stable and robust in SPX analyses, and in MaxEnt analyses performed with quite different default models (see the dashed lines in Fig. 5). This peak can be interpreted as the bound state peak. Similar behaviors can be seen in the peak around 43.0 GeV (the fourth peak in the reference spectrum) and the long tail. However, the peaks at intermediate frequency (the second and the third peaks in the reference spectrum) exhibit strong default model dependence. They could be a mixture of higher excited states or artifacts due to the finite lattice spacing and limited number of correlator points. Anyway, we believe the SPX method can capture the "stable" features of the spectral functions.

IV. DISCUSSIONS

A. Robustness with respect to noisy data

In the previous work, the robustness of the SPX method in the presence of noisy data from quantum Monte Carlo simulations has been demonstrated [46]. This study aims to reexamine the noise resilience of the SPX method when applied to noisy LQCD data. Let us take the 1BW model as an example (please refer to Sec. III B for the model parameters) to address this issue. The noise level is varied from $\delta = 10^{-8}$ to $\delta = 10^{-2}$. The analytic continuation results are displayed in Fig. 6. Just as expected, the SPX method is highly robust to variations in noise levels of the input Euclidean data. For low noise levels ($10^{-4} \le \delta \le 10^{-8}$), the calculated spectra closely approximate the exact solution and manifest minimal deviation. At a moderate noise level ($\delta = 10^{-3}$), the calculated spectrum shows a light fluctuation around $\omega = 0.5$ GeV, and the main peak is shifted slightly (approximately 0.1 GeV) toward the lower energy region. For a high noise level ($\delta = 10^{-2}$), three sharp peaks emerge in the calculated spectrum. Apart from the peak at 1.8 GeV, the other peaks at 0.5 GeV and 3.0 GeV are unphysical. In Fig. 6(h), a plot of $\log(\langle \chi^2 \rangle)$ against $\log(\delta^{-1})$ is shown. Initially, $\log(\langle \chi^2 \rangle)$ decreases linearly as $\log(\delta^{-1})$ increases from 2.0 to 5.0. Subsequently, it approaches to a constant value (approximately -8.1) when $\log(\delta^{-1}) \ge 6.0$. This benchmark suggests that the SPX method remains robust when applied to noisy LQCD data, even with a moderately elevated noise level. Nonetheless, minimizing the noise level can enhance the performance of the SPX method.

Fei *et al.* have pointed out that the NAC method requires high-precision input data to ensure the Pick criterion is not violated and the existence of the Nevanlinna interpolants [47,65]. Thus, in the previous calculations, we just assume that the input Euclidean data are noiseless for the NAC method ($\delta = 0.0$). Now let us examine the noise resilience of the NAC method for synthetic LQCD data. For the sake of simplicity, the 2BW
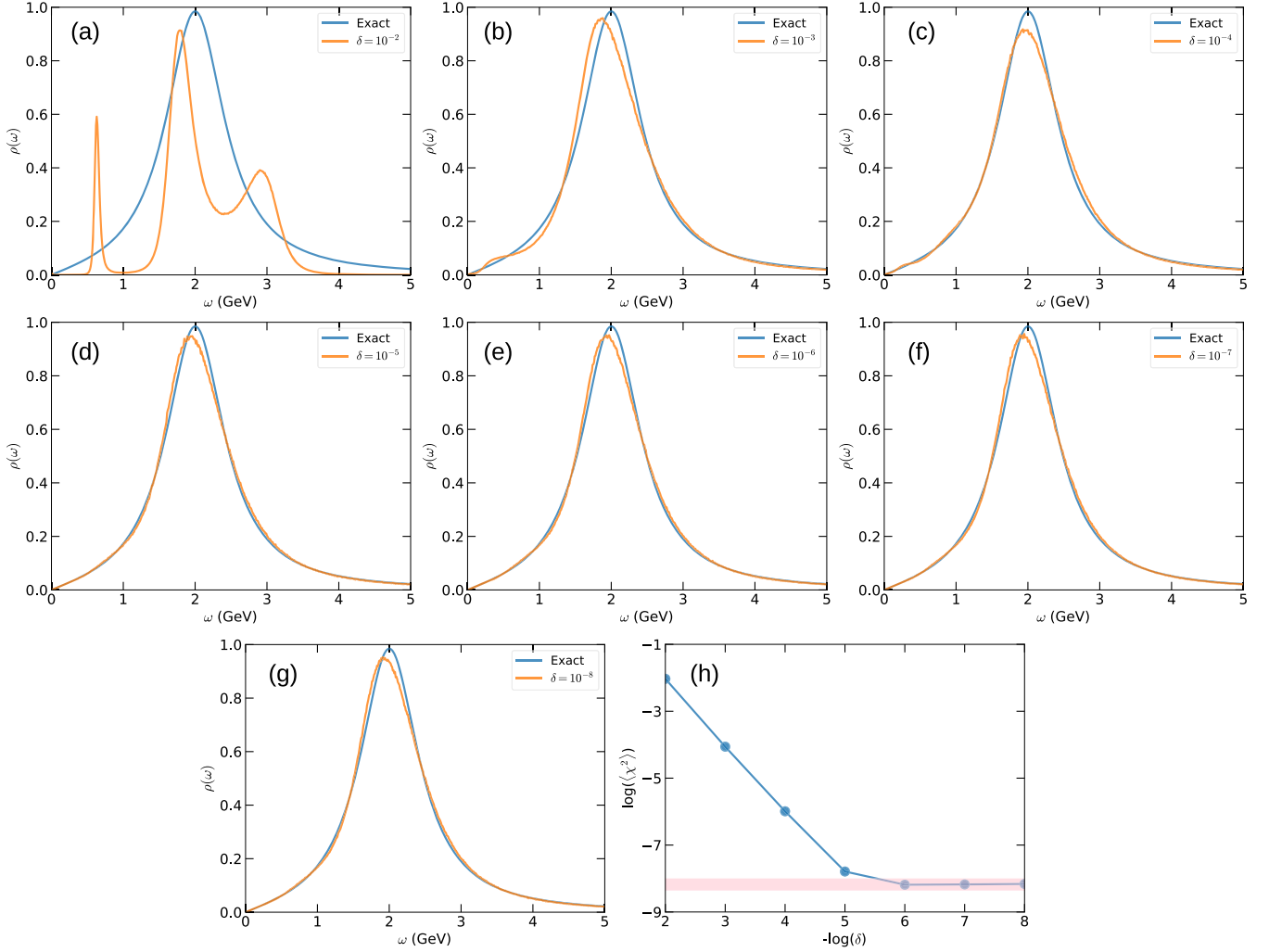
FIG. 6. Robustness of the SPX method with respect to the noisy LQCD data. Here we just consider the Breit-Wigner model (1BW model). The noise level δ is varied from 10^{-8} to 10^{-2}. The other model parameters can be found in Sec. III B. (a)–(g) Dependence on noise level δ of calculated spectral functions. (h) The goodness-of-fit function χ^2 as a function of the noise level δ. The horizontal bar indicates the asymptotic value of $\log(\langle\chi^2\rangle)$.

model is taken as a test bed. The model parameters are presented in Sec. III B. The noise level δ is fixed to be 10^{-8}. For each NAC run, the noisy part of the mock Euclidean data is always refreshed. We repeat the NAC simulations for 100 times with and without Hardy basis function optimization [60]. Then we collect the calculated spectra and evaluate their arithmetic average. The analytic continuation results are shown in Fig. 7. We confirm again that the NAC method is extremely sensitive to noise, irrespective of the Hardy basis function optimization algorithm. Small fluctuations in the input Euclidean data can lead to huge variations in the resulting spectra. Even though the noise level is quite small, the performance of the NAC method is not good. It fails to capture the major characteristics of the 2BW model, and produces some spurious peaks. If the noise level is further increased, its performance should deteriorate ulteriorly (not shown in Fig. 7).

B. Self-adaptive sampling algorithm for the SPX method

In the SPX method, the poles should be placed in a very dense frequency grid. In general, such a frequency grid can be either uniform or nonuniform. So, some *prior knowledge* about the spectrum and the physical system could be encapsulated in the form of the grid to improve the performance and usefulness of the SPX method. This has led to the development of the C-SPX method [46]. Previous works have suggested that by modifying boundaries and grid interval distribution of the grid, the SPX method is capable of capturing complicated features in the spectra. However, it can be observed in Fig. 4 that the C-SPX method, as well as the NAC and MaxEnt methods, fail to resolve the major characteristics of the bottomonium spectrum [63,69]. It implies that simple constraints on the spectral boundaries (or limitations on scopes of the poles)
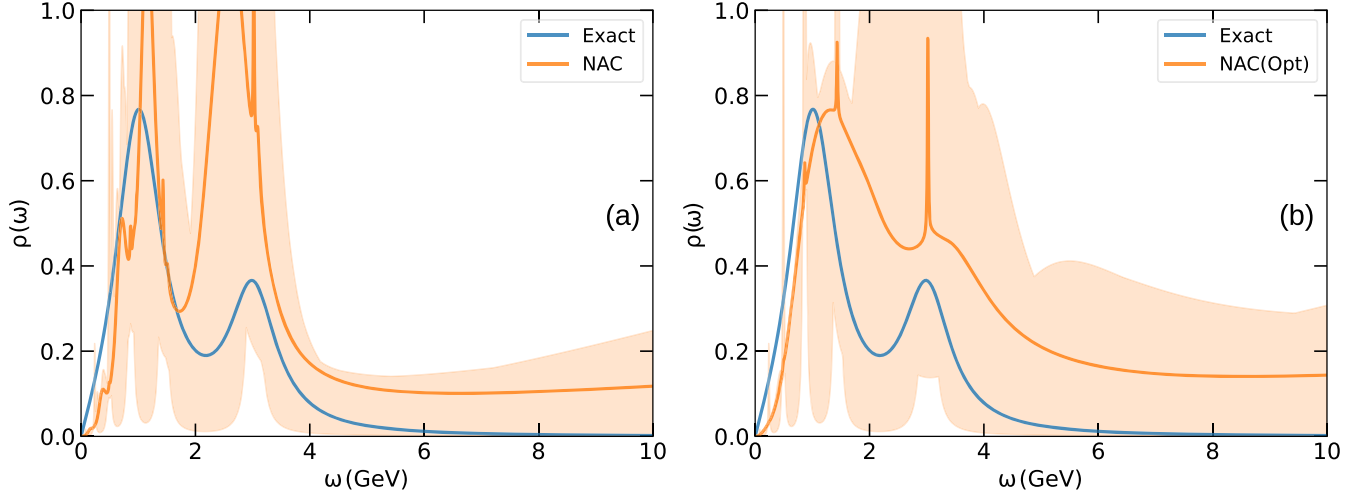
FIG. 7. Robustness of the NAC method with respect to the noisy LQCD data. Here we just consider the Breit-Wigner model (2BW model). The noise level $\delta = 10^{-8}$. The other model parameters can be found in Sec. III B. The darker shaded region denotes the window for the reconstructed spectral functions. The blue and yellow solid lines mean the exact spectrum and the averaged spectrum, respectively. (a) Without Hardy basis function optimization. (b) With Hardy basis function optimization.

are not enough. We have to figure out a systematic way to refine the probability distribution of the poles to approximate the true spectrum. Next, we will demonstrate how to achieve this goal by a combination of the self-adaptive sampling algorithm and the SPX method (dubbed SA-SPX) [46].

The main principle behind the SA-SPX method is to iteratively adjust the grid interval distribution. In consequence, the probability distribution of the poles is coordinated to approximate the true spectral density, and the corresponding goodness-of-fit functional [see Eq. (11)] is automatically minimized. This can be achieved by using the spectral density obtained from a previous SPX run or from other analytic continuation methods to update the grid. Now let us concentrate on the bottomonium model again [63,69]. Its parameters can be found in Sec. III E. Initially, we generate the first frequency grid for the poles using the spectral functions obtained by the NAC and MaxEnt methods. From the calculated spectra, one can conclude that there is likely a sharp resonance peak around $\omega = 9.6$ GeV (with a band edge at approximately 9.5 GeV) and a broad feature ranging from 10.0 to 16.0 GeV (see Fig. 4). Keeping these hints in mind, we try to design a pseudospectrum consisting of two Gaussian peaks:

$$\rho_{\text{pseudo}}(\omega) = \sum_{i=1}^{2} A_i \exp\left[\frac{(\omega - M_i)^2}{\Gamma_i}\right], \quad (30)$$

where $A_1 = 5.0$, $A_2 = 1.80$, $M_1 = 9.60$, $M_2 = 11.5$, $\Gamma_1 = 0.01$, and $\Gamma_2 = 5.0$. This pseudospectrum is illustrated in Fig. 8(a). The first peak at M_1 originates from the resonance peak identified by the NAC method, while the second peak at M_2 is inspired by the spectrum obtained by the MaxEnt method. It is worth noting that these spectral

parameters can be further adjusted to mimic more accurately the results obtained by the NAC and MaxEnt methods. This pseudospectrum serves as a reference model. To generate the frequency grid for the poles, we execute the following steps: (1) Calculate the integrated spectral function $\phi(\epsilon)$ via the equation:

$$\phi(\epsilon) = \int_{\omega_{\min}}^{\epsilon} \rho_{\text{pseudo}}(\omega) d\omega, \quad \epsilon \in [\omega_{\min}, \omega_{\max}]. \quad (31)$$

Here, $\omega_{\min}$ and $\omega_{\max}$ are the left and right boundaries of the spectrum, respectively. And ϵ represents a point within the interval $[\omega_{\min}, \omega_{\max}]$. (2) Evaluate the new frequency grid f_i by using the equation:

$$f_i = \phi^{-1}(\lambda_i), \quad i = 1, \ldots, N_f, \quad (32)$$

where λ_i is a linear mesh in the interval $[\phi(\omega_{\min}), \phi(\omega_{\max})]$, and N_f is the number of grid points. Now the boundaries for the grid are set as $\omega_{\min} = 9.5$ GeV and $\omega_{\max} = 16.0$ GeV. The resulting new grid, as displayed in Fig. 8(b), is compared with the standard linear grid. Next, the newly generated grid is utilized to perform a SPX simulation from scratch. The calculated spectrum is then used to generate a newer frequency grid [at this time, the $\rho_{\text{pseudo}}(\omega)$ in Eq. (31) should be replaced with the newly calculated spectrum $\rho(\omega)$], and the SPX simulation is repeated. This iterative procedure is carried out until the obtained spectrum and frequency grid are converged. In our experience, 5–10 iterations are typically sufficient for achieving convergence. Figure 8(c) shows the results obtained by using the C-SPX and SA-SPX methods, as well as the exact spectrum. It is evident that the spectrum obtained with the SA-SPX method comes closer to the exact spectrum than
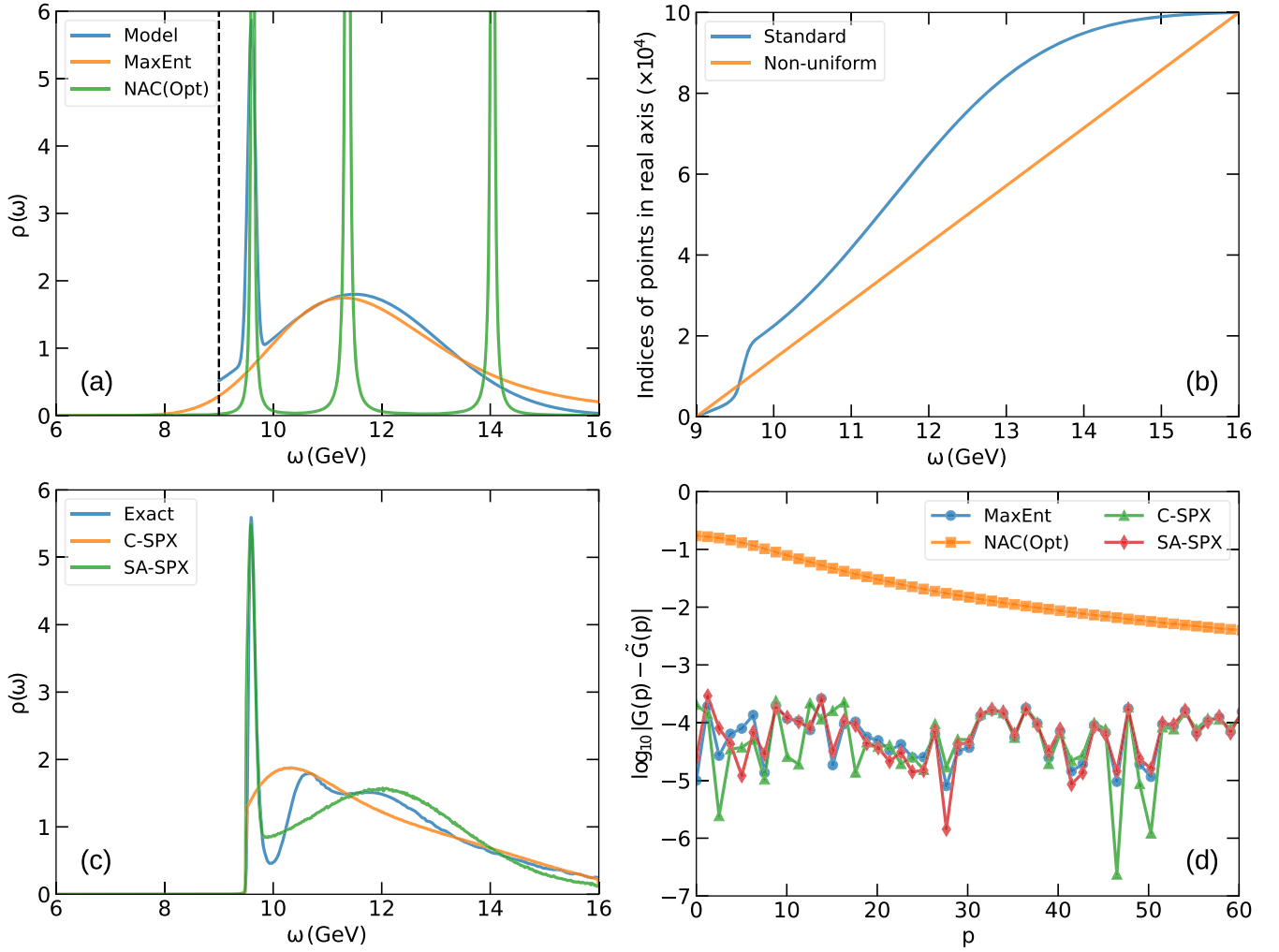
FIG. 8. Analytic continuations of the bottomonium correlation function. (a) Spectra obtained by the MaxEnt and optimized NAC methods. The information extracted from the two spectra is used to construct a reference model [the pseudospectrum $\rho_{\text{pseudo}}(\omega)$, see the solid blue line]. (b) Standard linear grid and nonuniform grid used in the C-SPX and SA-SPX simulations, respectively. Note that the nonuniform grid is constructed from the reference model via Eqs. (31) and (32). (c) Spectra obtained by the C-SPX and SA-SPX methods. (d) Error analysis for the reconstructed Euclidean data from the MaxEnt, optimized NAC, C-SPX, and SA-SPX methods.

the one obtained with the C-SPX method. The sharp resonance peak, the small bump near 12.0 GeV, and the long tail of the rise-and-decay feature are well reproduced by using the SA-SPX method. The only missing characteristic is the bump near 10.8 GeV. Additionally, an error analysis about the reconstructed Euclidean data is presented in Fig. 8(d). The goodness-of-fit function of the NAC method is the largest ($\chi^2 \approx 0.01$ to 0.1), while those of the MaxEnt, C-SPX, and SA-SPX methods are comparable ($\chi^2 \approx 10^{-6}$ to 10^{-4}). This indicates that the spectrum obtained by the NAC method for this particular case is not reliable.

C. Hardy basis function optimization for the NAC method

As mentioned before, the iterative interpolation algorithm for the NAC method allows for the selection of an

arbitrary contractive function, denoted as θ_{N_s+1}, at the final step. In the literature, Fei *et al.* proposed an optimized approach [47], in which θ_{N_s+1} is expanded using the Hardy basis and its conjugate generate functions [60], and the coefficients for the expansion are determined by minimizing a smoothness norm [see Eq. (16)]. They demonstrated that using a constant value for θ_{N_s+1} is apt to yield spectral functions with oscillations, whereas the optimized algorithm is useful for eliminating these oscillations and generating smoother spectral functions. Until now we just employ the optimized NAC approach for analytic continuations. However, we wonder whether the Hardy basis optimization is always better than the standard option. In order to answer this question, we perform additional tests for the Breit-Wigner model by using the NAC method. We compare the analytic continuation results obtained with a
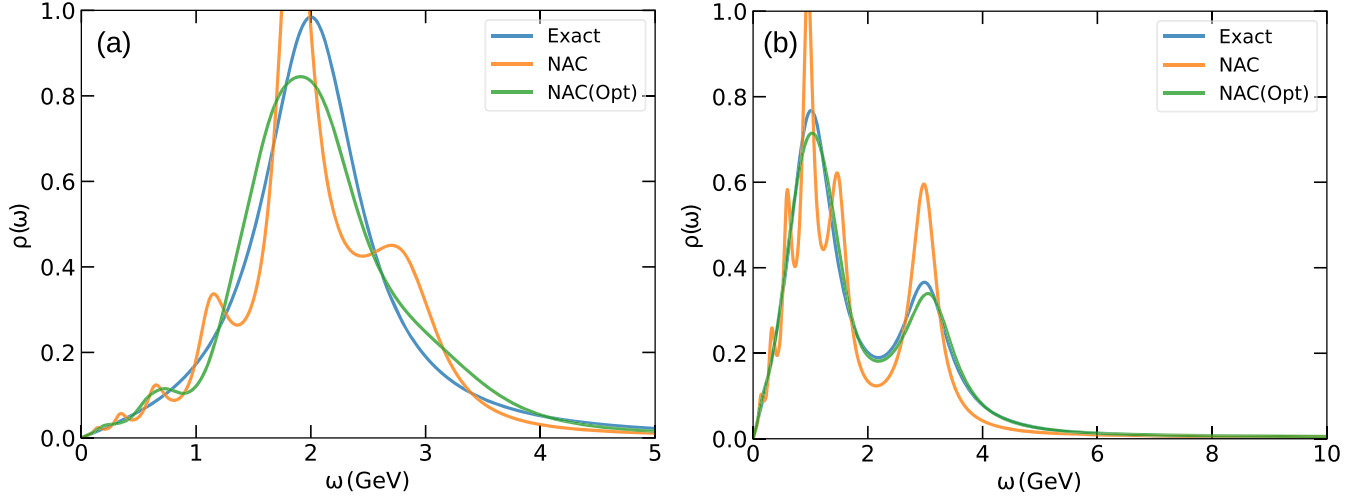
FIG. 9. A comparison for the NAC and optimized NAC methods. The Breit-Wigner model is used and the model parameters are the same as those presented in Sec. III B. The input Euclidean data are noiseless. (a) Results for the 1BW model. (b) Results for the 2BW model.

constant θ_{N_s+1} and the optimized θ_{N_s+1} (see Fig. 9). As anticipated, the optimized θ_{N_s+1} suppresses oscillations in a large degree and yields smoother spectra, especially for the 2BW model. However, we also notice that the performance of the Hardy basis function optimization algorithm strongly depends on the value of the λ parameter. The optimized NAC method tends to make a wrong estimation about the location of the high-energy peak if the λ parameter is not reasonable. Therefore, if we know nothing about the basic features of the spectra, perhaps a constant θ_{N_s+1} is a much safer choice.

V. CONCLUSION

In the present work, we conduct a systematic investigation of two newly developed methods, namely the SPX method and the NAC method, for analytically continuing for the mock and realistic LQCD data. We treat four exact spectral functions, which are derived from physically motivated models or realistic LQCD simulations, including the Breit-Wigner model, the Gaussian mixture model, the resonance-continuum model, and the bottomonium model. We use the exact spectral functions to build clean Euclidean data by numerical integration. And later the statistical noises are added. We also examine true Euclidean data. The charmonium correlation functions in the pseudoscalar channel obtained in LQCD calculations are extracted from literature. All the synthetic Euclidean correlators are used as input and then transformed back to the real axis using different analytic continuation methods. By comparing the results with the exact spectra, we are able to assess the accuracy of these methods.

The SPX method is generally capable of resolving the major features of the spectral functions involved in this study. However, it encounters difficulties when dealing with spectra that exhibit a wide platform, such as the continuum model (see Sec. III D), or when two features are too close together, such as the bottomonium spectrum (see Sec. IV B). We believe that these difficulties can be partially overcome with the help of the constrained sampling algorithm and the self-adaptive sampling algorithm. The SPX method demonstrates good noise tolerance and exhibits robustness with respect to moderate noise levels. Overall, the performance of the SPX method is comparable to that of the commonly used MaxEnt method. In cases where the spectral function is complicated, the SPX method could outperform the MaxEnt method due to its ability to incorporate prior information about the spectrum into the frequency grid for the poles. This grid could be iteratively refined to obtain better spectrum.

As for the NAC method, it is found to be numerically unstable even for input Euclidean data with extremely low noise level ($\delta = 10^{-8}$). This drawback greatly limits the application of the NAC method in the LQCD simulations. Additionally, we observe that the Hardy basis optimization for θ_{N_s+1} sometimes produces worse results when compared to those obtained with constant θ_{N_s+1}. Although the Hardy basis optimization can suppress possible oscillations in the spectrum, it tends to yield incorrect estimations for the positions of the high-energy peaks if the λ parameter is not optimal. Therefore, better basis functions for expanding θ_{N_s+1} are highly desirable. Or else we need a smart algorithm to determine the optimal λ. Nonetheless, the NAC method still proves its usefulness in analytic continuations of LQCD simulation data, as it allows for quick yet accurate estimations of the positions of the low-energy band edges and the resonance peaks. These important clues can then be used to construct a reference model for the

probability distribution of the poles, which is subsequently utilized by the C-SPX and SA-SPX methods.

ACKNOWLEDGMENTS

[1] *Lattice QCD for Nuclear Physics*, edited by Huey-Wen Lin and Harvey B. Meyer (Springer International Publishing, New York, 2015).

[2] Rajan Gupta, Introduction to lattice QCD, arXiv:9807028.

[3] Hartmut Wittig, QCD on the lattice, in *Particle Physics Reference Library: Volume 1: Theory and Experiments*, edited by Herwig Schopper (Springer International Publishing, Cham, 2020), pp. 137–262.

[4] John W. Harris and Berndt Müller, The search for the quark-gluon plasma, Annu. Rev. Nucl. Part. Sci. **46**, 71 (1996).

[5] Claudia Ratti, Lattice QCD and heavy ion collisions: A review of recent progress, Rep. Prog. Phys. **81**, 084301 (2018).

[6] Christoph Lehner, High-precision lattice QCD calculations of the muon anomalous magnetic moment, Nat. Rev. Phys. **4**, 14 (2022).

[7] Gert Aarts *et al.*, Phase transitions in particle physics: Results and perspectives from lattice quantum chromodynamics, Prog. Part. Nucl. Phys. **133**, 104070 (2023).

[8] H. Ohno, S. Aoki, S. Ejiri, K. Kanaya, Y. Maezawa, H. Saito, and T. Umeda (WHOT-QCD Collaboration), Charmonium spectral functions with the variational method in zero and finite temperature lattice QCD, Phys. Rev. D **84**, 094504 (2011).

[9] Gert Aarts, Chris Allton, Mehmet Buğrahan Oktay, Mike Peardon, and Jon-Ivar Skullerud, Charmonium at high temperature in two-flavor QCD, Phys. Rev. D **76**, 094513 (2007).

[10] A. Jakovác, P. Petreczky, K. Petrov, and A. Velytsky, Quarkonium correlators and spectral functions at zero and finite temperature, Phys. Rev. D **75**, 014506 (2007).

[11] Saumen Datta, Frithjof Karsch, Peter Petreczky, and Ines Wetzorke, Behavior of charmonium systems after deconfinement, Phys. Rev. D **69**, 094507 (2004).

[12] M. Asakawa and T. Hatsuda, J/ψ and η_c in the deconfined plasma from lattice QCD, Phys. Rev. Lett. **92**, 012001 (2004).

[13] Takashi Umeda, Kouji Nomura, and Hideo Matsufuru, Charmonium at finite temperature in quenched lattice QCD, Eur. Phys. J. C **39**, 9 (2005).

[14] G. Aarts, C. Allton, S. Kim, M. P. Lombardo, M. B. Oktay, S. M. Ryan, D. K. Sinclair, and J. I. Skullerud, What happens to the Υ and ηb in the quark-gluon plasma? Bottomonium spectral functions from lattice QCD, J. High Energy Phys. 11 (2011) 103.

[15] H.-T. Ding, A. Francis, O. Kaczmarek, F. Karsch, H. Satz, and W. Soeldner, Charmonium properties in hot quenched lattice QCD, Phys. Rev. D **86**, 014509 (2012).

[16] Gert Aarts, Chris Allton, Justin Foley, Simon Hands, and Seyong Kim, Spectral functions at small energies and the electrical conductivity in hot quenched lattice QCD, Phys. Rev. Lett. **99**, 022002 (2007).

[17] H.-T. Ding, A. Francis, O. Kaczmarek, F. Karsch, E. Laermann, and W. Soeldner, Thermal dilepton rate and electrical conductivity: An analysis of vector current correlation functions in quenched lattice QCD, Phys. Rev. D **83**, 034504 (2011).

[18] Min He, Hendrik van Hees, and Ralf Rapp, Heavy-quark diffusion in the quark–gluon plasma, Prog. Part. Nucl. Phys. **130**, 104020 (2023).

[19] Ralf-Arno Tripolt, Philipp Gubler, Maksim Ulybyshev, and Lorenz von Smekal, Numerical analytic continuation of Euclidean data, Comput. Phys. Commun. **237**, 129 (2019).

[20] John W. Negele and Henri Orland, *Quantum Many-Particle Systems* (CRC Press, Boca Raton, 1998).

[21] Massimo Di Pierro, From Monte Carlo integration to quantum chromo dynamics (an introduction), arXiv:hep-lat/0009001.

[22] Mark Jarrell and J. E. Gubernatis, Bayesian inference and the analytic continuation of imaginary-time quantum Monte Carlo data, Phys. Rep. **269**, 133 (1996).

[23] J. E. Gubernatis, Mark Jarrell, R. N. Silver, and D. S. Sivia, Quantum Monte Carlo simulations and maximum entropy: Dynamics from imaginary-time data, Phys. Rev. B **44**, 6011 (1991).

[24] M. Asakawa, Y. Nakahara, and T. Hatsuda, Maximum entropy analysis of the spectral functions in lattice QCD, Prog. Part. Nucl. Phys. **46**, 459 (2001).

[25] Alexander Rothkopf, Bayesian inference of real-time dynamics from lattice QCD, Front. Phys. **10**, 1028995 (2022).

[26] Yannis Burnier and Alexander Rothkopf, Bayesian approach to spectral function reconstruction for Euclidean quantum field theories, Phys. Rev. Lett. **111**, 182003 (2013).

[27] Anders W. Sandvik, Stochastic method for analytic continuation of quantum Monte Carlo data, Phys. Rev. B **57**, 10287 (1998).

[28] K. S. D. Beach, Identifying the maximum entropy method as a special limit of stochastic analytic continuation, arXiv:0403055.

[29] A. S. Mishchenko, N. V. Prokof'ev, A. Sakamoto, and B. V. Svistunov, Diagrammatic quantum Monte Carlo study of the Fröhlich polaron, Phys. Rev. B **62,** 6317 (2000).

[30] Anders W. Sandvik, Constrained sampling method for analytic continuation, Phys. Rev. E **94,** 063308 (2016).

[31] K. Vafayi and O. Gunnarsson, Analytical continuation of spectral data from imaginary time axis to real frequency axis using statistical sampling, Phys. Rev. B **76,** 035115 (2007).

[32] Sebastian Fuchs, Thomas Pruschke, and Mark Jarrell, Analytic continuation of quantum Monte Carlo data by stochastic analytical inference, Phys. Rev. E **81,** 056701 (2010).

[33] Hui Shao, Yan Qi Qin, Sylvain Capponi, Stefano Chesi, Zi Yang Meng, and Anders W. Sandvik, Nearly deconfined spinon excitations in the square-lattice spin-1/2 Heisenberg antiferromagnet, Phys. Rev. X **7,** 041072 (2017).

[34] Khaldoon Ghanem and Erik Koch, Average spectrum method for analytic continuation: Efficient blocked-mode sampling and dependence on the discretization grid, Phys. Rev. B **101,** 085111 (2020).

[35] Khaldoon Ghanem and Erik Koch, Extending the average spectrum method: Grid point sampling and density averaging, Phys. Rev. B **102,** 035114 (2020).

[36] Olav F. Syljuåsen, Using the average spectrum method to extract dynamics from quantum Monte Carlo simulations, Phys. Rev. B **78,** 174429 (2008).

[37] H.-T. Ding, O. Kaczmarek, Swagato Mukherjee, H. Ohno, and H.-T. Shu, Stochastic reconstructions of spectral functions: Application to lattice QCD, Phys. Rev. D **97,** 094503 (2018).

[38] Hui Shao and Anders W. Sandvik, Progress on stochastic analytic continuation of quantum Monte Carlo data, Phys. Rep. **1003,** 1 (2023).

[39] Lukas Kades, Jan M. Pawlowski, Alexander Rothkopf, Manuel Scherzer, Julian M. Urban, Sebastian J. Wetzel, Nicolas Wink, and Felix P. G. Ziegler, Spectral reconstruction with deep neural networks, Phys. Rev. D **102,** 096001 (2020).

[40] Shuzhe Shi, Lingxiao Wang, and Kai Zhou, Rethinking the ill-posedness of the spectral function reconstruction—Why is it fundamentally hard and how artificial neural networks can help, Comput. Phys. Commun. **282,** 108547 (2023).

[41] Meng Zhou, Fei Gao, Jingyi Chao, Yu-Xin Liu, and Huichao Song, Application of radial basis functions neutral networks in spectral functions, Phys. Rev. D **104,** 076011 (2021).

[42] Chen ShiYang, Hengtong Ding, Fei-Yi Liu, G. Papp, and Chun-Bin Yang, Machine learning hadron spectral functions in lattice QCD, Proc. Sci. LATTICE2021 (**2022**) 148.

[43] Samuel Offler, Gert Aarts, Chris Allton, Benjamin Jaeger, Seyong Kim, Maria Paola Lombardo, Ben Page, Sinead Ryan, Jon-Ivar Skullerud, and Thomas Spriggs, Reconstruction of bottomonium spectral functions in thermal QCD using Kernel Ridge regression, Proc. Sci. LATTICE2021 (**2022**) 509.

[44] Lingxiao Wang, Shuzhe Shi, and Kai Zhou, Reconstructing spectral functions via automatic differentiation, Phys. Rev. D **106,** L051502 (2022).

[45] Jan Horak, Jan M. Pawlowski, José Rodríguez-Quintero, Jonas Turnwald, Julian M. Urban, Nicolas Wink, and Savvas Zafeiropoulos, Reconstructing QCD spectral functions with Gaussian processes, Phys. Rev. D **105,** 036014 (2022).

[46] Li Huang and Shuang Liang, Stochastic pole expansion method for analytic continuation of the Green's function, Phys. Rev. B **108,** 235143 (2023).

[47] Jiani Fei, Chia-Nan Yeh, and Emanuel Gull, Nevanlinna analytical continuation, Phys. Rev. Lett. **126,** 056402 (2021).

[48] Jiani Fei, Chia-Nan Yeh, Dominika Zgid, and Emanuel Gull, Analytical continuation of matrix-valued functions: Carathéodory formalism, Phys. Rev. B **104,** 165111 (2021).

[49] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, Science **220,** 671 (1983).

[50] H. J. Vidberg and J. W. Serene, Solving the Eliashberg equations by means of N-point Padé approximants, J. Low Temp. Phys. **29,** 179 (1977).

[51] Johan Schött, Inka L. M. Locht, Elin Lundin, Oscar Grånäs, Olle Eriksson, and Igor Di Marco, Analytic continuation by averaging Padé approximants, Phys. Rev. B **93,** 075104 (2016).

[52] K. S. D. Beach, R. J. Gooding, and F. Marsiglio, Reliable Padé analytical continuation method based on a high-accuracy symbolic computation algorithm, Phys. Rev. B **61,** 5147 (2000).

[53] Žiga Osolin and Rok Žitko, Padé approximant approach for obtaining finite-temperature spectral functions of quantum impurity models using the numerical renormalization group technique, Phys. Rev. B **87,** 245135 (2013).

[54] François Gelis, *Quantum Field Theory: From Basics to Modern Topics* (Cambridge University Press, Cambridge, England, 2019).

[55] Michael E. Peskin and Daniel V. Schroeder, *An Introduction to Quantum Field Theory* (CRC Press, Boca Raton, 1995).

[56] Piers Coleman, *Introduction to Many-Body Physics* (Cambridge University Press, Cambridge, England, 2016).

[57] Jorge Nocedal and Stephen J. Wright, *Numerical Optimization* (Springer, New York, NY, 2006).

[58] V. M. Adamyan, J. Alcober, and I. M. Tkachenko, Reconstruction of distributions by their moments and local constraints, Appl. Math. Res. Express **2003,** 33 (2003).

[59] Georg Pick, Über die beschränkungen analytischer funktionen durch vorgegebene funktionswerte, Math. Ann. **78,** 270 (1917).

[60] Steven G. Krantz, Special classes of holomorphic functions, in *Handbook of Complex Variables* (Birkhäuser, Boston, Boston, MA, 1999), pp. 149–154.

[61] Thomas Bergamaschi, William I. Jay, and Patrick R. Oare, Hadronic structure, conformal maps, and analytic continuation, Phys. Rev. D **108,** 074516 (2023).

[62] Kosuke Nogaki and Hiroshi Shinaoka, Bosonic Nevanlinna analytic continuation, J. Phys. Soc. Jpn. **92,** 035001 (2023).

[63] G. Aarts, C. Allton, T. Harris, S. Kim, M. P. Lombardo, S. M. Ryan, and J.-I. Skullerud, The bottomonium spectrum at finite temperature from $N_f = 2 + 1$ lattice QCD, J. High Energy Phys. 07 (2014) 097.

[64] To prepare input data for the NAC method, one at first converts $\rho(\omega)$ to $\tilde{\rho}(\omega)$ via Eq. (18), and then calculates $\tilde{G}(i\nu_n)$ via Eq. (20).

[65] Zhen Huang, Emanuel Gull, and Lin Lin, Robust analytic continuation of Green's functions via projection, pole estimation, and semidefinite relaxation, Phys. Rev. B **107,** 075151 (2023).

[66] Li Huang, ACFlow: An open source toolkit for analytic continuation of quantum Monte Carlo data, Comput. Phys. Commun. **292,** 108863 (2023).

[67] Kosuke Nogaki, Jiani Fei, Emanuel Gull, and Hiroshi Shinaoka, Nevanlinna.jl: A Julia implementation of Nevanlinna analytic continuation, SciPost Phys. Codebases 19 (2023).

[68] Dominic Bergeron and A.-M. S. Tremblay, Algorithms for optimized maximum entropy and diagnostic tools for analytic continuation, Phys. Rev. E **94,** 023303 (2016).

[69] Alexander Rothkopf, Heavy quarkonium in extreme conditions, Phys. Rep. **858,** 1 (2020).

[70] The left and right boundaries for the C-SPX method can be deduced from the spectrum obtained by the NAC method. Actually, we perform analytic continuation by the NAC method at first. From the obtained spectrum, we can make a rough estimation about the boundaries of the true spectrum. Then this information can be utilized to customize the constraints for the C-SPX method. Please see Ref. [46] for more details.

[71] The center of the Gaussian peak is fixed to be 11.5 GeV, which is learnt from the spectrum obtained by the NAC method. Please refer to Ref. [70] as well.