

Noiseless limit and improved-prior limit of the maximum entropy method and their implications for the analytic continuation problem

Thomas Chuna^{1,2,*}, Nicholas Barnfield³, Paul Hamann^{2,4}, Sebastian Schwalbe^{1,2},
Michael P. Friedlander⁵, and Tobias Dornheim^{1,2}

¹*Center for Advanced Systems Understanding (CASUS), D-02826 Görlitz, Germany*

²*Helmholtz-Zentrum Dresden-Rossendorf (HZDR), D-01328 Dresden, Germany*

³*Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA*

⁴*Institut für Physik, Universität Rostock, D-18057 Rostock, Germany*

⁵*Department of Computer Science and Mathematics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4*



(Received 10 November 2025; accepted 11 May 2026; published 6 July 2026)

Quantum Monte Carlo (QMC) methods are uniquely capable of providing exact simulations of quantum many-body systems. Unfortunately, the applications of a QMC simulation are limited because extracting dynamic properties requires solving the analytic continuation problem. Across the many fields that use QMC methods, there is no universally accepted analytic continuation algorithm for extracting dynamic properties, but many publications compare to the maximum entropy method. We investigate when entropy maximization is an acceptable approach. We show that stochastic sampling algorithms reduce to entropy maximization when the Bayesian prior is near to the true solution. We investigate when is Bryan's controversial optimization algorithm [Bryan, *Eur. Biophys. J.* **18**, 165 (1990)] for entropy maximization (sometimes known as the maximum entropy method—MEM) appropriate to use. We show that Bryan's algorithm is appropriate when the noise is near zero or when the Bayesian prior is near the true solution. We also investigate the mean squared error, finding a better scaling when the Bayesian prior is near the true solution than when the noise is near zero. Together, these three results indicate that enhancing the quality of the Bayesian prior enhances the validity of the MEM and has better return on investment than reducing the noise. We point to examples of improved data-driven Bayesian priors that have already leveraged this advantage. We support these results by solving the double Gaussian problem using both Bryan's algorithm and the newly formulated dual approach to entropy maximization [Chuna *et al.*, *J. Phys. A: Math. Theor.* **58**, 335203 (2025)].

DOI: 10.1103/yqjh-ycbh

I. INTRODUCTION

Much of contemporary physics as well as quantum chemistry and materials science is concerned with describing dynamic properties of correlated quantum many-body systems. Quantum Monte Carlo (QMC) methods [1–3] are uniquely capable of providing approximation-free simulations of quantum many-body systems that interact via the Coulomb potential. Unfortunately, the applications of QMC simulations are limited in practice because extracting dynamic properties from a simulation is notoriously difficult. Simulations usually generate estimates of N -body imaginary-time correlation functions (ITCFs) [4–7], not real-time correlations. Dynamic properties, denoted by spectral function $S(\omega)$, are real-time (or real-frequency) quantities and can only be estimated by analytically continuing imaginary-time correlation functions

$F(\tau)$ back to real time [8]:

$$F(\tau) = \int_{-\infty}^{\infty} d\omega \mathcal{A}(\tau, \omega) S(\omega), \quad (1)$$

with $\mathcal{A}(\tau, \omega)$ being a known transformation. Here, we consider any $\mathcal{A}(\tau, \omega)$ that, when discretized, yields an ill-conditioned matrix. As examples, one might consider $\mathcal{A}(\tau, \omega) = e^{-\tau\omega}$ or $\mathcal{A}(\tau, \omega) = e^{-\tau\omega}/(1 \pm e^{-\beta\omega})$ [9–11]. The present authors typically consider $\mathcal{A}(\tau, \omega) = e^{-\tau\omega}$, which relates the density-density ITCF on the left-hand Side (LHS) with the dynamic structure factor under the integral. The latter is the key observable, e.g., in x-ray scattering experiments with warm dense matter [12–15] and neutron scattering experiments with ultracold atoms [16–19]. Additionally, $\mathcal{A}(\tau, \omega) = e^{-\tau\omega}\Theta(\omega)$ is typically recovered in the temperature $T = 1/\beta = 0$ limit of various analytic continuation (AC) problems [9,10], making it of broader interest.

The task at hand is thus the numerical inversion of Eq. (1). Across the many fields that use QMC methods, there is no universally best AC algorithm for extracting dynamic properties. Many methods have been developed for the AC problem. Three main camps are fitting parametrized models [20–28] (this includes neural nets, which are models with an enormous number of parameters), stochastic optimization [29–36],

*Contact author: t.chuna@hzdr.de

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

and regularized optimization [37–39]. The most ubiquitous method is regularized optimization, for example, solving the maximum entropy method (MEM) [8,9,18,40]

$$S = \max_x -\chi^2[x | b] + \alpha S_{SJ}[x | \mu]. \quad (2)$$

Here, Eq. (1) has been discretized such that $A_{ij} = \mathcal{A}(\tau_i, \omega_j)$, $x_j = S(\omega_j)$, and $b_i = F(\tau_i)$. The chi-squared goodness-of-fit metric $\chi^2[x | b] = \frac{1}{2} \|Ax - b\|_{C^{-1}}$ weights the residual between the QMC data for the ITCF and a proposed solution x by the statistical error of the data C . However, the condition number of A is large so the goodness-of-fit metric is not sufficiently constraining and the Shannon-Jaynes entropy $S_{SJ}[x | \mu] = \sum_i x_i - \mu_i - x_i \ln x_i / \mu_i$ weighted by the regularization parameter α penalizes solutions that deviate from the prior μ . Together, these terms create a well-posed strictly convex problem where x is constrained to be positive. However, the default model introduces a bias, which means that though the QMC data are approximation-free, the final AC estimate is not.

When new methods are developed, the publications typically contain comparisons with solving Eq. (2); we emphasize parts of this extensive literature. First, solving Eq. (2) is often a limiting case of a more general method. Specifically, Shi *et al.* [27] show that Eq. (2) is the limit of a single-layer neural network, Benedix Robles *et al.* [39] show that Eq. (2) is the Dirac delta limit of kernel methods, and Beach [41] shows that Eq. (2) follows from a “mean-field” expansion of stochastic methods. Second, the use of Eq. (2) is often debated because practitioners traditionally rely on Bryan’s modified Levenberg-Marquardt optimization algorithm [42,43], which is controversial because it neglects singular value decomposition (SVD) basis vectors corresponding to small singular values [10,27,44,45]. Rothkopf [46,47] has explored ways to improve the basis and recently Chuna *et al.* [48] have provided an algorithm that can keep all the basis vectors at a reduced computational cost by solving the dual optimization problem. Third, and likely related to the basis issues, Bryan’s algorithm typically produces smoother solutions when compared to stochastic algorithms [41,46,49,50]. Such comparisons give an intuition of the approach’s mean squared error (MSE), but there is no clear methodological victor [35,51,52] and very few MSE formulas are developed in the literature [53]. Fourth, publications are considering the noiseless limit as the best approach to analytic continuation [9,22,34,37,40,51,54,55]. Generally, these publications find that as noise decreases Bryan’s algorithm performs equally well as other numerical methods [e.g., neural nets [20–22], stochastic sampling and averaging (Figs. 6, 7, and 46 of Ref. [34]), and dual Newton MEM [48]].

In this work, we investigate when solving Eq. (2) constitutes an acceptable alternative to using stochastic methods. We find that when the Bayesian prior is near the true solution, then Beach’s mean-field approximation [41] is fulfilled and Eq. (2) is valid; this demonstrates that Beach’s approximation is satisfied by a “real-world” scenario. Additionally, we investigate when solving Eq. (2) via Bryan’s algorithm is acceptable. We find that when the noise is near zero or the Bayesian prior is near the true solution, then Bryan’s algorithm constitutes an appropriate choice. This is because the estimator of Eq.

(2) is linear in those limits and, per Rothkopf [44], Bryan’s controversial algorithm relies on the assumption that the estimator is linear. We also investigate whether practitioners get a better return on investment from improving the Bayesian prior or reducing the noise. We develop MSE formulas of Eq. (2) in these limits and find that the MSE in the noiseless limit scales with a numerically infinite coefficient, while the MSE for the improved-prior limit does not. We take this to imply that practitioners get a better return on investment from improving the Bayesian prior. The major results of the paper are the analytic arguments listed above, and we support these arguments with a numerical investigation of the double Gaussian problem suggested by Goulko *et al.* [51], applying the dual Newton algorithm and Bryan’s algorithm to the problem.

This paper is organized as follows. Section II formulates the mean squared error of Eq. (2) in the noiseless and improved-prior limits and numerically investigates synthetic data generated from the double-peak Gaussian problem [51]. Section III shows that the improved-prior limit implies the mean-field assumption that Beach used to reduce stochastic methods to entropy maximization. Section IV presents the conclusions.

II. THE MEAN SQUARED ERROR OF THE MAXIMUM ENTROPY ESTIMATE

A. Definition of the mean squared error in terms of the bias and variance

The MSE expresses, on average, the 2-norm distance of your estimator from the true solution. The MSE can be expressed as a bias-variance decomposition, and for a one-dimensional estimator, derivations are commonplace in introductory textbooks [56]. A multivariate estimator $\hat{x}$ has a bias-variance decomposition given by summing over its dimensions. This is expressed as

$$\text{MSE} \equiv \mathbb{E}[\|x_0 - \hat{x}\|_2^2] = \|\text{Bias}(\hat{x})\|_2^2 + \text{Tr}\{\text{Cov}(\hat{x})\}. \quad (3a)$$

The first term is essentially sum of squares over the componentwise bias,

$$\text{Bias}(x_i) = \mathbb{E}[(x_0)_i - \hat{x}_i], \quad (3b)$$

where x_0 is the true solution. The second term is the variance, given by the trace of the covariance matrix,

$$\text{Cov}(\hat{x})_{ij} = \mathbb{E}[(\hat{x}_i - \mathbb{E}[\hat{x}_i])(\hat{x}_j - \mathbb{E}[\hat{x}_j])]. \quad (3c)$$

The bias-variance decomposition breaks down the MSE into two components: the systematic error and the statistical error. Some additional commentary may be found online [57,58].

To use expressions (3b) and (3c), we must have a closed form of the estimator $\hat{x}$. However, the optimum of the entropic regularized cost function (2) is given by [41,53]

$$A^\top C^{-1} A \hat{x} - A^\top C^{-1} b + \alpha \ln \hat{x} / \mu = 0 \quad (4)$$

and, as pointed out by Rothkopf [44], the term $\ln \hat{x} / \mu$ prevents us from isolating $\hat{x}$ and linearly parametrizing the solution; expansions are needed. Previous investigations of the MSE expanded the solution $\hat{x} = x_0 + \Delta \hat{x}$ about the true solution x_0 [53]; we do not take that approach here. Instead, we consider, for fixed regularization weight α , the noiseless limit

of uncorrelated data with equal variance, i.e., $C^{-1} \approx \sigma^{-2}I$, where

$$\sigma^2 \ll 1 \quad (5)$$

and the limit that the Bayesian prior is near the true solution, which we refer to as the “improved-prior” limit

$$\frac{|\hat{x}_i - \mu_i|}{\mu_i} \ll 1. \quad (6)$$

B. MSE in the noiseless limit

To compute the MSE, we insert Eq. (5) into Eq. (4) yielding

$$\frac{1}{\sigma^2}(A^\top A \hat{x} - A^\top b) + \alpha \ln \hat{x}/\mu = 0. \quad (7)$$

In the noiseless limit, the first term dominates. We drop the logarithmic term to recover the ordinary least squares (OLS) estimate. By the Gauss-Markov theorem, this is the best linear unbiased estimator (BLUE), with variance

$$\text{Cov}(\hat{x}) = \sigma^2(A^\top A)^{-1}. \quad (8)$$

We use the singular value decomposition $A = U \Sigma V^\top$ and cyclicity of the trace to arrive at

$$\text{Tr Cov}(\hat{x}) = \sigma^2 \text{Tr}\{\Sigma^{-2}\}, \quad (9)$$

which exposes how the singular values affect the MSE.

C. MSE in the improved-prior limit

Evaluating Eq. (3) for an improved prior (6) requires a more detailed treatment. To start, we must compute $\hat{x}$ and $\mathbb{E}[\hat{x}]$. Beginning from Eq. (4), we rewrite $\hat{x}/\mu = 1 + (\hat{x} - \mu)/\mu$ and linearize the logarithm by Taylor expanding

$$A^\top C^{-1} A \hat{x} - A^\top C^{-1} b + \alpha \frac{\hat{x} - \mu}{\mu} \approx 0. \quad (10)$$

This estimator resembles the estimator arising from the common L_2 -distance regularization, but with $x - \mu$ weighted by $1/\mu$. Next, we substitute $b = b_0 + \varepsilon = Ax_0 + \varepsilon$ and insert $0 = A^\top C^{-1} A(\mu - \mu)$. After simplification, this produces

$$\hat{x} = \mu + M A(x_0 - \mu) + M \varepsilon, \quad (11a)$$

where

$$M = \left(A^\top C^{-1} A + \alpha \text{diag}\left(\frac{1}{\mu}\right) \right)^{-1} A^\top C^{-1}. \quad (11b)$$

Assuming $\mathbb{E}[\varepsilon] = 0$ and taking the expectation value yields

$$\mathbb{E}[\hat{x}] = \mu + M A(x_0 - \mu). \quad (12)$$

We may now compute the bias (3b) by substituting Eqs. (11) and (12). We arrive at the expression

$$\text{Bias}[\hat{x}] = \|(I - M A)(x_0 - \mu)\|_2^2. \quad (13)$$

Grouping $H = M A$ and distributing terms yields

$$\text{Bias}[\hat{x}] = \|x_0 - \mu\|_2^2 - 2\|x_0 - \mu\|_H^2 + \|x_0 - \mu\|_{H^\top H}^2. \quad (14)$$

To expose how the singular values affect the bias, we assume $C = \sigma^2 I$ and simplify via the SVD $A = U \Sigma V^\top$:

$$H = \sigma^{-2} V \left(\sigma^{-2} \Sigma^2 + \alpha V^\top \text{diag}\left(\frac{1}{\mu}\right) V \right)^{-1} \Sigma^2 V^\top, \quad (15)$$

$$H^\top H = \sigma^{-4} V \Sigma^2 \left(\sigma^{-2} \Sigma^2 + \alpha V^\top \text{diag}\left(\frac{1}{\mu}\right) V \right)^{-2} \Sigma^2 V^\top. \quad (16)$$

Notice that for $\sigma \ll 1$, we can drop the α terms, yielding $H = H^\top H = I$ and $\text{Bias}[\hat{x}] = 0$ as is expected.

Next, we compute the trace of the covariance. First, we substitute Eq. (11) into Eq. (3c) and cancel terms, which leaves behind only the noise term:

$$\text{Tr Cov}(\hat{x}) = \text{Tr}\{\text{Cov}[M \varepsilon]\}. \quad (17)$$

It is interesting to assume $C = \sigma^2 I$ and simplify via the SVD to expose how the singular values affect the statistical error. We pass M through the covariant operator and use cyclicity of the trace to produce

$$\text{Tr Cov}(\hat{x}) = \sigma^2 \text{Tr} \left\{ \Sigma^2 \left(\Sigma^2 + \sigma^2 \alpha V^\top \text{diag}\left(\frac{1}{\mu}\right) V \right)^{-2} \right\}. \quad (18)$$

Notice in the limit that $\sigma \ll 1$ we recover Eq. (9), which is a nice sanity check.

D. Numeric investigation of MSE

In this section, we support our analytic arguments with numeric results demonstrating that improving the prior has a better return on investment than reducing the noise and that in the noiseless limit and improved-prior limit Bryan’s algorithm becomes valid. The associated codes are freely available on GitHub [59,60].

1. Description of problem

We consider synthetic data produced from the double Gaussian problem presented in Ref. [51]:

$$x_0 = \sum_{i=1}^2 \frac{c_i}{\sigma_i} e^{-\frac{(\omega - z_i)^2}{2\sigma_i^2}}, \quad (19)$$

where the moments are defined as $c_1 = 0.62$, $\sigma_1 = 0.12$, $z_1 = 0.74$ and $c_2 = 0.41$, $\sigma_2 = 0.064$, $z_2 = 2.93$, the transformation kernel is defined as

$$A_{ij} = e^{-\tau_i \omega_j}, \quad (20)$$

and the grid size is $N_\omega = 150$, $\omega \in [4.0/N_\omega, 4.0]$ and $N_\tau = 30$, $\tau \in [0, 5]$. A plot of x_0 is given in Fig. 1. This problem has been of particular interest in the analytic continuation community [23,34]. We generate $N_s = 100$ samples of b by adding Gaussian noise with standard deviation σ_0 that is scaled to element b^0 . This is expressed as

$$b_i^s \sim \mathcal{N}(b_i^0, (b_i^0 \sigma_0)^2), \quad (21)$$

where $b_0 = A x_0$. From these samples, we estimate $b_i = \text{Avg}_s(b_i^s)$ and $C_{ij} = \text{Var}_s(b_i^s)/N_s \delta_{ij}$. So, on average, the data error is $\sigma_0/\sqrt{N_s}$. We select the regularization weight via the χ^2 -kink algorithm [61].

To investigate the impact of the Bayesian prior, we make the prior a convex combination of the true solution x_0 and the

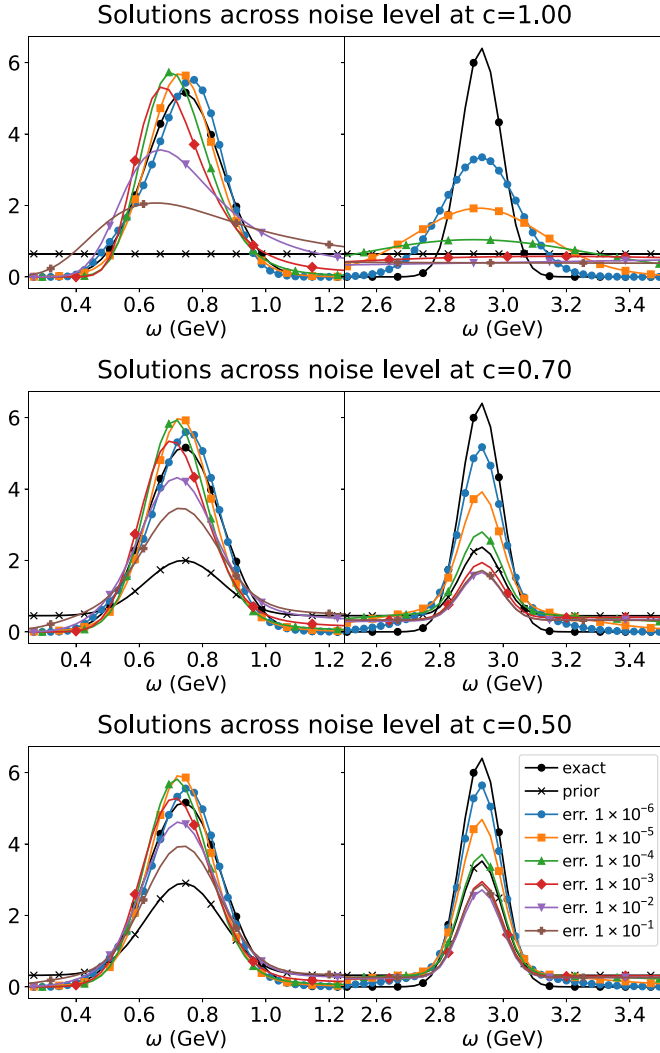


FIG. 1. Plot of the double Gaussian test problem from Goulko *et al.* [51], where the left plot is the first peak and the right plot is the second peak. For visualization, we neglect the large, flat middle region. From top to bottom, we improve the quality of the Bayesian prior, defined in Eq. (22), with $c = 1.0, 0.7, 0.5$. Within each plot, we present different noise levels.

uniform prior (equivalently flat) x_{flat} via

$$\mu(c) = (1 - c)x_0 + c x_{\text{flat}}, \quad c \in [0, 1]. \quad (22)$$

Essentially, c parametrizes the line through solution space from the flat model ($c = 1$) to the true solution ($c = 0$). Note that the randomly sampled data have no guarantee that its minimum is located at the true solution. In theory, for very small c , the ITCF data may actually make the result worse. As such, we consider $c \in [0.05, 1.0]$; plots of $\mu(c)$ at $c = 1.0, 0.7, 0.5$ are given in Fig. 1.

To investigate the noiseless limit $\tilde{\alpha} = \sigma^2 \alpha \rightarrow 0$, we logarithmically vary $\frac{\sigma_0}{\sqrt{N_S}} \in [10^{-6}, 10^{-1}]$. For context, in the literature the noiseless limit is typically considered to begin around 10^{-5} , while realistic data are considered 10^{-3} . Additionally, α is not a fixed value, but selected by the χ^2 -kink algorithm [61]. Therefore, it may be the case that as $\sigma^2 \rightarrow 0$ χ^2 kink selects $\alpha \rightarrow \infty$, violating the approximations made when comput-

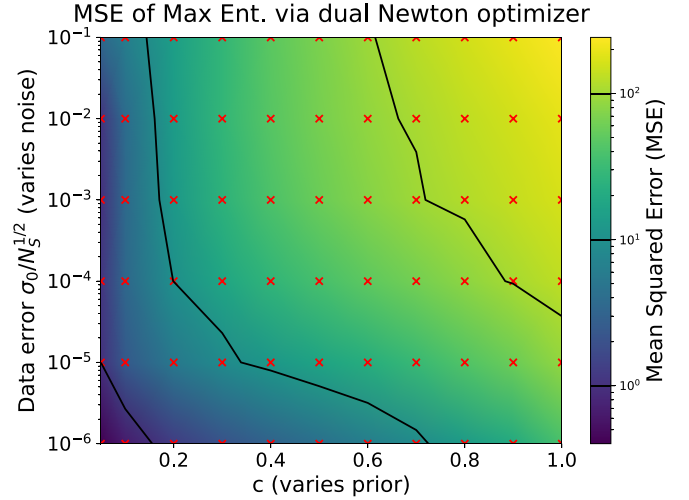


FIG. 2. Heatmap of the mean squared error for the solution produced by the χ^2 -kink algorithm with a Newton optimizer on the dual problem varied over noise [noise defined in Eq. (21)] and prior quality [c defined in Eq. (22)]. Contour levels are marked as a black horizontal line in the MSE colorbar. The grid of red $\times$'s indicates the c and σ_0 values that were used; Python's matplotlib smoothly interpolates between these values. Notice that the y axis is logarithmic, while the x axis is linear.

ing the noiseless limit in Sec. II. However, in practice, this does not happen. For example, for $c = 0.5$ and $\sigma_0/\sqrt{N_S} = 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}$, the corresponding values of $\sigma^2 \alpha$ are $10^{-9}, 10^{-8}, 10^{-6}, 10^{-3}, 10^{-2}, 10^0$, which monotonically also increase with σ_0 .

2. Numeric results

We investigate the double Gaussian problem using the dual formulation of the entropic regularization [48]. Each plot in Fig. 1 shows the estimate obtained for different quality priors $c = 1.0, 0.7, 0.5$. For each prior, we show the estimate converging toward the true solution in the noiseless limit in Fig. 1. We see that a flat prior $c = 1.0$ produces low-quality results across all noise levels, but with a small improvement in the prior $c = 0.7$, the estimate remains stable at large noise and yields quality results as the noise is reduced; this improvement continues as the prior improves.

Next, we investigate the mean squared error $\mathbb{E}[\|x_0 - \hat{x}\|_2^2]$. To estimate this quantity, we conduct $N_r = 100$ runs of the problem described above for given c and σ_0 , and then average over N_r . The MSE estimates are presented in Fig. 2. Since the y axis is logarithmic and the x axis is linear, improving the default model clearly has a greater impact on the error. This matches the intuition created by the analytic formulas.

Finally, in Fig. 3, we plot a heatmap of the deviation between the solution obtained using Bryan's algorithm and the solution obtained using the dual Newton optimizer. To estimate this deviation, we run both approaches on a given run and compute the 2-norm distance. Then, we average over the $N_r = 100$ differences to estimate the average deviation. We see that, as expected, in both the noiseless limit and the improved-prior limit there is diminishing deviation between

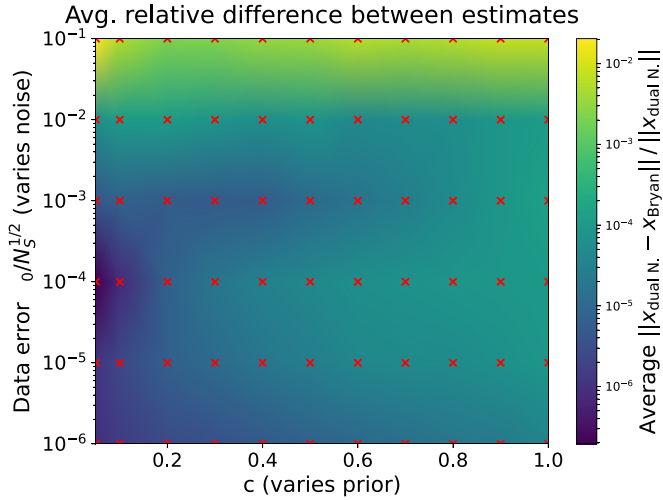


FIG. 3. Heatmap of the average relative distance between $x_{\text{dual } N}$ and x_{Bryan} , which are the solutions produced by χ^2 -kink algorithm with a Newton optimizer on the dual problem and the solution produced by the χ^2 -kink algorithm with Bryan's modified Levenberg-Marquardt optimizer on the primal problem. In the improved-prior limit (i.e., $c \rightarrow 0$) and the noiseless limit, the deviations between the estimates go to zero.

the solution produced using Bryan's algorithm and the solution produced using the dual Newton algorithm.

III. IDENTIFYING MEM AS A SPECIAL CASE OF STOCHASTIC METHODS IN THE IMPROVED-PRIOR LIMIT

Finally, we revisit Beach's seminal work [41] to show that the improved-prior limit satisfies Beach's mean-field approximation. Beach shows that the MEM cost function is recovered from the mean-field (i.e., saddle point) approximation. In the language of this paper, this amounts to assuming that the covariance (3c) scaled by the default model is negligible:

$$\mathbb{E} \left[\left(\frac{\mathbb{E}[x_i]}{\mu_i} - \frac{x_i}{\mu_i} \right) \left(\frac{\mathbb{E}[x_j]}{\mu_j} - \frac{x_j}{\mu_j} \right) \right] \approx 0. \quad (23)$$

Equation (23) is essentially the Ginzburg criterion (Sec. 2 in Ref. [62]) and the LHS can be rewritten in the typical representation of the covariance as

$$\mathbb{E} \left[\left(\frac{\mathbb{E}[x_i]}{\mu_i} - \frac{x_i}{\mu_i} \right) \left(\frac{\mathbb{E}[x_j]}{\mu_j} - \frac{x_j}{\mu_j} \right) \right] = \frac{\text{Cov}(x_i, x_j)}{\mu_i \mu_j}. \quad (24)$$

Following Asakawa, Hatsuda, and Nakahara's monkey argument (Appendix A in Ref. [9]), we assume the covariance between x_i and x_j is described by a multinomial distribution. Essentially, in fixed-grid stochastic sampling, there are M blocks to be distributed into N bins. For a particular arrangement of blocks x , where x_i is the number of blocks in bin i , the probability of a block being in bin i is $p_i = \hat{x}_i/M$ ($\mathbb{E}[x] = \hat{x}$) and the probability that the i th bin receives x_i blocks is given by the multinomial distribution. Thus, the covariance between bins is $\text{Cov}(x_i, x_j) = -M p_i p_j$; inserting this into Eq. (24)

yields

$$\mathbb{E} \left[\left(\frac{\mathbb{E}[x_i]}{\mu_i} - \frac{x_i}{\mu_i} \right) \left(\frac{\mathbb{E}[x_j]}{\mu_j} - \frac{x_j}{\mu_j} \right) \right] = \frac{-1}{\mu_i \mu_j} M^{-1} \hat{x}_i \hat{x}_j. \quad (25)$$

Notice now that Beach's mean-field approximation (23) is satisfied by assuming $M \rightarrow \infty$. However, we do not need this assumption in the improved-prior limit, where $\hat{x}/\mu \approx 1$ and Eq. (25) is reduced to M^{-1} . So, even small M values ($\mathcal{O}[10^2]$) satisfy the Ginzburg criterion (23).

Intuitively, we may have expected that the improved-prior limit reduces stochastic sampling to the MEM. In statistical mechanics textbooks, the mean-field approximation is the assumption that the variable of interest has only small perturbations about the mean (e.g., see Sec. 16 in Ref. [63]). We express this as

$$\frac{x}{\mu} = \frac{\hat{x}}{\mu} + \frac{\delta x}{\mu}. \quad (26)$$

This would precisely be the case if you initialize your stochastic sampling algorithm to a prior μ near the true solution. Essentially, the samples of x will only be small perturbations about μ because the sampler will reject large jumps away from this already small χ^2 location.

Finally, we suggest that the noiseless limit does not satisfy Beach's mean-field approximation. Notice in the noiseless limit that, because the least squares estimate is the BLUE, then $\hat{x} = x_0$, which does not immediately reduce Eq. (25) to M^{-1} . To add strength to the suggestion, the free term and the interaction term in Beach's seminal paper [Eqs. (23)–(25) of Ref. [41]] are both scaled by $1/\sigma^2$. Thus, the noiseless limit would not uniquely affect the interaction term and thus would not facilitate the expansion.

IV. CONCLUSIONS

This paper helps practitioners select the best method to tackle the analytic continuation problem based on the quality of their Bayesian prior and emphasizes to practitioners that improving the prior will yield better results than improving data. Our three major results are as follows: (1) If the Bayesian prior is near the true solution, then stochastic sampling methods reduce to solving the entropy regularized least squares problem (2); (2) if the Bayesian prior is near the true solution, then the estimator reduces to a linear estimate and thus Bryan's algorithm becomes valid; and (3) the best way to reduce the mean squared error of the estimate is to improve the Bayesian prior.

(1) In Sec. III, we demonstrated that the assumption of an improved prior (6) is sufficient to reduce stochastic methods to the MEM. Essentially, the improved-prior limit implies Beach's mean-field approximation. This result provides an intuitive situation in which the Beach's assumption is valid. Additionally, we suggest that the noiseless limit does not imply Beach's mean-field approximation.

(2) In Sec. II, we demonstrated that the noiseless and improved-prior limits reduced the MEM estimator $\hat{x}$ to a linear estimator. As per Rothkopf [44], this implies that Bryan's null-space trick is valid in these limits. We also observed this to be the case in our numeric investigation (see Fig. 3). This result explains a host of results in the literature where

MEM performs equally well as other methods in the noiseless limit [20–24]. The key takeaway is that other algorithms are expected to outperform the MEM when the signal-to-noise ratio is large because Bryan’s algorithm is not valid in that limit.

(3) In Sec. II, we have shown that in the noiseless limit, the MSE is $\sigma^2 \text{Tr } \Sigma^{-2}$. Thus, quadratic improvements may be found with vanishing noise, but for the Laplace transformation $\text{Tr } \Sigma^{-2}$ is numerically infinite. Given the extreme cost of quantum Monte Carlo simulations, this is not an appealing approach. By comparison, in the improved-prior limit the bias (13) and the variance (18) contain quadratically vanishing terms that are scaled by matrix M , which is windowed by $\alpha \text{diag}(1/\mu)$ (i.e., small singular values do not lead to numerical infinity). As a result, you can reap the benefits of the quadratically diminishing terms. The MSE was estimated numerically for the double Gaussian problem to verify these claims (see Fig. 2). This adds nuance to the conventional wisdom that the prior should be uninformative.

Future work from the stochastic community may wish to investigate the intuition discussed in Sec. III near Eq. (26). Future applications of the MEM ought to avoid using the flat Bayesian prior because an extreme distance between the solution and prior violates the assumptions underlying the MEM algorithm. If detailed information about the structure of the solution is not known, then stochastic methods are the best option. The present authors encourage using models (which are exact in some limit) as Bayesian priors [8,64,65] or using approximate, sometimes classical, real-time estimates as a Bayesian prior. This has already been done in ring polymer molecular dynamics [66] and is possible for high-energy-density systems where the results of time-dependent density functional theory molecular dynamics are often Laplace transformed in order to be compared to QMC ITCF data [67,68].

Furthermore, future applications of the MEM ought to consider data-driven Bayesian priors, e.g., preprocess the data to create a prior. Within Ref. [69], it is argued that using a data-driven prior Bayesian is similar in spirit to detrending approaches in data science [56,70–72]. Early examples of intentionally uninformative data-driven priors were based on moment matching [8,52,53,73], i.e., using a Gaussian whose moments were computed from the ITCF data. Here, we are encouraging the use of informative data-driven priors, which tend to be exact in some limit and are informed/fit to the

QMC data. For example, chemical models of the hydrogen dynamic structure factor have been fit to ITCF data [74], as have conductivity models been used for semiconductors [75] and conformal field theories [76]. Alternatively, the static approximation does not require fitting and has recently yielded promising results as a Bayesian prior; it has been developed independently in both lattice gauge theories [77] and the electron liquid [69,78]. In the electron liquid literature, the static approximation [79,80] derived from density response theory [13,81] is relatively common, but this is not the case in gauge theories.

ACKNOWLEDGMENTS

This work was partially supported by the Center for Advanced Systems Understanding (CASUS), financed by Germany’s Federal Ministry of Education and Research (BMBF) and the Saxon state government out of the State budget approved by the Saxon State Parliament. This work has received funding from the European Union’s Just Transition Fund (JTF) within the project Röntgenlaser-Optimierung der Laserfusion (ROLF), Contract No. 5086999001, cofinanced by the Saxon state government out of the State budget approved by the Saxon State Parliament. This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2022 research and innovation program (Grant Agreement No. 101076233, “PREXTREME”). T.D. gratefully acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) via Project No. DO 2670/1-1. Computations were performed on a Bull Cluster at the Center for Information Services and High-Performance Computing (ZIH) at Technische Universität Dresden and at the Norddeutscher Verbund für Hoch- und Höchstleistungsrechnen (HLRN) under Grant No. mv00024.

Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

DATA AVAILABILITY

The data that support the findings of this article are openly available [59].

- [1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, *Rev. Mod. Phys.* **73**, 33 (2001).
- [2] J. Anderson, *Quantum Monte Carlo: Origins, Development, Applications* (Oxford University Press, USA, 2007).
- [3] D. M. Ceperley, Path integrals in the theory of condensed helium, *Rev. Mod. Phys.* **67**, 279 (1995).
- [4] M. Boninsegni, N. V. Prokofev, and B. V. Svistunov, Worm algorithm and diagrammatic Monte Carlo: A new approach to continuous-space path integral Monte Carlo simulations, *Phys. Rev. E* **74**, 036701 (2006).
- [5] E. Rabani, D. R. Reichman, G. Krilov, and B. J. Berne, The calculation of transport properties in quantum liquids using the maximum entropy numerical analytic continuation method: Application to liquid para-hydrogen, *Proc. Natl. Acad. Sci. USA* **99**, 1129 (2002).
- [6] T. Dornheim, Z. Moldabekov, P. Tolias, M. Böhme, and J. Vorberger, Physical insights from imaginary-time density–density correlation functions, *Matter Radiat. Extremes* **8**, 056601 (2023).
- [7] T. Dornheim, Z. A. Moldabekov, and J. Vorberger, Nonlinear density response from imaginary-time correlation functions: *Ab initio* path integral Monte Carlo simulations of the warm dense electron gas, *J. Chem. Phys.* **155**, 054110 (2021).
- [8] M. Jarrell and J. Gubernatis, Bayesian inference and the analytic continuation of imaginary-time quantum Monte Carlo data, *Phys. Rep.* **269**, 133 (1996).

- [9] M. Asakawa, Y. Nakahara, and T. Hatsuda, Maximum entropy analysis of the spectral functions in lattice QCD, *Prog. Part. Nucl. Phys.* **46**, 459 (2001).
- [10] A. Rothkopf, Heavy quarkonium in extreme conditions, *Phys. Rep.* **858**, 1 (2020).
- [11] L. Huang and S. Liang, Stochastic pole expansion method for analytic continuation of the Green's function, *Phys. Rev. B* **108**, 235143 (2023).
- [12] S. H. Glenzer and R. Redmer, X-ray Thomson scattering in high energy density plasmas, *Rev. Mod. Phys.* **81**, 1625 (2009).
- [13] T. Dornheim, Z. A. Moldabekov, K. Ramakrishna, P. Tolias, A. D. Baczewski, D. Kraus, T. R. Preston, D. A. Chapman, M. P. Böhme, T. Döppner, F. Graziani, M. Bonitz, A. Cangi, and J. Vorberger, Electronic density response of warm dense matter, *Phys. Plasmas* **30**, 032705 (2023).
- [14] M. Schörner, M. Bethkenhagen, T. Döppner, D. Kraus, L. B. Fletcher, S. H. Glenzer, and R. Redmer, X-ray Thomson scattering spectra from density functional theory molecular dynamics simulations based on a modified Chihara formula, *Phys. Rev. E* **107**, 065207 (2023).
- [15] D. S. Bepalov, U. Zastrau, Z. A. Moldabekov, T. Gawne, T. Dornheim, M. Meshhal, A. Amouretti, M. Andrzejewski, K. Appel, C. Baetz, *et al.*, Momentum-resolved X-ray Thomson scattering benchmark of Electronic-response models in warm dense aluminium, *Phys. Rev. Lett.* **136**, 245102 (2026).
- [16] A. Filinov and M. Bonitz, Collective and single-particle excitations in two-dimensional dipolar Bose gases, *Phys. Rev. A* **86**, 043628 (2012).
- [17] T. Dornheim, Z. A. Moldabekov, J. Vorberger, and B. Militzer, Path integral Monte Carlo approach to the structural properties and collective excitations of liquid ^3He without fixed nodes, *Sci. Rep.* **12**, 708 (2022).
- [18] M. Boninsegni and D. M. Ceperley, Density fluctuations in liquid ^4He . Path integrals and maximum entropy, *J. Low Temp. Phys.* **104**, 339 (1996).
- [19] G. Ferré and J. Boronat, Dynamic structure factor of liquid ^4He across the normal-superfluid transition, *Phys. Rev. B* **93**, 104510 (2016).
- [20] H. Yoon, J.-H. Sim, and M. J. Han, Analytic continuation via domain knowledge free machine learning, *Phys. Rev. B* **98**, 245101 (2018).
- [21] R. Fournier, L. Wang, O. V. Yazyev, and Q. Wu, Artificial neural network approach to the analytic continuation problem, *Phys. Rev. Lett.* **124**, 056401 (2020).
- [22] X. Xie, F. Bao, T. Maier, and C. Webster, Analytic continuation of noisy data using Adams Bashforth residual neural network, *Discrete Contin. Dyn. Syst.* **15**, 877 (2022).
- [23] L. Wang, S. Shi, and K. Zhou, Reconstructing spectral functions via automatic differentiation, *Phys. Rev. D* **106**, L051502 (2022).
- [24] L. Wang, S. Shi, and K. Zhou, Unsupervised learning spectral functions with neural networks, *J. Phys.: Conf. Ser.* **2586**, 012158 (2023).
- [25] G. Aarts, K. Fukushima, T. Hatsuda, A. Ipp, S. Shi, L. Wang, and K. Zhou, Physics-driven learning for inverse problems in quantum chromodynamics, *Nat. Rev. Phys.* **7**, 154 (2025).
- [26] S. Groth, T. Dornheim, and J. Vorberger, *Ab initio* path integral Monte Carlo approach to the static and dynamic density response of the uniform electron gas, *Phys. Rev. B* **99**, 235122 (2019).
- [27] S. Shi, L. Wang, and K. 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).
- [28] T. Dornheim, S. Groth, J. Vorberger, and M. Bonitz, *Ab initio* path integral Monte Carlo results for the dynamic structure factor of correlated electrons: From the electron liquid to warm dense matter, *Phys. Rev. Lett.* **121**, 255001 (2018).
- [29] E. Vitali, M. Rossi, L. Reatto, and D. E. Galli, *Ab initio* low-energy dynamics of superfluid and solid ^4He , *Phys. Rev. B* **82**, 174510 (2010).
- [30] N. V. Prokof'ev and B. V. Svistunov, Spectral analysis by the method of consistent constraints, *JETP Lett.* **97**, 649 (2013).
- [31] S. Saccani, S. Moroni, and M. Boninsegni, Excitation spectrum of a supersolid, *Phys. Rev. Lett.* **108**, 175301 (2012).
- [32] N. S. Nichols, P. Sokol, and A. Del Maestro, Parameter-free differential evolution algorithm for the analytic continuation of imaginary time correlation functions, *Phys. Rev. E* **106**, 025312 (2022).
- [33] F. Bao, Y. Tang, M. Summers, G. Zhang, C. Webster, V. Scarola, and T. A. Maier, Fast and efficient stochastic optimization for analytic continuation, *Phys. Rev. B* **94**, 125149 (2016).
- [34] H. Shao and A. W. Sandvik, Progress on stochastic analytic continuation of quantum Monte Carlo data, *Phys. Rep.* **1003**, 1 (2023).
- [35] H.-T. Shu, H.-T. Ding, O. Kaczmarek, S. Mukherjee, and H. Ohno, A stochastic approach to the reconstruction of spectral functions in lattice QCD, *arXiv:1510.02901*.
- [36] 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).
- [37] J. Otsuki, M. Ohzeki, H. Shinaoka, and K. Yoshimi, Sparse modeling approach to analytical continuation of imaginary-time quantum Monte Carlo data, *Phys. Rev. E* **95**, 061302 (2017).
- [38] J. Otsuki, M. Ohzeki, H. Shinaoka, and K. Yoshimi, Sparse modeling in quantum many-body problems, *J. Phys. Soc. Jpn.* **89**, 012001 (2020).
- [39] A. Benedix Robles, P.-A. Hofmann, T. Chuna, T. Dornheim, and M. Hecht, PyLIT: Reformulation and implementation of the analytic continuation problem using kernel representation methods, *Comput. Phys. Commun.* **319**, 109904 (2026).
- [40] Y. Kora and M. Boninsegni, Dynamic structure factor of superfluid ^4He from quantum Monte Carlo: Maximum entropy revisited, *Phys. Rev. B* **98**, 134509 (2018).
- [41] K. Beach, Identifying the maximum entropy method as a special limit of stochastic analytic continuation, *arXiv:cond-mat/0403055*.
- [42] R. Bryan, Maximum entropy analysis of oversampled data problems, *Eur. Biophys. J.* **18**, 165 (1990).
- [43] S. F. Gull, Developments in maximum entropy data analysis, *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Springer, Dordrecht, 1989), pp. 53–71.
- [44] A. Rothkopf, Bryan's maximum entropy method—Diagnosis of a flawed argument and its remedy, *Data* **5**, 85 (2020).
- [45] M. Asakawa, Comment on “Heavy quarkonium in extreme conditions”, *arXiv:2001.10205*.
- [46] A. Rothkopf, Improved maximum entropy analysis with an extended search space, *J. Comput. Phys.* **238**, 106 (2013).

- [47] A. Kelly, A. Rothkopf, and J.-I. Skullerud, Bayesian study of relativistic open and hidden charm in anisotropic lattice QCD, *Phys. Rev. D* **97**, 114509 (2018).
- [48] T. Chuna, N. Barnfield, T. Dornheim, M. P. Friedlander, and T. Hoheisel, Dual formulation of the maximum entropy method applied to analytic continuation of quantum Monte Carlo data, *J. Phys. A: Math. Theor.* **58**, 335203 (2025).
- [49] S. Fuchs, T. Pruschke, and M. Jarrell, Analytic continuation of quantum Monte Carlo data by stochastic analytical inference, *Phys. Rev. E* **81**, 056701 (2010).
- [50] L. Huang, ACFlow: An open source toolkit for analytic continuation of quantum Monte Carlo data, *Comput. Phys. Commun.* **292**, 108863 (2023).
- [51] O. Goulko, A. S. Mishchenko, L. Pollet, N. Prokof'ev, and B. Svistunov, Numerical analytic continuation: Answers to well-posed questions, *Phys. Rev. B* **95**, 014102 (2017).
- [52] D. Bergeron and A.-M. S. Tremblay, Algorithms for optimized maximum entropy and diagnostic tools for analytic continuation, *Phys. Rev. E* **94**, 023303 (2016).
- [53] O. Gunnarsson, M. W. Haverkort, and G. Sangiovanni, Analytical continuation of imaginary axis data using maximum entropy, *Phys. Rev. B* **81**, 155107 (2010).
- [54] J. Fei, C.-N. Yeh, and E. Gull, Nevanlinna analytical continuation, *Phys. Rev. Lett.* **126**, 056402 (2021).
- [55] S. Kim, P. Petreczky, and A. Rothkopf, Quarkonium in-medium properties from realistic lattice NRQCD, *J. High Energy Phys.* **11** (2018) 088.
- [56] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, Springer Texts in Statistics (Springer, New York, 2013).
- [57] C. Yeh, Bias-variance decomposition of mean squared error, GitHub, 2019, <https://chrisyeh96.github.io/2019/07/18/bias-variance-decomposition.html>.
- [58] J. J. Meyer, The multivariate bias-variance decomposition, 2020, <https://johannesjakobmeyer.com/blog/005-multivariate-bias-variance-decomposition/>.
- [59] T. Chuna, InverseProblemSolvers, GitHub, 2025, <https://github.com/chunatho/InverseProblemSolvers>.
- [60] M. P. Friedlander, DualPerspective, GitHub, 2025, <https://github.com/MPF-Optimization-Laboratory/DualPerspective.jl>.
- [61] J. Kaufmann and K. Held, ana_cont: Python package for analytic continuation, *Comput. Phys. Commun.* **282**, 108519 (2023).
- [62] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2021).
- [63] M. E. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford University Press, Oxford, UK, 2023).
- [64] M. Jarrell, The maximum entropy method: Analytic continuation of QMC data, in *Correlated Electrons: From Models to Materials, Modeling and Simulation*, edited by E. Pavarini, E. Koch, F. Anders, and M. Jarrell (Forschungszentrum Jülich Zentralbibliothek, Verlag, Jülich, 2012), Vol. 2, Chap. 13.
- [65] N. G. Diamantis and E. Manousakis, Flat histogram quantum Monte Carlo for analytic continuation to real time, [arXiv:1411.5984](https://arxiv.org/abs/1411.5984).
- [66] S. Habershon, B. J. Braams, and D. E. Manolopoulos, Quantum mechanical correlation functions, maximum entropy analytic continuation, and ring polymer molecular dynamics, *J. Chem. Phys.* **127**, 174108 (2007).
- [67] Z. A. Moldabekov, S. Schwalbe, T. Gawne, T. R. Preston, J. Vorberger, and T. Dornheim, Applying the Liouville–Lanczos method of time-dependent density-functional theory to warm dense matter, *Matter Radiat. Extremes* **10**, 047601 (2025).
- [68] Z. A. Moldabekov, S. Schwalbe, U. Hernandez Acosta, T. Gawne, J. Vorberger, M. Pavanello, and T. Dornheim, Enhancing the efficiency of time-dependent density functional theory calculations of dynamic response properties, *npj Comput. Mater.* **12**, 168 (2026).
- [69] T. Chuna, N. Barnfield, J. Vorberger, M. P. Friedlander, T. Hoheisel, and T. Dornheim, Estimates of the dynamic structure factor for the finite temperature electron liquid via analytic continuation of path integral Monte Carlo data, *Phys. Rev. B* **112**, 125112 (2025).
- [70] Z. Wu, N. E. Huang, S. R. Long, and C.-K. Peng, On the trend, detrending, and variability of nonlinear and nonstationary time series, *Proc. Natl. Acad. Sci. USA* **104**, 14889 (2007).
- [71] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, 2nd ed. (Springer, New York, NY, 2002).
- [72] R. Adhikari and R. K. Agrawal, An introductory study on time series modeling and forecasting, [arXiv:1302.6613](https://arxiv.org/abs/1302.6613).
- [73] A. Lovato, S. Gandolfi, J. Carlson, S. C. Pieper, and R. Schiavilla, Electromagnetic and neutral-weak response functions of ^4He and ^{12}C , *Phys. Rev. C* **91**, 062501 (2015).
- [74] H. M. Bellenbaum, M. P. Böhme, M. Bonitz, T. Döppner, L. B. Fletcher, T. Gawne, D. Kraus, Z. A. Moldabekov, S. Schwalbe, J. Vorberger, and T. Dornheim, Estimating ionization states and continuum lowering from *ab initio* path integral Monte Carlo simulations for warm dense hydrogen, *Phys. Rev. Res.* **7**, 033016 (2025).
- [75] P. Buividovich, J. Ostmeyer, and A. Troisi, High-precision quantum Monte-Carlo study of charge transport in a lattice model of molecular organic semiconductors, *PoS LATTICE* **2024**, 067 (2024).
- [76] E. Katz, S. Sachdev, E. S. Sørensen, and W. Witczak-Krempa, Conformal field theories at nonzero temperature: Operator product expansions, Monte Carlo, and holography, *Phys. Rev. B* **90**, 245109 (2014).
- [77] P. Buividovich and B. Hind, Spectral reconstruction based on dimensional reduction in high-temperature gauge theories, [arXiv:2512.23560](https://arxiv.org/abs/2512.23560).
- [78] T. M. Chuna, J. Vorberger, P. Tolias, A. Benedix Robles, M. Hecht, P.-A. Hofmann, Z. A. Moldabekov, and T. Dornheim, Second roton feature in the strongly coupled electron liquid, *J. Chem. Phys.* **163**, 034117 (2025).
- [79] T. Dornheim, J. Vorberger, S. Groth, N. Hoffmann, Z. Moldabekov, and M. Bonitz, The static local field correction of the warm dense electron gas: An *ab initio* path integral Monte Carlo study and machine learning representation, *J. Chem. Phys.* **151**, 194104 (2019).
- [80] T. Dornheim, Z. A. Moldabekov, and P. Tolias, Analytical representation of the local field correction of the uniform electron gas within the effective static approximation, *Phys. Rev. B* **103**, 165102 (2021).
- [81] G. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2008).