Heating and cooling in self-consistent many-body simulations

Yang Yu [*], Sergei Iskakov [,†] and Emanuel Gull [‡]
Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA

We present a temperature-extrapolation technique for self-consistent many-body methods, which provides a causal starting point for converging to a solution at a target temperature. The technique employs the Carathéodory formalism for interpolating causal matrix-valued functions and is applicable to various many-body methods, including dynamical mean-field theory, its cluster extensions, and self-consistent perturbative methods such as the self-consistent GW approximation. We show results that demonstrate that this technique can efficiently simulate heating and cooling hysteresis at a first-order phase transition, as well as accelerate convergence.

I. INTRODUCTION

When conducting experiments in condensed-matter physics, it is common practice to investigate the temperature dependence of observables while keeping other parameters constant. A particular example are specific-heat or transport measurements, which are often used as a preliminary probe to identify intriguing temperature-dependent behavior. For example, in a first-order coexistence regime, heating and cooling curves may reveal history-dependent hysteresis.

In theoretical calculations using self-consistent finite-temperature field theories, changing the temperature of a system is generally not practical because it causes a shift of Matsubara frequencies [1,2]. Extrapolating the lowest Matsubara frequency during cooling can be problematic, resulting in noncausal solutions. Therefore, implementing heating and cooling protocols in self-consistent finite-temperature field theories, in analogy to heating and cooling measurements in experiment, is an unresolved issue.

Self-consistent finite-temperature simulation methods include nonperturbative techniques, such as the dynamical mean-field theory (DMFT) [3] and its cluster variants [4–7]; self-consistent perturbative methods, such as GW [8–13], second-order perturbation theory [14–17], fluctuation-exchange or T-matrix approximations [18,19], as well as bold-line diagrammatic Monte Carlo methods [20–23]; and combinations of embedding theories with perturbation theory [24–26].

In these methods, solutions are obtained through an iterative process that involves starting with an initial guess and continuing the process until self-consistency is achieved. The number of iterations required to reach convergence is directly related to how close the starting point is to the iteration fixed point. A "good" starting point can significantly reduce the computational effort required for the simulation, whereas a

"bad" starting point may result in iterations that diverge, iterate in limit cycles, or even converge to unphysical fixed points [27]. In parameter regimes where first-order coexistence occurs, multiple physical fixed points may exist [28–32].

This paper addresses the challenge of generating better starting points and implementing heating and cooling protocols in self-consistent many-body methods. It presents a solution that guarantees a causal starting point by using a converged solution at a different temperature. The proposed method relies on the Carathéodory formalism [33] for interpolating causal matrix-valued functions. Originally developed for analytic continuation of matrix-valued Matsubara functions to real frequencies, this formalism can be extended to temperature extrapolation, which involves evaluating an interpolant at different Matsubara frequencies. We demonstrate the effectiveness of this approach in obtaining improved starting points in the context of DMFT and real-material perturbation theory. Additionally, we examine heating and cooling hysteresis in the context of a first-order phase transition.

The paper proceeds as follows: In Sec. II we introduce the formalism. Section III contains a brief description of a pedagogical implementation of the method which we provide as a supplement. Section IV contains results for temperature extrapolation, convergence acceleration, and first-order hysteresis. Finally, Sec. V contains our conclusions.

II. METHOD

The central object of this paper is a causal matrix-valued fermionic Matsubara function which may represent a Green's function, a self-energy, or a cumulant [34]. The Matsubara function is expressed as a three-dimensional tensor, $G_{ij}(i\omega_n)$, which associates a matrix (identified by the indices i and j) with every Matsubara frequency $\omega_n = (2n+1)\pi/\beta$ (n denotes an integer, β the inverse temperature). Its continuation to the complex plane [33] is denoted $G_{ij}(\tilde{z})$. $iG(\tilde{z})$ is a Carathéodory function (up to a convention-dependent minus sign), such that $[iG(\tilde{z}) + (iG(\tilde{z}))^{\dagger}]/2$ is a positive semidefinite matrix for any $\tilde{z}$ in the upper half of the complex plane.

—————
[*]umyangyu@umich.edu
[†]siskakov@umich.edu
[‡]egull@umich.edu

The matrix structure of G depends on the application. Matrices may be scalar, diagonal, block-diagonal, or fully dense. The scalar case typically appears in single-site single-orbital cases, such as single-site DMFT. The diagonal case is often encountered in momentum-space simulations, such as cluster DMFT [4–7]. The general multi-orbital case with dense matrices typically appears in real-material *ab initio* calculations [25,26].

The frequency dependence of G is known at discrete Matsubara frequencies ω_n. In the simplest case, the frequencies are uniformly spaced positive fermionic Matsubara frequencies $\omega_n = (2n + 1)\pi/\beta$, $n = 0, 1, \ldots, N - 1$. More generally, data are provided on a nonuniform frequency grid using Chebyshev [35], intermediate representation (IR) [36], Legendre [37], spline [38], or discrete Lehmann representation [39] schemes with associated Fourier transforms [40,41]. For both uniform and nonuniform grids, the derivations in this section remain the same, with n denoting the index that enumerates the Matsubara frequencies ω_n. Bosonic Matsubara response functions $\Pi(\tilde{z})$ are not directly related to Carathéodory functions. However, $\Pi(\tilde{z})\tilde{z}$ and $\Pi(\tilde{z})/\tilde{z}^*$ correspond to Carathéodory functions [42], and Nogaki and Shinaoka [43] recently investigated a related mapping of bosonic functions.

Extrapolating in temperature requires the transformation of one Matsubara grid to another one corresponding to a different temperature. We propose to construct a causal matrix-valued interpolant through $G(i\omega_n)$ and evaluating it at the Matsubara points corresponding to the changed temperature grid to realize the temperature extrapolation. Our approach assumes that the causal matrix-valued functions $G(\tilde{z})$ at two proximate temperatures, β and β', share a similar analytical structure. This allows us to approximate the unknown causal function at β' using known causal function at β. Subsequently, the unknown Matsubara function at β' can then be approximated by evaluating the interpolant of the causal function at β on the Matsubara grid corresponding to β'. The approximated Matsubara function can later serve as the starting point for the self-consistent calculation at β'. From the expression

$$G(\tilde{z}) = \int d\omega \frac{A(\omega)}{\tilde{z} - \omega}, \qquad (1)$$

we observe that the temperature dependence of $G(\tilde{z})$ arises solely from the spectral function $A(\omega)$. Although extrapolating $A(\omega)$ along the temperature axis is possible, it is not implemented in our current approach. We find that for the Matsubara functions, the variation in Matsubara grids for different temperatures usually has a more significant impact on its value compared with the changes in the spectral function $A(\omega)$. As a result, our temperature extrapolation approach works effectively at least for nearby temperatures. In fact, empirically, our extrapolation approach performs well even for relatively large temperature differences, as long as the system remains in the same phase, as demonstrated in Sec. IV C. The method's algorithmic steps follow those presented for analytic continuation in Ref. [33], with the only distinction that the interpolant is evaluated on the imaginary axis, rather than just above the real axis. The major steps of the algorithm are described below.

For a Carathéodory function $\tilde{F}(\tilde{z}) = iG(\tilde{z})$, $\tilde{z}$ in the upper half of the complex plane, we assume the values of $\tilde{F}$ are known at a set of points $\{\tilde{z}_n | n = 0, 1, \ldots, N - 1\}$. A Möbius transform

$$z = \frac{\tilde{z} - i}{\tilde{z} + i} \qquad (2)$$

maps the upper half of the complex plane to the unit disk, the real axis to the unit circle, and the Matsubara points to points on the real axis. Since a Carathéodory function F is a function defined on an open subset of the complex plane and exhibits the property $(F + F^\dagger)/2$ being positive semidefinite, $C(z) = \tilde{F}(\tilde{z})$ then defines a Carathéodory function on the unit disk $|z| < 1$. The function $C(z)$ is known at a set of points $\{z_n | n = 0, 1, \ldots, N - 1\}$ with the values of

$$C_n \equiv C(z_n) = \tilde{F}(\tilde{z}_n). \qquad (3)$$

The Cayley transform [44] and its inverse, defined as

$$S(z) = [I - C(z)][I + C(z)]^{-1}, \qquad (4)$$

$$C(z) = [I + S(z)]^{-1}[I - S(z)], \qquad (5)$$

map the Carathéodory function $C(z)$ to a matrix-valued Schur function $S(z)$ [45] and vice versa. A Schur function is defined as a function $S(z)$ on the unit disk with the property $||S(z)|| \leqslant 1$, where the matrix norm $||\cdot||$ is defined as the largest singular value of S, or equivalently the largest eigenvalue of $(SS^\dagger)^{1/2}$.

The problem of interpolating the Carathéodory function $C(z)$ is thus converted into the problem of determining an interpolant for the Schur function $S(z)$, where $S(z)$ is known at N discrete points, given as $S_n \equiv S(z_n)$, with $n = 0, 1, \ldots, N - 1$.

To identify the interpolant of the Schur function $S(z)$, we proceed to find a set of Schur functions $S^0(z), S^1(z), \ldots, S^m(z), \ldots, S^N(z)$, where each Schur function $S^m(z)$ is known at $N - m$ points $z = z_m, z_{m+1}, \ldots, z_{N-1}$ with the values of $S_n^m \equiv S^m(z_n)$. $S^N(z)$ is an arbitrary Schur function without any constraints. We require $S_n^0 = S_n$ such that $S^0(z)$ serves as the desired interpolant for the Schur function $S(z)$.

To establish a connection between the two Schur functions $S^m(z)$ and $S^{m+1}(z)$, a function $B^m(z)$ is introduced, enabling a progressive determination of $S^0(z)$ starting from $S^N(z)$ [33]. The relationship between $B^m(z)$ and $S^m(z)$ is given by

$$\frac{|z_m|(z_m - z)}{z_m(1 - z_m^* z)} B^m(z) = L^m \big[S^m(z) - S_m^m \big]$$

$$\times \big[I - S_m^{m\dagger} S^m(z) \big]^{-1} R^m, \qquad (6)$$

with

$$L^m = \big[I - S_m^m S_m^{m\dagger} \big]^{-\frac{1}{2}}, \qquad (7)$$

$$R^m = \big[I - S_m^{m\dagger} S_m^m \big]^{\frac{1}{2}}. \qquad (8)$$

An equivalent form of this relation is

$$S^m(z) = \big[I + V^m(z) S_m^{m\dagger} \big]^{-1} \big[V^m(z) + S_m^m \big], \qquad (9)$$

with

$$V^m(z) = \frac{|z_m|(z_m - z)}{z_m(1 - z_m^* z)} (L^m)^{-1} B^m(z)(R^m)^{-1}. \qquad (10)$$

The relation between $B^m(z)$ and $S^{m+1}(z)$ reads

$$S^{m+1}(z) = [I - K^m K^{m\dagger}]^{-\frac{1}{2}} [B^m(z) - K^m]$$
$$\times [I - K^{m\dagger} B^m(z)]^{-1} [I - K^{m\dagger} K^m]^{\frac{1}{2}}, \quad (11)$$

where K^m is an arbitrary matrix with $||K^m|| < 1$.

The freedom in choosing arbitrary $K^0, K^1, \ldots, K^{N-1}$, along with $S^N(z)$, allows for a full coverage of all possible interpolants for $S(z)$. The seemingly undetermined quantities in the aforementioned equations are the values of $\{ S_m^m \mid m = 0, 1, \ldots, N-1 \}$. Starting from the given $\{ S_n^0 \mid n = 0, 1, \ldots, N-1 \}$, as S_n^{m+1} are entirely determined by S_n^m and S_n^n by proceeding from Eq. (6) to Eq. (11), all S_m^m can be acquired in the process of calculating $\{ S_n^m \mid m = 0, 1, \ldots, N-1; n = m, m+1, \ldots, N \}$ from $m = 0$ to $m = N-1$. After obtaining all S_m^m and the corresponding L^m and R^m, one can reverse the procedure, transitioning from Eq. (11) to Eq. (9), to compute $S^m(z)$ from $S^{m+1}(z)$ with a chosen $S^N(z)$ as the starting point, until the desired interpolant $S^0(z)$ is attained. Finally, an interpolant for $C(z)$ can be derived using Eq. (5), and by multiplying $-i$ and substituting z with $\tilde{z}$ through the Möbius transform, the interpolant for $G(\tilde{z})$ is obtained.

For the purpose of this work, we take $\{\tilde{z}_n\} = \{i\omega_n \mid n = 0, 1, \ldots, N-1\}$ along with the corresponding $G(i\omega_n)$ as the input. Setting $\tilde{F} = iG$, and with Eqs. (2)–(4), we obtain the constraints $\{(z_n, S_n)\}$ for the interpolation problem of a Schur function $S(z)$. By taking $K^m = 0$ for all m, from Eq. (11), we obtain $S^{m+1}(z) = B^m(z)$ which leads to

$$S_n^{m+1} = \frac{z_m(1 - z_m^* z_n)}{|z_m|(z_m - z_n)} L^m [S_n^m - S_m^m][I - S_m^{m\dagger} S_n^m]^{-1} R^m \quad (12)$$

according to Eq. (6). This equation offers a concrete way to calculate all S_n^m, starting from $S_n^0 = S_n$. Subsequently, we can obtain S_m^m and the corresponding L^m and R^m. Assume the desired new Matsubara points for a different temperature β' are $\{i\omega_n' \mid n = 0, 1, \ldots, N-1\}$. Using the Möbius transform, the corresponding points in the domain of $S(z)$ are obtained as z_n'. By setting $S^N(z) = I$, and using the obtained $\{S_m^m, L^m, R^m\}$, the values of $S_{n'} \equiv S^0(z_n')$ can be determined by proceeding from Eq. (11) to Eq. (9). With the inverse Cayley transform in Eq. (5) and $G = -i\tilde{F}$, the desired values of the causal interpolant $G(i\omega_n')$ at $i\omega_n'$ are obtained as output.

III. IMPLEMENTATION

We provide a straightforward pedagogical implementation of the algorithm in Ref. [46]. The code is written in the python programming language, using no dependencies other than numpy. Unlike related real-frequency analytic continuation codes [33,47], we find that an implementation in standard double precision was sufficient to perform all calculations, and that the temperature extrapolation of noisy Monte Carlo data presented no difficulties.

The implementation requires a Green's function (or, equivalently, self-energy or cumulant) $G[n, i, j] = G_{ij}(i\omega_n)$ as a three-dimensional tensor and $w[n] = i\omega_n$ as a one-dimensional vector. After reading the data, all preprocessing steps are performed until $\{S_m^m, L^m, R^m\}$ are obtained and stored, as explained in Sec. II.
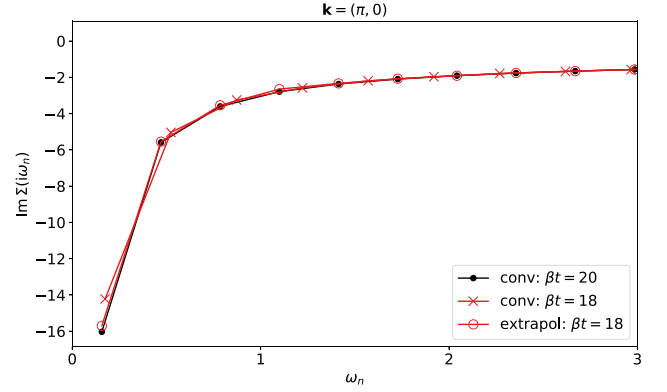

FIG. 1. Imaginary part of a cluster DMFT [7] self-energy, Im $\Sigma(i\omega_n)$, as a function of ω_n for the two-dimensional Hubbard model at the antinodal point $(\pi, 0)$ with $U = 6t$ (see text for calculation setup). Converged data at $\beta t = 18$ (red crosses) and $\beta t = 20$ (black dots) along with temperature extrapolation (red circle) from $\beta t = 18$ to $\beta t = 20$. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation.

The code also requires a set of frequencies $wp[n] = i\omega_n'$, provided as a one-dimensional vector. These frequencies should correspond to the Matsubara points at the new temperature β'. Using this data, the code evaluates the interpolated Green's function at the new Matsubara points and returns $G[n', i, j] = G_{ij}(i\omega_n')$ as a three-dimensional tensor.

The Ref. [46] contain detailed instructions and usage examples for the code.

IV. RESULTS

We showcase results for the temperature extrapolation technique applied to a range of typical self-consistent finite-temperature methods. Sec. IV A investigates the accuracy and convergence properties of the extrapolation technique. Section IV B examines the convergence acceleration resulting from an enhanced starting point provided by the extrapolation technique. This study is conducted specifically within the context of single-site DMFT and cluster DMFT calculations for the Hubbard model, as well as the self-consistent GW approximation for nickel oxide (NiO) real material calculations. Section IV C presents a comparison between the method introduced in this paper and other methods for generating starting points. Section IV D demonstrates the occurrence of hysteresis during heating and cooling processes at a first-order phase transition, utilizing the extrapolation technique.

A. Temperature extrapolation

In this section, we provide an illustration of temperature extrapolation in a cluster DMFT [7] calculation, as depicted in Fig. 1. Specifically, we present the self-energy at the antinodal point $(\pi, 0)$ for a simulation of the two-dimensional Hubbard model [48]. The simulation was performed on a 16-site (4×4) cluster, and the continuous-time auxiliary field impurity solver [49,50] was employed. The system was solved at half filling, in the paramagnetic state, with no next-nearest-neighbor hopping, and at an interaction strength
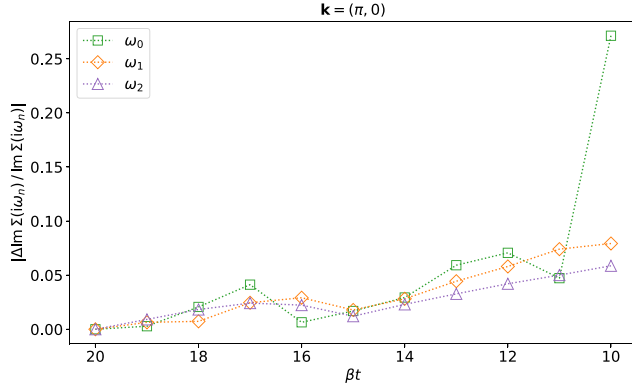
FIG. 2. Difference $|\Delta \text{Im}\Sigma(i\omega_n)/\text{Im}\Sigma(i\omega_n)|$ between the converged self-energy at $\beta t = 20$ and the extrapolation from the converged self-energy at higher temperature, shown as a function of the extrapolation inverse temperature for the three lowest Matsubara frequencies ω_n, with $n = 0, 1, 2$ (see text for calculation setup).

FIG. 3. Cluster DMFT convergence of the imaginary part of the local Green's function G_{loc} as a function of iteration for the three lowest Matsubara frequencies at $\beta t = 20$ (see text for calculation setup). Convergence to the fixed point is shown for the noninteracting starting point (filled symbols) and the starting point extrapolated from the converged solution at $\beta t = 18$ (open symbols). The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation.

of $U = 6t$, which corresponds to a Mott insulating regime. The self-energy in this regime exhibits a strong temperature dependence, and various phases such as metallic, insulating, superconducting, and pseudogapped phases are in close proximity [31,51,52]. For a comprehensive discussion of the physics of this self-energy, we refer the reader to the extensive literature on this system, including the reviews [7,48] and references therein.

Figure 1 presents the fully converged, self-consistent imaginary part of the self-energy, $\Sigma(i\omega_n)$, at $\beta t = 18$ (red crosses) and at $\beta t = 20$ (black dots). Red circles illustrate the extrapolation of the $\beta = 18t$ self-energy to $\beta = 20t$ using the Carathéodory formalism. The extrapolated data aligns with the converged points for all frequency points, except for the lowest one. The discrepancy between converged data and extrapolation data at $\beta = 20t$ serves as an indicator of additional correlations emerging in the system as it cools down.

Figure 2 demonstrates the deviation of the extrapolated self-energy at the three lowest Matsubara frequencies at $\beta t = 20$. The extrapolation is conducted using converged data at $\beta t = 10, 11, \ldots, 19$. Notably, while the lowest frequency of the self-energy exhibits significant deviations when extrapolated by a factor of two from $\beta t = 10$ to $\beta t = 20$, these deviations rapidly diminish if the extrapolation interval (the difference between the initial temperature and targeted temperature) is reduced.

Figures 1 and 2 demonstrate that the Carathéodory structure of many-body functions can be effectively utilized to extrapolate data in temperature.

B. Starting point and convergence

1. Convergence of the extrapolated starting point away from phase transitions

A temperature extrapolation of this type can be used to improve the starting point for the convergence of self-consistent many-body methods, such as DMFT [28,53,54] and its cluster extensions [4–7]. The DMFT equations are solved using a fixed-point iteration scheme, in which an initial estimate of the
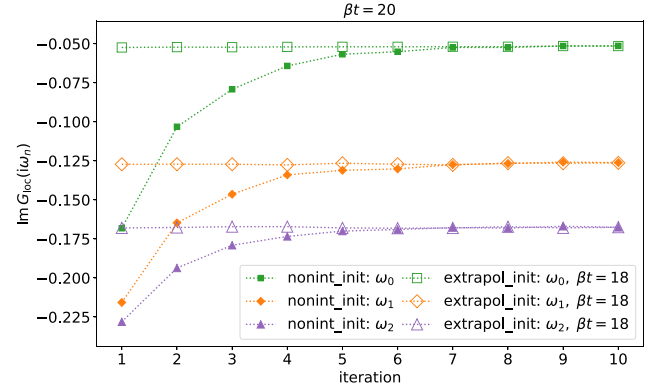
self-energy [commonly chosen as $\Sigma(i\omega_n) \equiv 0$] serves as an initial guess for the iteration. The equations typically converge rapidly if the initial starting point of the iteration is near the stationary point. Since the computational effort is proportional to the number of iterations required, a starting point closer to the stationary point reduces the calculation time.

Figure 3 displays the imaginary part of the lowest three frequencies of the local Green's function, $G_{\text{loc}}(i\omega_n)$, for a typical cluster DMFT calculation away from phase transitions, using noninteracting starting points (filled symbols) and starting points derived from temperature extrapolation (open symbols), as a function of iteration. The calculation setup is identical to the one in Sec. IV A (a 16-site cluster in the paramagnetic state with $\beta t = 20$, $U = 6t$, and at half filling). A naive noninteracting starting point converges in approximately seven iterations. Moreover, a few additional iterations are necessary to confirm that convergence has been achieved. In contrast, a starting point generated by extrapolating from higher to lower temperatures almost immediately yields a converged result.

2. Dynamical mean-field theory convergence near a phase transition

Self-consistent many-body methods are notoriously slow to converge in the vicinity of phase transitions. This can be illustrated at the example of a single-site DMFT calculation on an infinite coordination number Bethe lattice with bandwidth $W = 4t$. The single-site DMFT calculation exhibits a first-order phase transition between a paramagnetic metal at weak interaction and low temperature and a paramagnetic Mott insulator at large interaction and higher temperature [54–60], as depicted in the inset of Fig. 4. At an interaction strength $U = 4.7t$, the system is insulating at high temperature, metallic at low temperature, and both metallic and insulating solutions coexist in an intermediate temperature regime. We show the convergence of the imaginary-time Green's function in the middle of
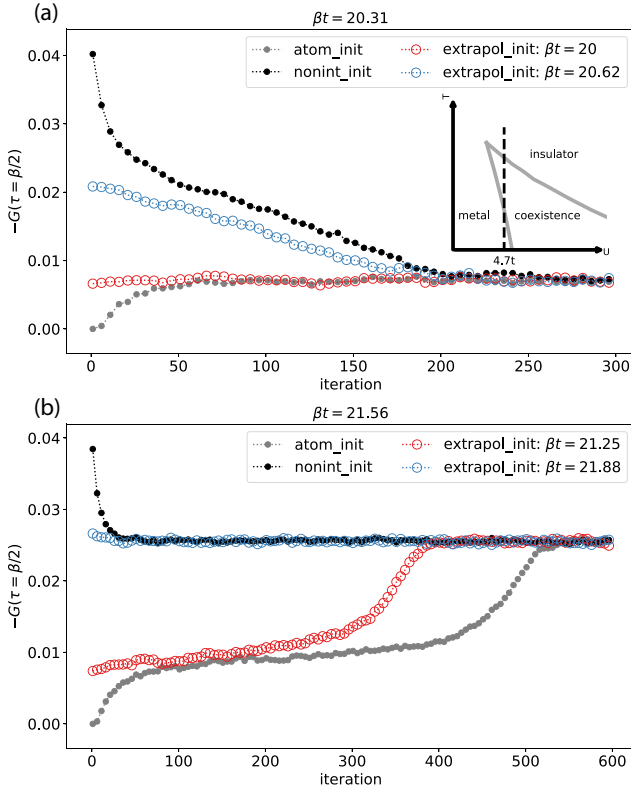
FIG. 4. DMFT convergence of $-G(\beta/2)$ on a half filled infinite-coordination Bethe lattice with $U = 4.7t$ and bandwidth $W = 4t$ as a function of iteration. The top and bottom panels display the insulating ($\beta t = 20.31$) and metallic ($\beta t = 21.56$) phases, respectively. Atomic starting point: gray dots. Noninteracting starting point: black dots. Extrapolated starting points: red circles and blue circles. Inset shows the phase diagram adapted from Ref. [55]. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation. Iteration points with an interval of five are shown for better visibility.
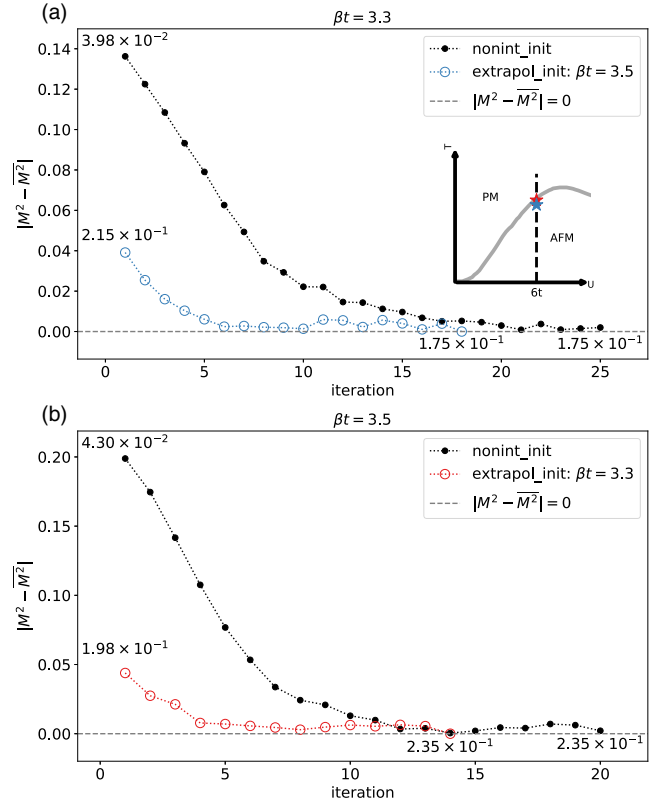
FIG. 5. Cluster DMFT convergence of the squared magnetic moment on the half filled three-dimensional Hubbard model (using a 34-site cluster with a doubled unit cell) at $U = 6t$ and $\beta t = 3.3$ (upper panel) and 3.5 (lower panel) as a function of iteration. Dots: noninteracting starting point. Circles: temperature extrapolation. Inset:] shows phase diagram adapted from Ref. [61]. M^2 at the first and last iteration indicated by numbers. $\overline{M^2}$ corresponds to the values at the last iteration of red or blue curves, which serves as a reference for converged values. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation.

the imaginary time interval, $-G(\beta/2)$, in Fig. 4, which is related to the spectral density at the Fermi surface in the low-temperature limit as $A(\omega = 0) \sim -\beta G(\beta/2)$, at temperatures close to the two boundaries of the coexistence region.

At a temperature of $\beta t = 20.31$, the system exhibits Mott insulating behavior. The upper panel of Fig. 4 shows that hundreds of iterations are required to achieve convergence when initiating from the noninteracting limit (black dots). Similarly, starting with an extrapolation from a lower temperature metallic solution in the coexistence region (blue circles) results in slow convergence toward the insulating fixed point. Conversely, the extrapolation from a higher-temperature insulating phase converges rapidly (red circles). A starting point derived from the atomic limit also leads to relatively fast convergence (gray dots).

Conversely, at a temperature of $\beta t = 21.56$ (Fig. 4, lower panel), the system is in a metallic state. Convergence from an "insulating" starting point requires hundreds of iterations to reach convergence (red circles for the starting point extrapolated from a higher-temperature insulating solution, and gray dots for the atomic limit starting point). Convergence from a

metallic solution, like the noninteracting limit solution (black dots), is substantially faster, and a starting point extrapolated from a converged metallic solution at lower temperature leads to even faster convergence (blue circles).

Figure 4 therefore shows that, even though all starting points, including atomic, noninteracting, and extrapolated, converge to the identical solution, extrapolation efficiently accelerates convergence, provided that the extrapolated solution is in the same phase as the solution at the target temperature. For extrapolation originating from a different phase (at the boundary of phase transitions, solutions with distinct properties become possible for two neighboring temperatures), the extrapolated starting points still perform better than the simplistic starting points corresponding to the "wrong" phase, specifically, the noninteracting solution or the atomic solution.

3. Cluster dynamical mean-field theory convergence near a phase transition

At second-order phase transitions, critical slowing down, rather than coexistence and hysteresis, is expected. This
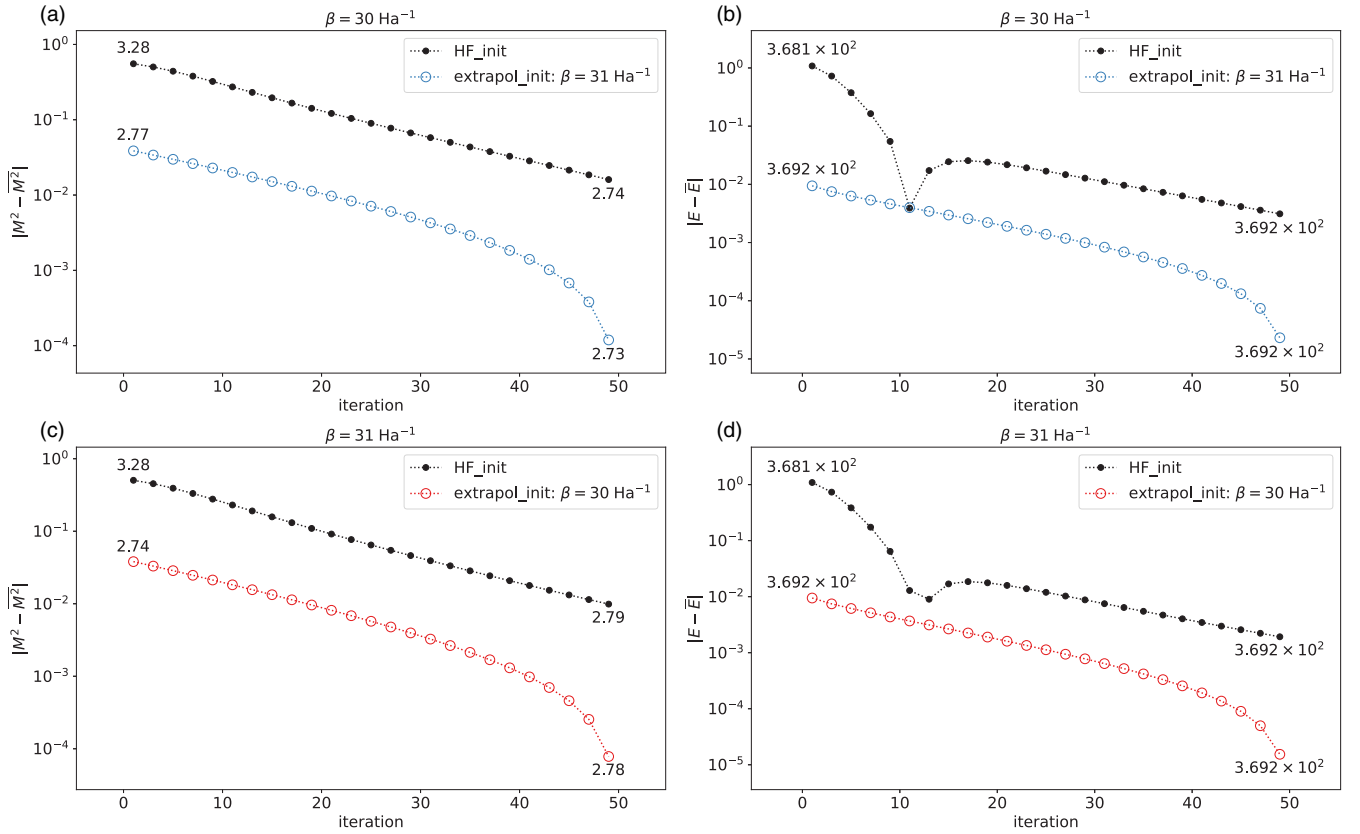
FIG. 6. GW convergence of the squared magnetic moment (left) and energy (right) near the AFM phase transition of NiO at $\beta = 30$ Ha^{-1} (upper panel) and $\beta = 31$ Ha^{-1} (lower panel) as a function of iteration. Calculations with unrestricted Hartree-Fock self-energy starting point are represented by black dots, while calculations with temperature-extrapolated self-energy are denoted by red circles (indicating extrapolation from higher temperature solutions) or blue circles (indicating extrapolation from lower-temperature solutions). The Hartree atomic units are used. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation. Iteration points with an interval of two are shown for better visibility. The values of M^2 and E at the first and last iterations are indicated by the numbers. The values of $\overline{M^2}$ and $\overline{E}$ are determined as the values at the last iteration of red or blue curves (the symbols for the last iteration are not shown).

phenomenon leads to a slow convergence of the fixed-point iteration. We illustrate this at the example of a half filled three-dimensional Hubbard model treated within cluster DMFT [62–64], calculated on a 34-site cluster with a doubled unit cell. This model exhibits a phase transition between a paramagnetic (PM) state at high temperature and an antiferromagnetic (AFM) insulator at low temperature, as depicted in the inset of Fig. 5. The model has been studied extensively within the context of ultracold atomic gases. In Fig. 5, we examine two points, $\beta t = 3.3$ and $\beta t = 3.5$, in the ordered phase at $U = 6t$, near the AFM phase transition (the critical point is at about $\beta t = 2.7$). A starting point from temperature extrapolation cannot completely overcome the slowing down (since the self-energy itself exhibits strong temperature dependence) but does lead to an acceleration of the convergence by at least a factor of two, as illustrated by the convergence of the squared magnetic moment shown in Fig. 5.

4. GW convergence of realistic many-body simulations

Finally, we turn to realistic simulations within a weak-coupling framework. We show examples for periodic solid NiO treated within the so-called GW approximation [8–13]. The GW approximation takes into account screening pro-

cesses via a renormalized, frequency-dependent interaction but neglects the second-order exchange diagram. It is therefore mostly used for weakly correlated systems such as semiconductors. In the calculation for NiO, a $4 \times 4 \times 4$ cluster with a doubled unit cell along the [1,1,1] direction and a fixed lattice constant $a = 4.1705$ Å [65] is utilized. The *gth-dzvp-molopt-sr* basis [66] along with the *gth-pbe* pseudopotential [3] is employed. For the density fitting of Coulomb integrals, the *def2-svp-ri* basis set is chosen as the auxiliary basis [67]. Finite-size errors in the GW exchange diagram are corrected using the Ewald probe-charge approach [68,69]. The Coulomb integrals and noninteracting matrix elements are obtained from PYSCF [70]. To decrease the number of frequencies utilized in the computation, the IR grid [36] is employed. A comprehensive description of the methods and implementation, in conjunction with the computational setup for evaluating NiO, is extensively detailed in Refs. [12,71].

When applied to the antiferromagnetic materials NiO, the GW method shows a continuous transition to an ordered state with a nonzero staggered magnetization at low temperature. Within the GW framework, the transition temperature is situated near $\beta = 25$ Ha^{-1}. As illustrated in Fig. 6, the convergence to this ordered state, initiated from an unrestricted Hartree-Fock solution (black dots), is relatively
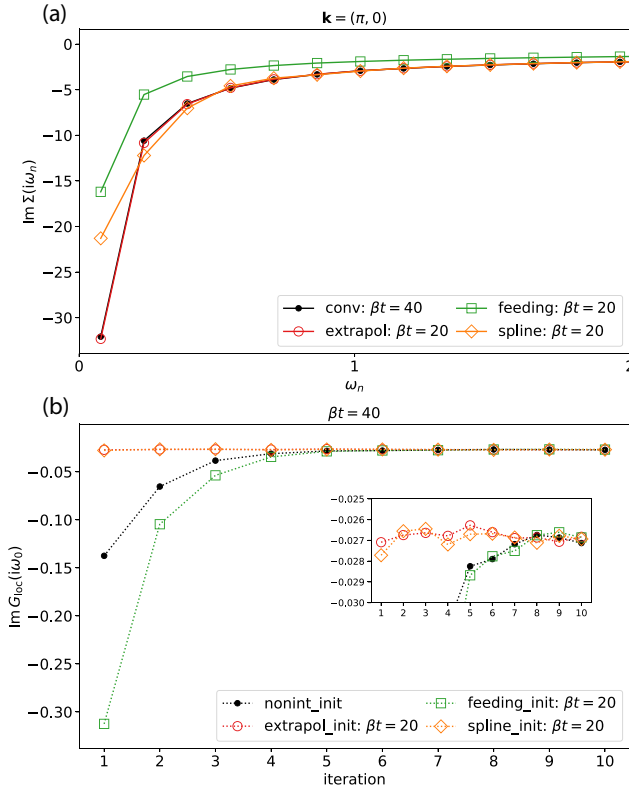
FIG. 7. Comparison for different acceleration methods for a 16-site cluster DMFT calculation for the Hubbard model at $U = 6t$. All temperature extrapolations are performed from $\beta t = 20$ to $\beta t = 40$. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation. (a) The imaginary part of the self-energy, Im $\Sigma(i\omega_n)$, at the antinodal point $(\pi, 0)$ as a function of ω_n. Converged data at $\beta t = 40$: black dots; extrapolated data via Carathéodory formalism: red circles; extrapolated data via temperature feeding: green squares; extrapolated data via cubic spline interpolation: orange diamonds. (b) The imaginary part of the local Green's function G_{loc} at the lowest Matsubara frequencies ω_0 as a function of iteration. Noninteracting starting point: black dots; Carathéodory starting point: red circles; temperature feeding starting point: green squares; cubic spline starting point: orange diamonds.

slow. This is evident in the convergence of both the squared magnetic moment for Ni and the total energy. However, when starting from extrapolated starting points (blue or red circles), the convergence occurs at a significantly faster pace, with an initial starting point akin to iteration 30 of the Hartree-Fock-initialized convergence.

C. Comparison with other methods

By assuming the causal matrix-valued functions $G(\tilde{z})$ at nearby temperatures having a close analytical structure, we introduce the temperature extrapolation for the Matsubara functions by evaluating the interpolant of $G(\tilde{z})$ of one temperature β on the Matsubara grid for another temperature β'. Besides the Carathéodory formalism, there are other interpolation methods available for the same purpose, warranting comparative analysis. We compare the Carathéodory formalism with a common alternative: cubic spline. For cubic spline,

additional constraints are needed at the boundaries to fix all the coefficients for cubic polynomials. After obtaining the spline, one could use the boundary cubic polynomial for extrapolating values outside the original scope or force the extrapolated segment to maintain the derivative values at the boundary (for instance, performing linear extrapolation if we set the boundary's second derivative to zero). For the Matsubara functions under study, we find the "not-a-knot spline" [72], which requires that the first two cubic polynomials at the boundary be identical, to be superior to other boundary constraints. Therefore, we only show this type of cubic spline for comparison, labeled as "spline." In addition, we also compare with a commonly used technique to accelerate temperature scanning, where the result at β is directly used as a starting point for β' despite the mismatching of the Matsubara grids. This approach is referred to as "temperature feeding" here and is labeled as "feeding." Given that the challenging part of simulations often manifests at lower temperatures, this section primarily focuses on the cooling process. Furthermore, as β' approaches β, all the discussed methods should provide reasonably good starting points and converge more rapidly than simpler methods, such as the noninteracting and Hartree-Fock starting points. Hence, we engage with a relatively large temperature difference here, setting β' to 2β.

Figure 7 utilizes the identical 16-site cluster DMFT calculation setup for the Hubbard model as detailed in Sec. IV B, albeit this instance involves an extrapolation from $\beta t = 20$ to $\beta t = 40$. In Fig. 7(a), the imaginary part of self-energy at the antinodal point is displayed. The extrapolated self-energy via the Carathéodory formalism (red circles) closely aligns with the fully converged self-energy (black dots) at all Matsubara frequencies. In contrast, extrapolated self-energy given by the cubic spline (orange diamonds) exhibits mismatches at the initial few low Matsubara frequencies. As anticipated, the temperature-feeding method demonstrates discrepancies across all Matsubara frequencies. Figure 7(b) illustrates the different methods' convergence behavior for the imaginary part of the local Green's function. Both the starting points of the cubic spline and Carathéodory methods display immediate convergence within the first two iterations, whereas temperature feeding worsens convergence compared with the noninteracting starting point. As observed in the inset, the advantages of the Carathéodory method's starting point over cubic spline's starting is about the same magnitude of Monte Carlo errors, which obscures the distinction in convergence between the two methods.

Figure 8 employs the same self-consistent GW calculation setup for NiO as detailed in Sec. IV B but, in this case, we extrapolate from $\beta = 100 \, \mathrm{Ha}^{-1}$ to $\beta = 200 \, \mathrm{Ha}^{-1}$. Figure 8(a) presents the imaginary time self-energy with fixed spin and orbital indices at the Γ point. This imaginary time self-energy is the self-consistently converged quantity utilized in our GW code. For the extrapolation stage, we carry out the extrapolation in frequency space and then transform back to the imaginary time space on the IR grid [36]. The Carathéodory method (red circles) gives nearly perfect extrapolation in comparison with the fully converged self-energy, while the cubic spline (orange diamonds) exhibits discrepancies at both $\tau = 0$ and $\tau = \beta$ (see inset). The temperature-feeding method, as anticipated, shows discrepancies over the entire time grid. We
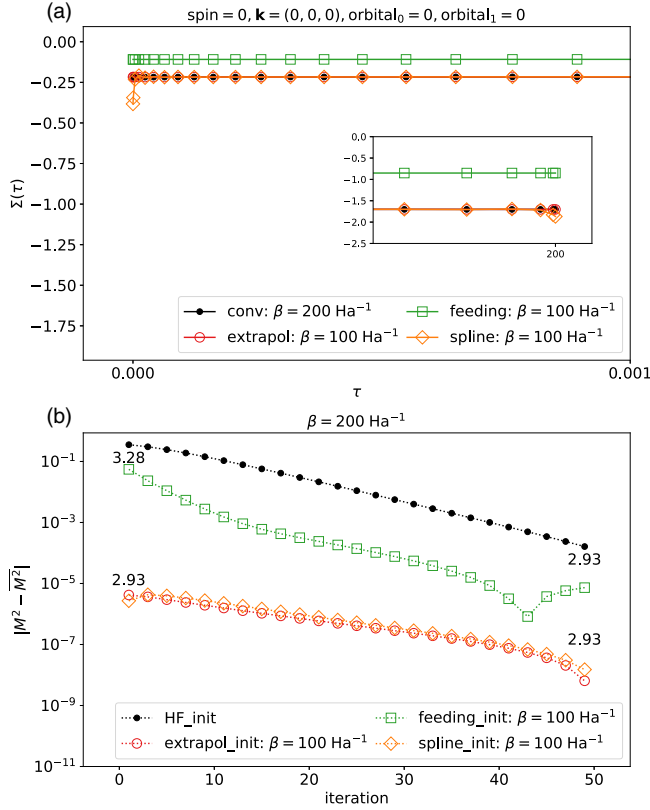
FIG. 8. Comparison for different acceleration methods for a self-consistent GW calculation for NiO. All temperature extrapolations are performed from $\beta = 100$ Ha^{-1} to $\beta = 200$ Ha^{-1}. The labels of the extrapolation results indicate the initial inverse temperature used for extrapolation. (a) Imaginary time self-energy, Im $\Sigma(\tau)$, with fixed spin, momentum, and orbital indices. Converged data at $\beta = 200$ Ha^{-1}: black dots; extrapolated data via Carathéodory formalism: red circles; extrapolated data via temperature feeding: green squares; extrapolated data via cubic spline interpolation: orange diamonds. (b) Convergence of the squared magnetic moment as a function of iteration. Hartree-Fock starting point: black dots; Carathéodory starting point: red circles; temperature feeding starting point: green squares; cubic spline starting point: orange diamonds. Iteration points with an interval of two are shown for better visibility. The values of M^2 at the first and last iterations of the red and black curves are indicated by the numbers. The value of $\overline{M^2}$ is determined as the value at the last iteration of the red curve (the symbol for the last iteration is not shown).

observe a similar behavior when performing extrapolation for the time-ordered Green's function $G(\tau)$, where the analytical properties $G_{ij}(0^+) + G_{ij}(\beta^-) = -\delta_{ij}$ (with fixed spin and momentum indices) are significantly violated by the starting points derived from the cubic spline and temperature feeding methods. In contrast, these properties are preserved to a high degree of precision for the starting point derived from the Carathéodory method. Figure 8(b) demonstrates the convergence behavior of different methods for the squared magnetic moment. Both the starting points of the cubic spline and Carathéodory methods exhibit superior convergence behavior compared with the Hartree-Fock starting point and the temperature feeding starting point, while the temperature feeding starting point outperforms the Hartree-Fock starting point. In
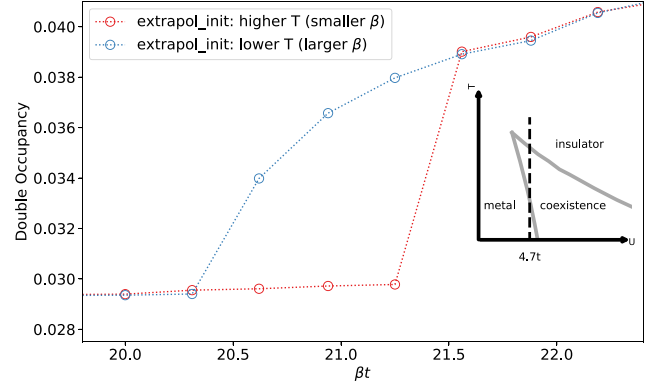

FIG. 9. DMFT analysis of a half filled Bethe lattice with $U = 4.7t$ and a bandwidth of $4t$, displaying the hysteresis near $\beta t = 21$. Points on the red curve utilize extrapolation from the nearest higher temperature converged results as the starting points for the calculation, while points on the blue curve employ extrapolation from the nearest lower temperature converged results as the starting points for the calculation. Inset shows phase diagram adapted from Ref. [55].

this scenario, it can be observed that the Carathéodory starting point converges faster than the cubic spline starting point by one or two iterations [the interval between points shown in Fig. 8(b) is two].

From the two test cases detailed above, we note that at relatively large temperature differences, temperature feeding can either accelerate or decelerate convergence. Both cubic spline interpolation and Carathéodory formalism, when used for extrapolation, are superior choices for the starting points, with the Carathéodory formalism demonstrating a slight advantage. We stress that systems with a strongly divergent self-energy will likely exhibit more evident superiority of the Carathéodory method in comparison to cubic spline interpolation. We also emphasize that, while cubic spline interpolation proves effective in the specific example we examined, its combination with IR coefficients results in a substantial loss of accuracy and likely failure of a self-consistent calculation.

D. Heating and cooling

The possibility of obtaining a starting point for fixed-point iteration by extrapolating from a nearby temperature point allows us to smoothly change temperature in subsequent simulation, in analogy to "heating" and "cooling" measurements in experiments. This capability is especially important at first-order phase transitions, where multiple stable fixed points, corresponding to the different phases, exist.

Such a heating and cooling process is shown in Fig. 9. Shown is a coexistence region and hysteresis curve at the example of the single-site DMFT on a Bethe lattice (the same calculation setup as the one mentioned in Sec. IV B). We plot the double occupancy, which is directly related to the potential energy, as a function of inverse temperature β. With starting points extrapolated from higher temperature converged results (red open circles), we find a transition from an insulating phase with small double occupancy to a metallic phase with large double occupancy around $\beta t = 21.56$. Conversely,

with starting points extrapolated from lower temperature converged results (blue open circles), we find a transition from a metallic phase to an insulating phase around $\beta t = 20.31$. Between those temperatures, both metal and insulator are stable solutions of the fixed-point equations, indicating phase coexistence [54–60,73].

V. CONCLUSIONS

The Carathéodory temperature extrapolation technique, introduced in this paper, provides an effective way to accelerate the convergence of self-consistent many-body calculations for both model Hamiltonians and realistic systems. It allows for smooth temperature variation in simulations, making it suitable for studying heating and cooling processes in many-body systems. The starting points provided by this method are generally superior, especially in systems with phase transitions and convergence issues. The Carathéodory technique therefore offers a versatile and efficient approach for studying temperature-dependent properties in self-consistent many-body calculations that should be adapted in any self-consistent finite-temperature many-body simulation.

We note that "convergence acceleration" techniques such as direct inversion in the iterative subspace (DIIS) [74–76] and Anderson acceleration [77–80] are complementary to the Carathéodory temperature extrapolation method employed in this study. Convergence acceleration techniques work by extrapolating from a set of initial iterations. The combination of such techniques with the better starting points provided by Carathéodory temperature extrapolation is therefore straightforward.

[1] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, revised English ed. (Dover Publications, New York, 1975).

[2] G. D. Mahan, *Many-Particle Physics*, 3rd ed., Physics of Solids and Liquids (Kluwer Academic/Plenum Publishers, New York, 2000).

[3] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

[4] M. H. Hettler, A. N. Tahvildar-Zadeh, M. Jarrell, T. Pruschke, and H. R. Krishnamurthy, Phys. Rev. B **58**, R7475 (1998).

[5] A. I. Lichtenstein and M. I. Katsnelson, Phys. Rev. B **62**, R9283 (2000).

[6] G. Kotliar, S. Y. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001).

[7] T. A. Maier, M. Jarrell, T. Pruschke, and M. Hettler, Rev. Mod. Phys. **77**, 1027 (2005).

[8] L. Hedin, Phys. Rev. **139**, A796 (1965).

[9] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[10] A. Kutepov, S. Y. Savrasov, and G. Kotliar, Phys. Rev. B **80**, 041103(R) (2009).

[11] A. L. Kutepov, Comput. Phys. Commun. **257**, 107502 (2020).

[12] C.-N. Yeh, S. Iskakov, D. Zgid, and E. Gull, Phys. Rev. B **106**, 235104 (2022).

[13] C.-N. Yeh, A. Shee, Q. Sun, E. Gull, and D. Zgid, Phys. Rev. B **106**, 085121 (2022).

[14] N. E. Dahlen and R. van Leeuwen, J. Chem. Phys. **122**, 164102 (2005).

[15] J. J. Phillips and D. Zgid, J. Chem. Phys. **140**, 241101 (2014).

[16] A. A. Rusakov and D. Zgid, J. Chem. Phys. **144**, 054106 (2016).

[17] S. Iskakov, A. A. Rusakov, D. Zgid, and E. Gull, Phys. Rev. B **100**, 085112 (2019).

[18] N. E. Bickers, D. J. Scalapino, and S. R. White, Phys. Rev. Lett. **62**, 961 (1989).

[19] S. Iskakov and E. Gull, Phys. Rev. B **105**, 045109 (2022).

[20] N. Prokof'ev and B. Svistunov, Phys. Rev. Lett. **99**, 250201 (2007).

[21] E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. B **82**, 075109 (2010).

[22] E. Gull, D. R. Reichman, and A. J. Millis, Phys. Rev. B **84**, 085134 (2011).

[23] G. Cohen, D. R. Reichman, A. J. Millis, and E. Gull, Phys. Rev. B **89**, 115139 (2014).

[24] S. Biermann, F. Aryasetiawan, and A. Georges, Phys. Rev. Lett. **90**, 086402 (2003).

[25] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).

[26] D. Zgid and E. Gull, New J. Phys. **19**, 023047 (2017).

[27] E. Kozik, M. Ferrero, and A. Georges, Phys. Rev. Lett. **114**, 156402 (2015).

[28] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[29] G. Kotliar and D. Vollhardt, Phys. Today **57**(3), 53 (2004).

[30] A. Macridin, M. Jarrell, and Th. Maier, Phys. Rev. B **74**, 085104 (2006).

[31] E. Gull, O. Parcollet, P. Werner, and A. J. Millis, Phys. Rev. B **80**, 245102 (2009).

[32] C. Walsh, P. Sémon, D. Poulin, G. Sordi, and A.-M. S. Tremblay, Phys. Rev. Lett. **122**, 067203 (2019).

[33] J. Fei, C.-N. Yeh, D. Zgid, and E. Gull, Phys. Rev. B **104**, 165111 (2021).

[34] T. D. Stanescu and G. Kotliar, Phys. Rev. B **74**, 125110 (2006).

[35] E. Gull, S. Iskakov, I. Krivenko, A. A. Rusakov, and D. Zgid, Phys. Rev. B **98**, 075127 (2018).

[36] H. Shinaoka, J. Otsuki, M. Ohzeki, and K. Yoshimi, Phys. Rev. B **96**, 035147 (2017).

[37] L. Boehnke, H. Hafermann, M. Ferrero, F. Lechermann, and O. Parcollet, Phys. Rev. B **84**, 075145 (2011).

[38] A. A. Kananenka, A. R. Welden, T. N. Lan, E. Gull, and D. Zgid, J. Chem. Theory Comput. **12**, 2250 (2016).

[39] J. Kaye, K. Chen, and O. Parcollet, Phys. Rev. B **105**, 235115 (2022).

[40] J. Li, M. Wallerberger, N. Chikano, C.-N. Yeh, E. Gull, and H. Shinaoka, Phys. Rev. B **101**, 035144 (2020).

[41] M. Kaltak and G. Kresse, Phys. Rev. B **101**, 205145 (2020).

[42] J. Fei, Ph.D. thesis, University of Michigan, 2021.

[43] K. Nogaki and H. Shinaoka, J. Phys. Soc. Jpn. **92**, 035001 (2023).

[44] A. Cayley, J. Reine Angew. Math. **1846**, 119 (1846).

[45] J. Schur, J. Reine Angew. Math. **148**, 122 (1918).

[46] See https://github.com/CQMP/pyCaratheodory for Python implementation of the algorithm, documentation, and examples.

[47] J. Fei, C.-N. Yeh, and E. Gull, Phys. Rev. Lett. **126**, 056402 (2021).

[48] M. Qin, T. Schäfer, S. Andergassen, P. Corboz, and E. Gull, Annu. Rev. Condens. Matter Phys. **13**, 275 (2022).

[49] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Europhys. Lett. **82**, 57003 (2008).

[50] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011).

[51] P. Werner, E. Gull, O. Parcollet, and A. J. Millis, Phys. Rev. B **80**, 045120 (2009).

[52] E. Gull, M. Ferrero, O. Parcollet, A. Georges, and A. J. Millis, Phys. Rev. B **82**, 155101 (2010).

[53] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989).

[54] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992).

[55] N. Blümer, Ph.D. thesis, Universität Augsburg, 2003.

[56] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992).

[57] M. J. Rozenberg, X. Y. Zhang, and G. Kotliar, Phys. Rev. Lett. **69**, 1236 (1992).

[58] X. Y. Zhang, M. J. Rozenberg, and G. Kotliar, Phys. Rev. Lett. **70**, 1666 (1993).

[59] M. J. Rozenberg, R. Chitra, and G. Kotliar, Phys. Rev. Lett. **83**, 3498 (1999).

[60] R. Bulla, Phys. Rev. Lett. **83**, 136 (1999).

[61] R. Staudt, M. Dzierzawa, and A. Muramatsu, Eur. Phys. J. B **17**, 411 (2000).

[62] P. R. C. Kent, M. Jarrell, T. A. Maier, and Th. Pruschke, Phys. Rev. B **72**, 060411(R) (2005).

[63] S. Fuchs, E. Gull, L. Pollet, E. Burovski, E. Kozik, T. Pruschke, and M. Troyer, Phys. Rev. Lett. **106**, 030401 (2011).

[64] S. Fuchs, E. Gull, M. Troyer, M. Jarrell, and T. Pruschke, Phys. Rev. B **83**, 235113 (2011).

[65] L. C. Bartel and B. Morosin, Phys. Rev. B **3**, 1039 (1971).

[66] J. VandeVondele and J. Hutter, J. Chem. Phys. **127**, 114105 (2007).

[67] C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).

[68] J. Paier, R. Hirschl, M. Marsman, and G. Kresse, J. Chem. Phys. **122**, 234102 (2005).

[69] R. Sundararaman and T. A. Arias, Phys. Rev. B **87**, 165122 (2013).

[70] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1340 (2018).

[71] S. Iskakov, C.-N. Yeh, E. Gull, and D. Zgid, Phys. Rev. B **102**, 085105 (2020).

[72] R. L. Burden, J. D. Faires, and A. M. Burden, *Numerical Analysis*, 10th ed. (Cengage Learning, Boston, 2016).

[73] A. Erpenbeck, W.-T. Lin, T. Blommel, L. Zhang, S. Iskakov, L. Bernheimer, Y. Núñez-Fernández, G. Cohen, O. Parcollet, X. Waintal, and E. Gull, Phys. Rev. B **107**, 245135 (2023).

[74] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[75] P. Pulay, J. Comput. Chem. **3**, 556 (1982).

[76] P. Pokhilko, C.-N. Yeh, and D. Zgid, J. Chem. Phys. **156**, 094101 (2022).

[77] G. B. Bacskay, Chem. Phys. **61**, 385 (1981).

[78] H. F. Walker and P. Ni, SIAM J. Numer. Anal. **49**, 1715 (2011).

[79] C. Evans, S. Pollock, L. G. Rebholz, and M. Xiao, SIAM J. Numer. Anal. **58**, 788 (2020).

[80] S. Pollock and L. G. Rebholz, IMA J. Numer. Anal. **41**, 2841 (2021).