Efficient and accurate tensor network algorithm for Anderson impurity problems

Zhijie Sun,[1] Zhenyu Li,[1,*] and Chu Guo[2,†]

[1] *State Key Laboratory of Precision and Intelligent Chemistry, University of Science and Technology of China, Hefei 230026, China*
[2] *Institute for Quantum Science and Technology, College of Science, National University of Defense Technology, Changsha 410073, China*

The Anderson impurity model (AIM) is of fundamental importance in condensed matter physics to study strongly correlated effects. However, accurately solving its long-time dynamics still remains a great numerical challenge. An emergent and rapidly developing numerical strategy to solve the AIM is to represent the Feynman-Vernon influence functional (IF), which encodes all the bath effects on the impurity dynamics, as a matrix product state (MPS) in the temporal domain. The computational cost of this strategy is basically determined by the bond dimension χ of the temporal MPS. In this work, we propose an efficient and accurate method which, when the hybridization function in the IF can be approximated as the summation of n exponential functions, can systematically build the IF as a MPS by multiplying $O(n)$ small MPSs, each with bond dimension 2. Our method gives a worst case scaling of χ as 2^{8n} and 2^{2n} for real- and imaginary-time evolution respectively. We demonstrate the performance of our method for two commonly used bath spectral functions, where we show that the actually required χs are much smaller than the worst case.

I. INTRODUCTION

The Anderson impurity model describes a localized electron that is coupled to a continuous noninteracting bath [1], which is a prototypical model to study a variety of quantum phenomena such as strongly correlated effects [2, 3], quantum transport [4, 5], or open quantum dynamics [6]. The AIM and its multi-orbital generalizations are also the central problem to solve in the dynamical mean field theory (DMFT) or non-equilibrium DMFT for high-dimensional quantum many-body problems [7–10]. Due to its profound importance, a plethora of numerical methods have been developed to solve the AIM, including exact diagonalization [11–17], continuous-time quantum Monte Carlo [18–27], numerical renormalization group [28–36], time-evolving MPS [37–48], hierarchical equation of motion (HEOM) [49–53]. However, although these methods could be highly accurate or efficient in certain regimes or for certain types of problems, for real- and long-time dynamics they are generally inefficient or inaccurate. Meanwhile, these methods are mostly heuristic, therefore they can not give a proper understanding of the complexity of the problem.

The AIM has a very special mathematical structure: the impurity electron is linearly coupled to a noninteracting bath. In case the interaction of the impurity electron vanishes, the whole model becomes quadratic and thus integrable. Even in the general case, it still allows a "partial analytic solution": the Feynman-Vernon IF, in which the bath degrees of freedom are integrated out, and only the impurity degrees of freedom on the time axis are left [54]. Roughly speaking, the Feynman-Vernon IF has turned the problem from $1+1$D into $0+1$D.

An emergent numerical strategy to solve the AIM is to represent the IF as an MPS, which makes full use of this partial analytic solution. This strategy has several distinguished features compared to existing approaches: (1) it is free of the bath discretization error as the bath has been integrated out; (2) it is free of the sign problem, as a general feature of MPS algorithms; (3) it is efficient as only a $0+1$D problem is solved. Along this line, two distinct approaches have been taken up to date. The first and pioneering approach represents the IF as an MPS in the Fock state basis, which has been applied for AIMs on the real-time [55–58] and imaginary-time axis [59]. The second approach represents the IF as a Grassmann MPS (GMPS) in the coherent state basis, which has been applied for AIMs and its multi-orbital extensions on the real-time [60, 61], imaginary-time [62] and the L-shaped Kadanoff contour [63]. A detailed comparison between these two approaches will be shown in the next section.

In this work, we propose an efficient and accurate method to construct the IF as a GMPS, following the second approach. The starting point of our method is to approximate the hybridization function in the IF as a summation of n exponential functions, similar to HEOM. The contribution of each exponential function can be written as a GMPS with bond dimension 2, using the standard WII algorithm [64], which is a first-order approximation. Then the IF can be constructed as a GMPS by multiplying $2n$ such small GMPSs for imaginary-time evolution, or $8n$ for real-time evolution. Crucially, we are able to prove that the WII approximation in this case is exact, due to the special properties of Grassmann algebra. Our method has a lower computational cost than existing alternatives, in the mean time it has very few error sources, which can all be well-controlled. Our method gives a worst-case scaling of χ as 2^{2n} or 2^{8n} for imaginary- or real-time evolution respectively. For real-time evolution, it has been proven that under mild assumptions about the bath spectral function (BSF), n only scales as

* zyli@ustc.edu.cn
† guochu604b@gmail.com

$n \sim \log(t)\log(1/\varepsilon)$ for a total evolution time t and a desired error ε [65, 66], which means that our method scales polynomially with t even in the worst case. We demonstrate the performance of our method on two commonly used BSFs: the semi-circular function and the Lorentzian function, with comparisons to several alternatives.

This paper is organized as follows. In Sec. II, we briefly present some necessary background for our method. In Sec. III, we show the details of our method, as well as its connections and differences to the alternatives. In Sec. IV, we perform numerical simulations with two commonly used BSFs to illustrate the efficiency and accuracy of our method. We conclude in Sec. V.

II. BACKGROUND

A. The path integral formalism

The total hamiltonian of the AIM can be written as

$$\hat{H} = \hat{H}_{\text{imp}} + \hat{H}_{\text{bath}} + \hat{H}_{\text{hyb}}, \tag{1}$$

where $\hat{H}_{\text{imp}} = \epsilon_d \sum_{\sigma=\uparrow\downarrow} \hat{a}_\sigma^\dagger \hat{a}_\sigma + U \hat{a}_\uparrow^\dagger \hat{a}_\downarrow^\dagger \hat{a}_\downarrow \hat{a}_\uparrow$ is the impurity Hamiltonian which describes a localized electron orbital with on-site energy ϵ_d and interaction strength U ($\hat{a}^\dagger$, $\hat{a}$ are the fermionic creation and annihilation operators), $\hat{H}_{\text{bath}} = \sum_{k,\sigma} \epsilon_k \hat{c}_{k,\sigma}^\dagger \hat{c}_{k,\sigma}$ is the bath Hamiltonian which describes itinerant electrons with energy band ϵ_k, $\hat{H}_{\text{hyb}} = \sum_{k,\sigma} \nu_k (\hat{a}_\sigma^\dagger \hat{c}_{k,\sigma} + \text{H.c.})$ is the hybridization Hamiltonian which describes the coupling between the localized electron and the itinerant electrons with strength ν_k. The BSF $J(\epsilon)$ is defined as

$$J(\epsilon) = \sum_k \nu_k^2 \delta(\epsilon - \epsilon_k), \tag{2}$$

which determines all the effects of $\hat{H}_{\text{hyb}}$ and $\hat{H}_{\text{bath}}$ on the impurity dynamics.

The linear coupling form of $\hat{H}_{\text{hyb}}$ (more precisely, linear in terms of the bath creation and annihilation operators) and the noninteracting form of $\hat{H}_{\text{bath}}$ ensure that the bath degrees of freedom can be analytically integrated out in the path integral (PI), leading to a PI for the impurity partition function Z_{imp} that contains only the impurity degrees of freedom on the time axis. For imaginary- and real-time evolution, the impurity partition functions are defined as $Z_{\text{imp}}(\beta) = \operatorname{Tr} e^{-\beta \hat{H}} / \operatorname{Tr} \hat{\rho}_{\text{bath}}^{\text{th}}$ and $Z_{\text{imp}}(t) = \operatorname{Tr} \hat{\rho}(t) / \operatorname{Tr} \hat{\rho}_{\text{bath}}^{\text{th}}$ respectively, where β is the inverse temperature, $\hat{\rho}_{\text{bath}}^{\text{th}}$ denotes the thermal state of the bath, $\hat{\rho}(t)$ denotes the total state at time t. In the following, we will illustrate the PI formalism for the case of imaginary-time evolution, and briefly mention the corresponding results for real-time evolution when necessary. For more detailed derivations and expressions on different contours one could refer to Refs. [67, 68].

The impurity partition function of the AIM for imaginary-time evolution can be written as:

$$Z_{\text{imp}}(\beta) = \int \mathcal{D}[\bar{\boldsymbol{a}}\boldsymbol{a}] \mathcal{K}[\bar{\boldsymbol{a}}\boldsymbol{a}] \prod_\sigma \mathcal{I}_\sigma[\bar{\boldsymbol{a}}_\sigma \boldsymbol{a}_\sigma], \tag{3}$$

where $\bar{\boldsymbol{a}}_\sigma = \{\bar{a}_\sigma(\tau)\}$, $\boldsymbol{a}_\sigma = \{a_\sigma(\tau)\}$ are Grassmann trajectories, and we have used $\bar{\boldsymbol{a}} = \{\bar{\boldsymbol{a}}_\uparrow, \bar{\boldsymbol{a}}_\downarrow\}$, $\boldsymbol{a} = \{\boldsymbol{a}_\uparrow, \boldsymbol{a}_\downarrow\}$ for brevity. The measure $\mathcal{D}[\bar{\boldsymbol{a}}\boldsymbol{a}]$ is defined as

$$\mathcal{D}[\bar{\boldsymbol{a}}\boldsymbol{a}] = \prod_{\sigma,\tau} \mathrm{d}\bar{a}_\sigma(\tau) \mathrm{d}a_\sigma(\tau) e^{-\bar{a}_\sigma(\tau) a_\sigma(\tau)}. \tag{4}$$

$\mathcal{K}$ denotes the contribution from the bare impurity dynamics, which is determined by $\hat{H}_{\text{imp}}$ and can be formally written as

$$\mathcal{K}[\bar{\boldsymbol{a}}\boldsymbol{a}] = e^{-\int_0^\beta \mathrm{d}\tau \mathcal{H}_{\text{imp}}(\tau)}, \tag{5}$$

where $\mathcal{H}_{\text{imp}}(\tau)$ is obtained from $\hat{H}_{\text{imp}}$ by making the substitutions $\hat{a}_\sigma \to a_\sigma(\tau)$, $\hat{a}_\sigma^\dagger \to \bar{a}_\sigma(\tau)$. The contribution from $\hat{H}_{\text{hyb}}$ and $\hat{H}_{\text{bath}}$ is encoded in the IF $\mathcal{I}_\sigma$, which can be written as

$$\mathcal{I}_\sigma[\bar{\boldsymbol{a}}_\sigma \boldsymbol{a}_\sigma] = e^{-\int_0^\beta \mathrm{d}\tau' \int_0^\beta \mathrm{d}\tau'' \bar{a}_\sigma(\tau') \Delta(\tau',\tau'') a_\sigma(\tau'')}. \tag{6}$$

Here $\Delta(\tau', \tau'')$ is the hybridization function, which can be calculated as

$$\Delta(\tau', \tau'') = \int \mathrm{d}\epsilon J(\epsilon) D_\epsilon(\tau', \tau''), \tag{7}$$

where $D_\epsilon(\tau', \tau'')$ is the free bath Matsubara Green's function, defined as

$$D_\epsilon(\tau', \tau'') = -[\Theta(\tau' - \tau'') - \bar{n}(\epsilon)] e^{-\epsilon(\tau' - \tau'')}, \tag{8}$$

with Θ the Heaviside step function and $\bar{n}(\epsilon) = (e^{\beta\epsilon} + 1)^{-1}$ the Fermi–Dirac distribution. The integrand of the impurity PI in Eq.(3) is often referred to as the augmented density tensor (ADT), which is denoted as

$$\mathcal{A}[\bar{\boldsymbol{a}}\boldsymbol{a}] = \mathcal{K}[\bar{\boldsymbol{a}}\boldsymbol{a}] \prod_\sigma \mathcal{I}_\sigma[\bar{\boldsymbol{a}}_\sigma \boldsymbol{a}_\sigma]. \tag{9}$$

The corresponding expressions for the real-time impurity partition function $Z_{\text{imp}}(t)$ can be simply obtained by changing the integrations in Eq.(5) and Eq.(6) from the imaginary contour to the Keldysh contour [69], and the free bath Matsubara Green's function in Eq.(8) into the free bath contour-ordered Green's function accordingly. Here we note that for both imaginary- and real-time evolution, the IF is time-translationally invariant (TTI), i.e., $\Delta(\tau', \tau'') = \Delta(\tau + \tau', \tau + \tau'')$ for any τ (which is essentially because that $\hat{H}_{\text{hyb}}$ and $\hat{H}_{\text{bath}}$ are time-independent), $\mathcal{K}$ will also be TTI if $\hat{H}_{\text{imp}}$ is time-independent.

B. Brief review of the GTEMPO method

The seminal works which directly turn the analytic expression of the IF into numerical algorithm is the quasi-adiabatic propagator path integral (QuAPI) method [70,

71]. QuAPI is first implemented for bosonic impurity problems, but the central idea can also be straightforwardly generalized to the fermionic case [68]. In the fermionic case and for imaginary-time evolution, one can first discretize the Grassmann trajectories into discrete Grassmann variables (GVs): $\bar{\boldsymbol{a}}_\sigma \to \{\bar{a}_{\sigma,M}, \cdots, \bar{a}_{\sigma,1}\}$ and $\boldsymbol{a}_\sigma \to \{a_{\sigma,M}, \cdots, a_{\sigma,1}\}$ (we will still use $\bar{\boldsymbol{a}}_\sigma$ and $\boldsymbol{a}_\sigma$ as brevities for the list of discrete GVs in the following), with an equidistant time step size $\delta\tau = \beta/(M-1)$. Then a first-order approximation of $\mathcal{I}_\sigma$ is

$$\mathcal{I}_\sigma[\bar{\boldsymbol{a}}_\sigma \boldsymbol{a}_\sigma] \approx e^{-\sum_{j,k=1}^{M} \bar{a}_{\sigma,j}\Delta_{j,k}a_{\sigma,k}}, \tag{10}$$

where

$$\Delta_{j,k} = \int_{j\delta\tau}^{(j+1)\delta\tau} \mathrm{d}\tau' \int_{k\delta\tau}^{(k+1)\delta\tau} \mathrm{d}\tau'' \Delta(\tau', \tau''). \tag{11}$$

One could also discretize the bare impurity dynamics part $\mathcal{K}$ in Eq.(5) to obtain a numerically exact expression up to the first-order Trotter decomposition error. As in this work we focus on the treatment of the IF (which is the computationally dominate part for the AIM), the detailed treatment of $\mathcal{K}$ will not be shown here, but can be found in Ref. [62]. In our implementation we adopt the ordering of the discete GVs as $a_M \bar{a}_M \ldots a_2 \bar{a}_2 a_1 \bar{a}_1$ when constructing each $\mathcal{I}_\sigma$ and $\mathcal{K}$ as GMPSs.

We can see that the discretized $\mathcal{I}_\sigma$ in Eq.(10) is a Grassmann tensor (GT) of $2M$ GVs. As all the terms in the exponent of the discretized $\mathcal{I}_\sigma$ commute with each other, a naive way to explicitly construct the GT is as follows: one could first build each term $e^{-\bar{a}_{\sigma,j}\Delta_{j,k}a_{\sigma,k}} = 1 - \bar{a}_{\sigma,j}\Delta_{j,k}a_{\sigma,k}$ as a two-variate GT, and then multiply M^2 such GTs together [67] (the GT multiplication corresponds to the element-wise multiplication of the normal tensors in the bosonic case [72]). Similarly, one could also construct $\mathcal{K}$ as a GT of $2M$ GVs, and then multiply $\mathcal{K}$ and $\mathcal{I}_\sigma$ together using Eq.(9) to obtain the ADT. Once the ADT has been built, one could calculate any (multi-time) observables of the impurity straightforwardly. For example, $Z_{\mathrm{imp}}(\beta)$ can be simply obtained by integrating out all the GVs with the measure in Eq.(4), and the Matsubara Green's function can be calculated as

$$\begin{aligned} G(j\delta\tau) &= \langle \hat{a}_{\sigma,j+1}\hat{a}^\dagger_{\sigma,1}\rangle \\ &= Z^{-1}_{\mathrm{imp}}(\beta) \int \mathcal{D}[\bar{\boldsymbol{a}}\boldsymbol{a}] a_{\sigma,j+1}\bar{a}_{\sigma,1}\mathcal{A}[\bar{\boldsymbol{a}}\boldsymbol{a}], \end{aligned} \tag{12}$$

where $\langle \cdots \rangle$ in the first line means the thermal average.

An obvious obstacle of the above approach is that the number of GVs grows linearly with the number of time steps M, therefore the size of the ADT will grow exponentially. To leverage this problem, a memory truncation scheme is introduced in QuAPI for real-time evolution, which only keeps the ADT within the time interval $[j\delta t, (j+M_c)\delta t]$, where the memory size M_c fixes the size of the ADT. However, such a hard memory cutoff could result in several issues. For example, it could introduce uncontrolled error, and it does not work on the imaginary-time axis as $\Delta(\tau', \tau'')$ does not necessarily decay with the distance $|\tau'' - \tau'|$ [73].

The time-evolving matrix product operator (TEMPO) method makes a crucial step forward on top of QuAPI [74]. Instead of using a single tensor to represent the ADT, it uses the MPS, and adjusts the tensor operations (such as the element-wise multiplication) into corresponding MPS operations. An immediate advantage of TEMPO is that one could use a very large M_c, as the computational cost of MPS algorithms only grows linearly with the number of sites (but grows cubically with the bond dimension). In fact one does not need to perform memory cutoff at all in TEMPO, although in the original work they still opt to make this approximation for efficiency. The Grassmann time-evolving matrix product operator (GTEMPO) method is a fermionic analog of TEMPO, which represents the discretized $\mathcal{K}$ and $\mathcal{I}_\sigma$ as GMPSs instead of normal MPSs [60]. The purpose of GMPS is that it can automatically take into account the sign factors occurred when swapping the GVs.

In the following, we discuss a subtle difference between TEMPO and GTEMPO, besides that they deal with different types of particles. In original QuAPI and TEMPO, a very special type of hybridization Hamiltonian has been considered, namely $\hat{H}_{\mathrm{hyb}} = \hat{X}\sum_k \nu_k(\hat{b}^\dagger_k + \hat{b}_k)$, referred to as "additive bath". As there is only a single Hermitian impurity operator $\hat{X}$ involved in $\hat{H}_{\mathrm{hyb}}$, one could choose the eigenbasis of $\hat{X}$ for the impurity PI and obtain the bosonic IF in a form similar to Eq.(6), that is,

$$\mathcal{I}[\mathbf{s}] = e^{-\int_0^\beta \mathrm{d}\tau' \int_0^\beta \mathrm{d}\tau'' s(\tau')\Delta(\tau',\tau'')s(\tau'')}, \tag{13}$$

where $s(\tau)$ is an eigenvalue of $\hat{X}$. This approach, however, would not work if there are multiple non-commutative impurity operators that are coupled to the bath, referred to as "nonadditive bath" in general. For example, $\hat{H}'_{\mathrm{hyb}} = \sum_k \nu_k(\hat{X}\hat{b}^\dagger_k + \hat{X}^\dagger\hat{b}_k)$ with $\hat{X}^\dagger \neq \hat{X}$ is a simple case of nonadditive bath. A solution to this more general situation is provided in Ref. [75], the central idea of which is to generalize the variable path $s(\tau)$ in Eq.(13) into an operator path $\hat{X}(\tau)$:

$$\mathcal{I}[\hat{\mathbf{X}}] = e^{-\int_0^\beta \mathrm{d}\tau' \int_0^\beta \mathrm{d}\tau'' \hat{X}(\tau')\Delta(\tau',\tau'')\hat{X}(\tau'')}, \tag{14}$$

which will be a matrix product operator (MPO) instead of an MPS after discretization. Then for nonadditive bath one can simply build such an MPO for each impurity operator that is coupled to the bath, and then multiply these MPOs together in the end using standard MPO-MPO multiplication [76]. As an example, to deal with $\hat{H}'_{\mathrm{hyb}}$, we can first rewrite it as

$$\hat{H}'_{\mathrm{hyb}} = \hat{U}\sum_k \nu_k(\hat{b}^\dagger_k + \hat{b}_k) + \hat{V}\sum_k \nu_k \mathrm{i}(\hat{b}^\dagger_k - \hat{b}_k), \tag{15}$$

where $\hat{U} = (\hat{X}^\dagger + \hat{X})/2$ and $\hat{V} = \mathrm{i}(\hat{X}^\dagger - \hat{X})/2$, then the IF can be obtained as

$$\mathcal{I}'[\hat{\mathbf{X}}\hat{\mathbf{X}}^\dagger] = \mathcal{I}[\hat{\mathbf{U}}]\mathcal{I}[\hat{\mathbf{V}}]. \tag{16}$$

One could also refer to Ref. [77] for an equivalent but perhaps more intuitive derivation of the IF for the case of nonadditive bath.

In the AIM (or its multi-orbital extensions), the hybridization Hamiltonian is naturally nonadditive, as the two impurity operators $\hat{a}_\sigma$ and $\hat{a}^\dagger_\sigma$ do not commute. In principle, one could follow the solution in the bosonic case in Eq.(16), by going to the Majorana representation for example. However, the Grassmann PI in Eq.(6) provides a more elegant solution in terms of GVs. The coherent state representation can also be explored in the bosonic case to deal with nonadditive baths, however, in the bosonic case the impurity could be very diverse (as well as the impurity operators that are coupled to the bath), so the treatment in Eq.(16) could be a more general and convenient choice.

To this end, we also note that Ref. [56] provides another very different solution to the nonadditive bath, which is different from both approaches described above. The central idea is to perform a change of variable $\bar{a}_{\downarrow,i} \to a_{\downarrow,M+i}$ in $\mathcal{I}_\downarrow$ and $a_{\uparrow,i} \to \bar{a}_{\uparrow,M+i}$ in $\mathcal{I}_\uparrow$ in Eq.(10), such that the transformed $\tilde{\mathcal{I}}_\downarrow$ only contains $\boldsymbol{a}_\downarrow$ and $\tilde{\mathcal{I}}_\uparrow$ only contains $\bar{\boldsymbol{a}}_\uparrow$. Then $\tilde{\mathcal{I}}_\sigma$ can be exactly mapped into Gaussian states, e.g.,

$$e^{-\sum_{j,k=1}^{M} \bar{a}_{\uparrow,j}\Delta_{j,k}\bar{a}_{\uparrow,M+k}} \to e^{-\sum_{j,k=1}^{M} \hat{a}^\dagger_{\uparrow,j}\Delta_{j,k}\hat{a}^\dagger_{\uparrow,M+k}}|\emptyset\rangle, \tag{17}$$

as the commutation relation in the exponents of both expressions are the same ($|\emptyset\rangle$ denotes the fermionic vacuum state). The effect of Eq.(17) is similar to Eq.(14). However, such an artificial particle-hole transformation has several drawbacks compared to (G)TEMPO: (1) the total number of GVs is doubled for the same step size; (2) it does not work for spinless fermions; (3) As the conjugate pairs of GVs for a single flavor are separated into two flavors (spin up and down), it is not straightforward to perform a partial integration of one flavor to get the reduced ADT of the other, which could greatly improve the computational efficiency for multi-orbital impurity problems [78]; (4) for superconducting baths which contains terms like $\hat{a}_i\hat{a}_j$ and $\hat{a}^\dagger_i\hat{a}^\dagger_j$, the exponent of the IF would contain terms like $\bar{a}_i\bar{a}_j$, a_ia_j on top of $\bar{a}_ia_j$, in which case the above particle-hole transformation can no longer separate $\bar{\boldsymbol{a}}$ and $\boldsymbol{a}$. This method will be referred to as the tensor network IF method in the next.

III. THE EXACT TTI IF METHOD

For the AIM, the computationally dominate step in both GTEMPO and the tensor network IF method is to build the IF as a MPS. In the tensor network IF method, the central target it to build the Gaussian state in Eq.(17) as a fermionic MPS [56], where a generalized Fishman-White algorithm is used to transform the operator $e^{-\sum_{j,k=1}^{M} \hat{a}^\dagger_{\uparrow,j}\Delta_{j,k}\hat{a}^\dagger_{\uparrow,M+k}}$ into a quantum circuit of nearest-neighbour gate operations [79]. The number of gates in the circuit would roughly scale as $O(M\log(M))$ if the bipartition entanglement entropy of the underlying Gaussian state grows very slowly, which is claimed to be the case for impurity problems [55]. As a result its computational cost would be $O(M\log(M)\chi^3)$ (one basically counts the number of singular value decompositions that need to be performed, each has a cost of $O(\chi^3)$).

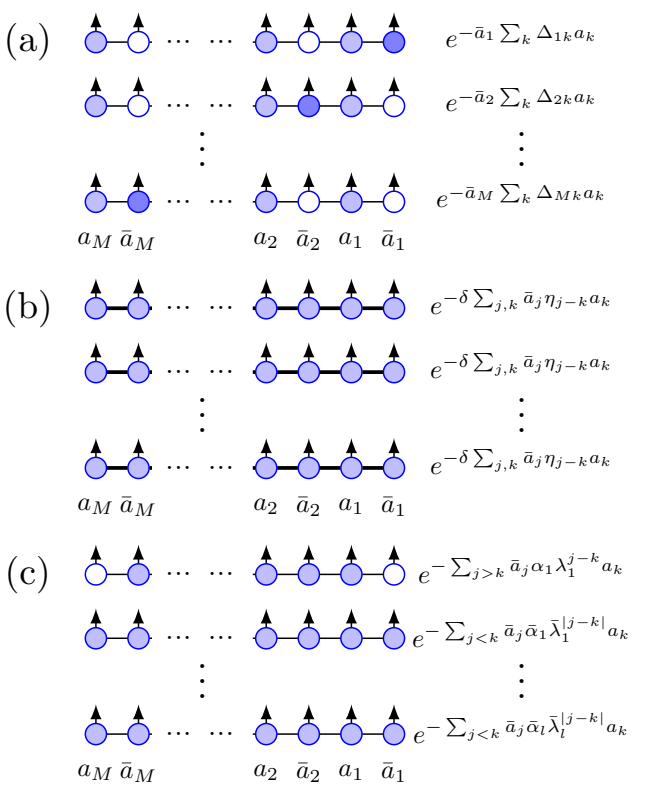

FIG. 1. Schematic illustration of (a) the partial IF method, (b) the TTI IF method and (c) the eTTI IF method. The empty circles mean that these sites (GVs) are not involved in the GMPS, while the light blue circles mean the opposite. The dark blue circles in (a) correspond to the first indices of the partial IF.

In GTEMPO, ones directly build the GT in Eq.(10) as a GMPS instead. Two approaches have been taken up to date: (i) the partial IF method [60] and (ii) the TTI IF method [72]. In the partial IF method, one rewrites

$$e^{-\sum_{j,k=1}^{M} \bar{a}_{\sigma,j}\Delta_{j,k}a_{\sigma,k}} = \prod_{j=1}^{M} e^{-\sum_{k=1}^{M} \bar{a}_{\sigma,j}\Delta_{j,k}a_{\sigma,k}}, \tag{18}$$

where each term in the product of the right hand side is referred to as a partial IF as they share the same GV $\bar{a}_{\sigma,j}$ and can be exactly written as a GMPS with bond dimension 2 [62]. The IF could be built by multiplying M such partial IFs and the computational cost is $O(M^2\chi^3)$. An important advantage of the partial IF method is that the only source of error in the method, on top of the Trotter error from QuAPI, is the MPS bond truncation

error, which can be well-controlled by χ. The partial IF method is schematically illustrated in Fig. 1(a).

In the TTI IF method, we first approximates the hybridization function as a sum of exponential functions

$$\eta_x = \Delta_{i+x,i} \approx \begin{cases} \sum_{l=1}^{n} \alpha_l \lambda_l^x & \text{if } x > 0; \\ \sum_{l=1}^{n} \bar{\alpha}_l \bar{\lambda}_l^{-x} & \text{if } x < 0, \end{cases} \tag{19}$$

where α_l, λ_l, $\bar{\alpha}_l$, $\bar{\lambda}_l$ are complex numbers in general. This decomposition can be implemented using the Prony method for example [80]. The error occurred in Eq.(19) is characterized by the mean square error between the hybridization function and the approximated one, denoted as ς. The exponent of $\mathcal{I}_\sigma$ can be rewritten as

$$\begin{aligned} \mathcal{F}_\sigma &= \sum_{j,k=1}^{M} \bar{a}_{\sigma,j} \Delta_{j,k} a_{\sigma,k} \\ &= \sum_{j>k} \bar{a}_{\sigma,j} \eta_{j-k} a_{\sigma,k} + \sum_{j<k} \bar{a}_{\sigma,j} \eta_{j-k} a_{\sigma,k} \\ &\quad + \sum_{j} \bar{a}_{\sigma,j} \Delta_{j,j} a_{\sigma,j}, \end{aligned} \tag{20}$$

where we have isolated the contributions with $j = k$ in the third line for later convenience. After substituting the approximation in Eq.(19) into the second line of Eq.(20), we can represent $\mathcal{F}_\sigma$ as a GMPS with bond dimension $2n+2$, using the standard technique of representing long-range Hamiltonians as MPOs [72]. After that, one could use the WI or WII approximation to build $e^{-\delta \mathcal{F}_\sigma}$ as a GMPS with bond dimension $2n+1$ [64], and obtain $\mathcal{I}_\sigma$ by multiplying $1/\delta$ such GMPSs. As all the GMPSs being multiplied are the same, a more efficient scheme could be used which requires only a constant number of GMPS multiplications [72]. The computational cost of the TTI IF method thus scales as $O(M\chi^4)$. We can see that the TTI IF method will be more efficient than the partial IF method for large M. However, it suffers from two additional errors compared to the partial IF method: (1) the error occurred in approximating the hybridization function as a sum of exponential functions in Eq.(19) (which is a well-defined mathematical problem itself, and has nothing to do with subsequent tensor network operations); (2) the error occurred in the WI or WII approximation of $e^{-\delta \mathcal{F}_\sigma}$, characterized by another step size δ (we note the δ is not directly related to β or t). The TTI IF method is schematically illustrated in Fig. 1(b).

Now we introduce an algorithm termed the exact TTI IF (eTTI IF) method, which is closely related to the TTI IF method, but is guaranteed to be more accurate and efficient than the latter. Substituting Eq.(19) and Eq.(20) into Eq.(10), we obtain

$$\begin{aligned} \mathcal{I}[\bar{\boldsymbol{a}}\boldsymbol{a}] \approx \prod_{l=1}^{n} e^{-\sum\limits_{j>k} \bar{a}_j \alpha_l \lambda_l^{j-k} a_k} \prod_{l=1}^{n} e^{-\sum\limits_{j<k} \bar{a}_j \bar{\alpha}_l \bar{\lambda}_l^{|j-k|} a_k} \times \\ e^{-\sum_j \bar{a}_j \Delta_{j,j} a_j}, \end{aligned} \tag{21}$$

where we have omitted the spin indices for brevity. The second line can be built as a GMPS with bond dimension 1 if $\bar{a}_j$ and a_j are represented by a single site, or a GMPS with bond dimension 2 if they are represented separately as two sites (for theoretical clearness we will use the former representation, but in our actual implementation we use the latter). As the cost of this local term is negligible, in the next we will focus on the two terms on the first line. For each l, the terms inside the two products can be written as GMPSs with site tensors

$$\begin{pmatrix} 1 & -\alpha\bar{a} & 0 \\ 0 & \lambda & \lambda a \\ 0 & 0 & 1 \end{pmatrix}; \quad \begin{pmatrix} 1 & \bar{\alpha}a & 0 \\ 0 & \bar{\lambda} & \bar{\lambda}\bar{a} \\ 0 & 0 & 1 \end{pmatrix}, \tag{22}$$

respectively, where we have omitted the time indices as the site tensors are TTI. The subscript l is also omitted for brevity. Taking the exponential of these two GMPSs (with step size 1), we can obtain the corresponding time-evolution operators via the WII approximation [64]:

$$\begin{pmatrix} 1 & -\alpha\bar{a} \\ \lambda a & \lambda - \alpha\lambda a\bar{a} \end{pmatrix}; \quad \begin{pmatrix} 1 & \bar{\alpha}a \\ \bar{\lambda}\bar{a} & \bar{\lambda} - \bar{\alpha}\bar{\lambda}a\bar{a} \end{pmatrix} \tag{23}$$

A crucial result of this work is the following theorem.

Theorem 1. The GMPSs with site tensors in Eq.(23) are the exact exponentials of the GMPSs with site tensors in Eq.(22). The detailed proof will be shown in the Appendix. A.

We note that the eTTI IF method, similar to the TTI IF method, can be directly used to build an infinite GMPS representation of the IF (infinite on the time axis). For finite t, one only needs to take the appropriate boundary conditions for each small GMPS [64, 72]. Here we also note that although the TTI property seems to be explicitly broken in the partial IF method as can be seen from Eq.(18), it can be restored by "rotating" the 2D tensor network in Fig. 1(a) by 45 degrees [81].

Combining Eq.(21) and Eq.(23), we can see that we can build the approximate $\mathcal{I}$ by multiplying $2n$ GMPSs, each with bond dimension 2 for the case of imaginary-time evolution. For real-time evolution, there are $8n$ GMPSs to be multiplied as each GV can live on 2 branches [60]. These results also imply that the bond dimension of the IF scale as 2^{2n} or 2^{8n} for imaginary- and real-time evolution respectively in the worst case. The eTTI IF method is schematically illustrated in Fig. 1(c). In case n scales as $\log(t)$ (which has been shown to generally be the case for the real-time evolution of quantum impurity problems [65]), the computational cost for the real-time evolution of the AIM would scale polynomially with t. Moreover, in practice, we find that the bond dimension can be significantly reduced by MPS bond truncation without affecting the accuracy, which will be shown in our numerical results. The computational cost of the eTTI IF method scales as $O(M\chi^3)$.

IV. NUMERICAL RESULTS

In this section, we will illustrate the efficiency and accuracy of the eTTI-IF method, by applying it to study the real-time evolution of the AIM. Although our method can directly target at the infinite time limit, we will focus on finite t in our numerical simulations. We choose the initial state as $\hat{\rho} = \hat{\rho}_{\rm imp} \otimes \hat{\rho}^{\rm th}_{\rm bath}$, where $\hat{\rho}_{\rm imp} = |\emptyset_{\rm imp}\rangle\langle\emptyset_{\rm imp}|$ ($|\emptyset_{\rm imp}\rangle$ denotes the vacuum state of the impurity). For this initial state, the lesser Green's function, defined as $G^<(t) = -{\rm i}\langle \hat{a}^\dagger \hat{a}(t)\rangle$, vanishes, so we focus on the greater Green's function, defined as $G^>(t) = -{\rm i}\langle \hat{a}(t)\hat{a}^\dagger\rangle$, which is equal to the retarded Green's function in this case. To benchmark the general performance of the eTTI-IF method, we consider two commonly used BSFs: (1) the semi-circular function

$$J_s(\epsilon) = \frac{\Gamma}{2\pi} D\sqrt{1-(\epsilon/D)^2}, \tag{24}$$

with $\Gamma = 0.1$; and (2) the Lorentzian function

$$J_l(\epsilon) = \frac{1}{\pi}\frac{D}{\epsilon^2 + D^2}. \tag{25}$$

In both cases we set $D = 2$. The inverse temperature will be chosen as $\beta = 10$. In all simulations, we use a discrete time step size $D\delta t = 0.1$ and a maximum bond dimension $\chi = 80$.

To achieve a compact approximation of the hybridization function in Eq.(19), we employ a scheme which combines the Prony algorithm with a subsequent least-square fitting (LSF) to minimize ς for a given n. Starting from $n = 1$, we first apply the Prony algorithm, and then we use the output as the initial guess for LSF to further reduce ς. If the final ς is still larger than the predefined tolerance, we increase n and repeat the previous procedure until the desired precision is obtained.

We first apply the eTTI-IF method to the non-interacting case ($U = 0$), for which an analytical solution is available. We will also choose $\mu = 0$. In Fig. 2(a, b), we show the imaginary part of $G^>(t)$ as a function of t, for the two BSFs in Eq.(24) and Eq.(25) respectively. The analytical solutions are shown in black solid lines. Three sets of GTEMPO results are shown, with IF built using the partial IF method (red solid line), using the eTTI IF method (the blue solid line), and using the TTI IF method with three different $m = \log_2(1/\delta)$ ($m = 3, 5, 7$ correspond to the three dashed green lines from lighter to darker). For the approximation in Eq.(19) which is used in the TTI-IF and eTTI-IF methods, we have set a tolerance $\varsigma = 10^{-5}$, for which we find that $n = 4$ suffices for both BSFs. The insets in Fig. 2(a, b) show the absolute errors of these GTEMPO results against the analytical solutions. We can see that all the GTEMPO results agree well with the analytical solutions for both BSFs, with errors on the order of 10^{-2} or less. The partial IF results are the most accurate, which is as expected as they suffer from the least sources of errors. Importantly, the TTI IF results become closer to the eTTI IF results

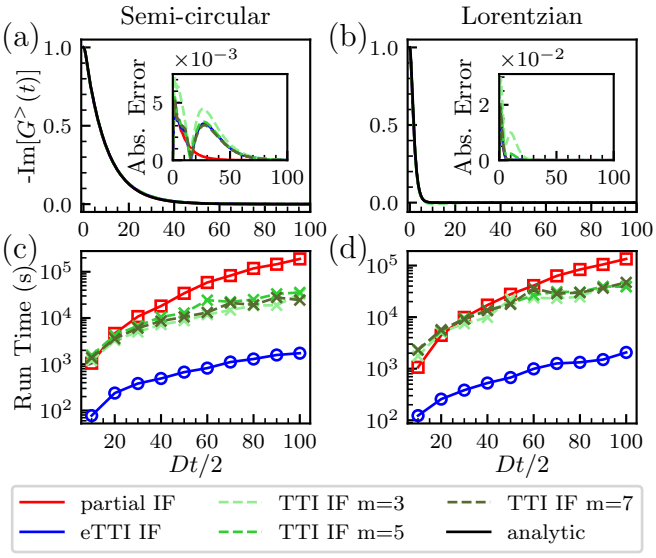

FIG. 2. (a,b) Imaginary part of $G^>(t)$ for the semi-circular (a) and Lorentzian (b) BSFs, calculated using the partial IF method (red solid line), the eTTI IF method (blue solid line), and the TTI-IF method with $m = 3, 5, 7$ (green dashed lines from lighter to darker). The analytical solutions are shown in black solid lines. The insets show the absolute errors of the results calculated using different methods against the analytical solutions. (c,d) The run time of different methods used to build the IF as GMPSs as a function of the total evolution time t for the semi-circular (c) and Lorentzian (d) BSFs.

with increasing m, which is also as expected as the former suffers from an additional time discretization error characterized by δ (or equivalently m) compared to the latter, and in the limit of large m these two sets of results should become equivalent. In Fig. 2(c,d), we show the run time to building the IF as GMPSs in these three sets of GTEMPO calculations as a function of the total evolution time t. We can see that the TTI-IF method is more efficient than the partial IF method at large t and is only mildly dependent on the choice of m. Crucially, the eTTI-IF method is more efficient than the TTI IF and partial IF methods by at least one order of magnitude.

The computational cost and accuracy of the eTTI-IF method are directly related to the number of exponential functions in Eq.(19), i.e., n. In Fig. 3(a), we show the value of n required to reach different tolerance ς for a fixed $t = 100$, where the blue solid line with circle and yellow solid line with x represent the results for the semi-circular and Lorentzian BSFs respectively. We observe an approximately linear relationship $n \propto 1/\log(\varsigma)$ for both BSFs (except for the Lorentzian BSF with very small $\varsigma < 10^{-7}$). In Fig. 3(c), we plot the maximum error of $G^>(t)$ calculated using the eTTI IF method, defined as the maximum value of the absolute error against the analytical solution on the whole time interval, as a function of the tolerance ς used in Eq.(19). We can see that the maximum errors saturate with $\varsigma = 10^{-5}$ and $\varsigma = 10^{-4}$ for the semi-circular and Lorentzian BSFs re-

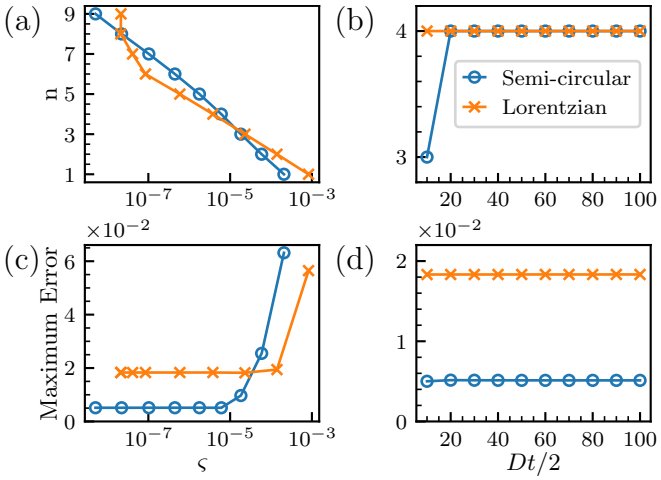

FIG. 3. (a) The number of exponential functions n, required to achieve a tolerance ς for the semi-circular (blue solid line with circle) and Lorentzian (orange line with x) BSFs, where we have fixed $t = 100$. (b) The number of exponential functions n to achieve a fixed tolerance $\varsigma = 10^{-5}$ as a function of t. (c,d) The maximum error of the imaginary part of $G^{>}(t)$ calculated using the eTTI IF method against the analytical solutions, as a function of the tolerance ς (c) and t (d), respectively.

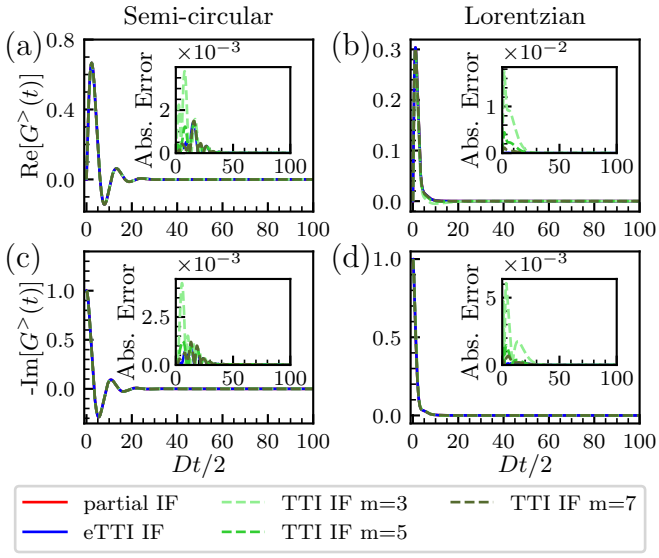

FIG. 4. (a,c) Real (a) and imaginary (c) parts of $G^{>}(t)$ for the AIM with the semi-circular BSF. (b,d) Real (b) and imaginary (d) parts of $G^{>}(t)$ for the AIM with the Lorentzian BSF. The blue solid lines represent the eTTI IF results, while the green dashed lines from lighter to darker are the TTI IF results with $m = 3, 5, 7$ respectively, similar to Fig. 2. The insets show the absolute error of the TTI-IF and eTTI-IF results against the partial IF results.

spectively. In Fig. 3(b), we plot n as a function of t, with a fixed tolerance of $\varsigma = 10^{-5}$ for both BSFs. We observe that n remains almost a constant throughout the time interval we have considered. Under the same setup as in Fig. 3(b), in Fig. 3(d) we plot the maximum error of $G^{>}(t)$ calculated by the eTTI-IF method against the analytical solution, as a function of t, we can see that the maximum error is on the order of 10^{-2} and keep almost constant throughout the whole time interval.

Finally, we apply the IFs calculated using different methods to compute $G^{>}(t)$ for the AIM with $U = 1, \mu = -0.5$. As no exact analytical solution exists for the interacting case, we use the partial IF results as a benchmark baseline to assess the accuracy of the eTTI IF method and the TTI IF method with $m = 3, 5, 7$, similar to Fig. 2. We plot the real and imaginary parts of $G^{>}(t)$ in Fig. 4(a,c) for the semi-circular BSF, and in Fig. 4(b,d) for the Lorentzian BSF, where the blue solid lines are the eTTI IF results, and the green dashed lines from lighter to darker are the TTI IF results with $m = 3, 5, 7$ respectively. The insets show the absolute errors of these results against the partial IF results. For both BSFs, we can see that the eTTI IF results are the closest to the partial IF results, while the TTI IF results get closer to the eTTI IF results for larger m, similar to the observations in the non-interacting case.

V. CONCLUSION

In summary, we have proposed an eTTI IF method to efficiently and accurately build the Feynman-Vernon IF as a Grassmann MPS, which is the computationally dominate step in the GTEMPO method to solve the AIM. Our method has a lower computational cost compared to the existing alternatives. The only additional source of error in the eTTI IF method, compared to the partial IF method, is the error occurred in approximating the hybridization function as a summation of exponential functions. Compared to the TTI IF method, the eTTI IF makes full use of the special properties of the Grassmann algebra, and is guaranteed to be more efficient and accurate as it suffers from fewer sources of error than the latter. Similar to the TTI IF method, the eTTI IF method can be straightforwardly applied to both finite- and infinite-time cases. In our numerical examples, we consider two commonly used bath spectral functions, and both the non-interacting and interacting cases. Our numerical results show that the eTTI IF method could reach a precision similar to the partial IF method, and that it is more accurate than the TTI IF method. Crucially, we observe that the eTTI IF method is at least one order of magnitude faster than the partial IF method and the TTI IF method. These results illustrate that our method can be used to greatly accelerate the solution of the AIM.

For multi-orbital extensions of the AIM, the eTTI IF method can still be used to accelerate the step of build-

ing the GMPS for the IF of each flavor. However, the computational cost may no longer be dominated by this step, but by calculating the ADT or the observables instead [62].

ACKNOWLEDGMENTS

Z. L. is partially supported by NSFC (22393913), by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0450101).

Appendix A: Proof of Theorem 1

As the two GMPSs in Eq.(22) have essentially the same form, we only need to prove Theorem 1 for one of them. In the following we will focus on the first product on the first line of Eq.(21). We perform a Taylor expansion of one term in this product,

$$e^{-\sum_{j>k}\bar{a}_j\alpha\lambda^{j-k}a_k} = 1-\sum_{j>k}\bar{a}_j\alpha\lambda^{j-k}a_k + \frac{1}{2!}\left(\sum_{j>k}\bar{a}_j\alpha\lambda^{j-k}a_k\right)^2-\cdots, \tag{A1}$$

where we have omitted the subscript l for brevity. It have been proven that the WII approximation contains all the terms in the Taylor expansion that overlap by at most one site [64]. Therefore the zeroth- and first-order terms on the first line of Eq.(A1) are already contained in the WII approximation. In the next we prove that all the terms in the Taylor expansion that overlap by more than one site exactly cancel each other, which completes the proof.

We impose a normal-ordering prescription similar to our ordering of the discrete GVs in our numerical implementation, in which the GVs are arranged according to their time step indices, and when two GVs share the same index, a will be conventionally placed before $\bar{a}$. For each pattern (normal ordered GVs), there could exist multiple combinations of the quadratic terms on the exponent that give rise to it. To account for all the possible combinations, one must systematically enumerate every possible pairing of $\bar{a}_i a_j$ with $i > j$ in the pattern. For a given pattern, different pairings give the same prefactor, and only the anti-commutation relations between GVs could introduce different sign factors. So hereafter we omit the scalar prefactor for simplicity. If two a (or $\bar{a}$) share the same index in one pattern, this pattern vanishes due to the Grassmann algebra $a^2 = \bar{a}^2 = 0$. So we will omit these trivially vanishing cases. We note that with our convention, a legal pattern must start with $\bar{a}$ and end with a.

As illustrative examples, we examine the second- and third-order contributions. For each order we enumerate all possible normal-ordered patterns, and then determine whether they yield non-vanishing contributions.

In the second order, there are only two pattern, i.e., $\bar{a}_i a_j \bar{a}_k a_l$ and $\bar{a}_i \bar{a}_j a_k a_l$. We categorize them into two classes:

- Set-1 $S_1 = \{\bar{a}_i a_j \bar{a}_k a_l\}$. The only possible combination in this pattern is of the form $\bar{a}_i a_j * \bar{a}_k a_l$ (denoted as $\overbrace{\bar{a}_i a_j \bar{a}_k a_l}$), which are already contained in the WII approximation as the two pairings overlap by at most one site if $j = k$.

- Set-2 $S_2 = \{\bar{a}_i \bar{a}_j a_k a_l\}$, which may arise from $\bar{a}_i a_k * \bar{a}_j a_l$ ($\overbrace{\bar{a}_i \bar{a}_j a_k a_l}$) or $\bar{a}_i a_l * \bar{a}_j a_k$ ($\overbrace{\bar{a}_i \bar{a}_j a_k a_l}$). These two combinations have opposite sign factors (the first combination requires a single swap operation, while the second combination requires two swap operations), so they cancel each other.

Similarly, in the third order, we categorize all the possible patterns into two classes:

- Set-1 $S_1 = \{\bar{a}_i a_j \bar{a}_k a_l \bar{a}_m a_n\}$. Again, the only possible combination in this pattern is of the form $\bar{a}_i a_j * \bar{a}_k a_l * \bar{a}_m a_n$ ($\overbrace{\bar{a}_i a_j \bar{a}_k a_l \bar{a}_m a_n}$), which is already contained in the WII approximation.

- Set-2 S_2 contains all the patterns except the one in S_1, which can be exhausted as

$$S_2 = \{\bar{a}_i a_j \bar{a}_k \bar{a}_l a_m a_n, \bar{a}_i \bar{a}_j a_k \bar{a}_l a_m a_n, \bar{a}_i \bar{a}_j \bar{a}_k a_l a_m a_n, \bar{a}_i \bar{a}_j a_k a_l \bar{a}_m a_n\}. \tag{A2}$$

  We analyze each pattern in S_2 in the following.

  - $\bar{a}_i a_j \bar{a}_k \bar{a}_l a_m a_n$. There are many possible combinations to generate this pattern. In order to enumerate all of them, we first group the last a with one $\bar{a}$ before it, and if left terms constitute a legal pattern, this combination is legal, and will be considered. This procedure gives two possible combinations: $\bar{a}_i a_j \bar{a}_l a_m * \bar{a}_k a_n$ ($\overbrace{\bar{a}_i a_j \bar{a}_k \bar{a}_l a_m a_n}$) and $\bar{a}_i a_j \bar{a}_k a_m * \bar{a}_l a_n$ ($\overbrace{\bar{a}_i a_j \bar{a}_k \bar{a}_l a_m a_n}$), which cancel with each other. Here we note that $a_j \bar{a}_k \bar{a}_l a_m * \bar{a}_i a_n$ is an illegal combination.

  - $\bar{a}_i \bar{a}_j a_k \bar{a}_l a_m a_n$. There are three possible combinations: $\bar{a}_j a_k \bar{a}_l a_m * \bar{a}_i a_n$ ($\overbrace{\bar{a}_i \bar{a}_j a_k \bar{a}_l a_m a_n}$), $\bar{a}_i a_k \bar{a}_l a_m * \bar{a}_j a_n$ ($\overbrace{\bar{a}_i \bar{a}_j a_k \bar{a}_l a_m a_n}$) and $\bar{a}_i \bar{a}_j a_k a_m * \bar{a}_l a_n$ ($\overbrace{\bar{a}_i \bar{a}_j a_k \bar{a}_l a_m a_n}$). The first two cancel each other. The last one contain $\bar{a}_i \bar{a}_j a_k a_m$, which is in the secnd order's S_2, so it has zero contribution upon further decomposing into $\bar{a}_j a_k * \bar{a}_i a_m * \bar{a}_l a_n$

$(\bar{a}_i\bar{a}_j a_k\bar{a}_l a_m a_n)$ and $\bar{a}_i a_k * \bar{a}_j a_m * \bar{a}_l a_n$ $(\bar{a}_i\bar{a}_j a_k\bar{a}_l a_m a_n)$.

- $\bar{a}_i\bar{a}_j\bar{a}_k a_l a_m a_n$. There are three possible combinations: $\bar{a}_j\bar{a}_k a_l a_m * \bar{a}_i a_n$ $(\bar{a}_i\bar{a}_j\bar{a}_k a_l a_m a_n)$, $\bar{a}_i\bar{a}_k a_l a_m * \bar{a}_j a_n$ $(\bar{a}_i\bar{a}_j\bar{a}_k a_l a_m a_n)$ and $\bar{a}_i\bar{a}_j a_l a_m * \bar{a}_k a_n$ $(\bar{a}_i\bar{a}_j\bar{a}_k a_l a_m a_n)$. They all contain the $\bar{a}\bar{a}aa$ sub-pattern, that belongs to the S_2 in second order, thus they all have zero contributions upon further decomposition.
- $\bar{a}_i\bar{a}_j a_k a_l\bar{a}_m a_n$. The only possible combination in this pattern is of the form $\bar{a}_i\bar{a}_j a_k a_l * \bar{a}_m a_n$ $(\bar{a}_i\bar{a}_j a_k a_l\bar{a}_m a_n)$, which also have zero contribution as it contains the $\bar{a}\bar{a}aa$ sub-pattern.

We therefore classify all the patterns in each order into two categories:

- Set-1 (S_1), which contains a single pattern with a strictly alternating ordering of the form $\bar{a}a\bar{a}a\ldots\bar{a}a$, which is already included in the WII approximation.
- Set-2 (S_2), which contains all the remaining patterns, characterized by the presence of multiple overlapping sites and are not accounted for in the WII approximation.

Lemma. At any order n, every pattern that belongs to Set-2 has zero contribution.

Proof of Lemma. The result holds for $n = 2$ as shown above. Let's assume that it holds for order n. And we consider a Set-2 pattern at order $n+1$. Grouping the last a with one preceding $\bar{a}$, if the remaining GVs form a Set-2 pattern at order n, the result vanishes by the induction hypothesis, if the remaining part is in Set-1 instead, then the pattern must necessarily be of the form

$$(\bar{a}a)^n\bar{a}\bar{a}a(\bar{a}a)^m a, \quad n, m \in \{0, 1, 2\ldots\}. \tag{A3}$$

There will always be two combinations that result in the sub-pattern $\bar{a}a\bar{a}a\ldots\bar{a}a$:

$$(\bar{a}a)^n\bar{a}\bar{a}a(\bar{a}a)^m a, \quad (\bar{a}a)^n\bar{a}\bar{a}a(\bar{a}a)^m a, \tag{A4}$$

which cancel each other exactly. Thus all Set-2 contributions vanish at order $n+1$, completing the induction.

As all patterns in Set-2 vanish, the WII approximation becomes exact for the GMPSs in Eq.(22), thereby completing the proof of Theorem 1. Finally, we note that a slightly counter-intuitive observation is that this proof does not seem to be generalizable to the bosonic case, as we can see from the proof that the anti-commutation relation of the GVs is crucial, which does not hold for bosons.

[1] P. W. Anderson, Localized magnetic states in metals, Phys. Rev. **124**, 41 (1961).
[2] G. D. Mahan, *Many-Particle Physics* (Springer; 3nd edition, 2000).
[3] A. Georges, The beauty of impurities: Two revivals of friedel's virtual bound-state concept, Comptes Rendus Physique **17**, 430 (2016), physique de la matière condensée au XXIe siècle: l'héritage de Jacques Friedel.
[4] F. Heidrich-Meisner, A. E. Feiguin, and E. Dagotto, Real-time simulations of nonequilibrium transport in the single-impurity anderson model, Phys. Rev. B **79**, 235336 (2009).
[5] F. Schwarz, I. Weymann, J. von Delft, and A. Weichselbaum, Nonequilibrium steady-state transport in quantum impurity models: A thermofield and quantum quench approach using matrix product states, Phys. Rev. Lett. **121**, 137702 (2018).
[6] U. Weiss, *Quantum Dissipative Systems* (World Scientific, Singapore, 1993).
[7] W. Metzner and D. Vollhardt, Correlated lattice fermions in $d = \infty$ dimensions, Phys. Rev. Lett. **62**, 324 (1989).
[8] A. Georges and G. Kotliar, Hubbard model in infinite dimensions, Phys. Rev. B **45**, 6479 (1992).
[9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. **68**, 13 (1996).
[10] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Electronic structure calculations with dynamical mean-field theory, Rev. Mod. Phys. **78**, 865 (2006).
[11] M. Caffarel and W. Krauth, Exact diagonalization approach to correlated fermions in infinite dimensions: Mott transition and superconductivity, Phys. Rev. Lett. **72**, 1545 (1994).
[12] E. Koch, G. Sangiovanni, and O. Gunnarsson, Sum rules and bath parametrization for quantum cluster theories, Phys. Rev. B **78**, 115102 (2008).
[13] M. Granath and H. U. R. Strand, Distributional exact diagonalization formalism for quantum impurity models, Phys. Rev. B **86**, 115111 (2012).
[14] Y. Lu, M. Höppner, O. Gunnarsson, and M. W. Haverkort, Efficient real-frequency solver for dynamical mean-field theory, Phys. Rev. B **90**, 085102 (2014).
[15] C. Mejuto-Zaera, L. Zepeda-Núñez, M. Lindsey, N. Tubman, B. Whaley, and L. Lin, Efficient hybridization fitting for dynamical mean-field theory via semi-definite relaxation, Phys. Rev. B **101**, 035143 (2020).
[16] R.-Q. He and Z.-Y. Lu, Quantum renormalization groups based on natural orbitals, Phys. Rev. B **89**, 085108 (2014).
[17] R.-Q. He, J. Dai, and Z.-Y. Lu, Natural orbitals renormalization group approach to the two-impurity kondo critical point, Phys. Rev. B **91**, 155140 (2015).

[18] A. N. Rubtsov and A. I. Lichtenstein, Continuous-time quantum monte carlo method for fermions: Beyond auxiliary field framework, J. Exp. Theor. Phys. Lett. **80**, 61 (2004).

[19] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, Continuous-time quantum monte carlo method for fermions, Phys. Rev. B **72**, 035122 (2005).

[20] P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A. J. Millis, Continuous-time solver for quantum impurity models, Phys. Rev. Lett. **97**, 076405 (2006).

[21] E. Gull, P. Werner, O. Parcollet, and M. Troyer, Continuous-time auxiliary-field monte carlo for quantum impurity models, EPL **82**, 57003 (2008).

[22] C.-K. Chan, P. Werner, and A. J. Millis, Magnetism and orbital ordering in an interacting three-band model: A dynamical mean-field study, Phys. Rev. B **80**, 235114 (2009).

[23] K. Haule, C.-H. Yee, and K. Kim, Dynamical mean-field theory within the full-potential methods: Electronic structure of CeIrIn5, CeCoIn5, and CeRhIn5, Phys. Rev. B **81**, 195107 (2010).

[24] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Continuous-time monte carlo methods for quantum impurity models, Rev. Mod. Phys. **83**, 349 (2011).

[25] L. Huang, T. O. Wehling, and P. Werner, Electronic excitation spectra of the five-orbital anderson impurity model: From the atomic limit to itinerant atomic magnetism, Phys. Rev. B **89**, 245104 (2014).

[26] H. Lu and L. Huang, Pressure-driven4flocalized-itinerant crossover in heavy-fermion compoundcein3: A first-principles many-body perspective, Phys. Rev. B **94**, 075132 (2016).

[27] C. Yue and P. Werner, Pairing enhanced by local orbital fluctuations in a model for monolayer fese, Phys. Rev. B **104**, 184507 (2021).

[28] K. G. Wilson, The renormalization group: Critical phenomena and the kondo problem, Rev. Mod. Phys. **47**, 773 (1975).

[29] R. Bulla, Zero temperature metal-insulator transition in the infinite-dimensional hubbard model, Phys. Rev. Lett. **83**, 136 (1999).

[30] R. Bulla, T. A. Costi, and T. Pruschke, Numerical renormalization group method for quantum impurity systems, Rev. Mod. Phys. **80**, 395 (2008).

[31] F. B. Anders, A numerical renormalization group approach to non-equilibrium green functions for quantum impurity models, J. Phys. Condens. Matter **20**, 195216 (2008).

[32] R. Žitko and T. Pruschke, Energy resolution and discretization artifacts in the numerical renormalization group, Phys. Rev. B **79**, 085106 (2009).

[33] X. Deng, J. Mravlje, R. Žitko, M. Ferrero, G. Kotliar, and A. Georges, How bad metals turn good: Spectroscopic signatures of resilient quasiparticles, Phys. Rev. Lett. **110**, 086401 (2013).

[34] K. M. Stadler, Z. P. Yin, J. von Delft, G. Kotliar, and A. Weichselbaum, Dynamical mean-field theory plus numerical renormalization-group study of spin-orbital separation in a three-band hund metal, Phys. Rev. Lett. **115**, 136401 (2015).

[35] S.-S. B. Lee and A. Weichselbaum, Adaptive broadening to improve spectral resolution in the numerical renormalization group, Phys. Rev. B **94**, 235127 (2016).

[36] S.-S. B. Lee, J. von Delft, and A. Weichselbaum, Doublon-holon origin of the subpeaks at the hubbard band edges, Phys. Rev. Lett. **119**, 236402 (2017).

[37] F. A. Wolf, I. P. McCulloch, O. Parcollet, and U. Schollwöck, Chebyshev matrix product state impurity solver for dynamical mean-field theory, Phys. Rev. B **90**, 115124 (2014).

[38] M. Ganahl, P. Thunström, F. Verstraete, K. Held, and H. G. Evertz, Chebyshev expansion for impurity models using matrix product states, Phys. Rev. B **90**, 045144 (2014).

[39] M. Ganahl, M. Aichhorn, H. G. Evertz, P. Thunström, K. Held, and F. Verstraete, Efficient dmft impurity solver using real-time dynamics with matrix product states, Phys. Rev. B **92**, 155132 (2015).

[40] F. A. Wolf, A. Go, I. P. McCulloch, A. J. Millis, and U. Schollwöck, Imaginary-time matrix product state impurity solver for dynamical mean-field theory, Phys. Rev. X **5**, 041032 (2015).

[41] D. J. García, K. Hallberg, and M. J. Rozenberg, Dynamical mean field theory with the density matrix renormalization group, Phys. Rev. Lett. **93**, 246403 (2004).

[42] S. Nishimoto, F. Gebhard, and E. Jeckelmann, Dynamical mean-field theory calculation with the dynamical density-matrix renormalization group, Physica B Condens. Matter **378-380**, 283 (2006).

[43] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, Variational matrix-product-state approach to quantum impurity models, Phys. Rev. B **80**, 165117 (2009).

[44] D. Bauernfeind, M. Zingl, R. Triebl, M. Aichhorn, and H. G. Evertz, Fork tensor-product states: Efficient multiorbital real-time dmft solver, Phys. Rev. X **7**, 031013 (2017).

[45] Y. Lu, X. Cao, P. Hansmann, and M. W. Haverkort, Natural-orbital impurity solver and projection approach for green's functions, Phys. Rev. B **100**, 115134 (2019).

[46] D. Werner, J. Lotze, and E. Arrigoni, Configuration interaction based nonequilibrium steady state impurity solver, Phys. Rev. B **107**, 075119 (2023).

[47] L. Kohn and G. E. Santoro, Efficient mapping for anderson impurity problems with matrix product states, Phys. Rev. B **104**, 014303 (2021).

[48] L. Kohn and G. E. Santoro, Quench dynamics of the anderson impurity model at finite temperature using matrix product states: entanglement and bath dynamics, J. Stat. Mech. Theory Exp. **2022**, 063102 (2022).

[49] Y. Tanimura and R. Kubo, Time evolution of a quantum system in contact with a nearly gaussian-markoffian noise bath, J. Phys. Soc. Jpn. **58**, 101 (1989).

[50] J. Jin, S. Welack, J. Luo, X.-Q. Li, P. Cui, R.-X. Xu, and Y. Yan, Dynamics of quantum dissipation systems interacting with fermion and boson grand canonical bath ensembles: Hierarchical equations of motion approach, J. Chem. Phys. **126**, 134113 (2007).

[51] J. Jin, X. Zheng, and Y. Yan, Exact dynamics of dissipative electronic systems and quantum transport: Hierarchical equations of motion approach, J. Chem. Phys. **128**, 234703 (2008).

[52] Y. Yan, J. Jin, R.-X. Xu, and X. Zheng, Dissipation equation of motion approach to open quantum systems, Front. Phys. **11**, 110306 (2016).

[53] J. Cao, L. Ye, R. Xu, X. Zheng, and Y. Yan, Recent advances in fermionic hierarchical equations of motion

method for strongly correlated quantum impurity systems, JUSTC **53**, 0302 (2023).
[54] R. P. Feynman and F. L. Vernon, The theory of a general quantum system interacting with a linear dissipative system, Ann. Phys. **24**, 118 (1963).
[55] J. Thoenniss, A. Lerose, and D. A. Abanin, Nonequilibrium quantum impurity problems via matrix-product states in the temporal domain, Phys. Rev. B **107**, 195101 (2023).
[56] J. Thoenniss, M. Sonner, A. Lerose, and D. A. Abanin, Efficient method for quantum impurity problems out of equilibrium, Phys. Rev. B **107**, L201115 (2023).
[57] N. Ng, G. Park, A. J. Millis, G. K.-L. Chan, and D. R. Reichman, Real-time evolution of anderson impurity models via tensor network influence functionals, Phys. Rev. B **107**, 125103 (2023).
[58] G. Park, N. Ng, D. R. Reichman, and G. K.-L. Chan, Tensor network influence functionals in the continuous-time limit: Connections to quantum embedding, bath discretization, and higher-order time propagation, Phys. Rev. B **110**, 045104 (2024).
[59] B. Kloss, J. Thoenniss, M. Sonner, A. Lerose, M. T. Fishman, E. M. Stoudenmire, O. Parcollet, A. Georges, and D. A. Abanin, Equilibrium quantum impurity problems via matrix product state encoding of the retarded action, Phys. Rev. B **108**, 205110 (2023).
[60] R. Chen, X. Xu, and C. Guo, Grassmann time-evolving matrix product operators for quantum impurity models, Phys. Rev. B **109**, 045140 (2024).
[61] R. Chen, X. Xu, and C. Guo, Real-time impurity solver using grassmann time-evolving matrix product operators, Phys. Rev. B **109**, 165113 (2024).
[62] R. Chen, X. Xu, and C. Guo, Grassmann time-evolving matrix product operators for equilibrium quantum impurity problems, New J. Phys. **26**, 013019 (2024).
[63] R. Chen and C. Guo, Solving equilibrium quantum impurity problems on the l-shaped kadanoff-baym contour, Phys. Rev. B **110**, 165114 (2024).
[64] M. P. Zaletel, R. S. K. Mong, C. Karrasch, J. E. Moore, and F. Pollmann, Time-evolving a matrix product state with long-ranged interactions, Phys. Rev. B **91**, 165112 (2015).
[65] J. Thoenniss, I. Vilkoviskiy, and D. A. Abanin, Efficient pseudomode representation and complexity of quantum impurity models (2024), arXiv:2409.08816 [cond-mat.str-el].
[66] I. Vilkoviskiy and D. A. Abanin, Bound on approximating non-markovian dynamics by tensor networks in the time domain, Phys. Rev. B **109**, 205126 (2024).
[67] X. Xu, C. Guo, and R. Chen, Grassmann time-evolving matrix product operators: An efficient numerical approach for fermionic path integral simulations, J. Chem. Phys. **161**, 151001 (2024).
[68] R. Chen, Path integral formalism for quantum open systems, Ann. Phys. **480**, 170083 (2025).
[69] L. V. Keldysh, Diagram technique for non-equilibrium processes, Soviet Physics JETP **20**, 1018 (1965).
[70] D. E. Makarov and N. Makri, Path integrals for dissipative systems by tensor multiplication. condensed phase quantum dynamics for arbitrarily long time, Chem. Phys. Lett. **221**, 482 (1994).
[71] N. Makri, Numerical path integral techniques for long time dynamics of quantum dissipative systems, J. Math. Phys. **36**, 2430 (1995).
[72] C. Guo and R. Chen, Efficient construction of the Feynman-Vernon influence functional as matrix product states, SciPost Phys. Core **7**, 063 (2024).
[73] C. Guo and R. Chen, Infinite grassmann time-evolving matrix product operator method for zero-temperature equilibrium quantum impurity problems, Phys. Rev. B **110**, 165119 (2024).
[74] A. Strathearn, P. Kirton, D. Kilda, J. Keeling, and B. W. Lovett, Efficient non-markovian quantum dynamics using time-evolving matrix product operators, Nat. Commun. **9**, 3322 (2018).
[75] D. Gribben, D. M. Rouse, J. Iles-Smith, A. Strathearn, H. Maguire, P. Kirton, A. Nazir, E. M. Gauger, and B. W. Lovett, Exact dynamics of nonadditive environments in non-markovian open quantum systems, PRX Quantum **3**, 010321 (2022).
[76] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states, Ann. Phys. **326**, 96 (2011).
[77] R. Chen, Path integral formalism of open quantum systems with non-diagonal system-bath coupling, Communications in Theoretical Physics **76**, 115701 (2024).
[78] Z. Sun, R. Chen, Z. Li, and C. Guo, Scalable tensor network algorithm for quantum impurity problems, Phys. Rev. B **112**, 155115 (2025).
[79] M. T. Fishman and S. R. White, Compression of correlation matrices and an efficient method for forming matrix product states of fermionic gaussian states, Phys. Rev. B **92**, 075132 (2015).
[80] S. L. Marple Jr, *Digital spectral analysis* (Courier Dover Publications, 2019).
[81] V. Link, H.-H. Tu, and W. T. Strunz, Open quantum system dynamics from infinite tensor network contraction, Phys. Rev. Lett. **132**, 200403 (2024).