



This is the accepted manuscript made available via CHORUS. The article has been published as:

## Tensor train continuous time solver for quantum impurity models

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 — Published 26 June 2023

DOI: [10.1103/PhysRevB.107.245135](https://doi.org/10.1103/PhysRevB.107.245135)

# A Tensor Train Continuous Time Solver for Quantum Impurity Models

A. Erpenbeck,<sup>1</sup> W.-T. Lin,<sup>1</sup> T. Blommel,<sup>1</sup> L. Zhang,<sup>1</sup> S. Isakov,<sup>1</sup> L. Bernheimer,<sup>2</sup>  
Y. Núñez-Fernández,<sup>3,4</sup> G. Cohen,<sup>2,5</sup> O. Parcollet,<sup>6,7</sup> X. Waintal,<sup>3</sup> and E. Gull<sup>1</sup>

<sup>1</sup>*Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA*

<sup>2</sup>*School of Chemistry, Tel Aviv University, Tel Aviv 6997801, Israel*

<sup>3</sup>*Université Grenoble Alpes, CEA, Grenoble INP, IRIG, Pheligs, F-38000 Grenoble, France*

<sup>4</sup>*Université Grenoble Alpes, CNRS, Institut Néel, F-38000 Grenoble, France*

<sup>5</sup>*The Raymond and Beverley Sackler Center for Computational Molecular  
and Materials Science, Tel Aviv University, Tel Aviv 6997801, Israel*

<sup>6</sup>*Center for Computational Quantum Physics, Flatiron Institute, New York, New York, 10010, USA*

<sup>7</sup>*Université Paris-Saclay, CNRS, CEA, Institut de Physique Théorique, 91191, Gif-sur-Yvette, France*

(Dated: June 7, 2023)

The simulation of strongly correlated quantum impurity models is a significant challenge in modern condensed matter physics that has multiple important applications. Thus far, the most successful methods for approaching this challenge involve Monte Carlo techniques that accurately and reliably sample perturbative expansions to any order. However, the cost of obtaining high precision through these methods is high. Recently, tensor train decomposition techniques have been developed as an alternative to Monte Carlo integration. In this study, we apply these techniques to the single-impurity Anderson model at equilibrium by calculating the systematic expansion in power of the hybridization of the impurity with the bath. We demonstrate the performance of the method in a paradigmatic application, examining the first-order phase transition on the infinite dimensional Bethe lattice, which can be mapped to an impurity model through dynamical mean field theory. Our results indicate that using tensor train decomposition schemes allows the calculation of finite-temperature Green’s functions and thermodynamic observables with unprecedented accuracy. The methodology holds promise for future applications to frustrated multi-orbital systems, using a combination of partially summed series with other techniques pioneered in diagrammatic and continuous-time quantum Monte Carlo.

## I. INTRODUCTION

The solution of strongly correlated quantum impurity models is one of the central tasks of condensed matter physics. Quantum impurity models describe the physics of an interacting ‘impurity’ or ‘quantum dot’ coupled to a large, typically infinite, number of noninteracting ‘bath’ or ‘lead’ degrees of freedom. Quantum impurity models were initially developed to describe the physics of magnetic impurities embedded in a non-magnetic host [1], but have since been adapted to describe quantum dots and molecular junctions [2], as well as atoms and molecules adsorbed on surfaces [3, 4]. Importantly, they also appear as auxiliary models in quantum embedding techniques such as dynamical mean field theory (DMFT) and self-energy embedding theory [5–8], which typically require the calculation of a finite-temperature Green’s function in the strongly correlated regime.

A reliable description of correlated systems requires methods that are numerically exact, in the sense that errors can be made arbitrarily small as a function of a control parameter. Among such methods, the continuous-time quantum Monte Carlo (CT-QMC) impurity solvers [9–12], which are based on the stochastic sampling of a perturbative expansion to all orders, have become ubiquitous in cluster [13] and real-materials DMFT applications [7]. Numerous variants [14–21], improvements, and open source implementations [22–30] exist.

CT-QMC methods provide exact results within Monte

Carlo confidence intervals. In particular, they do not require a discretization of bath degrees of freedom, like exact diagonalization methods [31–33], or of the time degrees of freedom, like Hirsch and Fye [34] or lattice Monte Carlo methods [35]. However, they suffer from the following two limitations: (1) Away from high symmetry points, ‘sign problems’ may cause the computational cost to grow exponentially as a function of system size, inverse temperature, or interaction strength, limiting calculations (with a few exceptions [20, 21]) in practice to systems with almost diagonal interactions and/or hybridizations, and (2) thermodynamic quantities such as the partition function and free energy are, in standard implementations, only available up to an unknown normalization constant [36].

The standard framework of CT-QMC and, more generally, of ‘diagrammatic’ [37–39] Monte Carlo methods is based on a perturbative series expansion that expresses observables of a quantum system in terms of an infinite series of high-dimensional integrals. This series is then summed to all orders in a stochastic sampling process, employing a Monte Carlo sampling procedure [12] that performs a random walk in diagram space. For a given number of samples  $n_s$ , this procedure produces unbiased stochastic estimates with errors that decrease rather slowly as  $\sim 1/\sqrt{n_s}$ . Notably, quasi-Monte Carlo methods can substantially accelerate this to  $\sim 1/n_s$  in at least some cases [40, 41]. In the context of correlated quantum transport, some of us [42] have recently shown that in the calculation of many high-dimensional integrals of the

perturbative series expansion, it may be advantageous to replace Monte Carlo integration by a decomposition of the integrand into a product of low-dimensional tensors, which can then be integrated separately [43]. The method is controlled in the sense that the exact result is recovered as the tensor rank increases.

The decomposition of tensors into approximate low-rank forms without needing to evaluate all their elements is based on tensor cross-interpolation (TCI) methods [44, 45]. The corresponding approximation of a high-dimensional tensor by the product of a sequence of low-rank tensors is known as a ‘tensor train’ in the applied mathematics and computer science literature [46]. Tensor trains are also known as matrix product states (MPS) pioneered in the density matrix renormalization group (DMRG) and related methods [47, 48].

In this paper, we adapt the tensor train approach of Núñez-Fernández *et al.* [42], to equilibrium imaginary-time quantum impurity problems. We benchmark the impurity solver for the analytically solvable noninteracting limit and demonstrate its accuracy for the single-site DMFT problem of a Bethe lattice in the infinite coordination number limit, where the DMFT is exact and the self-consistency condition becomes particularly simple [6]. We compute the Green’s function (GF) to high accuracy, investigate convergence with respect to the parameters that control the accuracy, and showcase that the tensor train methodology is often substantially more precise than quantum Monte Carlo for a given amount of computer time. In addition, we use the tensor train approach to compute the partition function and the impurity free energy. Unlike in CT-QMC, where normalization with respect to low order with quantum Wang–Landau [12, 36] or normalization with respect to a hypervolume [18, 49] are needed, thermodynamic quantities are directly accessible in the tensor train decomposition scheme. We then demonstrate the usefulness of our results at the example of the first-order Mott metal-to-insulator transition.

The paper proceeds as follows. Sec. II introduces the hybridization expansion and gives an overview of the tensor train methodology, describing the decomposition for different observables, as well as computational details. Sec. III illustrates results of the GF for the noninteracting and the DMFT case, and shows results for the free energy. Sec. IV presents conclusions.

## II. METHOD

This section introduces the methodology used in this work. The system and the hybridization expansion are presented in Sec. II A. The details of tensor train representation and its calculation are reviewed in Sec. II B. Details on how these two approaches are combined are presented in Sec. II C.

### A. Hybridization expansion formalism

We study a quantum impurity model described by the Hamiltonian  $H = H_I + H_B + H_{IB} \equiv H_0 + H_{IB}$ , consisting of an interacting impurity  $H_I$ , a noninteracting bath  $H_B$ , and the impurity–bath coupling or hybridization  $H_{IB}$ . For the single-site Anderson impurity model,

$$H_I = \epsilon_0(n_\uparrow + n_\downarrow) + Un_\uparrow n_\downarrow, \quad (1a)$$

$$H_B = \sum_{\sigma k} \epsilon_k c_{k\sigma}^\dagger c_{k\sigma}, \quad (1b)$$

$$H_{IB} = \sum_{k\sigma} (V_{k\sigma} d_\sigma^\dagger c_{k\sigma} + \text{h.c.}), \quad (1c)$$

where  $\epsilon_0$  denotes the on-site energy of the impurity and  $U$  is the Coulomb interaction between two electrons of opposite spin.  $k$  enumerates the (potentially infinite number of) bath states,  $\epsilon_k$  is the dispersion of the noninteracting bath, and  $V_{k\sigma}$  represents the coupling strength between the impurity and bath state  $k$ . The creation and annihilation operators associated with spin-orbital  $\sigma$  of the impurity are given by  $d_\sigma^{(\dagger)}$ , the operators  $c_{k\sigma}^{(\dagger)}$  denote the corresponding bath operators associated with state  $k$ , while  $n_\sigma = d_\sigma^\dagger d_\sigma$ .

The main observables of interest for this work are the partition function, which grants access to thermodynamic properties of the system, and the Green’s function (GF), which is of particular interest for quantum embedding schemes. The partition function is given by

$$Z = \text{Tr}\{e^{-\beta H}\}, \quad (2)$$

where  $\beta$  denotes the inverse temperature and  $\text{Tr}\{\dots\}$  is the trace over the impurity and bath degrees of freedom. We define the imaginary-time GF for electrons of spin  $\sigma$  as

$$G_\sigma(\tau) = -\langle T_\tau d_\sigma(\tau) d_\sigma^\dagger(0) \rangle, \quad (3)$$

where  $\langle \dots \rangle$  denotes the expectation value with respect to the Hamiltonian  $H$ ,  $T_\tau$  is the time-ordering operator, and  $d_\sigma(\tau) = e^{\tau H} d_\sigma e^{-\tau H}$ .

In the hybridization expansion formalism, Eqs. (2) and (3) are expanded in orders of the impurity–bath coupling  $H_{IB}$  [10]. This hybridization expansion is one of the standard techniques underlying Monte Carlo quantum impurity solvers [12] and provides the framework for many approximate and numerically exact methods [10, 14, 15, 18, 20, 50–55].

Expanding Eq. (2) in the impurity–bath coupling yields

$$Z = \sum_{k=0}^{+\infty} \int_0^\beta d\tau_1 \int_{\tau_1}^\beta d\tau_2 \cdots \int_{\tau_{k-1}}^\beta d\tau_k \langle H_{IB}(\tau_k) \cdots H_{IB}(\tau_1) \rangle_{H_0}, \quad (4)$$

with  $H_0 = H_I + H_B$ ,  $H_{IB}(\tau) = e^{H_0 \tau} H_{IB} e^{-H_0 \tau}$ .  $\langle \dots \rangle_{H_0}$  denotes the expectation value with respect to  $H_0$ . Inserting the explicit expression for the impurity–bath coupling

from Eq. (1c) and defining the time-ordered simplex  $S_0^\beta$  as the region of integration with  $0 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq \beta$ , the partition function can be reexpressed as

$$Z = \sum_{k=0}^{\infty} \int_{S_0^\beta} d\tau_1 \cdots d\tau_k \sum_{\sigma_1 \dots \sigma_k} z_{\sigma_1 \dots \sigma_k}^{(k)}(\tau_1, \dots, \tau_k). \quad (5)$$

with

$$z_{\sigma_1 \dots \sigma_k}^{(k)}(\tau_1, \dots, \tau_k) = \sum_{\phi_1 \dots \phi_k} \det \Delta \langle d_{\sigma_k}^{\phi_k}(\tau_k) \cdots d_{\sigma_1}^{\phi_1}(\tau_1) \rangle_{H_I}. \quad (6)$$

Here,  $\sigma_i \in \{\uparrow, \downarrow\}$  are spin-indices and  $\phi_i \in \{-, +\}$  are used to sum over all combinations of creation and annihilation operators with  $d_{\sigma_i}^- \equiv d_{\sigma_i}$  and  $d_{\sigma_i}^+ \equiv d_{\sigma_i}^\dagger$ . The influence of the bath on the impurity is encoded in the hybridization function  $\Delta(\tau)$ . The hybridization function can either be calculated explicitly for a given bath model—which is done for studies on quantum dots or molecular systems where  $\Delta(\tau) = -\langle T_\tau a_\sigma(\tau) a_\sigma^\dagger(0) \rangle$  with  $a_\sigma = \sum_k V_{k\sigma} c_{k\sigma}$ —or determined by a self-consistency condition. The latter scenario appears in quantum embedding schemes like DMFT. Given a hybridization function  $\Delta(\tau)$  and invoking particle-hole symmetry, the hybridization matrix  $\Delta$  entering Eq. (6) is

$$\Delta_{ij} = \begin{cases} \Delta(\tau_i - \tau_j) & \text{if } \sigma_i = \sigma_j \text{ and } \phi_i \neq \phi_j, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The hybridization expansion expressions for the GF can be obtained in a similar fashion [10, 12], and assume the form

$$G_\sigma(\tau) = \sum_{k=0}^{\infty} \int_{S_0^\beta} d\tau_1 \cdots d\tau_k \sum_{\sigma_1 \dots \sigma_k} g_{\sigma\sigma_1 \dots \sigma_k}^{(k)}(\tau, \tau_1, \tau_2, \dots, \tau_k). \quad (8)$$

Here,

$$g_{\sigma\sigma_1 \dots \sigma_k}^{(k)}(\tau, \tau_1, \dots, \tau_k) = -\frac{1}{Z} \sum_{\phi_1 \dots \phi_k} \det \Delta \times \langle d_{\sigma_k}^{\phi_k}(\tau_k) \cdots d_{\sigma_m}^{\phi_m}(\tau_m) d_\sigma(\tau) d_{\sigma_n}^{\phi_n}(\tau_n) \cdots d_{\sigma_1}^{\phi_1}(\tau_1) d_\sigma^\dagger \rangle_{H_I} \quad (9)$$

and  $\tau_m \geq \tau \geq \tau_n$ . While the main difference between the integrands  $z_{\sigma_1 \dots \sigma_k}^{(k)}$  and  $g_{\sigma\sigma_1 \dots \sigma_k}^{(k)}$  are the creation and annihilation operators at 0 and  $\tau$ , the additional operators further restrict the combinations of spin and creation and annihilation operators that give a nonzero contribution.

Eq. (6) and (9) contain sums over all possible creation and annihilation operator combinations  $\phi_i$ , which leads to an exponential number of possible combinations of operators. In the special case of density-density interactions, there is only a single non-zero contribution with alternating creation and annihilation operators for time-ordered arguments. This simplification in the density-density case is analogous to the simplification to the ‘segment’ picture in CT-HYB [10, 12].

The hybridization expansion presented here represents a *bare* expansion scheme. A variety of related partial summation schemes have also been successful [18, 20, 53, 54, 56, 57].

Eqs. (5) and (8) describe an infinite series of terms in a series expansion, where a contribution at order  $k$  consists of a  $k$ -dimensional integral. For finite systems at finite temperature, this series is convergent [9, 10]. However, since the largest contributions to the series typically comes from orders near  $\beta \langle H_{IB} \rangle$  [15], contributions at increasingly high orders are expected when the temperature is lowered. Traditionally, the expressions in Eqs. (5) and (8) are evaluated by Monte Carlo techniques, whereby the integrands  $z_{\sigma_1 \dots \sigma_k}^{(k)}$  and  $g_{\sigma\sigma_1 \dots \sigma_k}^{(k)}$  can be interpreted in terms of Feynman diagrams, which are then combined in a determinant and summed over in a statistical manner [10].

## B. Tensor Train Decomposition and TCI

In order to calculate observables such as the partition function or the GF within the hybridization expansion, the high-dimensional integrals in Eqs. (5) and (8) need to be evaluated. Traditionally, these integrals are calculated using Monte Carlo techniques [9, 10], which converge as  $\sim 1/\sqrt{n_s}$  for a given number of stochastic samples  $n_s$ . The tensor train methodology offers an alternative approach that has potential to converge faster than Monte Carlo. In the following, we summarize the principles underlying the tensor train representation and refer the reader to the applied mathematics literature for mathematical proofs [43, 45, 46]. The implementation of the tensor-fitting component of this work follows the paper of Núñez-Fernández *et al.* [42].

To motivate the construction of a tensor train representation for a given integrand, we first consider the general task of integrating a high-dimensional function  $f(\tau_1, \tau_2, \dots, \tau_k)$  over all its coordinates,

$$I = \int_{\tau \in [0, \beta]^k} d\tau_1 d\tau_2 \cdots d\tau_k f(\tau_1, \tau_2, \dots, \tau_k), \quad (10)$$

which is similar (though not yet equivalent) to evaluating Eqs. (5) and (8). If the integration variables are separable and independent, i.e.  $f(\tau_1, \tau_2, \dots, \tau_k) = f_1(\tau_1) \cdots f_k(\tau_k)$ , the integral can be reexpressed as  $k$  one-dimensional integrals,

$$I = \left[ \int_0^\beta d\tau_1 f_1(\tau_1) \right] \cdots \left[ \int_0^\beta d\tau_k f_k(\tau_k) \right], \quad (11)$$

which can be evaluated independently with standard quadrature rules. Assuming that the arguments  $\tau_i$ , with  $1 \leq i \leq k$ , are each represented on a quadrature grid with  $n_\tau$  points, the complexity of evaluating Eq. (10) is reduced from  $n_\tau^k$  to  $kn_\tau$  for Eq. (11) – a substantial improvement, especially for large values of  $k$ .

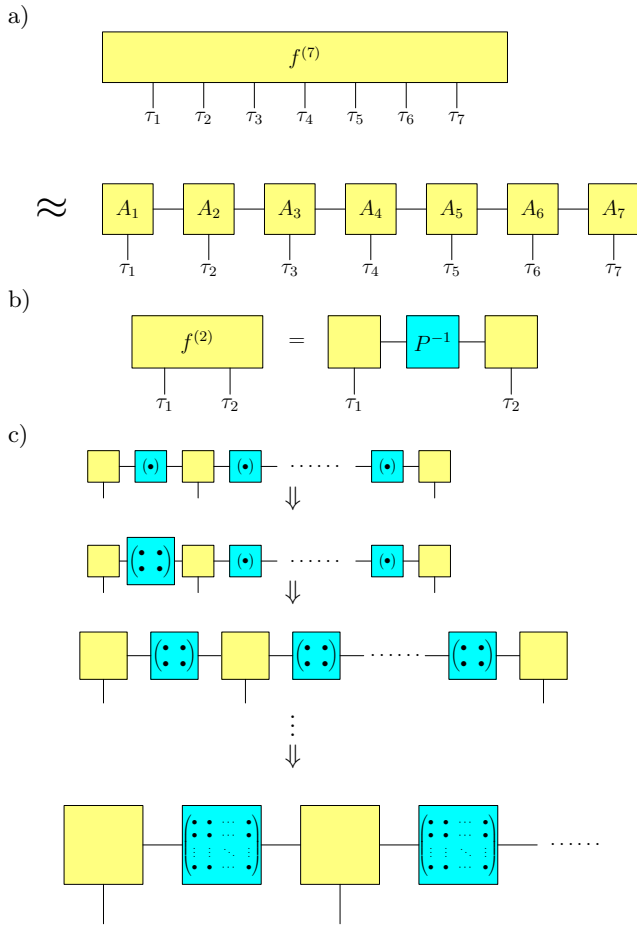


FIG. 1. a: Visualization of the tensor train representation of a function  $f^{(k)}(\tau_1, \tau_2, \dots, \tau_k)$  according to Eq. (12) for  $k = 7$ . Each block labeled  $A_i$  with  $1 \leq i \leq k$  is either a matrix valued function of  $\tau_i$ , or a tensor in case that the  $\tau_i$  are discretized on a grid. Horizontal lines between blocks imply matrix multiplication, vertical lines denote the dependence on  $\tau_i$ . b: Depiction of Eq. (13), where the pivot matrix is used to decompose a two-dimensional function into a tensor train. c: Visualization of the iterative scheme used to approximate a multi-dimensional integrand by a tensor train.

The tensor train representation can be viewed as a generalization of this special case. While an arbitrary function  $f^{(k)}(\tau_1, \tau_2, \dots, \tau_k)$  might not be separable in its arguments, we aim to construct an approximation

$$f^{(k)}(\tau_1, \tau_2, \dots, \tau_k) \simeq A_1(\tau_1)A_2(\tau_2) \cdots A_k(\tau_k), \quad (12)$$

where the  $A_i(\tau_i)$  are matrices of dimension  $r_{i-1} \times r_i$  for  $1 < i < k$  with  $r_1 = r_k = 1$  and matrix multiplication between the  $A_i(\tau_i)$  is implicit in this notation. This representation is schematically visualized in Fig. 1a. We note that selecting  $r_1 = r_k = 1$  is just one among several options available, and alternative approaches such as the tensor ring structure have demonstrated their advantages in other applications [48]. As the decomposition in Eq. (12) has the same temporal structure as Eq. (11), it allows for the same simplification when evaluating the

integral, the only difference being that the components are matrix valued functions of  $\tau_i$ . In the discretized case where the  $\tau_i$  are on a grid,  $A_i(\tau_i)$  can be interpreted as a tensor rather than a matrix valued function, for which the above statements also hold. We refer to a representation of the form in Eq. (12) as a tensor train representation.

A tensor train representation (or approximation) is said to be of rank  $r$  if the matrices (or tensors in the discretized case) in Eq. (12) are of dimension  $r \times r$  (or  $r \times n_\tau \times r$ ). The construction of low-rank tensor representations has been studied extensively [44–46, 58, 59]. Here, we use an extension by Dolgov and Savostyanov [43] of the algorithmic ideas of Oseledets and Tyrtyshnikov [46], which are based on the so-called cross interpolation [60], and which is applicable to matrices and tensors. The basic idea is that a set of points  $(\tau_1, \tau_2, \dots, \tau_k)$ , which are called pivots, defines an approximation for the function  $f^{(k)}$  at all possible values of the times. The approximation requires evaluating the function at the pivots themselves; and at all possible values of a certain coordinate, with other coordinates held constant. Essentially, evaluation is performed on sets of 1D lines in the high-dimensional hypercube of all coordinates, which cross through pivots.

To gain some intuition, it is useful to first consider the special case of matrices, i.e.  $k = 2$ . Let us assume that one knows the values in a matrix only at a certain subset of its rows and columns, defined by a set of pivot coordinates where they cross. It is possible to obtain an interpolation scheme based on this partial information. For a function  $f^{(2)}(\tau_1, \tau_2)$ , evaluating  $f^{(2)}$  at the set of pivots  $\tau_{1j}$  and  $\tau_{2j}$  with  $1 \leq j \leq r$  with the total number of pivots  $r$ , one can construct the pivot matrix  $P$  whose entries are given by  $P_{jj'} = f^{(2)}(\tau_{1j}, \tau_{2j'})$ . Using the inverse of the pivot matrix, one obtains the cross interpolation of the original function,

$$f^{(2)}(\tau_1, \tau_2) \simeq \sum_{j,j'=1}^r f^{(2)}(\tau_1, \tau_{2j}) [P^{-1}]_{jj'} f^{(2)}(\tau_{1j'}, \tau_2), \quad (13)$$

which can be visualized in tensor network form as in Fig. 1b. Eq. (13) represents an interpolation of  $f^{(2)}(\tau_1, \tau_2)$  in the sense that it is exact if  $\tau_1$  (or  $\tau_2$ ) belong to the set  $\tau_{1j}$  (or  $\tau_{2j}$ ). Moreover, if  $f^{(2)}$  is of rank  $r$ , i.e. it can be expressed as  $f^{(2)}(\tau_1, \tau_2) = \sum_{j,j'=1}^r f_{1j}(\tau_1) f_{2j'}(\tau_2)$  then Eq. (13) becomes exact when one uses  $r$  pivots provided that  $P$  remains invertible.

Generally, a given approximation of this type can be systematically improved by sequentially introducing more pivots into it. However, this rapidly becomes costly, and not all new pivots provide the same amount of information. While an optimal procedure remains unknown, there are well-established heuristic algorithms for systematically finding and incorporating pivots into the approximation in such a way that convergence occurs rapidly [61, 62]. In this work we have used the ‘maximum volume’ principle, in particular the rook pivoting procedure



described by Dolgov and Savostyanov [43]. In this procedure, we already have  $r$  pivot points  $\tau_{1j}$  and  $\tau_{2j}$  with  $1 \leq j \leq r$ , and wish to find a pair of new points  $\tau_{1,r+1}$  and  $\tau_{2,r+1}$  to add to these sets, as shown in Fig. 1c. We select this new pair by searching for a point that maximizes the approximation error. Because the TCI is an interpolation, the error at this pivot decreases to zero once it is included in the tensor train. The search is done by the so-called row-column alternating algorithm [43, 60], which starts with a random point  $\tau_1$ , and evaluates the interpolation error for each possible value of  $\tau_2$ . We then fix the value of  $\tau_2$  to be the value that maximizes the error, and evaluate the interpolation error for all possible values of  $\tau_1$  and select the point that maximizes the error. This iterative procedure continues until a fixed point is found, corresponding to a dominant pivot in its own row and columns. Compared to checking errors for each element of the matrix, this heuristic search has complexity  $\mathcal{O}(n_\tau)$ , instead of  $\mathcal{O}(n_\tau^2)$ , which substantially improves the efficiency for large number of  $\tau$ -points. We have found that this procedure manages to construct non-singular pivot matrices that represent the integrand with a reasonable number of function calls.

After establishing the TCI for a function of two variables, we outline the extension to functions with multiple discrete variables  $f^{(k)}(\tau_1, \tau_2, \dots, \tau_k)$ . The approximation is initialized by considering a single pivot, so that we have pivot matrices of size one (see top row of Fig. 1c). We then perform a search along the first two dimensions,  $\tau_1$  and  $\tau_2$ , for the next pivot to be added; all other coordinates are held constant at the value of the original pivot. The chosen pivot is used to enlarge the leftmost pivot matrix to  $2 \times 2$  (2nd row of Fig. 1c). In a manner reminiscent of the density matrix renormalization group algorithm, we subsequently sweep to the right, repeating this procedure by starting from existing pivots and modifying coordinates locally, until all pivot matrices are of rank 2 (3rd row of Fig. 1c). This is followed by another sweep to the left, resulting in rank 3 pivot matrices. This procedure repeats until the pivot matrices have reached an initially specified maximum rank (Fig. 1c, bottom column) [45, 46].

For multi-dimensional continuous variables  $\{\tau_1, \dots, \tau_k\}$ , we choose a set of collocation points for each  $\tau_i$ , such that the multivariate function is discretized into a multi-dimensional tensor and the discrete algorithm above can be applied.

### C. Applying the tensor train approximation to expressions from the hybridization expansion

The efficiency of a tensor train approximation for a given integration task depends on whether an accurate approximation for an integrand can be found for low rank  $r$ . In the following, we describe some of the technical aspects of our implementation of the tensor decomposition scheme to the hybridization expansion. We outline the

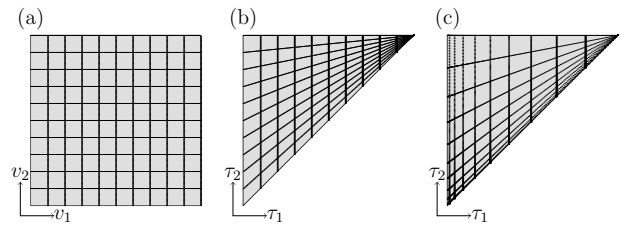


FIG. 2. Illustration of the mapping between the hypercube (panel a) and the simplex in two dimensions, using  $\chi(x) = x$  (panel b) and  $\chi(x) = x^2$  (panel c). Each intersection of lines in panel a is mapped into an intersection in panel b and c.

main aspects for the partition function, and then discuss the specific aspects that are needed to apply the method to the GF.

#### 1. Mapping the hypercube to the simplex

The expressions from the hybridization expansion framework require the integration over the simplex  $S_0^\beta$  see Eqs. (5). This is a consequence of the time-ordering. However, the TCI algorithm is defined on the hypercube. Simply extending the integral beyond the time-ordered region would introduce discontinuities that prevent an accurate low-rank tensor train approximation.

The change of variable proposed in Ref. [42], when applied to imaginary-time problems, extends the hybridization function  $\Delta(\tau)$  over the discontinuities at  $\tau = 0$  and  $\tau = \beta$  and is therefore also not suitable.

Instead, we use a change of variable in this work that maps the original simplex  $S_0^\beta$  to the hypercube  $[0, 1]^k$ .

Let  $\tau_i$  denote the variables within the simplex and  $v_i$  the corresponding variables in the hypercube with  $1 \leq i \leq k$ . The mapping  $h$  between the hypercube and the simplex used in this work is

$$\tau_1 = h(v_1) = \chi(v_1) \cdot \beta, \quad (14a)$$

$$\tau_j = h(v_j) = \tau_{j-1} + \chi(v_j) \cdot (\beta - \tau_{j-1}), \quad (14b)$$

with  $2 \leq j \leq k$  and  $\chi$  any differentiable monotonous function that maps the interval  $[0, 1]$  onto itself. A simple choice of  $\chi$  is the identity  $\chi(x) = x$  but a different choice of  $\chi$ , such as  $\chi(x) = x^2$ , may facilitate the TCI.

The two different mappings are visualized in Fig. 2. Panel a shows a uniform partitioning of the hypercube, which is mapped to the simplex by  $\chi(x) = x$  in panel b, and by  $\chi(x) = x^2$  in panel c.

#### 2. Change of variables

To proceed, we calculate the Jacobian of the change of variable defined by Eqs. (14a) and (14b). The Jacobian matrix  $J_{ij} = \partial\tau_i/\partial v_j$  is upper triangular so that its determinant is given by the product of its diagonal elements,

$$\det J(v_1, \dots, v_k) = \beta \chi'(v_1) \cdot \prod_{j=2}^k \chi'(v_j) (\beta - h(v_{j-1})). \quad (15)$$

While it is in principle possible to apply TCI separately to the Jacobian and the integrand  $\tilde{z}_{\sigma_1 \dots \sigma_k}^{(k)}(v_1, \dots, v_k)$ , we choose to apply TCI directly on their product, i.e. we apply the tensor train approximation to

$$\tilde{z}_{\sigma_1 \dots \sigma_k}^{(k)}(v_1, \dots, v_k) = \tilde{z}_{\sigma_1 \dots \sigma_k}^{(k)}(v_1, \dots, v_k) |\det J(v_1, \dots, v_k)|. \quad (16)$$

This approach proved to be the most efficient within the scope of this work, since the factorization of Eq. (16) by a tensor train allows for a direct calculation of the integral using one-dimensional quadrature rules.

The function  $\chi$  that enters Eqs. (16) through Eqs. (14a), (14b), and (15) controls two important aspects that influence the quality of low-rank tensor train approximations. First, it influences the spacing of the pivot points that are used within the tensor train decomposition. For example, the identity for  $\chi$  in Eqs. (14a) and (14b) shifts potential pivot points away from 0 and closer to  $\beta$ , which implies a bunching of pivot points close to  $\beta$ . This can be compensated for by a suitable choice of  $\chi$ . Second, as  $\chi$  enters the integrands, it can be used to ‘warp’ the integrand i.e. reshape it into a function that is easier to integrate. The approach of ‘warping’ the integrand was also used for facilitating integration using Quasi-Monte Carlo methods [40], and was used in Ref. [42], where changing to a representation in terms of relative time arguments rather than absolute ones was necessary to construct a tensor train approximation. For the results presented in this work, we use  $\chi(v_i) = v_i^2$ .

To illustrate the above method, we consider a simple example here. Let  $\Delta(t) = t(t-1)$  and consider the integral,

$$I_{\text{ex}} = \int_0^1 d\tau_3 \int_0^{\tau_3} d\tau_2 \int_0^{\tau_2} d\tau_1 e^{-(\tau_3 - \tau_2)} e^{-2(\tau_2 - \tau_1)} e^{-\tau_1} \times \Delta(\tau_3 - \tau_1) \Delta(\tau_2). \quad (17)$$

This integral has a structure similar to the functions in the hybridization expansion. It can be evaluated analytically on the simplex,  $I_{\text{ex}} = \frac{13}{4} - \frac{63}{4} e^{-2} - \frac{91}{30} e^{-1} \approx 2.56832 \times 10^{-3}$ . Using the transformation in Eqs. (14)–(15) with  $\beta = 1$ , the integral can be calculated with the tensor-train algorithm described in the previous section [63], which would produce the result given above with an error of less than  $10^{-5}$  within 10 iterations.

### 3. Summation over spin indices:

Eq. (5) features a sum over spin indices  $\sigma_i$ . The number of spin combinations that need to be considered grows

exponentially with the hybridization order  $k$ , which can render the calculation of higher orders in the hybridization expansion unfeasible. We have considered two different approaches for incorporating this aspect into the tensor train methodology.

The first approach is to use the sum over all spin combinations for the tensor train decomposition explicitly, that is applying the TCI algorithm on  $\sum_{\sigma_1 \dots \sigma_k} \tilde{z}_{\sigma_1 \dots \sigma_k}^{(k)}(v_1, \dots, v_k)$ . As this approach effectively averages over all spin combinations, which smooths to some extent the function that is approximated, we found that this method produces good approximations for relatively low tensor ranks. However, as the evaluation of the function that is approximated by the tensor train requires to explicitly perform the sum over all spin combinations which grows exponentially with the hybridization order, this approach becomes prohibitively expensive for high hybridization orders. As such, we deem this approach only feasible for high temperature and for systems that converge within hybridization orders of  $k \lesssim 15$ .

The second approach is to use the tensor train approximation not only for the arguments  $v_i$ , but also for the spin arguments  $\sigma_i$ . The sum over the spin indices is then performed along with the integration over the simplex. On a technical level, we achieve this by introducing a surrogate variable  $w_i \in [-1, 1]$  with  $v_i = |w_i|$  and  $\sigma_i = \uparrow$  if  $w_i \geq 0$  and  $v_i = |w_i|$  and  $\sigma_i = \downarrow$  if  $w_i < 0$ . Approximating  $\tilde{z}^{(k)}(w_1, \dots, w_k)$  by a tensor train and integrating over the variables  $w_i$  corresponds to summing over all spin indices and integrating over the original simplex. This approach overcomes the necessity to explicitly account for an exponential number of spin combinations and is therefore suitable for high hybridization orders. However, as the function that is being approximated in this case contains more information, the tensor rank that is required to obtain a certain accuracy increases as compared to the previous approach.

### 4. Decomposition schemes for the GF

The  $\tau$ -dependence of the GF is the result of the operator  $d_\sigma$  placed at time  $\tau$ , which is the main difference between the partition function and the GF. When performing the  $\tau_i$ -integrals in Eq. (8), this implies that the integrand  $g_{\sigma \sigma_1 \dots \sigma_k}^{(k)}(\tau, \tau_1, \dots, \tau_k)$  needs to consider configurations that have a variable number of up to  $k$  creation and annihilation operators to the left or to the right of  $\tau$ . When creation or annihilation operators move across  $\tau$ , where the operator  $d_\sigma(\tau)$  is located, the integrand drops to zero and a discontinuity occurs. Tensor train approximations converge slowly in the presence of such discontinuities. We therefore rewrite the integration over the

simplex in Eq. (8) as

$$G_\sigma(\tau) = \sum_{k=0}^{\infty} \sum_{l=0}^k \int_{S_0^\tau} d\tau_1 \dots d\tau_l \int_{S_\tau^\beta} d\tau_{l+1} \dots d\tau_k \sum_{\sigma_1 \dots \sigma_k} g_{\sigma\sigma_1 \dots \sigma_k}^{(k)}(\tau, \tau_1, \tau_2, \dots, \tau_k), \quad (18)$$

which fixes the number of creation and annihilation operators to the right and to the left of  $\tau$  to  $l$  and  $k-l$ , respectively, thus circumventing the emergence of discontinuities. Here,  $S_0^\tau$  and  $S_\tau^\beta$  are the time-ordered simplices between 0 and  $\tau$  and between  $\tau$  and  $\beta$ , respectively, with  $0 \leq \tau_1 \leq \dots \leq \tau_l \leq \tau$  and  $\tau \leq \tau_{l+1} \leq \dots \leq \tau_k \leq \beta$ .

To represent the  $\tau$ -dependence of the GF, we have explored two possible approaches. In the first one, we calculate the GF at every value of  $\tau$ . This approach scales linearly with the number  $\tau$ -points.

In the second one, we employ the tensor train approach to also interpolate the  $\tau$  dependence. On a technical level, this is done by adding  $\tau$  as a parameter to the function that is approximated by a tensor train. Generally, this requires a higher rank tensor approximation for comparable accuracy, whereby the actual increase in numerical effort depends on the details of the problem and the representation used for  $\tau$ . We chose the second method for the calculations reported below, where we find that we need about twice the rank to obtain results of comparable accuracy.

### 5. Numerical stability and current limitations

While the tensor train methodology is generally a powerful tool to obtain highly accurate results, we encountered cases where our implementation of the approach became unstable.

Unsurprisingly, we observed that the tensor train method fails to provide accurate results whenever it is applied to functions that are discontinuous. The issue can be avoided by ensuring that this case is not encountered for the observable of interest.

Second, we observed that the tensor train approach described in this work may become unstable for high hybridization orders,  $k \gtrsim 30$ , while at the same time assessing convergence for these high orders becomes challenging. These issues are likely caused by the selection of pivots. As only a fraction of the vast parameter space can be probed for high dimensional functions, pivots might be chosen in such a way that they do not provide a good representation of the function that is being interpolated. Moreover, the CI scheme relies on the inversion of the pivot matrix. At high orders, where the function approximated is essentially zero for extended regimes of the parameter space, picking near-singular pivots for the CI scheme may result in an imprecise tensor train approximations. Further investigation of these numerical aspects will likely resolve the issue.

## III. RESULTS

In this section, we present results for the single-impurity Anderson impurity model as described by Eqs. (1a)–(1c). The influence of the bath on the impurity is encoded in the hybridization function  $\Delta(\tau)$ .

We benchmark our method for the exactly solvable case of a noninteracting impurity in Sec. III A, which we use to assess the accuracy that can be obtained by the present method. In Sec. III B, we showcase the performance of the methodology for the paradigmatic metal-to-insulator transition observed in the infinite dimensional Bethe lattice. In particular, we show that the method not only provides accurate and noise-free results for the GF, but grants direct access to thermodynamic properties. We illustrate this with the metal-to-insulator transition.

### A. Noninteracting limit

We showcase the performance our method for the case of a noninteracting Anderson impurity model,  $U = \epsilon_0 = 0$ , which is also known as the resonant level model. The system is analytically solvable, see e.g. [64, 65], which allows us to benchmark the precision of the results obtained from a tensor train approximations of different ranks. Nevertheless, it is a challenging benchmark for hybridization expansions approaches as it performs an expansion around the ‘atomic’ limit of an isolated impurity [10].

We consider an impurity coupled to a bath whose dispersion  $\epsilon_k$  has a semi-elliptical form with bandwidth  $4t$ ; the associated density of states is  $D(\omega) = \frac{1}{\pi t^2} \sqrt{4t^2 - \omega^2}$  for  $-2t \leq \omega \leq 2t$  and the hybridization function is given by  $\Delta(\tau) = -\int d\omega D(\omega) \frac{e^{-\tau\omega}}{1+e^{-\beta\omega}}$ . This system presents a paradigmatic case studied in single-site DMFT, as it corresponds to an impurity embedded in an infinite-dimensional Bethe lattice [6].

Fig. 3a shows the convergence of the GF of the tensor train formalism described in this work to the exact result as a function of expansion order at inverse temperature  $\beta = 10/t$ . ‘Exact’ denotes the analytically known result. As is evident, the exact result is recovered (within the accuracy of this plot) as the expansion is increased beyond an hybridization expansion order  $k \sim 10$ .

In order to further assess the accuracy, we consider the deviation of the GF at  $\beta/2$ , *i.e.* in the middle of the interval. The left panel of Fig. 3b shows this deviation, as a function of expansion order, for different tensor decomposition ranks. We see that a maximum precision of  $10^{-4}$  can be reached for a decomposition rank of 30. Adding contributions at higher order does not make the result more precise, indicating that it is the tensor rank, rather than the truncation of the expansion at a given order, that limits this precision. This is corroborated by the curves for rank 60, rank 120 and rank 240, which systematically increase the precision of the GF to an accuracy of  $10^{-7}$ . Higher accuracy is reached by a combination of in-



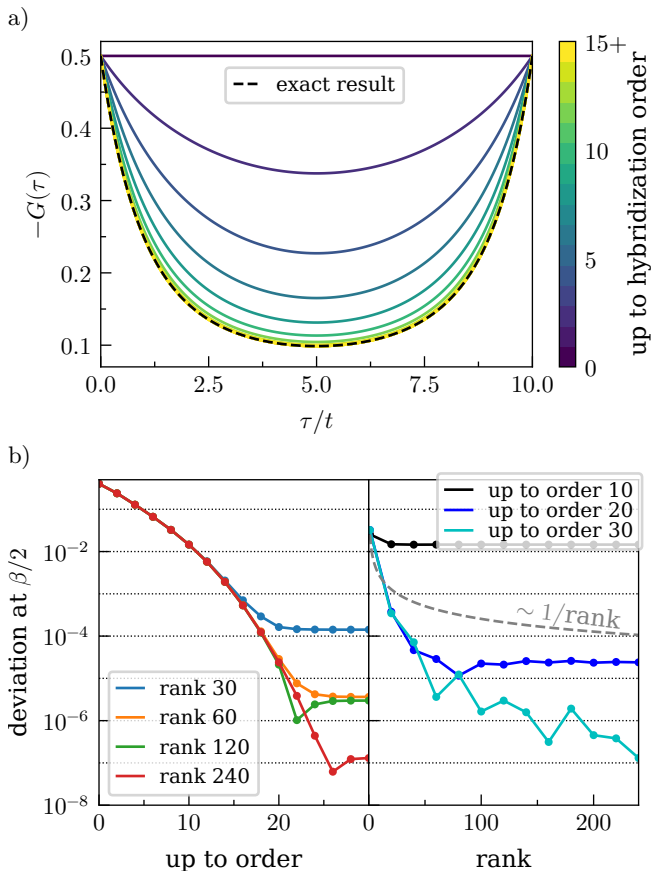


FIG. 3. Noninteracting impurity coupled to a semi-elliptic bath with bandwidth of  $4t$  at temperature  $\beta = 10/t$ . a: Contribution to GF from terms up to hybridization orders indicated by the color. b: Deviation of  $G(\beta/2)$  from the analytic result as a function of perturbative orders for a set of representative tensor ranks (left panel), and as a function of tensor rank for a set for representative perturbative orders (right panel).

creasing diagram order and increasing tensor rank; higher order hybridization contributions require higher ranks to be accurately approximated by a decomposition.

The right panel of Fig. 3b illustrates the same behavior as a function of tensor rank, evaluating contributions at up to hybridization order 10, 20, and 30. The grey dashed line indicates a convergence  $\sim 1/\text{rank}$  with respect to the tensor rank. In particular, it shows that the method converges faster than  $1/\text{rank}$  for higher orders. It is evident that while contributions up to order 10 are well described by an approximation tensor trains of rank less than 50, higher order contributions require substantially higher tensor ranks.

In practice, these results suggest a scheme where tensor train approximations for a fixed hybridization are performed for gradually increasing tensor ranks, until the integral values no longer change as a function of tensor rank. Note that it is difficult with existing CT-QMC techniques to reach a relative accuracy beyond  $10^{-5}$ ; the

tensor train methodology is therefore promising for obtaining high-precision data that could be used, for example, in analytic continuation [66].

## B. Dynamical Mean-Field Theory and Free Energy

In the following, we present results for interacting impurity models. The purpose of this section is two fold: First, we demonstrate the performance of our method for interacting system and its capability to generate results that are compatible with findings that are obtained using CT-QMC. Second, we show that our method grants direct access to thermodynamic observables, which are not straightforwardly available in standard CT-QMC methods.

For scope of this section, we study the paradigmatic example of the metal-to-insulator transition in an infinite dimensional Bethe lattice as described within DMFT [5, 67–72]. In its single-site formulation, DMFT approximates the momentum-dependent self-energy of an extended lattice problem by a local self-energy, and then provides a solution for the auxiliary impurity problem with a dynamically adjusted, self-consistently determined bath [5, 6]. For the infinite coordination-number Bethe-lattice, the self-energy is local and the methodology becomes exact [5, 6, 73, 74]. Below a critical temperature of  $\beta \sim 15/t$ , the paramagnetic version of the model is known to have a first-order Mott metal-to-insulator transition between a metallic state at weak interaction and an insulating state at large interaction, with an extended coexistence regime [5, 67–72].

We first study results for the GF obtained at different electron-electron interaction strength  $U$  at temperature  $\beta = 20/t$ . The GF for representative values of  $U$  is depicted in Fig. 4, as calculated by both standard CT-QMC and the tensor train approach. Both methods agree within their respective errors. For  $U \geq 5t$ , the spectral weight  $A(\omega = 0) \sim \beta G(\beta/2)$  becomes strongly suppressed, indicating the opening of the Mott gap and the qualitative difference between metallic and insulating solutions. We emphasize that both methods sample the same diagrammatic perturbation expansion [10], either by performing a stochastic random walk in diagram space or by calculating a tensor train approximation to the integrand at different orders.

Fig. 4 demonstrates that the tensor train approach can provide results that are compatible with findings obtained within CT-QMC schemes, and therefore establishes tensor train based schemes as an alternative to Monte Carlo based impurity solvers. As is evident from the data, the tensor train method does not suffer from stochastic noise. While it can generally be much more precise as compared to CT-QMC methods at similar computational cost, there are numerical aspects that influence the precision of the tensor train result. In particular, when decomposing the integrands for high hybridization orders where the parameter space is vast, the

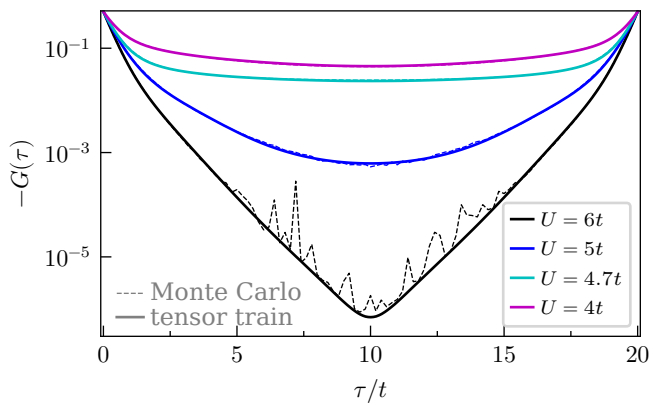


FIG. 4. DMFT GF for representative interaction strengths  $U$  at inverse temperature  $\beta = 20/t$ . The dashed lines represent results obtained by hybridization expansion CT-QMC [10, 26, 75] (measured in imaginary time as described in [10]). The full lines show results calculated within the tensor train approach. The tensor rank necessary for obtaining accurate results depends on interaction strength and ranges from 90 in the insulating regime to 200 in the metallic regime. Both results agree within their respective errors; differences between the two methods are visible at large  $U$  due to the logarithmic scale.

found that the quality of the tensor train approximation can become sensitive to the details of the pivots that are chosen. This is especially the cases when the integrand is essentially zero, or when only a small part of the parameter space contributes to the integral (see Sec. II C for more details). These cases require a careful analysis of the results obtained by the tensor train method, which in practice limits the feasibility of our current implementation of higher precision results for high hybridization orders. The situation is not unlike the one with ergodicity issues in Monte Carlo, where the choice of an initial state or of a few early moves may prevent the simulation from exploring the entirety of phase space. Further investigations – in particular with respect to how pivots are picked, how the inversion of the pivot matrix is performed, and what variables are used to represent the integrand at a specific hybridization order – are left for future work.

We now focus on the thermodynamic properties of the system at the metal-to-insulator transition. The first-order phase transition and the coexistence regime between metallic and insulating solutions in the single-site DMFT has been investigated in great detail [70–72, 76–79].

Thermodynamic quantities are directly accessible from the tensor train formalism, since the partition function can be obtained using Eq. (5). This is in contrast to CT-QMC, where diagrams are sampled with the probability that they contribute to the partition function, but an overall partition function normalization factor is typically not accessible (see Ref. [36] and Ref. [12] Sec. X.E on quantum Wang–Landau algorithms for sampling this

normalization in ‘bare’ expansions and Ref. [18] for normalizing to the hypervolume of the time integral in renormalized/inchworm perturbation theory), and in contrast to the Hirsch–Fye algorithm [34], where thermodynamic integration was used to delineate the phase boundary [72].

Fig. 5 shows the impurity free energy  $F_{\text{imp}} = -\log(Z)/\beta$ , as a function of interaction strength  $U$  for three representative temperatures, calculated directly from the partition function Eq. (5) using the tensor train methodology. The impurity free energy is closely related to the lattice free energy  $F_{\text{lattice}} = F_{\text{imp}} + \sum_{\sigma} \int_0^{\beta} \Delta^2(\tau)/t^2 d\tau$  [80, 81]. We perform the underlying DMFT calculation starting from two different reference systems: (i) the metallic system at small  $U$  where we successively increase the interaction strength and (ii) the insulating system at large  $U$  where we successively decrease the interaction strength. Below the critical temperature of  $\beta \sim 15/t$ , we find two coexisting solutions with differing free energy. This implies a coexistence regime where both metallic and insulating solutions can be stabilized, the extent of which increases with decreasing temperature.

#### IV. CONCLUSION

In conclusion, we presented a method for solving strongly correlated equilibrium quantum impurity problems by expressing the terms in a diagrammatic series expansion by an approximate tensor train form, so that integration over internal degrees of freedom becomes tractable. We tested the method on a typical problem in the field: the single-site Anderson impurity model, as it appears in the context of dynamical mean-field theory. Since the method is based on the hybridization expansion underlying commonly used CT-QMC algorithms, much of the knowledge and experience is directly transferable. We showed that CT-QMC and tensor train methods lead to consistent results. However, the tensor train results were more precise than CT-QMC results for the problems studied here, and do not suffer from any meaningful level of stochastic noise. We were able to converge the tensor train approximation in all cases shown here to a level of accuracy that is very costly to achieve in CT-QMC. Moreover, we showed that in contrast to CT-QMC approaches, the tensor train methodology allows for direct access to the partition function and thermodynamic properties. While we find that the current methodology does have limitations (such as, in certain cases, the selection of near-singular pivots for the tensor cross interpolation that may lead to imprecise tensor train approximations), we believe that further research into the numerics of tensor train approximations will overcome these issues.

Our work has demonstrated the significant potential of tensor train methods as solvers for equilibrium quantum impurity problems. We are confident that further

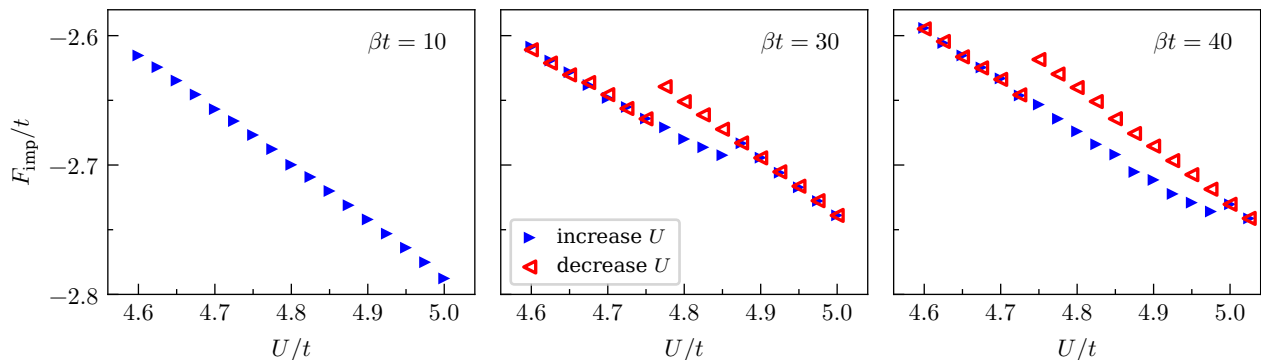


FIG. 5. Impurity free energy  $F_{\text{imp}} = -\log(Z)/\beta$ , which includes the contribution from both spins, as a function of interaction strength  $U$ . The three panels correspond to three representative temperatures. The metallic phase is found for small  $U$ , while large values of  $U$  correspond to the insulating phase. The two branches represent the DMFT results starting from the system in the metallic phase and successively increasing the interaction strength (indicated by the markers  $\blacktriangleright$ ), and DMFT results starting from the system in the insulating phase and successively increasing the interaction strength (indicated by the markers  $\blacktriangleleft$ ), respectively. For intermediate values of  $U \sim 4.8t$ , the two branches separate, indicating the existence of two co-existing phases. Upon comparing results for different fixed ranks, we estimate our relative error to be below  $2 \cdot 10^{-3}$ .

advancements and applications will expand the parameter space of impurity problems that can be reliably solved. One notable advantage of tensor train methods is their ability to address regimes that are traditionally difficult for QMC-based methods. These regimes include problems affected by the ‘sign problem’ that often restricts the applicability of QMC methods. Examples of such problems encompass systems away from equilibrium [18, 52, 82–84], lattice models away from half filling, multi-orbital impurity systems with frustrations, complex interactions, and general off-diagonal hybridizations [14, 15, 20, 21]. We anticipate that tensor train methods will prove to be valuable tools for investigating such multi-orbital systems. Nevertheless, since we employed a hybridization expansion [14], the overall computational cost scales exponentially with the number of impurity orbitals.

Tensor train methods also hold promise when applied to partial summation techniques, including both low-order schemes like the non-crossing approximation and its extensions [85–91], as well as numerically exact ‘bold’ methodologies [53–56, 92, 93] including the ‘inchworm’ scheme [18, 20, 94, 95], interaction expansion series [9, 11] such as equilibrium and bold-line nonequilibrium methods, as well as other types of ‘diagrammatic’ and ‘continuous-time’ formalisms that are traditionally treated evaluated using Monte Carlo methods [12, 57, 96, 97]. These formulations often rely on a low-order expressions or significantly reduce the hybridization orders that need to be considered. If our finding that the low-order contributions in the hybridization expansion can be effectively represented by tensors of order ten translates to partial summation schemes, it would imply that tensor train methods could yield results within minutes to hours on a standard desktop computer. The

potential to significantly accelerate these calculations ultimately enhances their accessibility and applicability in various research contexts.

## ACKNOWLEDGEMENTS

A.E. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 45364484. E.G. and W.-T. L. were supported by the Department of Energy via DE-SC-0022088. T.B. was funded by the Department of Energy via DE-SC0020347. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of Basic Energy Sciences, Scientific Discovery through Advanced Computing (SciDAC) program under Award Number DE-SC0022088. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 using NERSC award BES-ERCAP0021805. G.C. acknowledges support by the Israel Science Foundation (Grants No. 2902/21 and 218/19) and by the PAZY foundation (Grant No. 318/78). X.W. acknowledges support from the Plan France 2030 ANR-22-PETQ-0007 “EPIQ”. The Flatiron Institute is a division of the Simons Foundation.

## AUTHOR CONTRIBUTIONS

A.E. and W.-T. L. contributed equally to this paper.

- 
- [1] P. W. Anderson, *Phys. Rev.* **124**, 41 (1961).
- [2] R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, and L. M. K. Vandersypen, *Rev. Mod. Phys.* **79**, 1217 (2007).
- [3] R. Brako and D. M. Newns, *J. Phys. C: Solid State Phys.* **14**, 3065 (1981).
- [4] D. C. Langreth and P. Nordlander, *Phys. Rev. B* **43**, 2541 (1991).
- [5] A. Georges and G. Kotliar, *Phys. Rev. B* **45**, 6479 (1992).
- [6] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996).
- [7] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, *Rev. Mod. Phys.* **78**, 865 (2006).
- [8] D. Zgid and E. Gull, *New J. Phys.* **19**, 023047 (2017).
- [9] A. N. Rubtsov, V. V. Savkin, and A. I. Lichtenstein, *Phys. Rev. B* **72**, 035122 (2005).
- [10] P. Werner, A. Comanac, L. de' Medici, M. Troyer, and A. J. Millis, *Phys. Rev. Lett.* **97**, 076405 (2006).
- [11] E. Gull, P. Werner, O. Parcollet, and M. Troyer, *Europhys. Lett.* **82**, 57003 (2008).
- [12] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011).
- [13] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005).
- [14] P. Werner and A. J. Millis, *Phys. Rev. B* **74**, 155107 (2006).
- [15] K. Haule, *Phys. Rev. B* **75**, 155113 (2007).
- [16] L. Mühlbacher and E. Rabani, *Phys. Rev. Lett.* **100**, 176403 (2008).
- [17] E. Gull, D. R. Reichman, and A. J. Millis, *Phys. Rev. B* **82**, 075109 (2010).
- [18] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, *Phys. Rev. Lett.* **115**, 266802 (2015).
- [19] P. Gunacker, M. Wallerberger, E. Gull, A. Hausoel, G. Sangiovanni, and K. Held, *Phys. Rev. B* **92**, 155102 (2015).
- [20] E. Eidelstein, E. Gull, and G. Cohen, *Phys. Rev. Lett.* **124**, 206405 (2020).
- [21] J. Li, Y. Yu, E. Gull, and G. Cohen, *Phys. Rev. B* **105**, 165133 (2022).
- [22] B. Bauer, L. D. Carr, H. G. Evertz, A. Feiguin, J. Freire, S. Fuchs, L. Gamper, J. Gukelberger, E. Gull, S. Guertler, A. Hehn, R. Igarashi, S. V. Isakov, D. Koop, P. N. Ma, P. Mates, H. Matsuo, O. Parcollet, G. Pawłowski, J. D. Picon, L. Pollet, E. Santos, V. W. Scarola, U. Schollwöck, C. Silva, B. Surer, S. Todo, S. Trebst, M. Troyer, M. L. Wall, P. Werner, and S. Wessel, *J. Stat. Mech.* **2011**, P05001 (2011).
- [23] H. Shinaoka, M. Dolfi, M. Troyer, and P. Werner, *J. Stat. Mech.* **2014**, P06012 (2014).
- [24] O. Parcollet, M. Ferrero, T. Ayrál, H. Hafermann, I. Krivenko, L. Messio, and P. Seth, *Comput. Phys. Commun.* **196**, 398 (2015).
- [25] P. Seth, I. Krivenko, M. Ferrero, and O. Parcollet, *Comput. Phys. Commun.* **200**, 274 (2016).
- [26] A. Gaenko, A. Antipov, G. Carcassi, T. Chen, X. Chen, Q. Dong, L. Gamper, J. Gukelberger, R. Igarashi, S. Isakov, M. Könz, J. LeBlanc, R. Levy, P. Ma, J. Páki, H. Shinaoka, S. Todo, M. Troyer, and E. Gull, *Comput. Phys. Commun.* **213**, 235 (2017).
- [27] H. Shinaoka, E. Gull, and P. Werner, *Comput. Phys. Commun.* **215**, 128 (2017).
- [28] C. Yue, Y. Wang, J. Otsuki, and X. Dai, *Comput. Phys. Commun.* **236**, 135 (2019).
- [29] M. Wallerberger, A. Hausoel, P. Gunacker, A. Kowalski, N. Parragh, F. Goth, K. Held, and G. Sangiovanni, *Comput. Phys. Commun.* **235**, 388 (2019).
- [30] H. Shinaoka, Y. Nomura, and E. Gull, *Comput. Phys. Commun.* **252**, 106826 (2020).
- [31] M. Caffarel and W. Krauth, *Phys. Rev. Lett.* **72**, 1545 (1994).
- [32] E. Koch, G. Sangiovanni, and O. Gunnarsson, *Phys. Rev. B* **78**, 115102 (2008).
- [33] C. Mejuto-Zaera, L. Zepeda-Núñez, M. Lindsey, N. Tubman, B. Whaley, and L. Lin, *Phys. Rev. B* **101**, 035143 (2020).
- [34] J. E. Hirsch and R. M. Fye, *Phys. Rev. Lett.* **56**, 2521 (1986).
- [35] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, *Phys. Rev. D* **24**, 2278 (1981).
- [36] M. Troyer, S. Wessel, and F. Alet, *Phys. Rev. Lett.* **90**, 120201 (2003).
- [37] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, *J. Exp. Theor. Phys. Lett.* **64**, 911 (1996).
- [38] N. V. Prokof'ev, B. V. Svistunov, and I. S. Tupitsyn, *J. Exp. Theor. Phys. Sov. Phys.* **87**, 310 (1998).
- [39] N. V. Prokof'ev and B. V. Svistunov, *Phys. Rev. Lett.* **81**, 2514 (1998).
- [40] M. Maček, P. T. Dumitrescu, C. Bertrand, B. Triggs, O. Parcollet, and X. Waintal, *Phys. Rev. Lett.* **125**, 047702 (2020).
- [41] C. Bertrand, D. Bauernfeind, P. T. Dumitrescu, M. Maček, X. Waintal, and O. Parcollet, *Phys. Rev. B* **103**, 155104 (2021).
- [42] Y. Núñez-Fernández, M. Jeannin, P. T. Dumitrescu, T. Kloss, J. Kaye, O. Parcollet, and X. Waintal, *Phys. Rev. X* **12**, 041018 (2022).
- [43] S. Dolgov and D. Savostyanov, *Comput. Phys. Commun.* **246**, 106869 (2020).
- [44] S. A. Goreinov, *Dokl. Math.* **77**, 404 (2008).
- [45] D. V. Savostyanov, *Linear Algebra Its Appl.* **458**, 217 (2014).
- [46] I. Oseledets and E. Tyrtysnikov, *Linear Algebra Its Appl.* **432**, 70 (2010).
- [47] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [48] U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).
- [49] G. Cohen and E. Rabani, *Phys. Rev. B* **84**, 075150 (2011).
- [50] H. Keiter and J. C. Kimball, *Phys. Rev. Lett.* **25**, 672 (1970).
- [51] T. Pruschke and N. Grewe, *Z. Phys. B* **74**, 439 (1989).
- [52] L. Mühlbacher and E. Rabani, *Phys. Rev. Lett.* **100**, 176403 (2008).
- [53] E. Gull, D. R. Reichman, and A. J. Millis, *Phys. Rev. B* **82**, 075109 (2010).
- [54] G. Cohen, D. R. Reichman, A. J. Millis, and E. Gull, *Phys. Rev. B* **89**, 115139 (2014).
- [55] G. Cohen, E. Gull, D. R. Reichman, and A. J. Millis, *Phys. Rev. Lett.* **112**, 146802 (2014).



- [56] N. V. Prokof'ev and B. V. Svistunov, *Phys. Rev. B* **77**, 125101 (2008).
- [57] K. Van Houcke, F. Werner, E. Kozik, N. Prokof'ev, B. Svistunov, M. J. H. Ku, A. T. Sommer, L. W. Cheuk, A. Schirotzek, and M. W. Zwierlein, *Nat. Phys.* **8**, 366 (2012).
- [58] I. V. Oseledets, *Constr. Approx.* **37**, 1 (2013).
- [59] B. N. Khoromskij, *Tensor Numerical Methods in Scientific Computing*, Vol. 19 (Walter de Gruyter GmbH & Co KG, 2018).
- [60] E. Tyrtyshnikov, *Computing* **64**, 367 (2000).
- [61] S. Goreinov, N. Zamarashkin, and E. Tyrtyshnikov, *Mathematical Notes of the Academy of Sciences of the USSR* **62**, 515 (1997).
- [62] S. A. Goreinov, I. V. Oseledets, D. V. Savostyanov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, in *Matrix Methods: Theory, Algorithms and Applications* (World Scientific, Singapore, 2010) p. 247–256.
- [63] For an existing library, the reader can try `xfac` at <https://gitlab.kwant-project.org/ttd/xfac>.
- [64] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics: An Introduction* (Oxford Graduate Texts, Oxford, 2004).
- [65] H. J. W. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer Series in Solid-State Sciences, Berlin, Heidelberg, 2008).
- [66] J. Fei, C.-N. Yeh, and E. Gull, *Phys. Rev. Lett.* **126**, 056402 (2021).
- [67] M. Jarrell, *Phys. Rev. Lett.* **69**, 168 (1992).
- [68] M. J. Rozenberg, X. Y. Zhang, and G. Kotliar, *Phys. Rev. Lett.* **69**, 1236 (1992).
- [69] X. Y. Zhang, M. J. Rozenberg, and G. Kotliar, *Phys. Rev. Lett.* **70**, 1666 (1993).
- [70] M. J. Rozenberg, R. Chitra, and G. Kotliar, *Phys. Rev. Lett.* **83**, 3498 (1999).
- [71] R. Bulla, *Phys. Rev. Lett.* **83**, 136 (1999).
- [72] N. Blümer, *Metal-Insulator Transition and Optical Conductivity in High Dimensions* (Shaker Verlag, Aachen, 2003).
- [73] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989).
- [74] M. Eckstein, M. Kollar, K. Byczuk, and D. Vollhardt, *Phys. Rev. B* **71**, 235119 (2005).
- [75] H. Hafermann, P. Werner, and E. Gull, *Comput. Phys. Commun.* **184**, 1280 (2013).
- [76] J. Schlipf, M. Jarrell, P. G. J. van Dongen, N. Blümer, S. Kehrein, T. Pruschke, and D. Vollhardt, *Phys. Rev. Lett.* **82**, 4890 (1999).
- [77] W. Krauth, *Phys. Rev. B* **62**, 6860 (2000).
- [78] R. Bulla, T. A. Costi, and D. Vollhardt, *Phys. Rev. B* **64**, 045103 (2001).
- [79] J. Joo and V. Oudovenko, *Phys. Rev. B* **64**, 193102 (2001).
- [80] A. Georges, G. Kotliar, and Q. Si, *Int. J. Mod. Phys B* **06**, 705 (1992).
- [81] G. Kotliar, E. Lange, and M. J. Rozenberg, *Phys. Rev. Lett.* **84**, 5180 (2000).
- [82] M. Schiró and M. Fabrizio, *Phys. Rev. B* **79**, 153302 (2009).
- [83] P. Werner, T. Oka, and A. J. Millis, *Phys. Rev. B* **79**, 035320 (2009).
- [84] A. Erpenbeck, E. Gull, and G. Cohen, *Phys. Rev. Lett.* **130**, 186301 (2023).
- [85] N. E. Bickers, *Rev. Mod. Phys.* **59**, 845 (1987).
- [86] N. Grewe and H. Keiter, *Phys. Rev. B* **24**, 4420 (1981).
- [87] F. B. Anders and N. Grewe, *Europhys. Lett.* **26**, 551 (1994).
- [88] F. B. Anders, *Physica B* **206-207**, 177 (1995).
- [89] K. Haule, S. Kirchner, J. Kroha, and P. Wölfle, *Phys. Rev. B* **64**, 155111 (2001).
- [90] M. Eckstein and P. Werner, *Phys. Rev. B* **82**, 115115 (2010).
- [91] A. Erpenbeck, E. Gull, and G. Cohen, *Phys. Rev. B* **103**, 125431 (2021).
- [92] N. Prokof'ev and B. Svistunov, *Phys. Rev. Lett.* **99**, 250201 (2007).
- [93] G. Cohen, E. Gull, D. R. Reichman, A. J. Millis, and E. Rabani, *Phys. Rev. B* **87**, 195108 (2013).
- [94] Z. Cai, J. Lu, and S. Yang, *Numerical analysis for inchworm monte carlo method: Sign problem and error growth* (2020).
- [95] A. J. Kim, J. Li, M. Eckstein, and P. Werner, *Phys. Rev. B* **106**, 085124 (2022).
- [96] R. Rossi, *Phys. Rev. Lett.* **119**, 045701 (2017).
- [97] A. Moutenet, W. Wu, and M. Ferrero, *Phys. Rev. B* **97**, 085117 (2018).