



CHORUS

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

Efficient hybridization fitting for dynamical mean-field theory via semi-definite relaxation

Carlos Mejuto-Zaera, Leonardo Zepeda-Núñez, Michael Lindsey, Norm Tubman, Birgitta Whaley, and Lin Lin

Phys. Rev. B **101**, 035143 — Published 27 January 2020

DOI: [10.1103/PhysRevB.101.035143](https://doi.org/10.1103/PhysRevB.101.035143)

Efficient hybridization fitting for dynamical mean-field theory via semi-definite relaxation

Carlos Mejuto-Zaera,¹ Leonardo Zepeda-Núñez,² Michael Lindsey,²
Norm Tubman,³ Birgitta Whaley,¹ and Lin Lin^{2,4}

¹*Department of Chemistry, University of California, Berkeley, CA 94720*

²*Department of Mathematics, University of California, Berkeley, CA 94720*

³*Quantum Artificial Intelligence Lab. (QuAIL), Exploration Technology Directorate,
NASA Ames Research Center, Moffett Field, CA 94035, USA*

⁴*Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720*

We introduce a nested optimization procedure using semi-definite relaxation for the fitting step in Hamiltonian-based cluster dynamical mean-field theory (DMFT) methodologies. We show that the proposed method is more efficient and flexible than state-of-the-art fitting schemes, which allows us to treat as large a number of bath sites as the impurity solver at hand allows. We characterize its robustness to initial conditions and symmetry constraints, thus providing conclusive evidence that in the presence of a large bath, our semi-definite relaxation approach can find the correct set of bath parameters without needing to include *a priori* knowledge of the properties that are to be described. We believe this method will be of great use for Hamiltonian-based calculations, simplifying and improving one of the key steps in cluster dynamical mean-field theory calculations.

I. INTRODUCTION

Understanding the behavior of materials from first principles has been one of the most important challenges in the physical sciences since the advent of quantum mechanics. This is particularly difficult for so-called strongly correlated systems, for which a mean-field treatment is insufficient. In the pursuit of understanding emergent phenomena in the thermodynamic limit of such systems, embedding methods have played a major role. Some of the most prominent examples of embedding techniques are dynamical mean-field theory (DMFT) [1–3], density-matrix embedding theory (DMET) [4], and self-energy embedding theory (SEET) [5] among many others. These methods have been widely successful in unveiling the low energy properties of model Hamiltonians and real materials alike [6–14].

The embedding methods previously mentioned share the same framework; they substitute the computationally intractable original system with a simpler model system. This model system is usually composed of an interacting impurity, or *cluster*, and a non-interacting *bath*. The cluster commonly represents a subset of the original system, while its interplay with the rest of the original system is encoded by the bath. The objective is to find a model system whose low energy properties coincide with or indicate the properties of the original system of interest. Among these methods, Green’s function based approaches, such as Hamiltonian-based DMFT [15–20], are widely used. A common feature of this kind of embedding methods is the need to optimize, or fit, some function representing a parameterization of the bath.

Although this fitting step is just as important as the rest for the success of the embedding calculation, it has received comparatively little attention in the method development research and existing literature. Particularly in Hamiltonian-based DMFT approaches the fitting step plays a crucial role: it can be understood as the steering wheel that guides the exploration of parameter space that has to culminate in the identification of the appropriate model system. The performance of the fit can influence the final model system (the fixed point of the Hamiltonian parameters) and whether it is representative of the thermodynamic limit of the system under consideration. Unfortunately, this fitting problem is highly non-convex in the sense that the optimization often gets trapped in local minima, which strongly depend on the choice of the initial guess. Thus, typical approaches rely on running the optimization several (perhaps many) times, using different randomly generated initial guesses, and keeping the parameters that achieve the lowest cost [21]. Previous studies trying to alleviate this issue have met with partial success [22].

In this work, we propose a new efficient and empirically robust optimization algorithm that leverages semi-definite relaxation (SDR) for the fitting problem in Hamiltonian-based cluster DMFT approaches. The algorithm relies on convex relaxation [23–25], coupled to a novel nested approach that combines state-of-the-art conic solvers [26] with standard gradient-based optimization algorithms. It provides for a fitting routine that we show to be (a) systematically improvable, (b) more efficient than commonly used optimization methods, (c) empirically robust to initial conditions and symmetry constraints. One of the main advantages of this novel

method is the emergence of symmetries in the bath parameters reflecting the geometry of the cluster, which in previous methods had to be imposed in the optimization [27–29]. These symmetries generalize to large baths and clusters, which we can handle by using a state-of-the-art impurity solver, the adaptive sampling configuration algorithm (ASCI) [20, 30]. As a result, it can handle large bath sizes and asymmetric clusters with an unprecedented accuracy.

The structure of the paper is as follows: In section II we describe in detail the methodologies employed and introduced: section II A contains a detailed description of the DMFT method and the role of fitting within the approach, while in section II B we introduce and describe our SDR approach to bath fitting. Then section III presents the numerical results of this paper: section III A characterizes the properties of the SDR fit, and section III B presents its performance in DMFT calculations explicitly. The converged bath parameters for the results presented in section III B can be found in the Supplemental Material [31].

II. METHODS

A. Dynamical mean-field theory

DMFT and its cluster extensions form a family of numerical methods that treat strongly correlated many-body systems non-perturbatively [1–3]. The approach originates from the study of lattice Hamiltonian systems in infinite dimensions [32–35]. The main idea is to map a fully interacting lattice to a (many-site) Anderson impurity model, composed of an interacting cluster that conserves all original interactions of the system and a non-interacting bath. Doing so, one substitutes an intractable interacting many-body system with a simpler one that can be studied with existing numerical techniques, e.g. quantum Monte Carlo [36] or exact diagonalization approaches (ED) [15]. For this mapping to be useful, one seeks to define the impurity model such that its low energy physics coincides with that of the original system. The precise way in which this matching is defined and achieved is briefly summarized later. First, we outline the key differences between the two main solver types for DMFT: Monte Carlo and Hamiltonian-based methods. The results in this work are relevant for the latter.

The DMFT algorithm differs significantly between Monte Carlo-based solvers and Hamiltonian-based ones. The main advantage of using a Monte Carlo solver is that one can work in the infinite-bath limit, as required for formally exact implementation of DMFT. By contrast, Hamiltonian-based meth-

ods need to truncate the bath, adding an extra approximation to the calculation and raising questions about convergence with respect to the bath size. However, Monte Carlo methods are formally limited to computing quantities along the imaginary time/frequency axis, thus they need to perform an analytical continuation to obtain spectral quantities. Hamiltonian-based methods, meanwhile, can directly provide results on both the imaginary and real axes. Moreover, while Monte Carlo methods can handle the largest interacting clusters, they are limited by the sign problem to relatively high temperatures and low doping rates. In summary, both families contribute to the calculation of phase diagrams.

In this work, we address difficulties found in the finite-bath Hamiltonian approaches, where the bath truncation necessitates an optimization (fit) step. A brief description of the algorithmic structure follows. For a review of the DMFT algorithm using Hamiltonian-based methods, the reader is referred to [22].

In DMFT one first selects a subset of the original degrees of freedom, which are referred to as the cluster. The impurity model, defined below in Eq. (1), includes these degrees of freedom and all their interactions, while neglecting the rest of the original system. To account for the coupling between the cluster and the rest of the lattice, the impurity model is completed with a set of non-interacting degrees of freedom, which we call *baths*, which are coupled to the cluster sites. A possible physical interpretation considers the baths as paths outside the cluster that a particle may traverse in the original system. Thus a possible motion of a particle starting inside the cluster, leaving it, moving through the rest of the lattice, and returning to the cluster may be represented in the impurity model as the particle hopping from a cluster site to a bath site and back.

The Hamiltonian for the impurity model H_{imp} is given by

$$H_{imp} = H_C + \sum_{\ell=1}^{N_b} \epsilon_{\ell} d_{\ell}^{\dagger} d_{\ell} + \sum_{\ell=1}^{N_b} \sum_{\alpha=1}^{N_c} \left(V_{\alpha,\ell} d_{\ell}^{\dagger} c_{\alpha} + \text{h.c.} \right), \quad (1)$$

where N_c is the number of cluster degrees of freedom, N_b the number of baths, H_C is the original Hamiltonian restricted to the cluster degrees of freedom; c_{α} , and d_{ℓ} correspond to the cluster and bath annihilation operators respectively. The bath degrees of freedom are characterized by their single particle energies ϵ_{ℓ} and their couplings $V_{\alpha,\ell}$ to the cluster sites. The goal of DMFT is to determine the bath parameters $\{\epsilon_{\ell}\}_{\ell=1}^{N_b}$ and $\{V_{\alpha,\ell}\}_{\ell=1, \alpha=1}^{N_b, N_c}$ such that the low energy physics of H_{imp} and the original system H coincide. This means that the Green's function

$G_{imp}(\omega)_{\alpha,\beta}$ of H_{imp} should be the same as the corresponding submatrix of the Green's function of the global Hamiltonian H . This is achieved in DMFT by a self-consistent loop. To this effect, Hamiltonian-based methods, e.g., ED [15] or selective configuration interaction solvers [16, 19, 20], concentrate on the hybridization function Δ^{Bath} . This is the part of the non-interacting Green's function generated by the bath degrees of freedom. It can be shown [37] that the hybridization function follows the analytical expression

$$\Delta^{Bath}(\omega)_{\alpha,\beta} = \sum_{\ell=1}^{N_b} \frac{V_{\alpha,\ell} V_{\beta,\ell}^*}{\omega - \epsilon_\ell}. \quad (2)$$

Note that Δ^{Bath} equivalently encodes the fitting parameters $\{\epsilon_\ell\}_{\ell=1}^{N_b}$ and $\{V_{\alpha,\ell}\}_{\ell=1, \alpha=1}^{N_b, N_c}$ as a single rational function. The self-consistent loop is defined by matching G_{imp} , the impurity model's Green's function, to G_{latt} , the Green's function of the original system. Hence to close the loop we need a procedure to estimate G_{latt} from G_{imp} , which is detailed in the next paragraph. From G_{latt} we then extract the hybridization function Δ^{Calc} which should match Δ^{Bath} at self-consistency.

A given set of $\{\epsilon_\ell\}$ and $\{V_{\alpha,\ell}\}$ defines the impurity Hamiltonian H_{imp} , so that one can compute G_{imp} , at zero temperature, as follows,

$$G_{imp}(\omega)_{\alpha,\beta} = \langle \psi_0 | c_\alpha \frac{1}{\omega - (H_{imp} - E_0) + i\eta} c_\beta^\dagger | \psi_0 \rangle + \langle \psi_0 | c_\beta^\dagger \frac{1}{\omega + (H_{imp} - E_0) - i\eta} c_\alpha | \psi_0 \rangle, \quad (3)$$

where E_0 is the ground state energy, $|\psi_0\rangle$ is the ground state wavefunction, and η is an infinitesimal positive broadening factor. From this, one can compute the cluster self energy,

$$\Sigma_c(\omega)_{\alpha,\beta} = (\omega + \mu + i\eta)\delta_{\alpha,\beta} - h_{C,\alpha,\beta} - G_{imp}^{-1}(\omega)_{\alpha,\beta} - \Delta^{Bath}(\omega)_{\alpha,\beta}, \quad (4)$$

where h_C is the non-interacting part of H_C and μ is the chemical potential. The DMFT approximation amounts to assuming that one can estimate the full lattice Green's function by using the *local* cluster self energy, according to

$$G_{latt}(\mathbf{k}, \omega) = [(\omega + \mu + i\eta) - h(\mathbf{k}) - \Sigma_c(\omega)]^{-1}, \quad (5)$$

where $h(\mathbf{k})$ is the Fourier transform of h_C into momentum space. This recipe for computing G_{latt} corresponds to a block-diagonal ansatz for the global

self-energy in real space, with blocks corresponding to the clusters. Eq. (5) is exact in the infinite-dimensional limit for the Hubbard model [1], even when the cluster is only comprised of one single site. In finite dimensions, there are further corrections to Eq. (5); these terms incorporate the non-locality of the self-energy beyond the cluster. To improve the DMFT approximation one can consider larger and larger clusters, the exact result being recovered in the limit of infinite cluster size. The finite size scaling of the error is $\mathcal{O}\left(\frac{1}{N_c}\right)$ for cluster DMFT [2]. Other cluster methods have better scaling, such as $\mathcal{O}\left(\frac{1}{N_c^2}\right)$ for the dynamical cluster approximation (DCA), at the cost of more complex bath structures [27]. In particular, it can be shown that in cluster DMFT the only cluster sites that have non-zero couplings to the bath sites are those on the boundary of the cluster. This reduces the number of fit parameters and simplifies the structure of the Hamiltonian in Eq. (1). One can then Fourier-transform Eq. (5) to compute the original Green's function for the cluster degrees of freedom with the following expression,

$$G(\mathbf{R}_0, i\omega) = \frac{1}{V_{BZ}} \int_{BZ} d\mathbf{k} [(i\omega + \mu) - h(\mathbf{k}) - \Sigma_c(i\omega)]^{-1}, \quad (6)$$

where BZ stands for the first Brillouin zone and \mathbf{R}_0 is the center of the cluster and the origin of the superlattice of clusters. The hybridization function can be extracted by solving for Δ in Eq. (4) as

$$\Delta^{Calc}(\omega)_{\alpha,\beta} = (\omega + \mu + i\eta)\delta_{\alpha,\beta} - h_{C,\alpha,\beta} - G^{-1}(\mathbf{R}_0, i\omega)_{\alpha,\beta} - \Sigma_c(\omega)_{\alpha,\beta}. \quad (7)$$

If Δ^{Calc} coincides with Δ^{Bath} in Eq. (2), then the self-consistency loop terminates. Otherwise, one finds a new set of bath parameters $\{\epsilon_\ell\}$ and $\{V_{\alpha,\ell}\}$ by fitting Eq. (2) to Δ^{Calc} and starts a new iteration with a new impurity Hamiltonian. The cost function to minimize is usually defined in terms of a matrix norm of the difference between Δ^{Bath} and Δ^{Calc} , evaluated on some frequency grid $\{i\omega_n\}$ on the imaginary frequency axis. One avoids thus the poles of the Green's function and makes the fitting problem much easier. For simplicity, in this paper we will always consider the Frobenius norm, denoted $\|\cdot\|_F$, and a cost function of the form

$$\mathcal{J}(\{\epsilon_\ell\}, \{V_{\alpha,\ell}\}) = \frac{1}{N_\omega} \sqrt{\sum_{n=1}^{N_\omega} \|\Delta^{Calc}(i\omega_n) - \Delta^{Bath}(i\omega_n)\|_F^2}. \quad (8)$$

We point out that it is possible to add a frequency dependent weight function $g(\omega)$ to the sum. The objective of this weight function is to preferentially fit the frequencies close to the real axis. In this work, however, we emphasize the imaginary low-frequency region by choosing frequency grids $\{\omega_n\}_{n=1}^{N_\omega}$ with different point densities. Thus, we keep $g(\omega) = 1$ throughout the discussion. We will show that once enough bath parameters are provided, this fine tuning has no qualitative impact on our proposed optimization method.

Even when using the simplified cost function in Eq. (8), the optimization problem at hand is highly nontrivial and has been object of diverse studies [27, 28, 38]. We are tasked with minimizing the difference between two complex-valued, frequency-dependent $N_c \times N_c$ matrices using $N_b \cdot (N_c + 1)$ parameters. To maximize the model accuracy, N_c and N_b should be taken as large as our impurity solver allows. But the larger N_c and N_b are, the more difficult the fitting problem becomes, and in particular standard optimization procedures becomes increasingly susceptible to falling into local minima. In fact, when using a very efficient impurity solver, i.e., a solver capable of accommodating fairly large values of N_c and N_b , an inefficient fitting step can become the computational bottleneck [20]. In state-of-the-art Hamiltonian-based DMFT codes it is the norm to use elaborate fitting schemes in order to avoid instabilities and local minima [21, 22]. Furthermore, due to fitting on the imaginary axis, the result of the DMFT self-consistency can be significantly dependent on the fit frequency range [20], since physical information is distributed unevenly there (i.e. sum rules being encoded in the high frequency limit while the details of the state properties being encoded close to the real axis). Additionally, traditional approaches tend to require clusters of high symmetry in order to reduce the complexity of the fitting problem by block-diagonalizing the Green's function [27], which limits the size and shape of the clusters that can be studied.

In the following sections, we present a novel fitting method for impurity problems based on semi-definite relaxation (SDR) that addresses the issues mentioned in the previous section and substantially improves upon existing methodologies. Our method (i) implements the most general structure for a finite bath hybridization [39], (ii) shows fast timings allowing for treatment of large N_c and N_b , (iii) is systematically improvable in N_b , (iv) is empirically robust to starting conditions and symmetry constraints, and (v) shows rapid convergence for the low energy baths ϵ_ℓ . We demonstrate our claims by presenting results from zero-temperature cluster DMFT calculations using our fitting method and the

adaptive sampling configuration interaction (ASCI) algorithm as impurity solver [20, 30, 40, 41]

B. Formulation of the hybridization fitting

Motivated by the previous section, let $\Delta : \mathbb{C} \rightarrow \mathbb{C}^{N_c \times N_c}$ be a function that we seek to approximate in a discrete subset of the imaginary axis that we denote by Ω . In particular, we aim to find an approximation of the form

$$\Delta(i\omega) \approx \sum_{\ell=1}^{N_p} \frac{V_\ell V_\ell^*}{i\omega - \lambda_\ell}, \quad (9)$$

where $\omega \in \mathbb{R}$. We have introduced a shorthand notation omitting the cluster indices $\{\alpha, \beta\}$, so that the hybridization matrix is written simply as Δ , the coupling amplitudes are organized in vectors $V_\ell \in \mathbb{C}^{N_c}$, and the bath energies are still treated as scalars $\lambda_\ell \in \mathbb{R}$ for $\ell = 1, \dots, N_p$. In addition, for reasons to be clarified later, we denote the poles in Eq. (9) by λ_ℓ , not by ϵ_ℓ as in the last section.

To find the approximation in Eq. (9) we define the cost function,

$$\mathcal{J}(\{\lambda_\ell, V_\ell\}_{\ell=1}^{N_p}) = \frac{1}{N_\omega} \sqrt{\sum_{n=1}^{N_\omega} \left\| \Delta(i\omega_n) - \sum_{\ell=1}^{N_p} \frac{V_\ell V_\ell^*}{i\omega_n - \lambda_\ell} \right\|_F^2}, \quad (10)$$

where $\Omega = \{i\omega_n\}_{n=1}^{N_\omega}$. We aim to solve the optimization problem

$$\min_{\lambda_\ell, V_\ell} \mathcal{J}(\{\lambda_\ell, V_\ell\}_{\ell=1}^{N_p}), \quad (11)$$

whose minimizer provides the parameters for the ansatz in Eq. (9).

One natural approach to solve Eq. (11) is to use gradient based optimization to find the parameters, for a fixed number of poles, N_p . As already stated, the problem is highly non-convex in the sense that the optimization often gets trapped in local minima, which strongly depend on the choice of the initial guess. If one were to fix $\{V_\ell\}_{\ell=1}^{N_p}$, then the resulting problem could be reduced to a scalar-valued rational approximation problem that can be solved efficiently [42]. However, such algorithms, which are designed for scalar functions, fail to be directly applicable to our matrix-valued setting.

The main observation for solving Eq. (11) is that if we fix the poles $\{\lambda_\ell\}_{\ell=1}^{N_p}$, then the resulting problem can be reformulated using convex relaxation [23–25] into an SDR problem. The reformulated SDR problem is convex and can be efficiently solved using conic solvers [26], relying either on interior-point

methods [43] or first-order methods [44]. By exploiting this observation, we develop an algorithm with the following steps.

1. Relax the ansatz in Eq. (9) by replacing the rank-one matrices $V_\ell V_\ell^*$ by positive semi-definite matrices X_ℓ of arbitrary rank, resulting in a new loss function,
2. Optimize the poles and the matrices in an alternating fashion, which leverages state-of-the-art conic programming and quasi-Newton methods, and
3. Truncate the resulting positive semi-definite matrices and expanding them into sums of rank-one matrices, obtaining an approximation satisfying the ansatz in Eq. (9).

Finally, it should be noted that in cluster DMFT, only the N'_c sites in the boundary of the cluster have non-vanishing $V_{\ell,\alpha}$ [27]. In these cases, we can restrict the fit to a function $\Delta : \mathbb{C} \rightarrow \mathbb{C}^{N'_c \times N'_c}$.

1. Relaxation

We observe that the outer product $V_\ell V_\ell^*$ defines a positive semi-definite matrix of rank one. In a nutshell, the relaxation consists of replacing the rank-one matrices by more general positive semi-definite matrices X_ℓ without rank restriction, yielding the more general ansatz

$$\Delta(i\omega) \approx \sum_{\ell=1}^{N_p} \frac{X_\ell}{i\omega - \lambda_\ell}. \quad (12)$$

Analogous to the loss function defined in Eq. (10), we define the following loss function:

$$\mathcal{J}_{SDR}(\{\lambda_\ell, X_\ell\}_{\ell=1}^{N_p}) = \frac{1}{N_\omega} \sqrt{\sum_{n=1}^{N_\omega} \left\| \Delta(i\omega_n) - \sum_{\ell=1}^{N_p} \frac{X_\ell}{i\omega_n - \lambda_\ell} \right\|_F^2}. \quad (13)$$

We replace the optimization problem in Eq. (11) by a relaxed optimization problem

$$\min_{\lambda_\ell, X_\ell \succeq 0} \mathcal{J}_{SDR}(\{\lambda_\ell, X_\ell\}_{\ell=1}^{N_p}), \quad (14)$$

where $X_\ell \succeq 0$ means that X_ℓ is a Hermitian positive semi-definite matrix. This constraint is usually called a *conic* constraint, due to the fact that the set

of Hermitian positive semi-definite matrices forms a convex cone [25], i.e., if $A \succeq 0$ and $B \succeq 0$, then $aA + bB \succeq 0$, for any non-negative real numbers a and b .

From the standpoint of the size of the optimization space, we have increased the number of unknowns from $\mathcal{O}(N_p N_c)$ to $\mathcal{O}(N_p N_c^2)$. In addition, due to the rank relaxation, the minimizer from Eq. (14) does not immediately yield a function of the form given by Eq. (9). However, this shortcoming is not fundamental and can be easily addressed as follows at minimal additional cost.

Suppose that $\{\hat{\lambda}_\ell, \hat{X}_\ell\}_{\ell=1}^{N_p}$ are the minimizers of Eq. (14). By construction we have that $\hat{X}_\ell \succeq 0$. Then there exist matrices $\{U_\ell\}_{\ell=1}^{N_p}$ such that

$$\hat{X}_\ell = U_\ell U_\ell^* = \sum_{q=1}^{N_c} U_{\ell,q} U_{\ell,q}^*, \quad (15)$$

where $U_{\ell,q}$ is the q -th column of U_ℓ . Then we can expand

$$\sum_{\ell=1}^{N_p} \frac{\hat{X}_\ell}{i\omega - \hat{\lambda}_\ell} = \sum_{\ell=1}^{N_p} \sum_{q=1}^{N_c} \frac{U_{\ell,q} U_{\ell,q}^*}{i\omega - \hat{\lambda}_\ell} = \sum_{r=1}^{N_b^{\text{eff}}} \frac{\tilde{V}_r \tilde{V}_r^*}{i\omega - \tilde{\epsilon}_r}, \quad (16)$$

where in the last expression we have adopted appropriate definitions for \tilde{V}_r , $\tilde{\epsilon}_r$, and the index r . It is clear that Eq. (16) has the same form as the ansatz in Eq. (9), though here the *effective* number of baths, N_b^{eff} , is equal to $N_c N_p$. In particular, each bath energy $\tilde{\epsilon}_r$, is associated to N_c different baths.

2. Optimization

The minimization in Eq. (14) is non-trivial due mainly to the conic constraint. Unfortunately, popular methods such as the alternating direction method of multipliers (ADMM) [45], the method of alternating direction (AMA) [46], or backward-forward splitting (BFS) schemes [47] are not readily applicable to our case, given the non-convex nature of the cost function. Thus we propose a nested algorithm as follows.

Consider a loss function with respect to the poles as

$$\mathcal{J}_{\text{pol}}(\{\lambda_\ell\}_{\ell=1}^{N_p}) = \min_{X_\ell \succeq 0} \mathcal{J}_{SDR}(\{\lambda_\ell, X_\ell\}_{\ell=1}^{N_p}), \quad (17)$$

obtained by fixing the poles and solving an SDR problem using conic solvers [48, 49]. Using the optimality conditions for this problem the gradient can

be easily computed via

$$\frac{\partial}{\partial \lambda_k} \mathcal{J}_{\text{pol}}(\{\lambda_\ell\}_{\ell=1}^{N_p}) = \frac{\partial}{\partial \lambda_k} \left(\sqrt{\sum_{j=1}^{N_w} \left\| \Delta(i\omega_j) - \sum_{\ell=1}^{N_p} \frac{X_\ell}{i\omega_j - \lambda_\ell} \right\|_F^2} \right) \Big|_{X_\ell = X'_\ell} \quad (18)$$

where $X'_\ell = \arg \min_{X_\ell \geq 0} \mathcal{J}_{\text{SDR}}(\{\lambda_\ell, X_\ell\}_{\ell=1}^{N_p})$.

Using this newly defined cost function we solve

$$\{\hat{\lambda}_\ell\}_{\ell=1}^{N_p} = \arg \min_{\{\lambda_\ell\}_{\ell=1}^{N_p} \subset \mathbb{R}} \mathcal{J}_{\text{pol}}(\{\lambda_\ell\}_{\ell=1}^{N_p}) \quad (19)$$

using standard unconstrained gradient based methods. The full optimization loop is summarized in Alg. 1, which was implemented in MATLAB, where CVX [50] is used to solve Eq. (17), the derivatives of \mathcal{J}_{pol} are computed analytically using Eq. (18), and the outer optimization loop is solved with the BFGS method [51] included in the `fminunc` routine. We point out that the present algorithm is loosely related to BFS schemes [47, 52], and that it can be considered as an inexact alternating descent algorithm [53].

Algorithm 1 Pseudo-code for the nested optimization routine.

Input: Initial guess $\{\lambda_\ell\}_{\ell=1}^{N_p}$

Output: $\{\hat{\lambda}_\ell, \hat{X}_\ell\}_{\ell=1}^{N_p}$

while $\|\nabla \mathcal{J}_{\text{pol}}(\{\lambda_\ell\}_{\ell=1}^{N_p})\| > \epsilon_{\text{tol}}$ **do**

$X'_\ell \leftarrow \arg \min_{X_\ell \geq 0} \mathcal{J}_{\text{SDR}}(\{\hat{\lambda}_\ell, X_\ell\}_{\ell=1}^{N_p})$ by solving the SDR problem in Eq. (17)

$\lambda_\ell \leftarrow \text{BFGS}(\lambda_\ell, \nabla \mathcal{J}_{\text{pol}}(\{\lambda_\ell\}_{\ell=1}^{N_p}))$, where $\nabla \mathcal{J}_{\text{pol}}$ is computed using Eq. (18) and X'_ℓ

end while

$\hat{\lambda}_\ell \leftarrow \lambda_\ell$

$\hat{X}_\ell \leftarrow \arg \min_{X_\ell \geq 0} \mathcal{J}_{\text{SDR}}(\{\hat{\lambda}_\ell, X_\ell\}_{\ell=1}^{N_p})$

3. Truncation

As explained above, it is possible to post-process the output of Alg. 1 in order to obtain an approximation satisfying the ansatz in Eq. (9). However, this procedure increases the effective bath size significantly, which in turn increases the computational cost of the DMFT loop by enlarging the impurity problems. Fortunately, in many cases, the contributions of many of these effective baths are extremely small and can be discarded.

Such a truncation can be performed efficiently by solving one eigenvalue problem per SDR matrix. We

compute the eigenvalue decomposition of \hat{X}_ℓ as

$$\hat{X}_\ell = Q_\ell \Lambda_\ell Q_\ell^* \quad (20)$$

where Q_ℓ are unitary matrices and Λ_ℓ is a real diagonal matrix, containing the eigenvalues of X_ℓ ordered non-increasingly. For a fixed threshold δ , which we set to 10^{-5} in this work, we gather all the eigenvalues above this threshold and their corresponding eigenvectors in the matrices $\Lambda_\ell^\delta \in \mathbb{R}^{N_\ell^\delta \times N_\ell^\delta}$ and $Q_\ell^\delta \in \mathbb{C}^{N_c \times N_\ell^\delta}$, respectively, where N_ℓ^δ is the number of eigenvalues of \hat{X}_ℓ greater than δ , namely the ‘ δ -rank’ of \hat{X}_ℓ . Then we define

$$U_\ell^\delta = Q_\ell^\delta (\Lambda_\ell^\delta)^{1/2}, \text{ such that } \|U_\ell^\delta (U_\ell^\delta)^* - \hat{X}_\ell\|_F \lesssim \delta \quad (21)$$

Finally, we can follow Eq. (16) to obtain an approximation satisfying the ansatz in Eq. (9), but, hopefully, with somewhat smaller effective bath size, which is equal to the sum of the δ -ranks of the semi-definite matrices \hat{X}_ℓ . We point out that it should be possible to obtain a smaller effective bath size, by adding a rank-decreasing penalty, such as the nuclear norm, or trace, to the definition of \mathcal{J}_{pol} . However, we were not able to reliably achieve this without significantly decreasing the fit accuracy. To conclude, we remark that we have not explicitly said how we initialize the poles in Alg. 1. This will be stated in section III.

C. Impurity solver - ASCI

Before turning to the analysis of the results, a few words on the impurity solver are in order. As mentioned previously, to obtain accurate results in cluster DMFT it is important to raise N_c and N_b . In turn, this demands an impurity solver able to compute the Green’s function G_{imp} in Eq. (3) for a system with as many strongly-correlated degrees of freedom as possible. For zero temperature, Hamiltonian-based DMFT calculations, exact diagonalization (ED) [15] and truncation schemes [16, 19, 20] have been very popular [10, 21, 54]. The main idea behind truncation methods is the projection of the full system Hamiltonian into a subset \mathcal{T} of the complete Hilbert space \mathcal{H} . This subspace is chosen as small as possible, while still being able to represent the full ground state wavefunction accurately.

Among this kind of truncated ED schemes, the adaptive sampling configuration interaction (ASCI) algorithm [30, 40, 41] has been shown recently to be able to treat impurity models with large N_c and N_b efficiently [20]. The key aspect of ASCI is the identification of the optimal Hilbert space truncation of a given size to represent the ground state

wavefunction of a many-body system. This is done by introducing a ranking that allows one to estimate the most relevant Hilbert space degrees of freedom for the description of the ground state. ASCI is expected to perform particularly well in systems with (a) a high Hamiltonian-connectivity and (b) a ground state mainly composed by a few states in some single-body basis. In this work, using ASCI as impurity solver allows to use N_p up to 6 in the SDR fit, making the study of the large bath behavior of the fit possible, which is the regime where SDR becomes the more useful.

III. RESULTS

A. Characterization of the fit

In this section, we study in detail the performance of the SDR fit in Alg. 1 for impurity models, outside of the full DMFT calculation. We use as input data sample hybridization function matrices Δ from cluster DMFT calculations in (2×2) , (3×3) and (4×4) clusters of the two dimensional square lattice one-band Hubbard model. The Hamiltonian of this system is given by

$$H = -t \sum_{\langle \alpha, \beta \rangle, \sigma} c_{\alpha, \sigma}^\dagger c_{\beta, \sigma} - \mu \sum_{\alpha} n_{\alpha} + U \sum_{\alpha} n_{\alpha, \uparrow} n_{\alpha, \downarrow}, \quad (22)$$

where t is the hopping amplitude, $\langle \alpha, \beta \rangle$ denotes nearest neighbours, $\sigma = \{\uparrow, \downarrow\}$ is the spin degree of freedom, μ the chemical potential, U the local Coulomb interaction, $n_{\alpha, \sigma}$ is the number counting operator for site α , and spin species σ and $n_{\alpha} = n_{\alpha, \uparrow} + n_{\alpha, \downarrow}$. Unless otherwise specified, $U / t = 8$ throughout the calculations. We aim to showcase the following properties:

1. The quality of the SDR fit, as measured by the error cost function, can be systematically improved by increasing the number of poles.
2. The SDR fitting method is significantly faster than other non-linear fitting procedures when applied to this kind of impurity models.
3. The pole energies obtained from the SDR fit are robust with respect to the initial conditions. Furthermore, there is an easy diagnostic for identifying poor initial conditions, namely the rank of the X_{ℓ} matrices corresponding to each pole.
4. The fitting results can be further improved by taking the SDR solution in the form given by Eq. (16) and using it as an initial guess for

the optimization in Eq. (11). This last optimization is implemented via gradient-based methods.

5. The pole energies close to zero converge fairly rapidly with the number of poles.

The first four points are to be desired from a numerical point of view, since they identify the SDR method as efficient and robust. The last point is of particular importance from a physical perspective, since it helps justify the bath truncation in Hamiltonian-based approaches to DMFT. Indeed, if the low energy baths converge rapidly with the number of poles, this means that the low energy behavior of our embedding Hamiltonian also converges rapidly with the number of baths. Eventually, adding extra bath degrees of freedom will exclusively affect the high energy features of the model, which may not be the main interest of a low temperature (zero temperature) study. This allows us to justify *a posteriori* the bath truncation. Moreover, using SDR allows us in principle to determine the minimal number of bath orbitals required to converge the spectral properties of the embedding model up to the scale of interest. Whether the resulting bath size is then amenable to be treated with available impurity solvers, like the ASCI algorithm, is then something to be considered on a case by case basis.

The fits performed in this subsection are done over 50 imaginary frequency points in a linear span between $i\omega / t = 0$ and $i\omega / t = 40$. A detailed study of the effect that the frequency grid has on the fit is deferred to section III B. Since we have introduced several notation elements, we summarize the main abbreviations in Tab. I.

1. Fitting error

First, we want to examine how the fitting error progresses as a function of the number of poles for the different hybridization functions. We report the error and effective bath sizes in Tab. II.

From Tab. II it becomes clear that adding poles systematically improves the fit. The dominant contribution to the imaginary part of all hybridization functions considered here is the diagonal, whereas the dominant contribution of the real part is the off-diagonal. Adding extra poles improves the fit especially for the non-dominant parts of the hybridization function, as one would expect. This is exemplified in Fig. 1, where we show the fit result for diagonal and off-diagonal terms for the imaginary part of a (2×2) hybridization function as a function of the number of poles.

Symbol	Meaning
N_c	Number of cluster sites in the cluster DMFT calculation.
N'_c	Number of sites in the cluster boundary. Only this have nonzero coupling elements $V_{\alpha,k}$ [27].
N_p	Number of poles in the SDR fit.
N_b	Number of bath sites in the cluster DMFT calculation.
N_b^{eff}	Number of effective bath sites in the SDR fit. This number is determined after the truncation step, its maximal value being $N'_c \cdot N_p$.

TABLE I: Overview of notation.

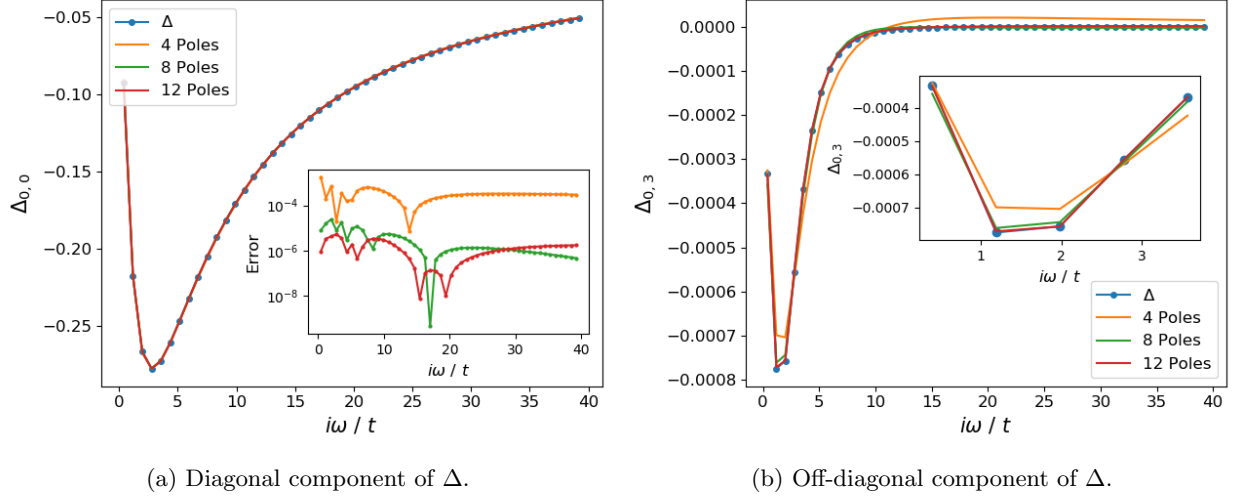


FIG. 1: Fit results for the hybridization function Δ of a (2×2) cluster impurity model for different number of poles. Shown are the imaginary parts of the $(0,0)$ diagonal and the $(0,3)$ off-diagonal components of the hybridization function to be fit (blue) and the resulting fits using 4 poles (orange), 8 poles (green) and 12 poles (red). The inset in the left subplot shows the absolute difference between the three fits and the computed hybridization function.

N_p	(2×2)		(3×3)		(4×4)	
	N_b^{eff}	Error	N_b^{eff}	Error	N_b^{eff}	Error
2	8	2.18e-3	16	2.16e-3	24	1.56e-2
3	12	1.59e-3	24	1.57e-3	36	8.44e-3
4	16	2.88e-4	32	3.97e-4	47	2.09e-3
5	20	2.64e-4	40	2.93e-4	56	1.31e-3
6	24	1.75e-5	48	5.99e-5	63	9.55e-4
7	28	1.14e-5	52	4.04e-5	72	8.37e-4
8	32	2.50e-6	55	5.41e-5	80	7.20e-4
9	35	1.56e-6	62	5.26e-5	86	7.16e-4
10	38	1.50e-6	66	3.08e-5	88	7.07e-4

TABLE II: Error and effective bath sizes N_b^{eff} for hybridization function fits on three different cluster sizes a function of the number of poles N_p using the SDR fit.

Eventually, the error of the fit reaches the 10^{-5} threshold, which is the value we chose for the trun-

cation of the X_ℓ matrices. Thus, after this limit is reached, the SDR method stops adding the maximal number of effective baths per pole (the X_ℓ matrices stop being full rank). One can further improve the quality of the fit by decreasing this threshold, at the price of increasing the number of effective baths. Conversely, we can use the SDR method to decide the number of effective baths needed to reach a particular quality of hybridization fit. For the (4×4) case, it seems that before the fit error reaches the 10^{-5} threshold the X_ℓ matrices stop being full rank.

2. Timings

To show the superior timing of the SDR fitting procedure, we compare it to the BFGS algorithm [51] as supported by the `fminunc` routine in MATLAB and the BOBYQA method (a derivative-free optimization method) implemented in the `nlopt`

	(2×2)						(3×3)					
	SDR		BFGS		BOBYQA		SDR		BFGS		BOBYQA	
N_p	[s]	Error	[s]	Error	[s]	Error	[s]	Error	[s]	Error	[s]	Error
2	5	2.18e-3	4	1.09e-1	28	2.18e-3	8	2.16e-3	87	9.24e-2	19729	1.17e-2
3	17	1.59e-3	5	9.98e-2	989	1.47e-3	16	1.57e-3	183	2.72e-2	21429	1.15e-2
4	8	2.88e-4	33	4.90e-3	2717	1.29e-4	14	3.97e-4	262	5.10e-3	31889	1.15e-2
5	7	2.64e-4	48	3.30e-3	5535	1.43e-4	24	2.93e-4	363	1.50e-3	40528	1.15e-2
6	27	1.75e-5	56	3.43e-4	1060	1.14e-4	29	5.99e-5	458	1.20e-3	> 24h	–
7	46	1.14e-5	69	2.27e-4	1629	9.64e-5	68	4.04e-5	522	1.10e-3	75922	1.15e-2
8	28	2.50e-6	79	9.92e-5	1790	7.83e-5	102	5.41e-5	617	1.10e-3	> 24h	–
9	96	1.56e-6	87	9.87e-5	2633	5.35e-5	56	5.26e-5	651	1.10e-3	45882	1.15e-2
10	39	1.50e-6	113	6.09e-5	1844	7.41e-5	129	3.08e-5	798	1.10e-3	84600	1.15e-2

TABLE III: Timings in seconds and final fit errors for the hybridization function fit using the SDR, quasi-Newton (BFGS), and derivative-free (BOBYQA) methods as a function of the number of poles N_p . When using BFGS or BOBYQA, the number of baths N_b corresponds to N_p times the number of sites in the cluster boundary N'_c . The same set of initial bath energies were used for all fits. The BFGS method was allowed to perform 2000 iterations.

library [55, 56] applied directly to Eq. (11). The timings in seconds for a number of SDR, BFGS, and BOBYQA fits are reported in Tab. III for different numbers of effective baths in both the (2×2) and (3×3) clusters. The number of baths used for the BFGS and BOBYQA optimizations is equal to $N_b = N_p \cdot N'_c$ (N'_c being the number of boundary sites in the cluster), which is the maximal number of effective baths used for the SDR with a given fixed number N_p of poles. Additionally, we report the final errors of all three methods. To provide a fair comparison, all fits were started with the same initial guesses for the bath energies and had the same error convergence tolerance of 10^{-5} . For the initial pole energies, we chose values symmetrically distributed around zero in a spread of $\pm 10 t$, e.g. in the $N_p = 2$ case we chose $\epsilon_p = \pm 2 t$. In the cases with odd N_p we set one of the initial poles at 0 t. Each of these poles corresponded to N'_c baths of the same energy for the BFGS and BOBYQA methods. The BFGS optimizations were limited to 2000 iterations. Increasing this number would eventually improve the error at the cost of longer runs, but we comment that in practice the possibility for further improvement is not substantial.

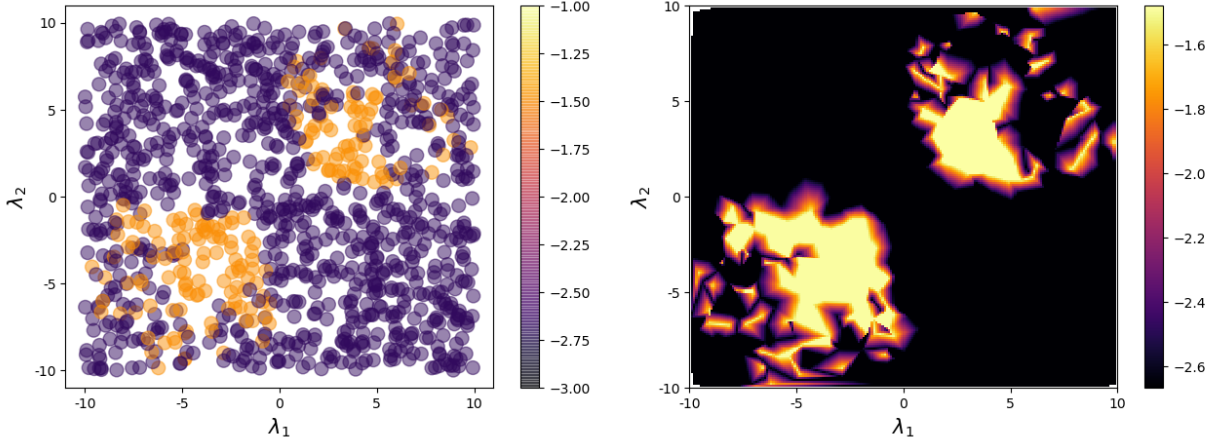
From Table III, the advantage of the SDR fit becomes clear. It is in general faster and, in the large bath limit, more effective than both BFGS and BOBYQA. For the small N_p values, i.e. $N_p \leq 4$, BOBYQA seems to be slower than BFGS, but for the (2×2) cluster it reaches a fit quality at least as good as the SDR fit. The BFGS optimizations in the (2×2) cluster with $N_p = 2, 3$ (i.e., 8 and 12 baths, respectively) converge to an error almost two orders of magnitude larger than the errors for SDR and BOBYQA. They are stuck in local minima.

The rest of the BFGS calculations were stopped after performing 2000 iterations. It is important to note that we are not claiming that the errors reported in Table III represent the best possible fit that the BFGS or BOBYQA algorithm can provide for the problem at hand. In fact, a clever choice of the initial guess and exploitation of symmetries presented in the system can greatly improve the optimization. The important conclusion to draw is that when used plainly, without such case-specific improvements, the SDR method clearly outperforms common optimization routines in the hybridization fitting problem. In what follows we show that the SDR fit is empirically robust with respect to these specializations.

3. Robustness to initial guess

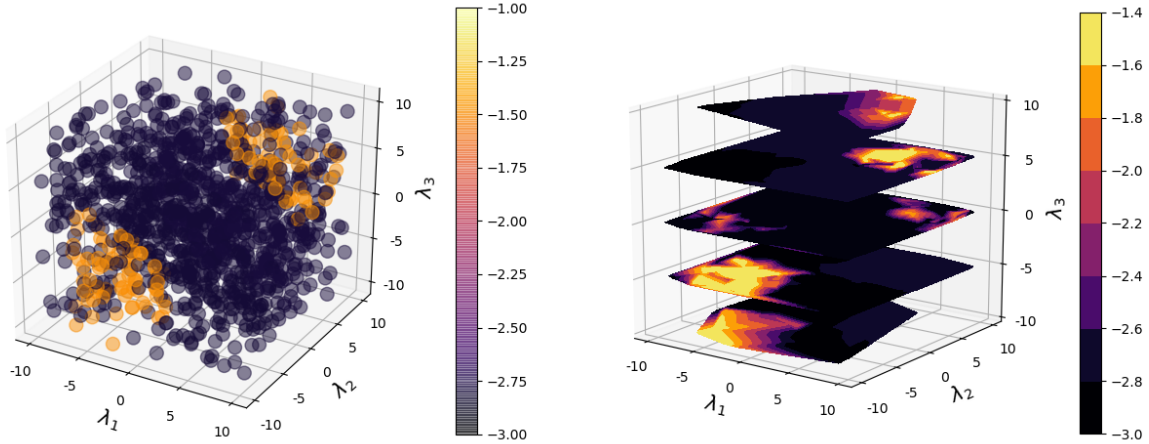
The goal of DMFT is to describe a many-body system, with particular interest in determining phase diagrams. When there are several competing orders this is a highly non-trivial task. In particular, for an iterative self-consistent method like DMFT, it is important to avoid falling into unphysical fixed points of the self-consistent loop. The fit plays a very important role here, and it is desirable for it to find the right set of bath parameters as independently as possible from the initial guess parameters and the frequency grid.

Here, we want to show empirically that the SDR fit is robust with respect to the initial guess. The effect of other parameters, like the frequency grid or the imposition of symmetries, is discussed in section III B. While different initial guesses can result in different pole energies using the SDR fit, for all the



(a) $N_p = 2$, fit error vs. initial pole energies. The best fits have the darkest color.

(b) $N_p = 2$, fit error vs. initial pole energies, linear interpolation. The best fits are colored black.



(c) $N_p = 3$, fit error vs. initial pole energies.

(d) $N_p = 3$, fit error vs. initial pole energies, linear interpolation.

FIG. 2: Logarithm of the fitting error for a (3×3) hybridization function using $N_p = 2$ (panels 2a and 2b) and $N_p = 3$ (panels 2c and 2d) vs. the initial pole energies. Presented are results for over 1000 initial points for each N_p , with the initial poles being drawn randomly from a uniform distribution. We report both the actual data as scatter plots and linearly interpolated heat maps. See text for details.

impurity models studied in this work, empirically we find it easy to distinguish between effective and non-effective fits. Over the set of test systems studied in this work, all X_ℓ matrices are of full rank in the case of the best fits. On the other hand, when the SDR method provides X_ℓ matrices of smaller rank, we are able to change the initial guess and find a better fit with full rank matrices, which reduces the error by one order of magnitude. Hence the robustness of the SDR fit is evidenced by the empirical realization that all effective fits for a given hybridization function had numerically equal pole energies and bath couplings. In other words, all initial parameters that produced an effective fit, i.e. which maximized the

ranks of the X_ℓ matrices, resulted in the same pole energies and couplings. It is this remarkable coincidence which makes us confident, at least empirically, of the robustness of the method. To the best of our knowledge, this property sets the SDR fit apart from all other non-linear optimization methods used for fitting of impurity models.

We exploited this property when choosing the initial set of pole energies for the input of the SDR fit. Originally, we would choose random energies from a normal distribution with average at negative the chemical potential $-\mu$. While this initial guess almost always results in an effective fit, for some examples, particularly with dense frequency grids or

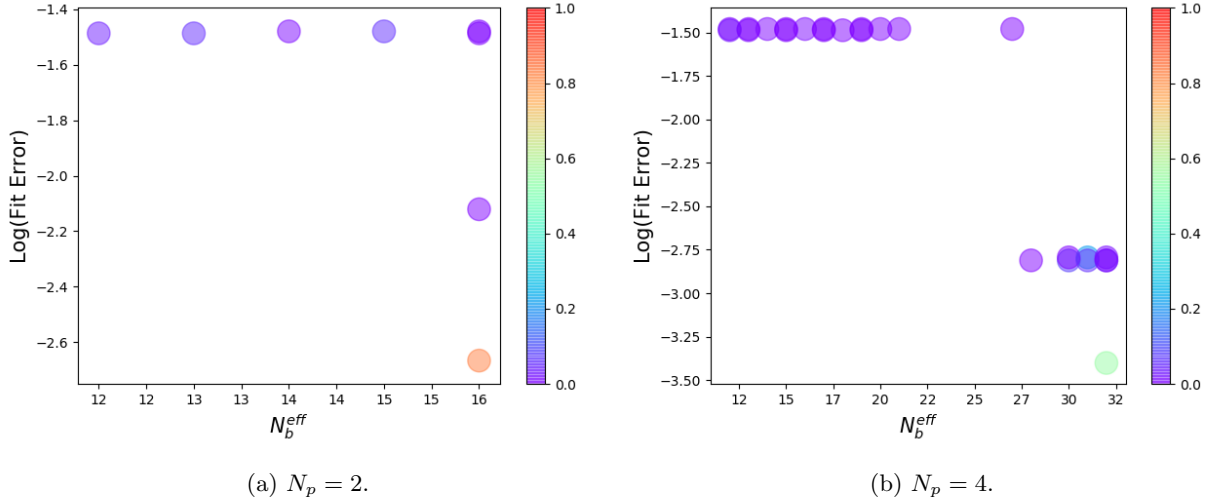


FIG. 3: Logarithm of the fitting error for a (3×3) hybridization function using $N_p = 2$ (left) and $N_p = 4$ (right) vs. the effective bath size N_b^{eff} . Presented are results for over 1000 initial points for each N_p , with the initial poles being drawn randomly from a uniform distribution. The color-code corresponds to the relative occurrence of each (error, N_b^{eff}) pair. See text for details.

	$N_p = 2$	$N_p = 3$	$N_p = 4$
Nr. of samples	1184	1195	1469
Rate of Success	83%	88% (43% + 45%)	55%
$\max[\sigma(\epsilon_\ell)]$	6e-6	5e-4	2e-4

TABLE IV: Statistical analysis of the robustness to the initial pole energies for the SDR fit on a (3×3) cluster hybridization function. Reported are the number of different initial conditions sampled, represented graphically in Fig. 2. The different initial pole energies were drawn randomly from a uniform distribution. The rate of success is defined as the number of optimal fits divided by the number of samples. The distribution of the errors for the non-optimal fits are presented in Fig. 3. In the case of $N_p = 3$, there were two optimal sets of bath parameters, occurring approximately with a 1:1 probability, see text for details. The last row reports the maximal standard deviation of each of the pole energies in the optimal fit parameters. This represents the spread of the pole energies in the optimal fit parameters.

with $N_p > 8$ in Fig. 5a, this kind of initial guess could lead to a poor fitting result. In these cases, we chose initial pole energies symmetrically distributed around zero in a spread from $\pm 8 t$. This always achieves an effective fit. In each case where this was necessary, we would test at least three different such initial guesses to assure the robustness of the fit parameters. We achieve an effective fit for

most of our practical initializations and ineffective fits can be fixed easily with small modifications to the initialization.

We illustrate this property in Fig. 2 and 3, and in Table IV. In Fig. 2 we present the logarithmic error of the fit for different values of initial poles in $N_p = 2$ (top) and $N_p = 3$ (bottom), for over 1000 different initial poles uniformly distributed in the $[-10; 10]^{\otimes N_p}$. The data suggests that most of the initial guesses result in an optimal fit (the exact proportion is given in Table IV). The initial guesses that fail are those that have no spread in the energy of the poles, i.e. they are concentrated in the corners of the square (cubic) interval. In Fig. 3 we report the logarithm of the final error vs the obtained effective bath size N_b^{eff} for $N_p = 2, 4$, also for over 1000 different initial poles each. The color-code corresponds to the relative occurrence of each (error, N_b^{eff}) pair. First, the optimal fit, significantly better than the next best one, is obtained always with the maximal N_b^{eff} and the highest occurrence. While there exist fits of maximal effective bath size and non-optimal error, these occur less than 5% of the times. For the $N_p = 2$ case the optimal fit is found with over 80% probability, in the $N_p = 4$ case with 55%. Thus, further increasing the number of poles probably makes randomly finding an effective starting guess more difficult, but the results in Fig. 2 suggest that an initial guess with some degree of spreading between the pole energies will likely give a good result. Even if that were not the case, the efficient timings already presented in Table III suggest that

it is always possible to perform some initial statistics in the first DMFT iteration, trying several initial pole energies in a reasonable time. This should be enough to effectively find a good fit. Since in the following iterations one can use the previous fit result as initial guess, the fit should be more likely to converge to the optimal solution in those cases.

Table IV summarizes the statistics of the robustness tests for $N_p = 2, 3$ and 4. The rate of success, defined as the ratio between the number of times the optimal fitting parameters were found and the number of attempts, decreases drastically from $N_p = 3$ to $N_p = 4$, but is still reasonably high for a random set of initial guesses. The most remarkable fact is still the uniqueness of the optimal solution, as discussed in the previous paragraphs. All optimal fits converge to the same pole energies, with minimal standard deviations as reported in Table IV. The case with odd N_p is an exception, in which there are actually two optimal sets of pole energies. These are approximately related to each other by a sign change. Such a sign relation is the symmetry suggested by the converged baths in Fig. 5 and 6, so the existence of this pair of optimal fit parameters for $N_p = 3$ may well be an even-odd artifact rather than an actual feature of the SDR fit.

4. Further optimization on top of SDR

We want to point out that the error achieved by the SDR fit is not necessarily the global minimum, in the sense that it can be improved further among hybridization functions of the form of Eq. (9) with a fixed number $N_b = N_p \cdot N_c$ of baths. By simply breaking the pole degeneracy, one can further reduce the fitting error. After having performed an SDR fit, the bath parameters obtained can then be used as starting guesses for a further optimization procedure, this time using all bath energies and couplings as free parameters at the same time. We used the BFGS algorithm and were able to improve the error by a factor between 1 and 4 depending on the number of poles N_p and the size of the cluster. Performing this gradient-based optimization on top of the SDR increases the total fitting time by an amount which depends on the number of baths and the cluster size. We present the factor of error improvement and timing for different fitting problems in Table V.

Considering the timings and factors of improvement in Table V, it is not always advantageous to perform the additional optimization, and one should consider whether the small error reduction is worth the computational cost. The degree to which the pole energies spread out after a gradient-based step is shown in Fig. 4 on the example of an SDR fit with

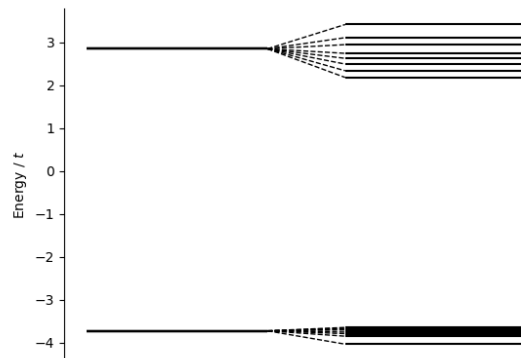


FIG. 4: Pole energies computed by the SDR fit for a (3×3) cluster hybridization with $N_p = 2$ before (left) and after (right) an additional BFGS optimization. Since the BFGS step does not maintain the degenerate pole structure, instead using all bath energies and couplings independently and simultaneously, the pole energies spread. See text for details.

$N_p = 2$ for a (3×3) hybridization function.

5. Pole energy convergence

In the previous subsections we have characterized in detail the important numerical properties of the SDR fit, independently of its influence on the full DMFT loop. Before describing how the SDR fit can improve the DMFT iteration, there is a final convergence property, of particular importance to justify the bath truncation in Hamiltonian-based DMFT, that we would like to address: the convergence of the pole energies with respect to the number of poles. To this end, in Fig. 5 we report the pole energies for the (2×2) and (3×3) hybridizations as a function of the number N_p of poles used in the SDR fit. (For simplicity, we only report even numbers of poles.) The pole energies seem to be symmetric around the zero energy for both the (2×2) and (3×3) clusters for $N_p = 2, 4, 6$. It should be noted that this is not a constraint on the SDR fit and that this symmetry is found numerically by the method without imposing explicitly any symmetry constraint. The “emergence” of other such symmetries during the DMFT loop with the SDR fit will be discussed in detailed in the section III B.

Fig. 5 suggests that the low energy poles converge quickly. By this we mean that adding extra poles does not significantly alter the energies of the previous ones. Effectively this means that that even

N_p	(2×2)		(3×3)	
	Error Reduction	Time [s]	Error Reduction	Time [s]
2	1.0	0.4	1.2	14
3	1.0	2	1.2	78
4	2.9	3	3.6	96
5	4.1	17	3.6	104
6	1.1	1	1.5	148
7	4.2	25	1.2	151
8	1.4	2	2.0	195
9	1.1	1	2.1	219
10	1.0	1	1.5	223

TABLE V: Error reduction (multiplicative improvement) and timing for optimizations using the SDR fit and a consecutive BFGS step. The timings refer only to final BFGS optimization.

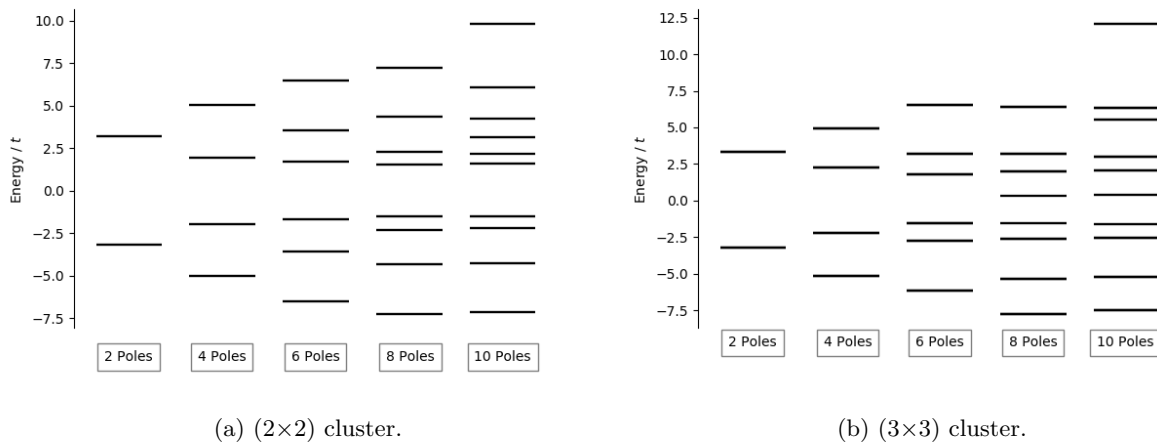


FIG. 5: Pole energies resulting from the SDR fits for the (2×2) (left) and (3×3) (right) clusters for $N_p = 2, 4, 6, 8, 10$. We see that the pole energies are fairly symmetric around zero energy and that the low energy poles converge very quickly. See text for details.

fits with relatively few poles can describe the low energy properties of the embedded system and that additional poles eventually only play a role for the simulation of the high energy features. This feature is of great importance for Hamiltonian-based implementations of DMFT, where a truncation of the bath is necessary. The results in Fig. 5 effectively corroborate the physical intuition that such a bath truncation is a viable approximation when one is only interested in describing the low energy features of the model. It should be noted that these are results for fits on a single hybridization function. Physically, one would need to converge a full DMFT calculation for each number of poles (each bath size) and compare those converged pole energies to make a definitive statement about the energy convergence in the embedded model. We present such a study in the following section, where the performance of

the SDR fit inside full cluster DMFT calculations is presented in detail.

B. Performance in DMFT calculation

Having characterized the performance of the SDR fit for impurity problems in detail, we now turn to showing how its properties affect full cluster DMFT calculations. There are four points that we would like to stress here, beyond those already presented in the previous section:

1. The pole energy convergence, which we have demonstrated for a single hybridization fit, translates into convergence of the bath energies of fully converged DMFT calculations, as the number of poles is increased.

2. The SDR fitting procedure finds, up to a negligible numerical error, symmetries present in the system without a need for imposing them.
3. Given a large enough number of poles, the SDR fit result is largely unaffected by the imaginary frequency range used to sample the hybridization function.
4. Given its timing and flexibility, the SDR fit works for both highly symmetric clusters where a block-diagonalization of the hybridization function is possible [27] and those clusters where this is not feasible.

As in the previous section, all calculations presented here are based on a two-dimensional square-lattice one-band Hubbard model at half-filling.

We now expound upon the advantage of the SDR fit in DMFT for clusters of low symmetry. The calculations of the previous section treat highly symmetric clusters, in particular square clusters embedded in the full square lattice. For these symmetric clusters, one can usually block-diagonalize the hybridization function, making the fitting problem much easier [27, 28]. For less symmetric clusters, however, this is not necessarily possible, and a flexible and efficient fitting method like SDR is important. To show that the SDR fit can handle lower-symmetry clusters with the same ease as fully symmetric ones, we present in this section results from DMFT calculations in (2×3) , (2×4) and (1×6) clusters of the two-dimensional square-lattice Hubbard model at half-filling and $U/t = 8$. Being able to treat a pseudo-one-dimensional cluster, or stripe, like the (1×6) cluster opens up interesting avenues of research, in light of the ongoing quest for stripe order in the under doped Hubbard model [9, 57–59]. All converged bath parameters used to produce the data in this section are collected in tables in the supporting information.

1. Convergence of the bath energy

In the previous section we showed that, for a single hybridization fit, the low energy baths converge rapidly with increasing number of poles (see Fig. 5) which is an important property for the bath truncation in Hamiltonian-based DMFT approaches. The results presented in Fig. 5 originated from one single hybridization fit, and were computed using numbers of poles well beyond the capabilities of state-of-the-art zero temperature impurity solvers. It is fundamental to test the same convergence behavior when considering fully converged cluster DMFT calculations for the bath sizes that can be handled. Fig. 6

shows the converged bath energies for different cluster DMFT calculations on (2×2) , (2×3) , (2×4) and (1×6) clusters with $N_p = 2, 4, 6$. The initial guess for the poles in the fits DMFT loop was chosen in the same way described in section III A 3, i.e. by choosing random pole energies centered around the negative chemical potential $-\mu$. For the subsequent DMFT iterations, the previous pole energies were used as initial guess. Numerical convergence of the pole energies was obtained after 10 to 20 DMFT iterations.

The most apparent property is the near-perfect symmetry of the bath energies around 0 in both cases. This is not explicitly imposed in the SDR procedure and may be a consequence of the particle-hole symmetry present in the one-band Hubbard model at half filling. As a consequence, calculations with an odd number of poles recovered one bath energy close to zero. The authors have not observed this behavior with other fitting methods, like the BOBYQA optimization [55, 56]. Meanwhile, the conclusion from the previous section still holds: the low-lying pole energies converge rapidly as the number N_p of poles is increased. We have omitted the case of odd N_p since there are intrinsic even-odd disparities that do not reflect the actual convergence properties of the pole energies. While the impurity solver imposes a strict limit on the total number of bath sites (in particular, preventing exploration beyond $N_p = 6$), it seems fair to say that the two lowest-lying pairs of baths have converged their energies at $N_p = 6$, and probably adding further poles will only yield higher-energy baths. The positive interpretation for the validity of bath truncation schemes remains.

2. Emergence of symmetry

Imposing the symmetries present in the cluster in the fitting process is a necessary step with most currently employed optimization schemes. In such schemes, this step both simplifies the fit and assures that the right physical fixed point in the DMFT self-consistent cycle is obtained [27, 28]. While it is possible to adopt these strategies when using the SDR approach, we have observed that this is not necessary since the bath couplings $V_{\alpha,\ell}$ computed using the SDR fit already show symmetry, at least qualitatively. By this we mean that all symmetrically equivalent cluster sites experience equivalent baths. To illustrate this point, in Fig. 7 we report the absolute value of the bath couplings for all cluster and bath sites of a (2×2) cluster DMFT calculation with 24 baths ($N_p = 6$). In this figure, the (2×2) cluster is represented by the (2×2) grid of subplots, each subplot corresponding to one of the cluster sites.

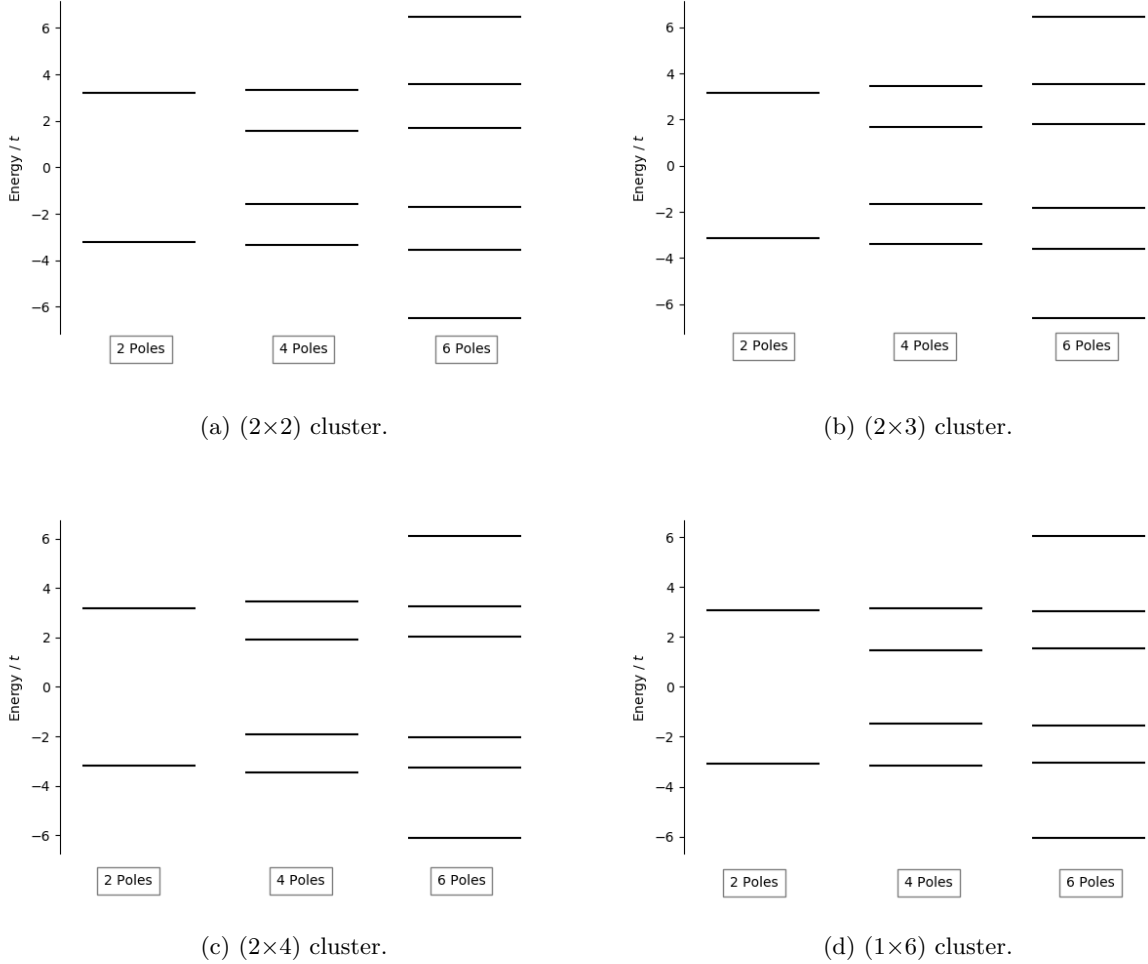


FIG. 6: Pole energies computed by the SDR fit for clusters of different size with $N_p = 2, 4, 6$. These are converged in cluster DMFT calculations at $U / t = 8$.

The circles then correspond to the 24 baths, their energies shown in their y-axis position and their absolute value of the coupling $|V_{\alpha,\ell}|$ encoded in the color map.

From Fig. 7 it becomes apparent that all four cluster sites experience an equivalent bath. For each given bath energy, the four cluster sites couple to baths that present the same four absolute coupling amplitudes, albeit in a different order depending on the cluster site. This permutation is physically inconsequential however, since the bath energies are degenerate. The full coupling values $V_{\alpha,\ell}$ can differ by a sign, but it is only their amplitude square, understood as a transition probability in a Fermi-golden-rule argument, that bears physical meaning. We want to stress that we do not explicitly impose any symmetry constraint on the SDR fit and that

it arrives at this symmetric bath configuration automatically. The only user-provided input is the initial guess for the bath parameters for the first DMFT iteration, which is chosen by drawing random numbers around the chemical potential μ for the pole energies and between 0 and 1 for the bath couplings. For this initial bath guess, we do impose symmetry by assigning the same bath couplings to symmetrically equivalent cluster sites. Nonetheless, this aspect of the initial guess only enters via the computation of the Hamiltonian; the SDR uses a random initial guess for the bath couplings. Furthermore, the SDR procedure relaxes the symmetries of the initial bath guess: the bath couplings to symmetrically equivalent sites are not identical, but merely equivalent up to the permutations as previously discussed and demonstrated in Fig. 7. We have not



FIG. 7: Absolute value of the bath couplings $|V_{\alpha,\ell}|$ for a (2×2) cluster DMFT calculation with 24 baths ($N_p = 6$). Each subplot corresponds to one site in the (2×2) cluster, the circles representing the different bath sites. The energy of each bath site is represented on the y-axis, while the color encodes the absolute value of the coupling $|V_{\alpha,\ell}|$ between the baths and the cluster sites.

observed such symmetry emergence from other fitting methods like the BOBYQA algorithm.

Such behavior is not limited to the fully symmetric (2×2) cluster, but in fact also holds for rectangular clusters that do not enjoy all the symmetry of the full square lattice. To demonstrate this, we report in Fig. 8 the absolute values for the couplings of a (2×3) cluster with 24 baths ($N_p = 4$). Here, an interesting thing occurs. From the full symmetry group of a (2×3) slab, one would expect to only have two symmetrically inequivalent sites: the four corners and the two center sites. This is due to the C_2 axis and the two reflection planes perpendicular to the cluster plane. This observation notwithstanding, as can be seen in Fig. 8, the SDR fit seems to break the reflection symmetries and identify three different types of sites, leaving only the C_2 perpendicular rotation symmetry. The small number of poles may cause the SDR routine to break the symmetry in

order to improve the fit. It is plausible that increasing the number of poles would remedy this loss of symmetry, but the large number of bath degrees of freedom that this analysis would need exceeds the limitations of most impurity solvers.

3. Effect of the imaginary frequency grid

The fact that the DMFT self-consistency loop is performed along the imaginary frequency axis has both advantages and disadvantages. On the one hand, it simplifies the calculations and fits, since the Green's function is smooth along the imaginary axis and decays as $\propto |\frac{1}{\omega}|$. On the other hand, it increases the difficulty of extracting physical information contained in Green's functions near the real axis. Indeed, the real-frequency physical details of the system, e.g., the existence or absence of a gap,

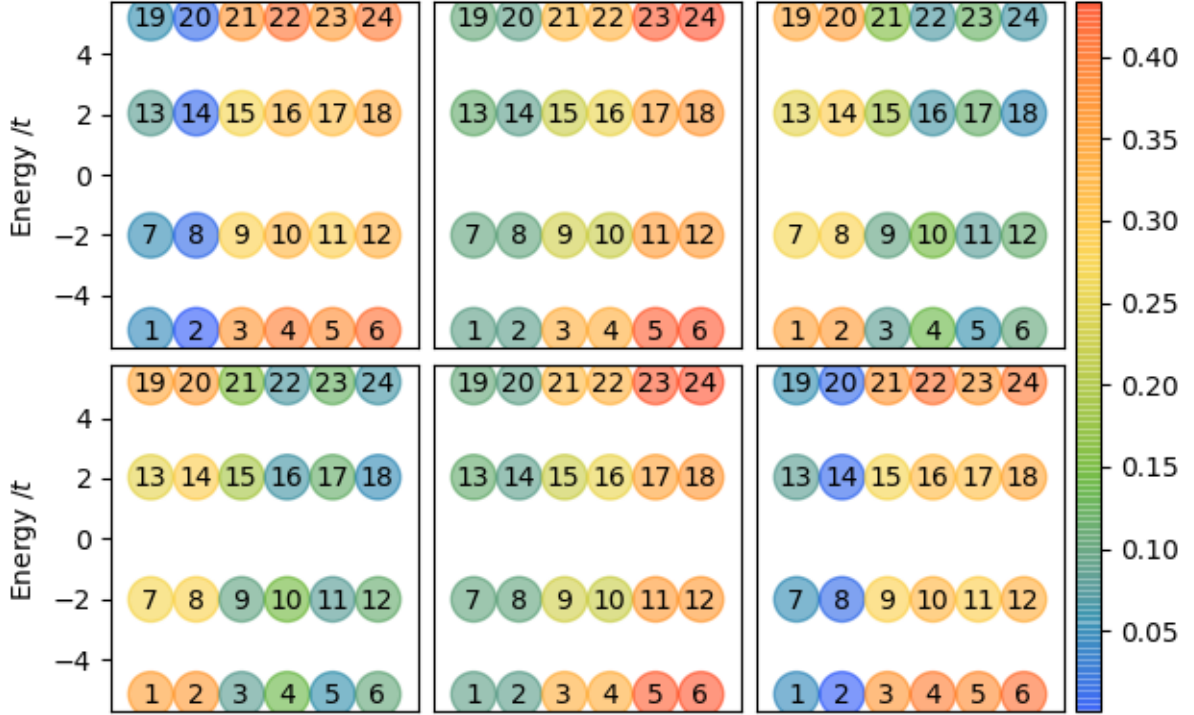


FIG. 8: Absolute value of the bath couplings $|V_{\alpha,\ell}|$ for a (2×3) cluster DMFT calculation with 24 baths ($N_p = 4$). Each subplot corresponds to one site in the (2×3) cluster, the circles representing the different bath sites. The energy of each bath site is represented on the y-axis, while the color encodes the absolute value of the coupling $|V_{\alpha,\ell}|$ between the baths and the cluster sites.

are encoded close to the real axis with small imaginary frequency values, and the long frequency tail mainly includes information about sum rules. To understand this last point, it suffices to consider that the Green's function along the imaginary frequency axis depends on its value on the real axis according to

$$G(i\omega_n) = -\frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{\text{Im}(G(\omega))}{i\omega_n - \omega}. \quad (23)$$

Sum rules are written as integrals over the full real frequency axis, e.g., the relation

$$-\frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \text{Im}(G(\omega)) = 1. \quad (24)$$

Such non-local information in the frequency space is stored across the entire imaginary frequency axis,

and to recover it one needs to consider the large-frequency regime as well as the low-frequency one. This is manageable at finite temperature, where the Green's function is only defined on a finite set of imaginary frequencies (the Matsubara frequencies). However, at zero temperature, one needs in principle the full positive imaginary frequency axis to get the full physics, while on the real axis it usually suffices to consider features close to a non-interacting Fermi energy.

Maybe the most immediate and palpable consequence for zero temperature Hamiltonian-based DMFT calculations is the fact that the frequency grid used for the fit can have a large impact on the reliability of the results. How closely this grid approaches the real axis and how far into the high frequency limit it extends can change the qualitative behavior of the real-frequency results. Of course, one cannot make the grid arbitrarily vast and dense,

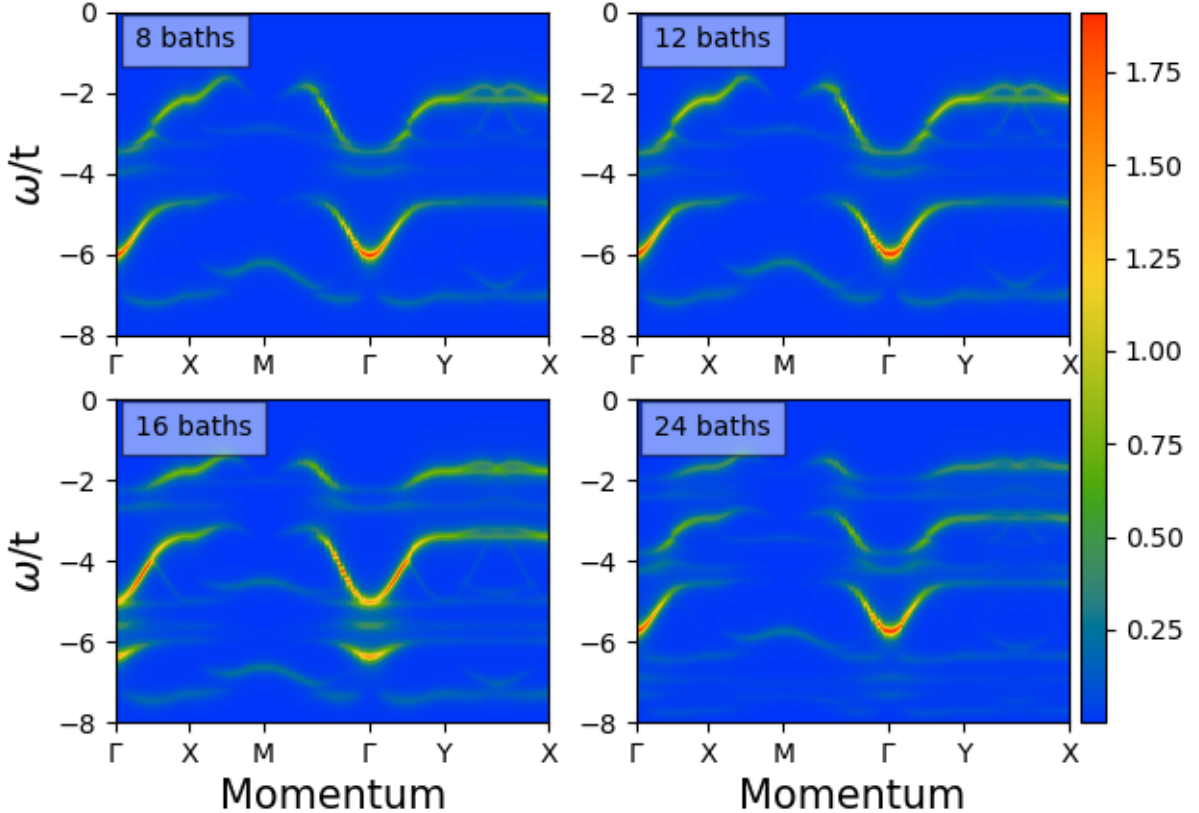


FIG. 9: Spectral weights for (2×2) cluster DMFT calculations on the square-lattice one-band Hubbard model at half-filling with $U/t = 8$ and different numbers of baths. SDR fit performed using a linear imaginary frequency grid with frequency cutoff at $i\omega_{max} = 40 t$.

since doing so increases the difficulty of the optimization step. As a consequence, it requires a significant amount of physical intuition about the problem to decide the limits and density of this frequency grid. In some circumstances, it may even be preferable to choose a dense grid, very close to the real axis but ignoring the high frequency limit, to achieve satisfactory results.

In a previous work [20], such behavior was observed using the BOBYQA algorithm for the fit in a series of (2×2) cluster DMFT calculations on the two-dimensional square-lattice one-band Hubbard model at half-filling and $U/t = 8$, when increasing the number of bath sites. In a series of cluster DMFT calculations with 8, 12, 16 and 24 baths, we observed a drastic change in the high frequency “hole-bands” of the spectral weight when increasing the number of baths from 12 to 16, and in turn to 24. We argued that the drastic change was due to overfitting, in the sense that the BOBYQA algorithm

was reducing the cost function in Eq. (8) in the high frequency tail at the expense of a poorer fit close to the real axis. However, as discussed above, this is where the details of the physics of the system are encoded. After limiting the fitting frequency grid to the immediate vicinity of the real axis, the drastic change in the large bath limit disappears, and thus we concluded that this change is not physical, but rather an artifact of the fit. Here, we perform a similar study using the SDR fit.

First, for comparison with the results obtained in [20], we report in Fig. 9 the spectral weights $A(\mathbf{k}, \omega)$, defined as

$$A(\mathbf{k}, \omega) = -\frac{1}{\pi} \text{Im}(G_{latt}^p(\mathbf{k}, \omega)), \quad (25)$$

for the same (2×2) cluster DMFT calculations described in the previous paragraph (i.e., using $N_p = 2, 3, 4, 6$), using the customary linear frequency grid of 50 points along the imaginary frequency axis from

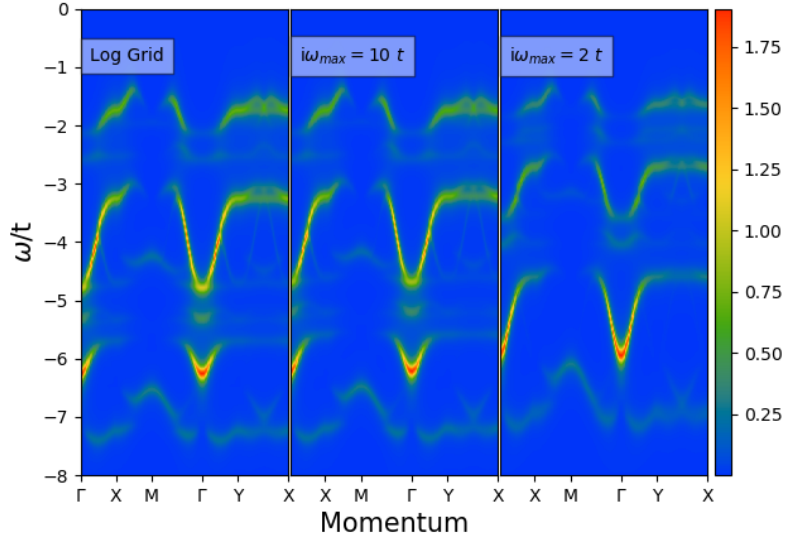


FIG. 10: Spectral weights for (2×2) cluster DMFT calculations with 16 baths ($N_p = 4$) on the square-lattice one-band Hubbard model at half-filling with $U/t = 8$. SDR fit performed using different imaginary frequency grids (see text for details).

$i\omega = 0 t$ to $i\omega = 40 t$ for the DMFT self-consistency. The superscript p in the lattice Green's function $G_{latt}^p(\mathbf{k}, \omega)$ accounts for the fact that one has to recover the translational symmetry broken in the cluster DMFT treatment. In this work we employ the following periodization scheme [19, 60]

$$G_{latt}^p(\mathbf{k}, \omega) = \frac{1}{N_c} \sum_{\alpha, \beta=1}^{N_c} e^{i\mathbf{k}(\mathbf{r}_\alpha - \mathbf{r}_\beta)} G_{latt}(\mathbf{k}, \omega)_{\alpha, \beta}, \quad (26)$$

where $G_{latt}(\mathbf{k}, \omega)$ is the unperiodized lattice Green's function in Eq. (5) evaluated on the real frequency axis, and \mathbf{r}_α are the relative position vectors of the lattice sites inside the cluster. Now Fig. 9 shows the same drastic change in the high-frequency bands when changing from 12 to 16 baths as the one observed in [20]. However, this change seems to be rectified in the 24 bath calculation. Before getting into this significant distinction between the BOBYQA and SDR results, we want to show explicitly that changing the frequency range indeed gets rid of the strange behavior for the 16 baths ($N_p = 4$) calculation.

In Fig. 10 we report the spectral weights for the (2×2) cluster DMFT with 16 baths that result from using different frequency grids in the fit. In particular we show a logarithmic grid of 50 points from $i\omega = 0.01 t$ to $i\omega = 40 t$ (left panel) and two linear grids of Matsubara frequencies $\omega_n = \frac{(2n+1)\pi}{\beta}$ with $\beta = 600$, with maximal frequencies of $10 t$ (center

panel) and $2 t$ (right panel). These three grids were chosen to increase the relative weight of points close to the real frequency axis in the fit, by choosing either the logarithmic spacing or a dense grid of points with a hard cutoff at some small frequency. For reference, the imaginary part of the diagonal component of the hybridization function along the imaginary frequency axis is shown for all these calculations in Fig. 11.

From Fig. 11 it appears that there is no difference between the hybridization functions between the three grids (in fact, the average absolute difference between the two Matsubara grids for $i\omega < 2 t$ is approximately 2×10^{-4}), and yet the real-axis properties show an appreciable difference (see Fig. 10). The reason for this is as follows. Computing the hybridization function with bath parameters obtained from the fit which stops at $i\omega = 2 t$ for imaginary frequencies *beyond* $i\omega = 2 t$ unveils significant differences between the hybridization functions (see inset in Fig. 11). This implies that the high frequency behavior of the hybridization function computed by fitting the smallest of our frequency ranges is not appropriately converged. Indeed, if one adds but five points to the Matsubara frequency range ending at $i\omega = 2 t$, these points equally distributed from $i\omega = 2 t$ to $i\omega = 40 t$, and performs a new DMFT calculation starting from the “converged” parameters of the small frequency range calculation, the bath parameters rapidly change to improve the high frequency hybridization function. These become equiv-

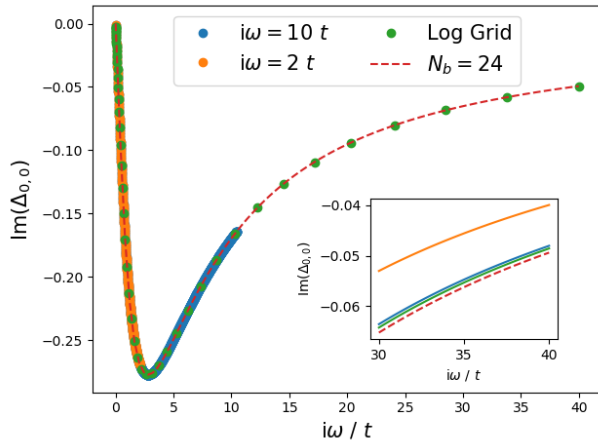


FIG. 11: Imaginary part of diagonal component of the hybridization function Δ along the imaginary frequency axis for DMFT calculations on a (2×2) cluster with 16 baths ($N_p = 4$) and three different frequency grids (see text for details) and with 24 baths ($N_p = 6$). The fit with 24 baths was performed on the linear grid with 50 points from $i\omega/t = 0$ to $i\omega/t = 40$. The inset shows the same Δ , computed in a dense grid in the high frequency range using the bath parameters of the four different fits.

alent to those from the larger frequency range calculation, or the logarithmic scale one, and the spectral weights recover the structure in the left panel of Fig. 10. This confirms the fact that to obtain the right physical behavior with a moderate number of baths, there is the need to choose the frequency grid appropriately. This requires some *a priori* knowledge of the solution, and it would be desirable that such fine-tuning would not be necessary. In particular, the hope is that in the large bath limit this bias of the frequency grid becomes unnecessary.

Here, the calculation with 24 baths ($N_p = 6$) shows that the SDR fit fulfills this expectation. The spectral weights for the 24 bath calculation shown in Fig. 9 (lower right panel) coincide with the 8 and 12 bath calculations, and with the 16 bath calculations with accurately fit low frequency behavior (see right panel of Fig. 10 and Fig. 12 for a comparison of the spectral weights for all bath size with the 16 bath calculation using the shorter frequency grid). This finding suggests that the 24 bath calculation has fit the low frequency behavior of the hybridization function to the 10^{-4} accuracy mentioned above. Moreover, the red curve in the inset of Fig. 11 shows the high frequency behavior of the hybridization function as computed with the converged bath parameters of the 24 bath calculation. Thus, it turns out that with 24 baths at its disposal, the SDR fit is capable of fitting

both the high and low frequency behaviors, recovering the right physical behavior without any need for contrived constraints on the frequency grid in the fit. We have also performed a 24 bath calculation with the logarithmic frequency scale and the Matsubara scale ending at $i\omega = 2t$. The results did not change, adding further support to our claim.

As an additional note, it is also possible to add more weights to the low frequency range along the imaginary frequency axis. This can be done by adding a frequency dependent weight function $g(\omega)$ to the cost function in Eq. (8). The effect of such a cost function is equivalent to choosing frequency grids of inhomogeneous densities, and therefore their effect does not qualitatively change the conclusions presented above.

We now summarize the main points of this subsection. The physical disadvantages of fitting on the imaginary frequency axis with a finite number of baths raise the important and nontrivial issue of deciding the frequency range for fitting. This choice may influence whether the results of the DMFT calculation are reliable and representative of the true physics of the system. We have found that one way to reach said physical results without relying on *a priori* knowledge to pick the best frequency range is to have a large number of baths and a corresponding fit that is able to treat such large bath effectively and in an unbiased way. We have shown that the SDR fit fulfills these conditions with a moderate number of poles ($N_p = 6$), and that paired with state-of-the-art impurity solvers which can handle the large bath limit (ASCI in this paper), it can reach the physical set of bath parameters without choosing a tailored frequency range.

IV. CONCLUSION

We have introduced an optimization method using semi-definite relaxation for the hybridization fitting step in the implementation of cluster DMFT. This optimization routine offers a straightforward approach to extrapolations into the infinite bath limit in Hamiltonian-based DMFT calculations. There has been development of other approaches with the same objective, notably an algorithm working directly on the real frequency axis [61] for single-site DMFT calculations, combining orbital rotations in the space of bath degrees of freedom with the substitution of the fitting step by a purely linear-algebraic calculation to treat hundreds of baths. There have also been advances toward employing compact Green's function representations on the imaginary frequency axis [62] to extrapolate to the infinite bath limit in pure exact diagonalization DMFT.

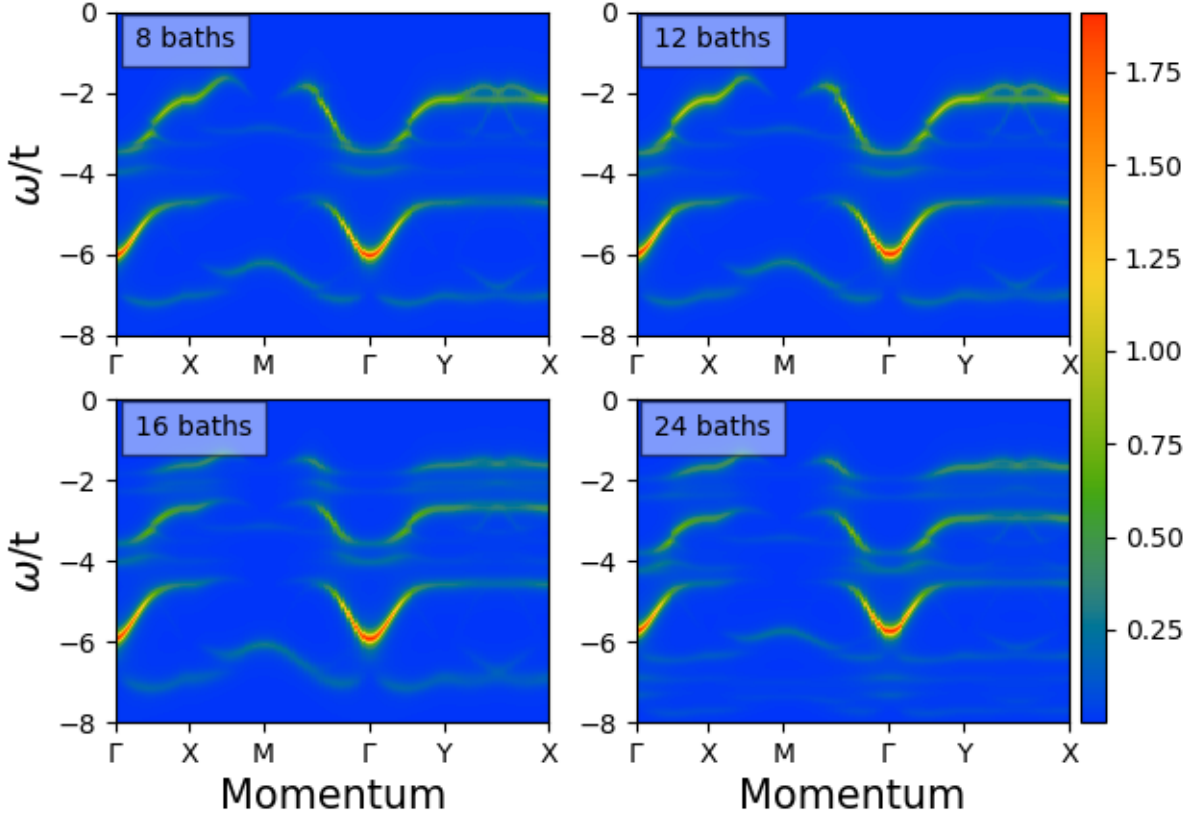


FIG. 12: Spectral weights for (2×2) cluster DMFT calculations on the square-lattice one-band Hubbard model at half-filling with $U/t = 8$ and different numbers of baths. SDR fit performed with a linear imaginary frequency grid with frequency cutoff $i\omega_{max} = 40t$ for $N_b = 8, 12, 24$ and $i\omega_{max} = 2t$ for $N_b = 16$.

Our method is conceptually simpler than these approaches, mainly implementing an optimization routine to treat the most general bath structure effectively. We have shown that it is an efficient and systematically improvable method, outperforming standard optimization routines and able to deal with a large number of bath parameters effectively. Furthermore, it has several important properties motivating its use in Hamiltonian-based cluster DMFT calculations: empirical robustness with respect to the initial guess, rapid convergence of low energy poles, apparent recognition of system symmetries, independence of the fitting frequency grid in the large bath limit, and versatility in the treatment of clusters with low symmetry. This method offers a systematic, natural, and easy-to-use approach to the fitting problem in DMFT, which despite its difficulty is perhaps the least-documented step in the self-consistency cycle.

ACKNOWLEDGMENTS:

This work was partially supported by the Department of Energy under Grant No. DE-SC0017867,

No. DE-AC02-05CH11231, by the Air Force Office of Scientific Research under award number FA9550-18-1-0095 (L.L. and L.Z.), by the National Science Foundation Graduate Research Fellowship Program under grant DGE-1106400 (M.L.), and by a *Obra Social "La Caixa"* graduate fellowship (C.M.Z.). Computational resources provided by the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation Grant No. OCI-1053575, are gratefully acknowledged. We thank Garnet Chan, Anil Damle, Alexandre Foley, Antoine Georges, Gabriel Kotliar, Olivier Parcollet, Reinhold Schneider, Lan Tran, Dominika Zgid and Jianxin Zhu for helpful discussions.

-
- [1] 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).
- [2] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, “Quantum cluster theories,” *Rev. Mod. Phys.* **77**, 1027 (2005).
- [3] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. Marianetti, “Electronic structure calculations with dynamical mean-field theory,” *Rev. Mod. Phys.* **78**, 865 (2006).
- [4] G. Knizia and G. K.-L. Chan, “Density matrix embedding: A simple alternative to dynamical mean-field theory,” *Phys. Rev. Lett.* **109**, 186404 (2012).
- [5] E. Kanonenka, A. A. Gull and D. Zgid, “Systematically improvable multiscale solver for correlated electron systems,” *Phys. Rev. B* **91**, 121111 (2015).
- [6] Q. Chen, G. H. Booth, S. Sharma, G. Knizia, and G. K.-L. Chan, “Intermediate and spin-liquid phase of the half-filled honeycomb hubbard model,” *Phys. Rev. B* **89**, 165134 (2014).
- [7] B.-X. Zheng and G. K.-L. Chan, “Ground-state phase diagram of the square lattice hubbard model from density matrix embedding theory,” *Phys. Rev. B* **93**, 035126 (2016).
- [8] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jiménez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof’ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull (Simons Collaboration on the Many-Electron Problem), “Solutions of the two-dimensional hubbard model: Benchmarks and results from a wide range of numerical algorithms,” *Phys. Rev. X* **5**, 041041 (2015).
- [9] B.-X. Zheng, C.-M. Chung, P. Corboz, G. Ehlers, M.-P. Qin, R. M. Noack, H. Shi, S. R. White, S. Zhang, and G. K.-L. Chan, “Stripe order in the underdoped region of the two-dimensional hubbard model,” *Science* **358**, 1155 (2017).
- [10] S. Sakai, M. Civelli, and M. Imada, “Direct connection between mott insulators and d-wave high-temperature superconductors revealed by continuous evolution of self-energy poles,” *Phys. Rev. B* **98**, 195109 (2018).
- [11] A. Paul and T. Birol, “Applications of dft+dmft in materials science,” arXiv preprint arXiv:1809.09246 (2018).
- [12] A. A. Markov, G. Rohringer, and A. N. Rubtsov, “Fate of a topological invariant for correlated lattice electrons at finite temperature,” ArXiv e-prints (2019), arXiv:1904.05087 [cond-mat.str-el].
- [13] C. Walsh, P. Semon, D. Poulin, G. Sordi, and A.-M. S. Tremblay, “Local entanglement entropy and mutual information across the mott transition in the two-dimensional hubbard model,” *Phys. Rev. Lett.* **122**, 067203 (2019).
- [14] C. Walsh, P. Semon, D. Poulin, G. Sordi, and A.-M. S. Tremblay, “Thermodynamic and information-theoretic description of the mott transition in the two-dimensional hubbard model,” *Phys. Rev. B* **99**, 075122 (2019).
- [15] M. Caffarel and W. Krauth, “Exact diagonalization approach to correlated fermions in infinite dimensions: Mott transition and superconductivity,” *Phys. Rev. Lett.* **72**, 1545 (1994).
- [16] D. Zgid, E. Gull, and G. K.-L. Chan, “Truncated configuration interaction expansions as solvers for correlated quantum impurity models and dynamical mean-field theory,” *Phys. Rev. B* **86**, 165128 (2012).
- [17] F. A. Wolf, I. P. McCulloch, and U. Schollwöck, “Solving nonequilibrium dynamical mean-field theory using matrix product states,” *Phys. Rev. B* **90**, 235131 (2014).
- [18] 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).
- [19] A. Go and A. J. Millis, “Adaptively truncated hilbert space based impurity solver for dynamical mean-field theory,” *Phys. Rev. B* **96**, 085139 (2017).
- [20] C. Mejuto-Zaera, N. M. Tubman, and K. B. Whaley, “Next Generation Dynamical Mean-Field Theory Simulations with the Adaptive Sampling Configuration Interaction Method,” *Phys. Rev. B* **100**, 125165 (2019).
- [21] A. Go and A. J. Millis, “Spatial correlations and the insulating phase of the high- T_c cuprates: Insights from a configuration-interaction-based solver for dynamical mean field theory,” *Phys. Rev. Lett.* **114**, 016402 (2015).
- [22] D. Zgid and G. K.-L. Chan, “Dynamical mean-field theory from a quantum chemical perspective,” *J. Chem. Phys.* **134**, 094115 (2011).
- [23] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics* **9**, 717 (2009).
- [24] E. J. Candès and T. Tao, “The power of convex relaxation: Near-optimal matrix completion,” *IEEE Transactions on Information Theory* **56**, 2053–2080 (2010).
- [25] L. Vandenberghe and S. Boyd, “Semidefinite programming,” *SIAM Review* **38**, 49–95 (1996), <https://doi.org/10.1137/1038003>.
- [26] S. R. Becker, E. J. Candès, and M. C. Grant, “Templates for convex cone problems with applications to sparse signal recovery,” *Mathematical Programming Computation* **3**, 165 (2011).
- [27] E. Koch, G. Sangiovanni, and O. Gunnarsson, “Sum rules and bath parametrization for quantum cluster theories,” *Phys. Rev. B* **78**, 115102 (2008).
- [28] A. Liebsch and H. Ishida, “Temperature and bath size in exact diagonalization dynamical mean field theory,” *J. Phys.-Condens. Mat.* **24**, 053201 (2011).
- [29] A. Foley, S. Verret, A.-M. S. Tremblay, and D. Senechal, “Coexistence of superconductivity

- and antiferromagnetism in the hubbard model for cuprates,” *Phys. Rev. B* **99**, 184510 (2019).
- [30] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, “A deterministic alternative to the full configuration interaction quantum monte carlo method,” *J. Chem. Phys.* **145**, 044112 (2016).
- [31] See Supplemental Material at [URL will be inserted by publisher] for tables containing the converged bath parameters used to produce the figures in section III B of the paper. These tables are also provided as a .zip file. We further include comparisons of cluster DMFT calculations using our method with other ED and configuration interaction based solvers.
- [32] W. Metzner and D. Vollhardt, “Correlated lattice fermions in $d = \infty$ dimensions,” *Phys. Rev. Lett.* **62**, 324 (1989).
- [33] E. Müller-Hartmann, “Correlated fermions on a lattice in high dimensions,” *Z. Phys. B Con. Mat.* **74**, 507–512 (1989).
- [34] E. Müller-Hartmann, “The hubbard model at high dimensions: some exact results and weak coupling theory,” *Z. Phys. B Cond. Mat.* **76**, 211–217 (1989).
- [35] A. Georges and G. Kotliar, “Hubbard model in infinite dimensions,” *Phys. Rev. B* **45**, 6479 (1992).
- [36] 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).
- [37] P. Phillips, *Advanced Solid State Physics, 2nd Edition, Ch. 7* (Cambridge University Press, 2012).
- [38] D. Senechal, “Bath optimization in the cellular dynamical mean-field theory,” *Phys. Rev. B* **81**, 235125 (2010).
- [39] C. J. Bolech, S. S. Kancharla, and G. Kotliar, “Cellular dynamical mean-field theory for the one-dimensional extended hubbard model,” *Phys. Rev. B* **67**, 075110 (2003).
- [40] N. M. Tubman, C. D. Freeman, D. S. Levine, D. Hait, M. Head-Gordon, and B. Whaley, “Modern Approaches to Exact Diagonalization and Selected Configuration Interaction with the Adaptive Sampling CI Method,” *ArXiv e-prints* (2018), arXiv:1807.00821 [quant-ph].
- [41] N. M. Tubman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, “An efficient deterministic perturbation theory for selected configuration interaction methods,” *arXiv preprint arXiv:1808.02049v1* (2018).
- [42] A. Damle, G. Beylkin, T. Haut, and L. Monzn, “Near optimal rational approximations of large data sets,” *Applied and Computational Harmonic Analysis* **35**, 251 – 263 (2013).
- [43] J. F. Sturm, “Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones,” *Optimization Methods and Software* **11**, 625–653 (1999), <https://doi.org/10.1080/10556789908805766>.
- [44] B. O’Donoghue, E. Chu, N. Parikh, and S. Boyd, “Conic optimization via operator splitting and homogeneous self-dual embedding,” *J. Optimiz. Theory App.* **169**, 1042–1068 (2016).
- [45] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
- [46] P. Tseng, “Applications of a splitting algorithm to decomposition in convex programming and variational inequalities,” *SIAM Journal on Control and Optimization* **29**, 119–138 (1991), <https://doi.org/10.1137/0329006>.
- [47] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, edited by H. H. Bauschke, R. S. W. S. Bauschke, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz (Springer New York, New York, NY, 2011) pp. 185–212.
- [48] E. D. Andersen, C. Roos, and T. Terlaky, “On implementing a primal-dual interior-point method for conic quadratic optimization,” *Math. Program.* **95**, 249–277 (2003).
- [49] R. H. Tütüncü, K. C. Toh, and M. J. Todd, “Solving semidefinite-quadratic-linear programs using sdpt3,” *Math. Program.* **95**, 189–217 (2003).
- [50] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx> (2014).
- [51] C. G. Broyden, “The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations,” *IMA Journal of Applied Mathematics* **6**, 76–90 (1970), <http://oup.prod.sis.lan/imamat/article-pdf/6/1/76/2233756/6-1-76.pdf>.
- [52] P. Lions and B. Mercier, “Splitting algorithms for the sum of two nonlinear operators,” *SIAM Journal on Numerical Analysis* **16**, 964–979 (1979), <https://doi.org/10.1137/0716071>.
- [53] J. Ortega and W. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Society for Industrial and Applied Mathematics, 2000) <https://epubs.siam.org/doi/pdf/10.1137/1.9780898719468>.
- [54] R. Wang, A. Go, and A. Millis, “Weyl rings and enhanced susceptibilities in pyrochlore iridates: k p analysis of cluster dynamical mean-field theory results,” *Phys. Rev. B* **96**, 195158 (2017).
- [55] M. J. D. Powell, “The BOBYQA algorithm for bound constrained optimization without derivatives,” Department of Applied Mathematics and Theoretical Physics, Cambridge England, technical report NA2009/06 (2009).
- [56] S. G. Johnson, The NLOpt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>.
- [57] G. Hager, G. Wellein, E. Jeckelmann, and H. Fehske, “Stripe formation in doped hubbard ladders,” *Phys. Rev. B* **71**, 075108 (2005).
- [58] E. W. Huang, C. B. Mendl, S. Liu, S. Johnston, H.-C. Jiang, B. Moritz, and T. P. Devereaux, “Numerical evidence of fluctuating stripes in the normal state of high- t_c cuprate superconductors,” *Science* **358**, 1161 (2017).

- [59] E. W. Huang, C. B. Mendl, H.-C. Jiang, B. Moritz, and T. P. Devereaux, “Stripe order from the perspective of the hubbard model,” *npj Quantum Materials* **3**, 22 (2018).
- [60] A. Go and G. S. Jeon, “Phase transitions and spectral properties of the ionic hubbard model in one dimension,” *Phys. Rev. B* **84**, 195102 (2011).
- [61] 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).
- [62] Y. Nagai and H. Shinaoka, “Smooth self-energy in the exact-diagonalization-based dynamical mean-field theory: Intermediate-representation filtering approach,” *J. Phys. Soc. Jpn.* **88**, 064004 (2019).