Truncated configuration interaction expansions as solvers for correlated quantum impurity models and dynamical mean-field theory

Dominika Zgid,1 Emanuel Gull,2 and Garnet Kin-Lic Chan1,∗

1Department of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853, USA
2Max Planck Institut für Physik komplexer Systeme, Dresden, Germany

(Received 12 March 2012; published 22 October 2012)

The development of polynomial cost solvers for correlated quantum impurity models, with controllable errors, is a central challenge in quantum many-body physics, where these models find applications ranging from nanoscience to the dynamical mean-field theory (DMFT). Here, we describe how configuration interaction (CI) approximations to exact diagonalization (ED) may be used as solvers in DMFT. CI approximations retain the main advantages of ED, such as the ability to treat general interactions and off-diagonal hybridizations and to obtain real spectral information, but are of polynomial cost. Furthermore, their errors can be controlled by monitoring the convergence of physical quantities as a function of the CI hierarchy. Using benchmark DMFT applications, such as single-site DMFT of the one-dimensional Hubbard model and 2 × 2 cluster DMFT of the two-dimensional Hubbard model, we show that CI approximations allow us to obtain near-exact ED results for a tiny fraction of the cost. This is true over the entire range of interaction strengths including “difficult” regimes, such as in the pseudogap phase of the two-dimensional Hubbard model. We use the ability of CI approximations to treat large numbers of orbitals to demonstrate convergence of the bath representation in the 2 × 2 cluster DMFT using a 24-bath orbital representation. CI approximations thus form a promising route to extend ED to problems that are currently difficult to study using other solvers such as continuous-time quantum Monte Carlo, including impurity models with large numbers of orbitals and general interactions.

DOI: 10.1103/PhysRevB.86.165128 PACS number(s): 71.10.Fd, 71.30.+h, 71.20.—b, 71.15.—m

I. INTRODUCTION

Quantum impurity models describe a finite set of interacting “impurity” orbitals coupled to a large number of noninteracting “bath” or “lead” states. They were originally designed to describe the effect of magnetic impurities embedded in a nonmagnetic host material,1 but have since found a wide variety of applications ranging from nanoscience,2 where they are used to describe quantum dots and molecular conductors, to surface science,3 for the description of molecule adsorption on a substrate, to research in quantum field theories.4,5 In recent years, they have gained an increasingly important role in condensed matter and materials science, where they appear as auxiliary models in the simulation of correlated lattice models within the so-called dynamical mean-field theory6−8 (DMFT) approximation and its extensions.9,10 A general quantum impurity model is described by the Hamiltonian

\[ H = H_{\text{loc}} + H_{\text{bath}} + H_{\text{hyb}}, \]  

\[ H_{\text{loc}} = \sum_{pq} t_{pq} d_{p}^{\dagger} d_{q} + \sum_{pqrs} t_{pqrs} d_{p}^{\dagger} d_{q}^{\dagger} d_{r} d_{s}, \]  

\[ H_{\text{bath}} = \sum_{ki} \varepsilon_{ki} c_{ki}^{\dagger} c_{ki}, \]  

\[ H_{\text{hyb}} = \sum_{kip} V_{kip} c_{ki}^{\dagger} d_{p} + \text{H.c.} \]

\[ H_{\text{loc}} \] describes the “impurity” itself, \( H_{\text{bath}} \) a set of noninteracting “bath” or “lead” sites, and the impurity-bath coupling or “hybridization” is contained in \( H_{\text{hyb}} \). The operators \( d^{\dagger} \) and \( c^{\dagger} \) create and annihilate impurity and lead electrons, \( t \) and \( I \) describe the impurity hopping and interaction terms, \( \varepsilon \) is a bath dispersion, and \( V \) the impurity-bath hybridization strength.

The finite number of impurity interactions makes quantum impurity models numerically tractable. The development of accurate and reliable numerical solvers for correlated quantum impurity models is therefore one of the central challenges of computational many-body physics. Many different approaches have been proposed. Among those that can be made exact with sufficient computational effort, at least for some classes of models, are quantum Monte Carlo methods, such as the continuous-time quantum Monte Carlo (CT-QMC),11 renormalization group (RG) methods, including numerical12,13 and density matrix RG;14,15 and exact diagonalization (ED).16 All these techniques have different strengths and weaknesses. For example, CT-QMC is formulated in imaginary time, and real-frequency data at high frequencies, obtained with analytic continuation, is notoriously unreliable, while numerical renormalization group (NRG) has limited resolution in spectral quantities far from the Fermi surface and can not be reliably extended beyond two impurity orbitals. ED does not suffer from the above two difficulties, but introduces a finite-size error associated with a discrete bath representation. For some special Hamiltonians, such as those with density-density interactions and diagonal hybridizations, CT-QMC has no sign problem, and thus affords a polynomial time solution of the impurity problem. However, for general Hamiltonians, all the above techniques including CT-QMC exhibit an exponential scaling with the number of impurity orbitals and, in the case of ED, with the number of bath orbitals. Consequently, there is an urgent need to develop controlled approximate solvers for general impurity models, where the exponential scaling is ameliorated or eliminated.

The dynamical mean-field theory and its cluster variants provide an ideal test bed for numerical quantum impurity solvers. DMFT is now established as a powerful theoretical
framework for describing interacting quantum solids, both in
the context of single-site multiorbital and single-orbital cluster
model Hamiltonians, as well as with realistic interactions
within the DFT + DMFT (Refs. 9 and 17) framework. In
DMFT, the bulk quantum problem is mapped onto a self-
consistent quantum impurity model. Depending on the lattice
model parameters, regions of weak, intermediate, and large
correlation strengths can be accessed, and the wealth of
previously computed data and the well-understood physics
makes reliable comparisons possible.

Our present work presents controlled polynomial cost
approximations to ED, using the idea of configuration
interaction (CI) that has long been studied in quantum chem-
istry. Recall that, at zero temperature, the Green’s function is

\[ \langle ig_{ij}(\omega) = \langle \Psi | a_i^\dagger (\omega - H + E - i\eta)^{-1} a_j | \Psi \rangle + \langle \Psi | a_j (\omega - H + E + i\eta)^{-1} a_i | \Psi \rangle, \quad (5) \]

where \( E \) and \( \Psi \) are the ground-state energy and wave function of \( H \). In ED, the true ground-state wave function \( \Psi \) is expanded
in the complete space of Slater determinants, and the size of
the complete space, which scales exponentially as a function
of the number of impurity and bath orbitals, is the primary
limitation of the calculation. Even using state-of-the-art ED
(Lanczos) codes, no more than 16 electrons in 16 orbitals
(32 spin orbitals) can be treated. CI approximates ED
by solving for \( \Psi \) within a restricted variational space of
Slater determinants. This variational space is constructed
by including determinants based on their excitation level relative to a single, or multiple, physically motivated reference
determinants. The various CI methods form a convergent
hierarchy of approximations, where the variational space
is systematically increased, and thus their error, relative to
the theoretical ED limit, can be controlled by monitoring
the convergence of the hierarchy. Furthermore, because CI
methods exhibit a polynomial scaling with respect to the
number of impurity and bath orbitals, they have the potential
to treat much larger systems than ED. Indeed, in quantum
chemistry, CI calculations with 1000 orbitals are routine.19

The central question to answer in the context of correlated
quantum impurity models is whether or not CI approximations
form a sufficiently rapidly convergent hierarchy for the physical
quantities of primary interest. If so, the ability to treat large
numbers of orbitals and off-diagonal hybridizations, while
retaining the strengths of ED, can be expected to be of great
utility in revealing the physics of complex quantum impurity
models. In Ref. 20, two of us (Zgid and Chan) demonstrated
that a very approximate CI solver could reproduce exact
diagonalization results in a simple quantum impurity problem
arising from the DMFT approximation to the cubic hydrogen
solid. However, in that work our focus was not on the quality
of the solver, but rather on chemical aspects of DMFT, such
as the use of realistic Hamiltonians which do not suffer from
double counting. In the current work, we return to a systematic
study of the CI approximations themselves. Since our target is
to assess the quality of our approximations, we concentrate
here on well-studied DMFT benchmark problems whose
physics is understood, including single-site DMFT of the
one-dimensional (1D) Hubbard model and 2 × 2 cluster DMFT
of the two-dimensional (2D) Hubbard model. The double-
counting issue9,17 does not arise in these systems. As we will
demonstrate, in these systems the CI approximations allow us
to reproduce the ED calculations at a small fraction of the cost.
Furthermore, because we can treat a larger number of orbitals
than with ED, we will demonstrate that we can converge these
models with respect to their bath representation. In the cases of
the 2 × 2 cluster DMFT approximation to the Hubbard model,
this has previously not been possible with ED (Lanczos).

The structure of this paper is as follows. In Sec. II, we
first describe the theory behind CI approximations, including
a detailed description of the excitation space, single-reference
and multireference CI approximations, complete active spaces,
and natural orbitals. In Sec. III, we briefly describe some
technical details of the implementation. In Sec. IV, we describe
our application of CI solvers in model DMFT problems
described above, using ED as a comparison where possible,
and we demonstrate further the ability of CI to converge
systems with large numbers of bath orbitals. Finally, we
describe perspectives and conclusions in Sec. V.

II. CONFIGURATION INTERACTION APPROXIMATIONS

Configuration interaction (CI) wave functions \( |\psi_C\rangle \) are a set of systematic approximations to the ED wave function
\( |\psi_{ED}\rangle \). Using CI wave functions, the ground state of the impurity
model is determined in a truncated subset of the complete set
of Slater determinants. Once a ground-state wave function is
obtained, the impurity Green’s function and self-energy, the
central quantities in DMFT, are evaluated through Eq. (5). CI
truncations rely on an a priori ranking of the importance of
the determinants, in terms of excitation character relative to a
single starting determinant (single-reference CI), or multiple
starting determinants (multireference CI). This ranking is
motivated by ordinary and degenerate perturbation theory,
although CI approximations are not perturbative approxima-
tions per se. Note that the accuracy of CI truncations of the
determinant space depends on the choice of orbital basis, and
this is also an important consideration in a CI calculation.

We first motivate single-reference CI approximations using
the Anderson model in the limit of small \( U \). Here, \( \Psi \) is close
to the Hartree-Fock (HF) determinant \( \Phi \). Consequently, we take \( \Phi \) to be
the (single-) reference determinant in the CI. We
next change from the site basis \( (d,d^\dagger,c,c^\dagger) \) to the basis
of HF orbitals \( (a,a^\dagger) \), which are a mixture of impurity and
bath orbitals. The determinants in the HF basis can be labeled
in terms of excitation or particle-hole character relative to
\( \Phi \). For example, a singly excited determinant \( \Phi_{ij} = a_j^\dagger a_i (\Phi) \)
has one particle and one hole relative to the HF determinant;
the doubly excited determinant \( \Phi_{iab} = a_j^\dagger a_i a_k^\dagger a_l^\dagger (\Phi) \)
has two particles and holes, and so forth. To construct a CI
approximation, we truncate the complete determinant space
based on the maximum excitation character of determinants in
the expansion of \( \Psi \). For example, in a CI singles and doubles
(CISD) approximation, we approximate \( \Psi \) with an expansion
with at most doubly excited determinants (see Fig. 1)

\[ |\Psi\rangle \approx \sum_{i,a} c_{ia}^p |\Phi_{ia}^p\rangle + \sum_{ij,ab} c_{ijab}^h |\Phi_{ijab}^h\rangle, \quad (6) \]
More accurate CI approximations, with up to triple (CISDT), quadruple (CISDTQ), and higher excitations can be formulated in a similar way, and these systematically approach the full ED solution. However, for small $U$, we can expect CISD to already be a good approximation because it contains all classes of determinants that couple with $\Phi$ through first-order perturbation theory in $U$. Similarly, CISDTQ contains all classes of determinants that couple through second order, and so on.

The above CI approximations (CISD, CISDT, etc.) are termed single reference because the truncation of the determinant space is based on excitations relative to a single-reference determinant $\Phi$. We expect this hierarchy of truncations to be rapidly convergent for small $U$, but for large $U$, multiple determinants can become degenerate with $\Phi$ on the scale of $U$ and contribute with similar weights to $\Psi$. For example, in the single-site Anderson model at large $U$, $\Psi$ is qualitatively given by a superposition of two determinants describing a “Kondo” singlet coupling of electrons of opposite spin on the impurity orbital and in the bath. In such cases, it is more reasonable to rank determinants with respect to a set of near-degenerate determinants. This is the basis of multireference CI approximations. Denoting the near-degenerate determinants as $\Phi(I)$ (where $I$ ranges over the degenerate set), then for each $\Phi(I)$, we can define singly, doubly, and higher excited determinants

$$
\Phi^a_{ij}(I) = a_i^a a_j^a |\Phi(I)\rangle,
$$

and

$$
\Phi^{ab}_{ij}(I) = a_i^a b_j^b |\Phi(I)\rangle,
$$

and so on. In the multireference CI singles-doubles approximation, $\Psi$ is expanded in

$$
|\Psi\rangle \approx \sum_I c_0(I) |\Phi(I)\rangle + \sum I \sum_{i,a} c^a_i(I) |\Phi^a_{ij}(I)\rangle
+ \sum I \sum_{i,j,ab} c^{ab}_{ij}(I) |\Phi^{ab}_{ij}(I)\rangle.
$$

Multireference approximations including triples and higher excitations can be defined analogously.

One drawback of multireference CI calculations is that they are more difficult to set up and describe compactly because of the need to identify the near-degenerate set of determinants $\Phi(I)$. A much simpler task is to specify only a set of near-degenerate orbitals, and to assume that all determinants obtained by considering different occupancies of the near-degenerate orbitals (with the remaining orbitals held empty or doubly occupied) constitute a near-degenerate set of references. This is the basis of the complete active space (CAS) specification of the references $\Phi(I)$. A CAS($n,m)$ set of references is obtained by identifying $m$ near-degenerate orbitals, and constructing the determinants consisting of all distributions of $n$ particles across the $m$ orbitals. Once the CAS space is constructed, we can define a CASCI as above. For example, the CASCISD approximation consists of expanding $\Psi$ using a space of singles and doubles excitations out of the CAS($n,m)$ space as in Eq. (8) (see Fig. 2), and higher analogs are similarly defined. In this work, we will exclusively use CAS spaces when defining our multireference CI calculations.

### A. Improved orbitals

It is clear that the accuracy of a CI approximation is dependent on the orbitals used to specify the determinants. For small $U$, the HF orbital basis performs well. However, this is not always the best choice when $U$ is large. The orbital basis which leads to the most rapid convergence of a complete CI expansion, as measured by one-particle density matrix norm, is called the *natural orbital* basis. Natural orbitals are eigenfunctions of the one-particle density matrix $D_{ij} = \langle \Psi | a_i^\dagger a_j | \Psi \rangle$, where $\Psi$ is the exact or CI wave function; the corresponding eigenvalues are the natural occupancies. The natural orbital basis is a commonly used basis for CI calculations. Of course, the natural orbitals themselves are defined using $\Psi$, which is not known until the CI calculation is performed. Consequently, a natural-orbital-based CI calculation is usually carried out in two steps. First, a CI calculation in the Hartree-Fock basis is performed to obtain the density matrix. This is then diagonalized to obtain the natural orbitals, and the CI calculation is repeated in this basis. In principle, this procedure can be iterated, although we have not done so in the calculations in this work. These procedures are standard in modern quantum chemistry and are commonly used to treat finite molecular systems. Detailed descriptions can be found in Refs. 18 and 22.
III. IMPLEMENTATION

We have implemented the CI-based approximations described above, as well as ED, within the context of single-site and cluster-based DMFT solvers. Our code uses a modified version of the efficient string-based CI algorithm in the DALTON quantum chemistry package. The solution of the CI and ED eigenvalue problems is carried out using iterative Davidson diagonalization, while the determination of the Green’s function in Eq. (5) is carried out using the Lanczos algorithm. For the hybridization and bath fitting necessary in the DMFT context, we have employed the procedures described in Ref. 20. The DMFT self-consistency was carried out until convergence in the self-energy was reached with a tolerance of 5%. All calculations were performed at $T = 0$, and all energies are in units of $t = 1$ unless otherwise specified. The $\beta$ used for fitting the dynamical mean-field parameters were $20/t$ (single-site DMFT) and $12.5/t$ (plaquette).

IV. RESULTS

We now assess the performance of CI approximations as quantum impurity solvers using established benchmark problems. To recapitulate, the two central questions are as follows: How rapidly do the CI approximations converge to ED, for example, as a function of excitation level or orbital basis, and, do CI approximations allow us to accurately treat a larger number of orbitals than ED? We found the impurity models occurring in the dynamical mean-field context to be more difficult to solve than simple impurity sites coupled to an analytically constructed density of states, and we therefore focus our presentation on impurity models obtained within this context. We first study an impurity model without self-consistency imposed. We then examine two DMFT models. The first is the single-site DMFT approximation to the 1D Hubbard model. Here, ED calculations can be converged with respect to the number of bath orbitals, which allows us to compare ED and CI approximations in the limit of a converged bath representation. Our second model is a 2×2 plaquette (four-site) cellular dynamical mean-field calculation for the 2D Hubbard model. Such four-site cluster models have been extensively studied with ED (Refs. 26–31) as well as with CT-QMC, and provide a standard calibration point. We begin by comparing ED and CI approximations using an ED parametrization with eight bath orbitals. Next, we demonstrate the ability of CI methods to treat large numbers of orbitals by converging the four-site cluster model with respect to the number of bath orbitals in the parametrization. Our largest calculation involves 28 orbitals, significantly larger than can be treated with ED. In the Appendix, we present a three-orbital single-site DMFT calculation of a model relevant for the physics of the $t_{2g}$ bands in transition-metal oxides. This model uses a Slater-Kanamori form of the impurity interaction. We show that CI approximations can be used with a general impurity Hamiltonian with non-density-density interactions and demonstrate convergence of the bath representation with up to 24 orbitals.

A. Anderson impurity model

As a test case for a quantum impurity model, we present in Table I the parametrization for typical hybridization strengths and energy-level parameters as they arise in the DMFT context, for $U/t = 4$ at half-filling. The Hamiltonian of this impurity

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\epsilon_i$</th>
<th>$V_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.55882</td>
<td>0.55326</td>
</tr>
<tr>
<td>2</td>
<td>-0.55882</td>
<td>0.55326</td>
</tr>
<tr>
<td>3</td>
<td>4.45759</td>
<td>0.54136</td>
</tr>
<tr>
<td>4</td>
<td>-4.45759</td>
<td>0.54136</td>
</tr>
<tr>
<td>5</td>
<td>-1.47891</td>
<td>0.48852</td>
</tr>
<tr>
<td>6</td>
<td>1.47891</td>
<td>0.48852</td>
</tr>
<tr>
<td>7</td>
<td>-0.18540</td>
<td>0.38319</td>
</tr>
<tr>
<td>8</td>
<td>0.18540</td>
<td>0.38319</td>
</tr>
<tr>
<td>9</td>
<td>0.03177</td>
<td>0.23349</td>
</tr>
<tr>
<td>10</td>
<td>-0.03177</td>
<td>0.23349</td>
</tr>
<tr>
<td>11</td>
<td>0.00000</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

Fig. 3. (Color online) Spectral function $-\frac{1}{\pi} \text{Im} G(\omega)$ (upper panel) and the imaginary part of the self-energy $\text{Im} \Sigma(\epsilon_n)$ (lower panel) for an impurity model using the bath parametrization in Table I. Solid lines (red online): ED. Light dashed line (green online): CISD. Dark dashed line (blue online): CISDT. Double-dotted line (black online): CAS(2,2)CISD. Dotted line (magenta online): CAS(2,2)CISDT.
model is

\[ H = U \left( n_\uparrow n_\downarrow - \frac{n_\uparrow + n_\downarrow}{2} \right) + \sum_{i\sigma} \epsilon_i c_i^{\dagger}\sigma c_{i\sigma} \]
\[ + \sum_{i\sigma} V_{i\sigma} c_i^{\dagger}\sigma d_{i\sigma} + \text{H.c.} \]  

(8)

The impurity model has 1 impurity site and 11 bath sites. In the particle-hole-symmetric case, the choice of the active orbitals is motivated by the energetic degeneracy of the eigenvalues present already in the noninteracting Hamiltonian. The active orbitals for the CAS calculation are the orbitals 6 and 7 of the natural orbitals displayed (as obtained in ED) in Table II, which are singly occupied. We first remark on the sizes of the CI determinant spaces and the corresponding runtimes, which are given in Table III. We see that all the CI approximations involve only a small fraction of the full ED determinantal space and take a much shorter amount of time to run. All of these calculations are doable within minutes on a desktop PC.

Figure 3 shows results for the spectral function (upper panel) and the imaginary part of the self-energy (lower panel) for the methods of Table III. All methods recover both the high- and the low-energy parts of the self-energy to high accuracy. Differences in the spectral function are visible for \( \omega > 2 \), where higher excitations that are not contained within the approximations become important. Note that we intentionally use only a small imaginary broadening so as to preserve as much structure as possible and emphasize the difference between different approximations. This is why the spectral functions do not appear smooth.

Figure 3 shows that for simple Anderson impurity models, the truncated CI expansions are extremely robust, and even low excitation levels can recover the proper self-energy. In the

<table>
<thead>
<tr>
<th>Orbital number</th>
<th>1–4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9–12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U/t = 4 )</td>
<td>2.00</td>
<td>1.895</td>
<td>1.020</td>
<td>0.980</td>
<td>0.105</td>
<td>0.000</td>
</tr>
<tr>
<td>( U/t = 6 )</td>
<td>2.00</td>
<td>1.738</td>
<td>1.009</td>
<td>0.991</td>
<td>0.262</td>
<td>0.000</td>
</tr>
<tr>
<td>( U/t = 8 )</td>
<td>2.00</td>
<td>1.502</td>
<td>1.001</td>
<td>0.998</td>
<td>0.498</td>
<td>0.000</td>
</tr>
<tr>
<td>( U/t = 20 )</td>
<td>2.00</td>
<td>2.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

TABLE II. Orbital occupancies in the natural orbital basis, for the impurity model of Table I. The choice of the active space is motivated by the partially occupied natural orbitals.

FIG. 4. (Color online) Single-site DMFT approximation to the 1D Hubbard model at half-filling, using 11 bath orbitals: spectral function (DOS) \( A(\omega) = -\frac{1}{\pi} \text{Im} \tilde{G}(\omega) \). Solid lines (red online): ED. Light dashed line (green online): CISD. Dark dashed line (blue online): CISDT. Double dotted line (black online): CAS(2,2)CISD. Dotted line (magenta online): CAS(2,2)CISDT. (a) \( U/t = 4 \), (b) \( U/t = 6 \), (c) \( U/t = 8 \), (d) \( U/t = 20 \). Note that panel (a) is similar (but not completely identical) to Fig. 3, where all methods used the converged ED parameters of Table I for solving the impurity Hamiltonian.
TABLE III. Size of determinant space for 12 electrons and 12 orbitals, and runtimes in the solution of $\Psi$, using various CI approximations and ED. $t_{GS}/t_{EDGS}$ ($t_{GF}/t_{EDGF}$): runtime of ground-state (Green’s function) calculation with respect to ED ground-state (Green’s function) calculation.

<table>
<thead>
<tr>
<th>Method</th>
<th>Space size</th>
<th>$t_{GS}/t_{EDGS}$</th>
<th>$t_{GF}/t_{EDGF}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CISD</td>
<td>1819</td>
<td>0.0026263</td>
<td>0.019265</td>
</tr>
<tr>
<td>CISDT</td>
<td>18819</td>
<td>0.012848</td>
<td>0.057334</td>
</tr>
<tr>
<td>CAS(2,2)CISD</td>
<td>6044</td>
<td>0.012242</td>
<td>0.035215</td>
</tr>
<tr>
<td>CAS(2,2)CISDT</td>
<td>49644</td>
<td>0.12739</td>
<td>0.13240</td>
</tr>
<tr>
<td>ED</td>
<td>853776</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Dynamical mean-field context, an additional complication arises: the self-consistency condition and the bath-fitting procedure lead to an amplification of differences that make the final result more sensitive to differences in the impurity self-energy.

B. Single-site DMFT for the 1D Hubbard model

1. Particle-hole-symmetric case

We carried out single-site DMFT calculations for the 1D Hubbard model using an 11-orbital bath parametrization (12 orbitals in total). We used CISD, CISDT, CAS(2,2)CISD, CAS(2,2)CISDT approximations as well as ED to obtain the spectral functions and impurity self-energies for $U/t = 4, 6, 8, 20$. All calculations are performed in the natural orbital basis, as described in Sec. II A. Shown is the converged, self-consistent dynamical mean-field solution.

The spectral functions at half-filling are shown in Fig. 4. We observe good qualitative agreement of all CI methods with ED for all values of $U/t$. If we include triple excitations [CISDT or CAS(2,2)CISDT], the CI spectral functions become indistinguishable to the eye from ED. In the case of CISDT, this is achieved using only about 2% of the complete determinant space of ED. Perhaps surprisingly, multireference CI approximations are not necessary to obtain good agreement even for large $U/t$ where $\Psi$ contains large weights from determinants other than the Hartree-Fock determinant. This reflects the simplicity of the 1D Hubbard model: the two main determinantal contributions to $\Psi$ at large $U$ differ in the occupancies of only two electrons, which can be adequately described using doubles excitations. It also reflects the nonperturbative nature of CI: so long as the determinants of interest are within the CI space, they can assume arbitrarily large weights in $\Psi$, and strongly interacting (large-$U$) systems can be treated.

The corresponding self-energies of the various approximations are shown in Fig. 5 (we plot only the imaginary part

![Fig. 5](image-url)
Im Σ(iω_n)]. Again, good qualitative agreement between all the CI methods and ED is observed for all values of U/t. Indeed, for U/t = 4, 6, 20, even the simplest CI approximation (CISD) yields an essentially indistinguishable self-energy from ED. Only at U/t = 8 [Fig. 5(c)] do we see appreciable differences. Here, we need to use CAS(2,2)CISDT to achieve less than 1% error in the self-energy. Of course, CAS(2,2)CISDT is also the most accurate approximation to ED as measured by the size of the excitation space.

As discussed in Sec. II A, the accuracy of the CI expansions can be improved by working in the natural orbital basis. Examining the ED calculations at half-filling we find that across the range of different U/t, only four natural orbitals have occupancies appreciably different from 0 and 2. Consequently, we choose these four active orbitals for an active space calculation in the natural orbital basis. In Fig. 6, we show the spectral functions at half-filling using the CAS(4,4) approximation, in the natural orbital basis of the ED calculation. Note that the CAS(4,4) wave function involves only 16 determinants, but the spectral functions are still remarkably similar to the ED spectral functions. In fact, they are of similar quality to the CAS(2,2)CISD spectral functions (also in the natural orbital basis). This demonstrates the compactness of the natural orbital description.

2. Away from particle-hole symmetry

We next consider the 1D Hubbard model away from half-filling. The corresponding imaginary parts of the self-energies, for U/t = 6 and dopings of 5%–30%, are shown in Fig. 7 for CISD, CISDT, and CAS(2,2)CISD. To better illustrate the differences between the methods, here we plot the percentage error in the imaginary part of the self-energies, relative to ED. While all the CI approximations yield qualitatively reasonable self-energies, we see that when we include triple excitations, the errors become significantly less than 1%. This is consistent with our expectation that away from half-filling, the wave function of this model becomes more single determinantal and therefore it is more advantageous to base the description on a single-reference determinant (in this case, the HF determinant) than to include multiple determinant reference wave functions as in CAS(2,2)CISD.

C. Four-site cellular DMFT approximation to the 2D Hubbard model

We now turn to cluster dynamical mean-field theory, and in particular the 2×2 cellular DMFT approximation of the 2D Hubbard model. We begin with a 12-orbital quantum impurity model using 8 bath orbitals, corresponding to 2 bath...
FIG. 7. (Color online) Single-site DMFT approximation to the 1D Hubbard model away from half-filling, using 11 bath orbitals: percent error in the imaginary part of the self-energy $\text{Im}\Sigma(i\omega_n)$ (relative to ED) for CISD (solid lines, red online), CISDT (dashed lines, green online), and CAS(2,2)CISD (dotted lines). (a) 5% doping, (b) 10% doping, (c) 15% doping, (d) 30% doping. $\omega_n = (2n + 1)\pi/\beta$, $\beta = 20$.

orbitals per impurity site, a model that has been extensively studied in previous ED calculations. In common with these studies, we use the fourfold symmetry of the $2 \times 2$ cluster and calculate the Green’s function and self-energies in the symmetry-adapted basis of the cluster. In this basis, the Green’s functions and self-energies become diagonal. Labeling the sites of the $2 \times 2$ cluster as $1 \equiv (0,0), 2 \equiv (1,0), 3 \equiv (0,1), 4 \equiv (1,1)$, the symmetry orbitals $\Gamma, M, X$ (doubly degenerate) are given by

$$\psi_{\Gamma} = \frac{1}{2} (\phi_1 + \phi_2 + \phi_3 + \phi_4),$$

$$\phi_M = \frac{1}{2} (\phi_1 - \phi_2 - \phi_3 + \phi_4),$$

$$\phi_X = \frac{1}{2} (\phi_1 + \phi_2 - \phi_3 - \phi_4),$$

$$\phi_X' = \frac{1}{2} (\phi_1 - \phi_2 + \phi_3 - \phi_4).$$

The symmetry orbitals $\phi_X$ and $\phi_X'$ form a degenerate pair. For a detailed description of this model see, e.g., Refs. 30 and 31.

We carried out CISD, CISDT, CAS(2,2)CISD, CAS(2,2)CISDT, CAS(2,2)CISDTQ, and ED calculations of the spectral functions and self-energies at half-filling. The CI calculations were carried out in the natural orbital basis of a CAS(2,2)CISD calculation in the Hartree-Fock basis.

The local spectral functions $-\frac{1}{4\pi} \text{Tr} \text{Im} G(\omega)$ are shown in Fig. 8. The imaginary parts of the $X$ self-energy, $\text{Im}\Sigma_X(\omega)$, corresponding to Eq. (11), are shown in Fig. 9. Similar to the 1D case, we find good agreement between all the CI methods and ED for all studied values of $U/t$, although there are some visible differences between CAS(2,2)CISD and ED. Once triple and higher excitations are included, however, the spectral functions become indistinguishable to the eye. The same conclusion can be drawn from analyzing the self-energies. While CAS(2,2)CISD is qualitatively similar to ED, the self-energy for $U/t = 4$ is shifted from the ED self-energy, with the errors largest at small frequencies. Once triples are included, the agreement becomes much better, and with quadruples the self-energy is indistinguishable from that of ED. If we consider CAS(2,2)CISDT as yielding quantitative agreement, then this is achieved using 49 644 determinants in the CI expansion, or only about 6% of the ED determinantal space.

In this model, we find that the most difficult values of $U/t$ to achieve agreement between the CI methods and ED are for $U/t = 4$ and 5. Here, the form of the self-energy is that of a correlated Fermi liquid with a large effective mass. This behavior appears in the vicinity of the first-order cluster DMFT metal-insulator transition which, in CT-QMC simulations, is
FIG. 8. (Color online) Cellular dynamical mean-field approximation to the 2D Hubbard model at half-filling on a $2 \times 2$ cluster using eight bath orbitals: local spectral function \( A(\omega) = -\frac{1}{\pi} \text{Tr} \text{Im} G(\omega) \). For a description of the methods, see text. (a) \( U/t = 4 \), (b) \( U/t = 5 \), (c) \( U/t = 6 \), (d) \( U/t = 8 \).

near \( U/t = 5.4, 34, 37 \) (Note that in this pseudogap region, ED calculations can actually converge to two different correlated metallic solutions, depending on the initial guess for the bath parametrization, a feature which is repeated in the CI calculations. We have chosen to present the more insulating solution in Fig. 9.)

We now briefly turn to some calculations on this model which can not be performed using ED. An essential weakness of ED (and CI) solvers in the DMFT context is the need to parametrize the bath using a finite number of bath orbitals. If the number of bath orbitals is too small, the resolution of the spectral function and other quantities is very low, and furthermore, artifacts can appear in the ED calculations due to a large fitting error at low frequencies. \( 31, 38-40 \) CI approximations, however, allow us to treat larger numbers of orbitals, and thus potentially alleviate the bath parametrization problem by allowing us to use a sufficient number of bath orbitals. We now demonstrate this for the $2 \times 2$ cluster. In Fig. 10, we plot the self-energies for $U/t = 8$ from CAS(2,2)CISD calculations [using the CAS(2,2)CISD natural orbital basis] and for 8, 16, and 24 bath orbitals. For 12 bath orbitals, we used CAS(4,4)CISD rather than CAS(2,2)CISD for technical reasons due to the degeneracy of the reference wave function. The largest calculation with 24 bath orbitals (or a total of 28 orbitals in the impurity model) is roughly twice the size of what can be treated with ED. Our studies confirm that convergence in this model is achieved fairly rapidly, but that there are nonetheless quantitative differences between the standard 8-bath orbital parametrization and larger bath representations, particularly for small frequencies. The 16-bath orbital and 24-bath orbital parametrizations are indistinguishable, indicating that full convergence has been reached. We have also carried out calculations for other values of $U/t$, where we observe similar convergence behavior. The convergence of the bath parametrization appears slower for $U/t = 5$ and 6, which may once again be related to the proximity to a metal-insulator transition.

V. CONCLUSIONS

In this work, we have described how configuration interaction (CI) approximations to exact diagonalization (ED) can be used as solvers for quantum impurity models, such as those encountered in dynamical mean-field theory (DMFT). CI solvers form a controlled hierarchy of polynomial cost approximations that retain the main advantages of ED, such
as the ability to treat general interactions and obtain real spectral information. As we have demonstrated in this work, the convergence of the CI hierarchy is sufficiently rapid that in many cases, they almost exactly approximate the ED results, at a small fraction of the cost. This is true even in “difficult,” “strongly correlated” regimes, such as the pseudogap regime of the $2 \times 2$ cluster DMFT of the Hubbard model. In addition, this great increase in computational efficiency potentially allows us to treat considerably larger quantum impurity models than have been considered in ED. In this work, we used this ability to demonstrate bath convergence for the $2 \times 2$ cluster DMFT of the Hubbard model in a calculation with 28 orbitals, for a case where previously only 12 orbitals (Lanczos) were accessible.

Here, we have focused on well-studied DMFT problems in order to benchmark the CI approximations. In future work, we plan to apply these CI approximations to study problems where existing solvers have difficulties. Some of these include impurity models with a large number of orbitals and with general interactions and off-diagonal hybridizations, for which CT-QMC methods encounter a severe sign problem. Another interesting direction to explore will be to examine more sophisticated quantum chemistry approximations to ED. For example, for weak interactions, coupled-cluster approximations are known to be far superior to configuration interaction approximations for a given computational cost. These and other directions are currently being pursued.
ACKNOWLEDGMENTS

D. Zgid acknowledges helpful discussions with A. J. Millis, D. R. Reichman, A. I. Lichtenstein, L. de Medici, and A. Liebsch. D. Zgid and G. K.-L. Chan acknowledge support from the Department of Energy, Office of Science, through Award No. DE-FG02-07ER46432. E. Gull was partially supported by NSF-DMR-1006282.

APPENDIX: THREE-ORBITAL MODEL WITH ROTATIONALLY INVARIANT INTERACTIONS

As a further application, we present here results for a three-orbital model with general, rotationally invariant interactions. Problems of this type have been notoriously difficult to solve, or suffer from a severe fermionic sign problem, even at half-filling. So far, only the continuous-time hybridization expansion and exact diagonalization methods have been able to access this regime, but the exponential scaling of the local impurity Hilbert space size makes five- and seven-orbital systems inaccessible without severe truncations or fitting errors. The three-orbital model with the Slater-Kanamori form

$$H_{\text{loc}} = H_{\text{SK}} = U \sum_a n_{a\uparrow} n_{a\downarrow} + (U - 2J) \sum_{a\neq b} n_{a\uparrow} n_{b\downarrow} + (U - 3J) \sum_{a>b} n_{aa} n_{bb} - J \sum_{a\neq b} (d_{a\uparrow}^\dagger d_{a\downarrow}^\dagger d_{b\downarrow} d_{b\uparrow} + d_{a\downarrow}^\dagger d_{a\uparrow}^\dagger d_{b\uparrow} d_{b\downarrow}) \quad (A1)$$

The imaginary part of the self-energy at half-filling (in the Mott insulating phase) for Fig. 11 is a toy model that has been well studied on a Bethe lattice and we find no additional complications caused by the more general interaction structure.