Next Generation Dynamical Mean-Field Theory Simulations with the Adaptive Sampling Configuration Interaction Method

Carlos Mejuto-Zaera, Norm M. Tubman*, K. Birgitta Whaley
*Corresponding author: ntubman@berkeley.edu

University of California, Berkeley, California 94720, United States

(Dated: September 5, 2018)

In the pursuit of accurate descriptions of strongly correlated quantum many-body systems, dynamical mean-field theory (DMFT) has been an invaluable tool for elucidating the spectral properties and quantum phases of both phenomenological models and \textit{ab initio} descriptions of real materials. Key to the DMFT process is the self-consistent map of the original system into an Anderson impurity model, the ground state of which is computed using an impurity solver. The power of the method is thus limited by the complexity of the impurity model the solver can handle. Simulating realistic systems generally requires many correlated sites. By adapting the recently proposed adaptive sampling configuration interaction (ASCI) method as an impurity solver, we enable much more efficient zero temperature DMFT simulations. The key feature of the ASCI method is that it selects only the most relevant Hilbert space degrees of freedom to describe the ground state. This reduces the numerical complexity of the calculation, which will allow us to pursue future DMFT simulations with more correlated impurity sites than in previous works. Here we present the ASCI-DMFT method and applications to the one-dimensional and two-dimensional Hubbard models that exemplify its efficient convergence and timing properties. We show that the ASCI approach is several orders of magnitude faster than the current best published ground state DMFT simulations, which allows us to study the bath discretization error in simulations with small clusters, as well as to address cluster sizes beyond the current state of the art. Our approach can also be adapted for other embedding methods such as density matrix embedding theory and self-energy embedding theory.

The systematic study of the properties of strongly correlated many-electron systems remains one of the main areas of research in condensed matter physics. In this regard, dynamical mean-field theory (DMFT) has been successfully applied to study metal-insulator phase transitions [1, 2], exotic quantum phases of matter [3, 4], critical exponents in quantum field theories [5], the volume expansion in Plutonium [6] and high temperature superconductivity [7, 8] among other. DMFT works by mapping the many body system of interest self-consistently into an Anderson impurity Hamiltonian [9]. A central feature of the method is the impurity solver, which finds the ground state of the impurity model. Numerically exact solvers based on Monte Carlo approaches [10, 12] can be used in the finite temperature case, but these have difficulties converging to the $T = 0$ regime.

One of the greatest challenges to expand the applicability of zero temperature DMFT is to find an efficient representation of the ground state wave function of the impurity system [13, 22]. Simulations with configuration interaction (CI) [23] based approaches have been used in zero temperature studies. These approaches attempt to identify a subspace in which to find the ground state and have had some success in treating strongly correlated systems [24, 27]. One of these methods, CI singles and doubles (CISD), has been used in an attempt to increase the size of the systems that can be simulated [19, 20]. Additionally an iterative CISD has been considered as a DMFT solver [21, 22], which has also been successful, but somewhat computationally expensive. This iterative CISD approach developed recently, which is called adaptive configuration interaction and is unrelated to the method used in this work [21], misses some of the key features that are important for an efficient DMFT impurity solver as will be described below.

The CI methods currently used in the DMFT literature are not representative of modern CI techniques [28–33]. Selected CI (SCI) methods have recently been shown to be much more efficient than previous CI methods. Recently, the adaptive sampling CI method (ASCI) was introduced as a modern approach to SCI, and since then the ASCI method and other approximate SCI methods have been rapidly developing further [34–39]. The key idea that allows this method to be more efficient than traditional CI methods is to remove the active space and instead to identify the most relevant degrees of freedom in Hilbert space to describe the ground state. The ASCI method has been shown to successfully and efficiently treat strongly correlated electronic systems known for their difficulty, for example the Cr$_2$ dimer [33, 40].

In this work, we implement the ASCI method as the impurity solver for cluster DMFT (CDMFT) simulations. CDMFT, originally proposed for studying lattice models [41–44], works by self consistently finding bath parameters for a set of sites in a sublattice (the cluster). Correlation effects are taken into account [9, 44], allowing for quantum fluctuations between the cluster sites and the rest of the system. This is done by mapping the original system self-consistently into an Anderson impurity model. The parameters of this map are the number of cluster sites $N_c$, the number of bath sites $N_b$, the bath energies and the coupling terms between the bath and cluster sites.

The self-consistent map begins by a choice of the bath
parameters. Given the bath parameters, the ground state of the impurity Hamiltonian is computed using the impurity solver. The Green’s function can then be computed from the wave function $|\Psi\rangle$ and from that one can access the self-energy $\Sigma(\omega)$. These Green’s function and self-energy are local quantities defined only on the cluster. The full lattice Green’s function can be computed from these local quantities. The self-consistent condition then amounts to equating the Green’s function of the Anderson model to the full lattice Green’s function, and this is solved by iteratively fitting the bath parameters. For more details and specific prescriptions on the CDMFT calculation, consult ref [46] and [47].

To proceed, we present ASCI as an impurity solver for the cluster Hamiltonian $H$. There has been a lot of interest in developing CI methods to treat DMFT impurity systems, especially recently [19,22]. CI methods work as an impurity solver by diagonalizing a Hamiltonian in a basis of many-fermion states (determinants). However, traditional CI methods that have been previously considered for this purpose are not the most effective to treat strongly correlated systems since they are missing several important aspects that are central to SCI. A key feature of SCI methods is to identify the most relevant determinants needed to describe the ground state wave function. In particular, ASCI does so by ranking the determinants according to their coefficient in a trial ground state wave function and the Hamiltonian matrix elements [34,40]. The method proceeds iteratively, improving the subspace onto which $H$ is projected. This subspace is referred to as the target space and characterized by the number of determinants included, $\text{tdets}$.

ASCI starts with a guess for the target space, denoted as $\{\text{tdets}\}$, e.g., the Hartree Fock determinant plus some set of low rank excitations (singles, double, triples,...). The ground state energy and wave function of $H$ are then computed in the space $\{\text{tdets}\}$ (e.g., by Lanczos). After diagonalization, the wave function is defined by its expansion coefficients $C_j$. The method then proceeds to update the target space by choosing a new set of determinants (a set of size $\text{tdets}$) that better describes the ground state. This update is done by searching all the singly and doubly excited determinants from a subset of $\{\text{tdets}\}$, which we denote as $\{\text{dsearch}\}$. The size of $\{\text{dsearch}\}$ is a parameter that we choose, $\text{edets}$, and its influence on the simulation is described in detail in ref [33,40].

The set $\{\text{dsearch}\}$ contains the determinants corresponding to the largest coefficients $|C_j|$ from the ground state wave function. We denote all determinants found in the search as the set $\{\text{SD}\}$. This set can have many orders of magnitude more elements than $\{\text{tdets}\}$.

After the search, we calculate

$$A_i = \frac{\sum_j H_{i,j} C_j}{H_{i,i} - E_0},$$

for all determinants in $\{\text{SD}\}$, which provides an estimate of their importance in the ground state wave function. The prime in the sum indicates a sum over $\{\text{tdets}\}$, $H_{i,j} = \langle i | H | j \rangle$, and $E_0$ is the current best estimate of the ground state energy.

A new target space is built by ranking elements of the old target space together with the new singles and doubles, according to the absolute value of their coefficients $C_j$ and $A_i$, respectively, and selecting the $\text{tdets}$ determinants with the largest coefficients. $H$ can then be generated in the new target space and its ground state computed. This process is then repeated until convergence.

This method is advantageous for systems in which the ground state can be described with enough accuracy using a small subset of the total Hilbert space. The required accuracy is application-dependent, but previous work has shown that the ASCI method can treat strongly correlated systems generally accepted as difficult, with higher accuracy and less resources [33,10,48]. The speed of convergence of the ASCI method can be greatly influenced by the correct choice of $\text{tdets}$, with larger $\text{tdets}$ yielding higher accuracy but requiring a longer time for each iteration. When dealing with a new system, one begins with a modest size of the target space and ramps this up until the ground state energy is converged to the desired precision. For a more in depth discussion of the ASCI algorithm and its other parameters, see reference [33,40].

In this Letter, we adapt the ASCI algorithm to provide an impurity solver for CDMFT simulations. For this we need to perform an additional step in order to calculate the Green’s function efficiently, namely to compute a basis for the Hilbert spaces of single particle and hole excitations on top of the ground state target space $\{\text{tdets}\}$. Naively, one would construct these spaces simply by applying the corresponding creation (annihilation) operators on the converged $\{\text{tdets}\}$, in order to be able to represent single particle (hole) excitations. In this basis, we denote the coefficients of the particle (hole) excitation on top of the ground state as $C_i^{N \pm 1}$. This basis would be enough to represent the impurity Hamiltonian in those spaces to the accuracy of the ASCI wave function. However, to compute the impurity Green’s function we need to invert the Hamiltonian (the explicit equations and details of the algorithm are discussed in [47]). To represent the inverse of the Hamiltonian to the accuracy of the wave function, we need more states because by inverting we shuffle all matrix elements communicating with the target space $\{\text{tdets}\}$. The solution to this problem is to add states connected to the single particle (hole) space that cannot be accessed by acting with the creation (annihilation) operator on $\{\text{tdets}\}$. The coefficients of these states in the expansion of the corresponding single particle (hole) states are thus exactly zero. In the spirit of ASCI, we need to search in the space of these zero states to find those relevant for describing the Green’s function. In order to do this, we introduce a ranking to decide which are the most important zero states to get an accurate Green’s function. For each zero state, we
Here the primed sum means that we only add over the states in the naive basis of single particle (hole) excitations from \( \{ D_{t,\text{det}} \} \), \( H_{ij,\pm 1} \) are the Hamiltonian matrix elements in the single particle (hole) space, and \( A_{ij,\pm 1} \) are the coefficients of \(|x\rangle = H^{-1} c^\dagger_{i,\sigma} |GS\rangle \) \((|x\rangle = H^{-1} c_{i,\bar{\sigma}} |GS\rangle)\), \(|GS\rangle\) being the ASCI ground state, in the same basis. If all zero states have been included, then the right hand side of Eq. \( (2) \) should be exactly zero. The absolute value of the right hand side is thus a ranking for the zero states. This approximation introduces a new parameter in the simulation: the number of zero states to add for the Green’s function calculation. For the calculations shown here, we include up to one million zero states, which is always at least one order of magnitude larger than the size of \( \{ D_{t,\text{det}} \} \). This includes all the zero states for the simulations with a 2x2 cluster.

To demonstrate the efficiency of the ASCI method as an impurity solver in CDMFT calculations, we consider here the one-dimensional (1d) and two-dimensional (2d) square lattice Hubbard models. The Hubbard model is characterized by the Hamiltonian

\[
H_{\text{Hub}} = -t \sum_{\langle i,j \rangle, \sigma} (c_{i,\sigma}^\dagger c_{j,\sigma} + h.c.) - \mu \sum_i(n_{i,\uparrow} + n_{i,\downarrow}) + U \sum_i n_{i,\uparrow} n_{i,\downarrow},
\]

with hopping amplitude \( t \), chemical potential \( \mu \), Coulomb interaction strength \( U \) and spin label \( \sigma = \{\uparrow, \downarrow\} \). At half filling we have \( \mu = U/2 \), for other particle fillings one would need to determine the chemical potential and the number of electrons in the impurity model self-consistently according to \[19\]. By optimizing the target space in the ASCI method, we can show that we reproduce the results in the literature with drastically reduced computational resources.

In the supporting information \[17\] we present detailed convergence benchmarks in the case of the 1d Hubbard model. Here we just summarize the main results: A small fraction of the total Hilbert space is indeed enough to obtain accurate representations for the self energy along the imaginary and real axis. In the case of a 1d Hubbard model calculation with \( N_c = 1 \) and \( N_b = 11 \) one only needs about 10000 states out of the total 853776. For \( N_c = 4 \) and \( N_b = 8 \) we reach similarly good results with the same number of states. We compare our results to equivalent zero temperature CI based DMFT algorithms \[21\]. The timings of the ASCI-DMFT algorithm are very favorable, with 30 minutes on a single core of an Intel Core i5 CPU (2.9 GHz) being enough to converge the calculations described above, and a 1d calculation with \( N_c = 8 \) and \( N_b = 16 \) taking 5 hours on the same system.

Having established the timing and convergence properties of the ASCI-DMFT algorithm in the 1d Hubbard model example, we now show its viability for study of the two-dimensional square lattice Hubbard model. Here, we choose \( U/t = 8 \) and different electron fillings. When away from half-filling, one needs to undertake a self-consistent determination of the chemical potential and number of electrons that (a) minimize the energy, and (b) represent the desired lattice filling \[19\]. Using current CI based DMFT methods for this self-consistent calculation is excessively expensive in time, and most benchmarking has therefore been done at half-filling. To show that ASCI-DMFT allows the self-consistent treatment away from half-filling, we report here "single point" calculations with 2x2, 3x3 and 4x4 clusters, where "single point" means that we chose a specific number of electrons and value of chemical potential, corresponding to one step in the self-consistent determination of these. We also present half-filling results.

In the case of the small 2x2 cluster, the compact wave function representation of the ASCI impurity solvers allows us to study the effect of the bath discretization error by performing simulations with \( N_b = 8 \) and 16 bath sites. The spectral weights for these simulations are presented in Figs. \[4\] and \[5\]. The \( N_b = 8 \) calculations can be done in 1 hour on a single core and show excellent agreement with previous literature \[21\], while the \( N_b = 16 \) calculations required 48 hours.

In the \( N_b = 8 \) calculation, the system at half-filling is insulating, while away from half filling it becomes metallic, as can be seen from the hole bands that cross the chemical potential. The role of bath discretization error is illustrated by comparison of this with the results for \( N_b = 16 \) in Fig. \[2\]. We see that the spectral weights change significantly from those obtained with \( N_b = 8 \). In general, the peaks are now less sharp and the spectral weights are more smoothly distributed in frequency space at each k-point. Physically, this indicates that the hole quasi-particles are less well defined than the chemical potential, corresponding to one step in the self-consistent determination of these. We also present half-filling results.

The spectral weights change significantly from those obtained with \( N_b = 8 \). In general, the peaks are now less sharp and the spectral weights are more smoothly distributed in frequency space at each k-point. Physically, this indicates that the hole quasi-particles are less well defined than the chemical potential, corresponding to one step in the self-consistent determination of these. We also present half-filling results.

Using ASCI as an impurity solver also allows us to make first calculations with cluster sizes larger than the current state of the art for CI based DMFT methods. Fig. \[6\] presents spectral weights for single point calculations at three different fillings for simulations with a
FIG. 1. Spectral weights for the two-dimensional square lattice Hubbard model in a calculation with $U/t = 8$, $N_c = (2 \times 2)$ and $N_b = 8$ for different electron fillings. The percentage given corresponds to the number of electrons in the simulation relative to the number of cluster sites. We show a particular cut through the first Brillouin zone. The dashed line represents the chemical potential.

FIG. 2. Spectral weights for the two-dimensional square lattice Hubbard model in a calculation with $U/t = 8$, $N_c = (2 \times 2)$ and $N_b = 16$ for different electron fillings. The percentage given corresponds to the number of electrons in the simulation relative to the number of cluster sites. We show a particular cut through the first Brillouin zone. The dashed line represents the chemical potential.

3x3 cluster and $N_b = 19$, and one single point calculation with a 4x4 cluster and $N_b = 24$ [17] at 1/8 filling. The 3x3 calculations took between two and four days, and the 4x4 needed a full week on 28 cores. The computational bottleneck for these calculations is the fitting step for the bath parameters, which we implement here using the BOBYQA implementation in the nlopt library.

FIG. 3. Spectral weights for the two-dimensional square lattice Hubbard model in a calculation with $U/t = 8$, $N_c = (3 \times 3)$ and $N_b = 19$ for electron fillings $n = 50\%$, 46\% and 14\%, and $N_c = (4 \times 4)$, $N_b = 24$ and $n = 13\%$. The percentage given corresponds to the number of electrons in the simulation relative to the number of cluster sites. We show a particular cut through the first Brillouin zone. The dashed line represents the chemical potential. The black dots represent the results for a 2x2 cluster with $N_b = 8$ at the corresponding filling.

[50, 51]. Increasing the number of cluster and bath sites dramatically increases the number of fitting parameters, which makes the non-linear fitting process expensive and will need to be addressed in future refinements.

The results for 3x3 with $N_b = 19$ show interesting differences from the 2x2 results. At half-filling the system shows the insulating behavior seen for the 2x2 $N_b = 8$ calculations. However, consistent with the larger bath size as noted above, the peaks are significantly more diffuse, just as in the 2x2 $N_b = 16$ calculations. The 3x3 calculations away from half-filling share this general loss of sharpness in the spectral weights. This is also observed in the 4x4 single point calculation (bottom right panel in Fig. 3) which appears to show development of flat bands below the chemical potential. However since is a single point calculation [19] it is not necessarily indicative of the true ground state at this filling.

We have presented a CDMFT implementation using ASCI as the impurity solver, and shown that the superior efficiency of this approach allows study of both the bath discretization error in small clusters and exploration of cluster sizes beyond the current state of the art for CI based DMFT methods. The results provide strong motivation for undertaking further application of modern CI techniques to DMFT. Additional tools available with ASCI include many-body perturbation theory corrections [52] and heat-bath extensions [34], both of which can be expected to accelerate these algorithms for applications of DMFT. A new suite of algorithms for increasing the efficiency of ASCI on modern computers will im-
prove the timings presented here even further [33]. ASCI can also be readily applied to other embedding techniques such as density matrix embedding theory [33], [53].

We demonstrated the effect of the main parameter of the ASCI method in the truncation approach, the size $tdets$ of the active space. The physical properties of the system along the imaginary frequency axis, represented by the cluster self energy, converge very quickly with a modest number of determinants. The convergence on the real frequency axis requires a larger space, but is also fairly rapid, as can be seen from the local densities of states [47]. A reasonable strategy is thus to use small to moderate target space sizes for the DMFT iteration loops, which occur along the imaginary frequency axis, and then to increase the size on the real axis for computing the measurable physical properties of the system. Parameters may be further tuned during the iterations along the imaginary frequency axis by beginning with a small number of determinants, computing a few DMFT loops to bring the bath parameters into the correct range, and then increasing $tdets$ to achieve high accuracy in the fits. This is particularly important for calculations away from half-filling, in which the number of electrons and the chemical potential have to be determined in a self-consistent procedure. Thus, at the beginning of the self-consistent method one could start with a small number of determinants, which can then be increased once the desired particle filling is reached.

Application of ASCI to molecular Hamiltonians, has allowed simulation on the order of 50 electrons in 100-200 sites/orbitals [33]. For DMFT applications, we expect that future iterations of ASCI will allow calculations for 50 to 60 degrees of freedom, of which 20-30 are correlated cluster sites. This will allow the study of complicated systems, e.g., many-band Hubbard models for the study of transitions between exotic phases of matter [56] or realistic many body Hamiltonians, when our impurity solver is combined with existing ab initio methods such as GW-EDMFT [57]. This new efficient impurity solver will also be useful for benchmarking the solution of embedding Hamiltonians with hybrid quantum-classical algorithms realized on quantum computers [58], [59].

ACKNOWLEDGEMENTS

We acknowledge helpful discussions with James LeBlanc. CMZ thanks the Fundación Bancaria La Caixa for an Obra Social “La Caixa” graduate fellowship. NMT was supported through the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. 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.

[1] M. Capone, M. Civelli, S. S. Kancharla, C. Castellani, and G. Kotliar, Phys. Rev. B 69, 195105 (2004).
[2] H. Park, K. Haule, and G. Kotliar, Phys. Rev. Lett. 101, 186403 (2008).
[3] R.-Q. He and Z.-Y. Lu, Phys. Rev. B 86, 045105 (2012).
[4] J. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G. K.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, and C. A. Jimenez-Hoyos, Phys. Rev. X 5, 041041 (2015).
[5] O. Akerlund, P. de Forcrand, A. Georges, and P. Werner, Phys. Rev. D 88, 125006 (2013).
[6] G. Kotliar and E. Abrahams, Nature 410, 793 (2001).
[7] A. I. Lichtenstein and M. Katsnelson, Phys. Rev. B 62, R9283 (2000).
[8] A. Go and A. J. Millis, Phys. Rev. Lett. 114, 016402 (2015).
[9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. 68, 13 (1996).
[10] P. Werner, A. Comanac, L. deMedici, M. Troyer, and A. J. Millis, Phys. Rev. Lett. 97, 076405 (2006).
[11] K. Haule, Phys. Rev. B 75, 155113 (2007).
[12] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. 83, 349 (2011).
[13] M. Caffarel and W. Krauth, Phys. Rev. Lett. 72, 1545 (1994).
[14] R. Bulla, Phys. Rev. Lett. 83, 136 (1999).
[15] D. J. Garcia, K. Hallberg, and M. J. Rozenberg, Phys. Rev. Lett. 93, 246403 (2004).
[16] S. Nishimoto, F. Gebhard, and E. Jeckelmann, J. Phys.-Condens. Mat. 16, 7063 (2004).
[17] M. Ganahl, M. Aichhorn, H. G. Evertz, P. Thunström, K. Held, and F. Verstraete, Phys. Rev. B 92, 155132 (2015).
[18] F. A. Wolf, A. Go, I. P. McCulloch, A. J. Millis, and U. Schollwöck, Phys. Rev. X 5, 041032 (2015).
[19] D. Zgid and G. K.-L. Chan, J. Chem. Phys. 134, 094115 (2011).
[20] D. Zgid, E. Gull, and G. K.-L. Chan, Phys. Rev. B 86, 165128 (2012).
[21] A. Go and A. J. Millis, Phys. Rev. B 96, 085139 (2017).
[22] Y. Lu, M. Höpfler, O. Gunnarsson, and M. W. Haverkort, Phys. Rev. B 90, 085102 (2014).
[23] A. Szabo and N. S. Ostlund, Modern quantum chemistry: introduction to advanced electronic structure theory (Courier Corporation, 2012).
[24] C. F. Bender and E. R. Davidson, Phys. Rev. 183, 23 (1969).
[25] G. W. Bryant, Phys. Rev. Lett. 59, 1140 (1987).
[26] I. Boustani, W. Peseworf, P. Fantucci, V. Bonacic-Koutecky, and J. Koutecky, Phys. Rev. B 35, 9437 (1987).
[27] V. BonacicKoutecky, L. epiva, P. Fantucci, and J. Koutecky, J. Chem. Phys. 98, 7981 (1993).
[28] M. Caffarel, T. Appliencourt, E. Giner, and A. Scemama, J. Chem. Phys. 144, 151103 (2016).
[29] E. Giner, A. Scemama, and M. Caffarel, Can. J. Chem. 91, 879 (2013).
[30] P. J. Robinson, S. D. P. Flores, and E. Neuscamman, J. Chem. Phys. 147, 164114 (2017).
[31] J. J. Eriksen, F. Lipparrini, and J. Gauss, J. Phys. Chem. Lett. 8, 4633 (2017), pMID: 28892390.
[32] R. Chakraborty, P. Ghosh, and D. Ghosh, International Journal of Quantum Chemistry, e25509.
[33] N. M. Tubman, C. D. Freeman, D. S. Levine, D. Hait, M. Head-Gordon, and B. Whaley, ArXiv e-prints (2018), arXiv:1807.00821 [quant-ph].
[34] A. A. Holmes, N. M. Tubman, and C. Umrigar, J. Chem. Theory Comput. 12, 3674 (2016).
[35] J. B. Schriber and F. A. Evangelista, J. Chem. Phys. 144, 161106 (2016).
[36] P. M. Zimmerman, J. Chem. Phys. 146, 104102 (2017).
[37] P. M. Zimmerman, J. Chem. Phys. 146, 224104 (2017).
[38] W. Liu and M. R. Hoffmann, J. Chem. Theory Comput. 12, 1169 (2016), pMID: 26765279.
[39] J. B. Schriber and F. A. Evangelista, J. Chem. Theory Comput., null (2017) pMID: 28892621.
[40] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, J. Chem. Phys. 145, 044112 (2016).
[41] W. Metzner and D. Vollhardt, Phys. Rev. Lett. 62, 324 (1989).
[42] E. Müller-Hartmann, Z. Phys. B Con. Mat. 74, 507 (1989).
[43] E. Müller-Hartmann, Z. Phys. B Cond. Mat. 76, 211 (1989).
[44] A. Georges and G. Kotliar, Phys. Rev. B 45, 6479 (1992).
[45] G. Kotliar, S. Y. Savrasov, G. Palsson, and G. Biroli, Phys. Rev. Lett. 87, 186401 (2001).
[46] C. J. Bolech, S. S. Kancharla, and G. Kotliar, Phys. Rev. B 67, 075110 (2003).
[47] See supplemental material.
[48] S. Lehtola, N. M. Tubman, K. B. Whaley, and M. Head-Gordon, J. Chem. Phys. 147, 154105 (2017).
[49] 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), Phys. Rev. X 5, 041041 (2015).
[50] S. G. Johnson, The NLopt nonlinear-optimization package, http://ab-initio.mit.edu/nlopt.
[51] 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).
[52] N. M. Tubman, D. S. Levine, D. Hait, M. Head-Gordon, and K. B. Whaley, arXiv preprint arXiv:1808.02049v1 (2018).
[53] G. Knizia and G. K.-L. Chan, Phys. Rev. Lett. 109, 186404 (2012).
[54] E. Kananenka, A. A. Gull and D. Zgid, Phys. Rev. Lett. 91, 121111 (2015).
[55] D. Medvedeva, ArXiv e-prints (2017), arXiv:1709.09176 [cond-mat].
[56] R. T. Clay, A. W. Sandvik, and D. K. Campbell, Phys. Rev. B 59, 4665 (1999).
[57] F. Nilsson, L. Boehnke, P. Werner, and F. Aryasetiawan, Phys. Rev. Materials 1, 043803 (2017).
[58] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, and M. Troyer, Phys. Rev. X 6, 031045 (2016).
[59] N. C. Rubin, ArXiv e-prints (2016), arXiv:1610.06910 [quant-ph].
[60] E. Jones, T. Oliphant, P. Peterson, et al., “SciPy: Open source scientific tools for Python,” (2001–).
[61] A. G. and G. S. Jeon, Phys. Rev. B 84, 195102 (2011).
SUPPORTING INFORMATION

Here we present bullet point flow charts for the DMFT and ASCI algorithms. "Ground state" is shortened as GS. A brief physical motivation of the impurity model is also in order: In DMFT, quantum fluctuations outside from the cluster sites is accounted for by introducing the fermionic bath site degrees of freedom. The possibility for a particle to leave the cluster, move through the rest of the lattice and finally return to the cluster is thus represented by the couplings to the single bath sites. From this interpretation it becomes clear that an infinite number of bath degrees of freedom is formally needed to recover the thermodynamic limit behavior of the system. Since exact diagonalization or configuration interaction solver force finite baths, studying the effect of this bath discretization error becomes fundamental to evaluate the validity of DMFT calculations.

**TABLE I. DMFT method [19]**

| Input: | Number of cluster sites $N_c$, Number of bath sites $N_b$, Hamiltonian $H$. |
| Output: | Bath parameters $E_p$, $V_{p,\alpha}$ ($p \in [1, N_b], \alpha \in [1, N_c]$). |
| Algorithm: | 1. Initial guess for bath parameters. Usually this comes from a low level calculation, like Hartree-Fock. |
|           | 2. Compute GS wave function ($GS$) and energy $E_{GS}$ of current impurity model with the impurity solver. |
|           | 3. Compute cluster Green’s function $G_{c}(i\omega)$ and self energy $\Sigma_{c}(i\omega)$. These are $(N_c \times N_c)$ matrices defined as $G_{c,\alpha,\beta}(i\omega) = (GS|c_{\omega+i\mu} \leftrightarrow H + \Sigma_{c}(i\omega)) c_{\beta}(GS) + (GS|c_{\omega+i\mu} \leftrightarrow (H + \Sigma_{c}(i\omega)) c_{\alpha}(GS) and \Sigma_{c}(i\omega) = G_{0}^{-1}(i\omega) - G_{c}(i\omega)$, where $c_{\alpha}, c_{\beta}$ are the annihilation and creation operators for cluster sites $\alpha$ and $\beta$ respectively, $G_{0}(i\omega)$ is the non-interacting Green’s function of the impurity model and $\mu$ is the chemical potential. |
|           | 4. Compute local full lattice Green’s function from the cluster Green’s function by Fourier transforming $G(i\omega, R) = \frac{1}{V_{BZ}} \int_{BZ} dk \exp(i\mathbf{k} \cdot \mathbf{R})[(i\omega + \mu) - h(k) - \Sigma_{c}(i\omega)]^{-1}$. Here $h(k)$ is the Fourier transform of the non-interacting part of $H$ with respect to the unit cell defined by the cluster and $BZ$ stands for the Brillouin zone defined by the same unit cell. To recover the Green’s function on the cluster, we set $R = R_{0} \equiv 0$. |
|           | 5. Impose self-consistency $G_{c}(i\omega) = G(i\omega, R_{0})$. This means that we can express $G(i\omega, R_{0})$ with the bath parameters. |
|           | 6. Find new bath parameters by fitting $G(i\omega, R_{0})$. |
|           | 7. If bath parameters converged, finish. If not, go to step 2. |

**TABLE II. ASCI method [33, 40]**

| Input: | Size $tdets$ of the target space $\{D_{tdets}\}$, size $cdets$ of the core space $\{D_{search}\}$, Hamiltonian $H$, basis (usually localized orbital basis). |
| Output: | Optimal target space of size $tdets$, ground state energy $E_{GS}$ and wavefunction $|GS\rangle$. |
| Algorithm: | 1. Initial guess for target space. This can be a Hartree-Fock solution plus single and double excitations. |
|           | 2. Compute GS in current target space. |
|           | 3. Find connected singles and doubles to the $cdets$ most important target space states (the core space). Most important means largest coefficient in the GS wavefunction. |
|           | 4. Rank all states, target space plus the singles and doubles from the core space, according to Eq. [3]. |
|           | 5. Update the target space by choosing the top $tdets$ states of the ranking in step 4. |
|           | 6. Compute GS in the new target space. |
|           | 7. If GS energy is converged, finish. If not, go to step 3. |

**Convergence benchmark and timings**

To show the effectiveness of ASCI as an impurity solver, we perform convergence tests with respect to the $tdets$ parameter.
We first show that ASCI can indeed identify the most important states to describe the ground state wavefunction for the typical impurity models that are encountered in DMFT calculations. For that, we present the sorted absolute values of the ground state wavefunction coefficients for two systems that can be solved with exact diagonalisation (ED), both at half filling. i) A 1D Hubbard model DMFT calculation with \( N_e = 1 \) and \( N_b = 11 \), Fig. 1. ii) A 2D square lattice Hubbard model cluster DMFT calculation with \( N_e = 2 \times 2 \) and \( N_b = 8 \), Fig. 2. The ED coefficients are shown as blue dashed lines, and the coefficients computed in ASCI calculations with different \( t\text{dets} \), namely \( t\text{dets} = 250, 500, 1000 \) and 3000, are shown as red dots. As a reference, the total number of states in the full Hilbert space is 853776. Additionally, we report the estimated coefficients \( A_{ij} \) for the search set \( \{D_{SD}\} \) in Eq. 1 of the main text as orange circles. In both figures, the ability of ASCI to select the most important \( t\text{dets} \) determinants becomes completely evident, and the coefficients computed with ASCI show excellent agreement with the ED results. The only
FIG. 1. Sorted absolute values of the ground state wavefunction coefficients for the final iteration in a 1D Hubbard model DMFT calculation with $U/t = 8$, $N_c = 1$, $N_b = 11$. The determinant order is determined by the exact diagonalisation calculation, represented in all sub-figures by the blue dashed line. Each sub-figure presents the corresponding wavefunction coefficients for ASCI-DMFT calculations using different $tdets$, namely $tdets = 250, 500, 1000$ and $3000$ as red dots. The estimated coefficients as computed according to Eq. 1 are presented as orange circles.

minor discrepancies arise for the states with smallest coefficients when the target space includes more 3000 states for the 2D system. In this case, ASCI seems to have more difficulties to adapt to the abrupt decrease in the coefficients from the $10^4$-th state onwards. These difficulties arise probably from the higher degree of strong correlation in this two dimensional, cluster calculation. The estimated $A_i$ coefficients, shown as orange circles in the Figures, have a greater discrepancy with the ED results which in some cases can be of some orders of magnitude. However, these estimates follow the general shape of the ED coefficients well and allow for the efficient and accurate identification of the most relevant states. ASCI is shown thus to be able to select the most important states to describe the ground state for the kind of impurity models that arise in DMFT calculations.

Having established the convergence properties of ASCI as an impurity solver, we now show the convergence of the ASCI-DMFT algorithm with increasing $tdets$ for a 1-site calculation on the 1d Hubbard model, with $U/t = 8$ and 11 bath sites ($N_c = 1$, $N_b = 11$). We present results for five values of $tdets$: 250, 500, 1000, 3000 and 10000. At half filling with 12 electrons, half spin up and half spin down, the full Hilbert space has 853776 determinants. In this subsection we are interested in studying the convergence of the full DMFT calculation. This is affected by (a) the convergence properties of the ASCI impurity solver, and (b) how sensitive the fitting step in the DMFT routine is to slight changes in the Green’s function. We have already addressed point (a) in the previous paragraph. We now consider the effect of the fitting process. We compare two different fitting routines, namely the non-linear least squares optimization in SciPy [60] (scipy.optimize.least_squares) and the BOBYQA implementation of the nlopt library [50, 51]. In both cases, we minimize the cost function

$$
\chi = \sum_{\omega_n} \frac{1}{|\omega_n|} \sqrt{\sum_{\mu\nu} \left[ \Delta_{new}(i\omega_n) - \Delta_{baths}(i\omega_n) \right]_{\mu\nu}^2},
$$

(1)

where $N_c$ is the number of sites in the cluster, $N_\omega$ is the number of Matsubara frequencies, $\Delta_{new}(i\omega_n)$ is the new hybridization function computed in the current DMFT iteration and $\Delta_{bath,\mu\nu}(i\omega_n) = \sum_p N_b \frac{V_{p\mu} V_{p\nu}}{E_p - i\omega_n}$ is the hybridization...
FIG. 2. Sorted absolute values of the ground state wavefunction coefficients for the final iteration in a 2D Hubbard model cluster DMFT calculation with $U/t = 8$, $N_c = 2 \times 2$, $N_b = 8$. The determinant order is determined by the exact diagonalisation calculation, represented in all sub-figures by the blue dashed line. Each sub-figure presents the corresponding wavefunction coefficients for ASCI-DMFT calculations using different $tdets$, namely $tdets = 250$, 500, 1000 and 3000 as red dots. The estimated coefficients as computed according to Eq. 1 are presented as orange circles.

function computed with the bath parameters, see [19].

We have to study the convergence properties along both the imaginary and real frequency axes. For this, we report the self-energy along the imaginary frequency axis for calculations using the SciPy fit in Fig. 3 and for calculations using the nlopt fit in Fig. 4. Similar convergence plots for the density of states (DOS) along the real frequency axis are shown for calculations using the SciPy fit in Fig. 5 and for calculations using the nlopt fit in Fig. 6. We report the self-energies along the imaginary frequency axis for both the first and last iteration in the ASCI-DMFT calculation.

By looking at the self-energies as a function of $tdets$ for the first iteration in either Fig. 3 or 4 we confirm the fast converging properties of the ASCI impurity solver that were discussed above. The only appreciable differences with respect to the ED results occur at small frequencies and only for the more modest $tdets$ values, namely 1000 and below. By looking at the self-energies at the last iteration, we can appreciate the differences generated by the fitting routine, see, e.g., the right sub-figures of Fig. 3 and 4. Indeed, although the ED self-energies agree very well between the two different fitting methods, the convergence behavior for the calculations with truncated Hilbert spaces differ noticeably. This is an artefact of the sensitivity of the fitting methods to slight changes in the computed self-energies, and not related to the impurity solver.

This dependence on the fitting method is further exemplified when looking at the DOS along the real frequency axis in Fig. 5 and 6. The SciPy least squares fit gives a fairly consistent convergence behavior as a function of $tdets$. In Fig. 5 almost all peaks fall on top of each other, with the only persistent changes being the continuous decrease of a small satellite peak around $\omega/t = 0$. On the other hand, calculations using the BOBYQA implementation in the nlopt library, see Fig. 6, show more discrepancies between the DOS computed from different values of $tdets$. Although all calculations share the same qualitative structure, which is similar to the one obtained with the SciPy fit, the position of the main peaks can oscillate by almost 0.3 $t$ between calculations (see inset in Fig. 6). There is also a significant change in intensity for the peaks at $\omega/t \approx 4$ and $\omega/t \approx 6$. These results emphasize the importance of the fitting step for achieving quantitatively accurate finite bath DMFT calculations, independently of the quality of the impurity solver used. Indeed, as we point out in the main body of this Letter, for more complicated calculations, such as the 2D square lattice Hubbard model cluster DMFT with $N_c = 4 \times 4$ and $N_b = 24$, the fitting step can be the bottleneck in the ASCI-DMFT method. To further scale zero temperature CI based DMFT studies to larger clusters, the fitting
FIG. 3. Imaginary part of the self-energy along the imaginary frequency axis for a 1d Hubbard model. ASCI-DMFT results with $U/t = 8$, $N_c = 1$, $N_b = 11$ and different sizes of the target space using the SciPy least squares optimization routine. Presented are $tdets = 250$, 500, 1000, 3000 and 10000 in different scales of red, and the exact diagonalisation results in green. We present results for the first and final iteration in the ASCI-DMFT procedure.

The procedure will need to be developed further. For the cluster DMFT calculations presented in this Letter, we chose the BOBYQA implementation of the nlopt library as a fitting method. The reason for this is the superior timing of this implementation compared to that of the SciPy least squares fit, as the number of fit parameters increases. This becomes crucial when treating the 2D Hubbard model with clusters larger than 3x3. In these cases, it becomes necessary to trade the better convergence properties of the SciPy fit for the timing of the nlopt routine in order to converge the DMFT calculations in a reasonable time. Even the more efficient nlopt implementation can require up to 48 hours on a single core per fit for the 4x4 cluster presented in the body of the paper.

It is worth noting that the equally fast convergence in both real and imaginary frequency axes seen in the above results is a peculiarity of the 1 site DMFT calculations. Larger clusters require a significantly bigger value of $tdets$ along the real frequency axis, where the Green’s functions and self energies have a more detailed structure. In our calculations we perform the DMFT iterations with a small $tdets$ parameter until the bath parameters have converged, and then increase $tdets$ by a factor of 2 or 3 in order to compute physical quantities such as the local density of states or the k-resolved spectral weights. A good indicator for an insufficiently large target space along the real frequency is given by the appearance of non-causal self energies [21].

We now give a short timing analysis of the ASCI-DMFT algorithm. The average timing per iteration for the three main steps in the ASCI-DMFT method (ASCI algorithm, computation of the cluster Green’s function and fit using SciPy) for 1-site and 4-site calculations are presented in Fig. 7 as a function of $tdets$. One can see that changing the value of $tdets$ most affects the timing of the ASCI method, followed by the computation of the Green’s function. The fitting is, as expected, nearly independent of $tdets$. For any reasonable size of the target space, the impurity solver step, performed in this work with the ASCI algorithm, is the most time consuming. Thus, the impurity solver is the main bottleneck. For bigger cluster sizes ($N_c = 8$ for example), the computation of the Green’s function can take a comparable time to the impurity solver, especially along the real frequency axis. Additionally, at intermediate cluster sizes ($N_c = 4$) the fitting procedure can be the most time consuming part of the DMFT calculation, which again shows the efficiency of the ASCI algorithm as an impurity solver, see the right panel of Fig. 7.

To further demonstrate the accuracy of our calculations, we present the spectral weights of the 1d Hubbard model for the case with $U/t = 8$, $N_c = 4$ and $N_b = 8$ in Fig. 8. The spectral weights were computed using the prescription given in [61] to unfold the cluster Brillouin zone. This calculation, performed on a target space of 10000 determinants,
FIG. 4. Imaginary part of the self-energy along the imaginary frequency axis for a 1d Hubbard model. ASCI-DMFT results with $U/t = 8$, $N_c = 1$, $N_b = 11$ and different sizes of the target space using the BOBYQA implementation in the nlopt library. Presented are $t\text{dets} = 250$, 500, 1000, 3000 and 10000 in different scales of red, and the exact diagonalisation results in green. We present results for the first and final iteration in the ASCI-DMFT procedure.

was completed in about 3 hours on a single core of an Intel Core i5 CPU (2.9 GHz). A more complicated calculation with $N_c = 8$ and $N_b = 16$ on the same system took about 8 hours on one core. All these timings are to our knowledge two orders of magnitude below the best timings of current methods [21]. The comparison with existing calculations, already tested against the Bethe ansatz [21], is very good.
FIG. 5. Local density of states for a 1d Hubbard model. ASCI-DMFT results with $U/t = 8$, $N_c = 1$, $N_b = 11$ and different sizes of the target space using the SciPy least squares optimization routine. Presented are $tdets = 250$, 500, 1000, 3000 and 10000 in different scales of red and the exact diagonalisation results in green. The inset zooms on the central region of the plot.
FIG. 6. Local density of states for a 1d Hubbard model. ASCI-DMFT results with $U/t = 8$, $N_c = 1$, $N_b = 11$ and different sizes of the target space using the BOBYQA implementation in the nlopt library. Presented are $tdets = 250$, $500$, $1000$, $3000$ and $10000$ in different scales of red and the exact diagonalisation results in green. The inset zooms on the peak at $\omega/t \approx 2$.

FIG. 7. Iteration-average timings of the different parts of the DMFT code as a function of the target space size $tdets$ for a 1d Hubbard model with $U/t = 8$. The cluster sizes are $(N_c = 1, N_b = 11)$ (left) $(N_c = 4, N_b = 8)$ (right). The timing of the ASCI algorithm is presented in red, that of the computation of the cluster Green’s function (GF) in blue and that of the SciPy least squares fit in green.
FIG. 8. Spectral weights for the 1 dimensional Hubbard model in a calculation with $U/t = 8$, $N_c = 4$ and $N_b = 8$. These are computed using the prescription in [61] to unfold the cluster Brillouin zone.