



From Real Materials to Model Hamiltonians With Density Matrix Downfolding

Huihuo Zheng¹, Hitesh J. Changlani², Kiel T. Williams³, Brian Busemeyer³ and Lucas K. Wagner^{3*}

¹ Argonne Leadership Computing Facility, Argonne National Laboratory, Lemont, IL, United States, ² Department of Physics and Astronomy and Institute for Quantum Matter, Johns Hopkins University, Baltimore, MD, United States, ³ Department of Physics and Institute for Condensed Matter Theory, University of Illinois at Urbana-Champaign, Urbana, IL, United States

Due to advances in computer hardware and new algorithms, it is now possible to perform highly accurate many-body simulations of realistic materials with all their intrinsic complications. The success of these simulations leaves us with a conundrum: how do we extract useful physical models and insight from these simulations? In this article, we present a formal theory of downfolding—extracting an effective Hamiltonian from first-principles calculations. The theory maps the downfolding problem into fitting information derived from wave functions sampled from a low-energy subspace of the full Hilbert space. Since this fitting process most commonly uses reduced density matrices, we term it density matrix downfolding (DMD).

OPEN ACCESS

Edited by:

Chandan Setty,
University of Illinois at
Urbana-Champaign, United States

Reviewed by:

Ke-Qiu Chen,
Hunan University, China
Martin Kiffner,
Centre for Quantum Technologies,
Singapore

*Correspondence:

Lucas K. Wagner
lkwagner@illinois.edu

Specialty section:

This article was submitted to
Condensed Matter Physics,
a section of the journal
Frontiers in Physics

Received: 01 December 2017

Accepted: 24 April 2018

Published: 11 May 2018

Citation:

Zheng H, Changlani HJ, Williams KT,
Busemeyer B and Wagner LK (2018)
From Real Materials to Model
Hamiltonians With Density Matrix
Downfolding. *Front. Phys.* 6:43.
doi: 10.3389/fphy.2018.00043

Keywords: downfolding, effective model, strongly correlated systems, quantum Monte Carlo, machine learning

1. INTRODUCTION TO DOWNFOLDING THE MANY ELECTRON PROBLEM

In multiscale modeling of many-particle systems, the effective Hamiltonian (or Lagrangian) is one of the most core concepts. The effective Hamiltonian dictates the behavior of the system on a coarse-grained level, where “sub-grid” effects are folded into the parameters and form of the effective Hamiltonian. Many concepts in condensed matter physics can be viewed as statements about the behavior of the effective Hamiltonian. In particular, identification of “strongly correlated” materials as materials where band theory is not an accurate representation of the system is a statement about effective Hamiltonians. Effective Hamiltonians at different length scales also form the basis of the renormalization group [1]. A major goal in condensed matter physics is to determine what effective Hamiltonians apply to different physical situations, in particular quantum effective Hamiltonians, which lead to large-scale emergent quantum phenomena.

The dominant effective model for quantum particles in materials is band structure, and for metals, Fermi liquid theory. However, a major challenge is how this paradigm should be altered when it is no longer a good description of the physical system. Examples of these include the high- T_c cuprates and other transition metal oxides, which do not appear to be well-described by these simple effective Hamiltonians. For these systems, many models have been proposed, such as the Hubbard [2], Kanamori [3], t - J [4], and Heisenberg models. While these models have been extensively studied analytically and numerically, and have significantly enhanced our understanding of the physics of correlated electrons, their effectiveness for describing a real complex system of interest is often unclear. At the same time, more complex effective models can be commensurately more difficult to solve, so one would like to also find an accurate effective model that is computationally tractable.

To address the need for a link between *ab initio* electron-level models and larger scale models, downfolding has most commonly been carried out using approaches based on density functional theory (DFT). The effective one-particle hopping terms are obtained by projecting the DFT bands onto localized Wannier functions [5]. The interactions are then estimated based on certain models of screening of the Coulomb interactions [e.g., constrained DFT, Random Phase Approximations (RPA)]. Since effects of interactions between the orbitals of interest have already been accounted for by DFT, a double counting correction is required to obtain the final downfolded Hamiltonian. The approach has been developed and widely applied [5–8]; but remains an active area of research [9]. There are other downfolding approaches that include the traditional Löwdin method, coupled to a stochastic approach [10, 11] and the related method of canonical transformations [12, 13]. While they have many advantages, it is typically not possible to know if a given model *ansatz* was a good guess or not, and it is very rare for a technique to provide an estimate of the quality of the resultant model.

The situation described above stands in contrast to the derivation of effective classical models. For concreteness, let us discuss classical force fields computed from *ab initio* electronic structure calculations. Typically, a data set is generated using an *ab initio* calculation in which the positions of the atoms and molecules are varied, creating a set of positions and energies. The parameters in the force field *ansatz* are varied to obtain a best-fit classical model. Then, using standard statistical tools, it is possible to assess how well the fit reproduces the *ab initio* data within the calculation, without appealing to experiment. While translating that error to error in properties is not a trivial task, this approach has the important advantage that in the limit of a high quality fit and high quality *ab initio* results, the resultant model is predictive.

Naïvely, one might think to reconcile the fitting approach used in classical force fields with quantum models by matching eigenstates between a quantum model and *ab initio* systems, varying the model parameters until the eigenstates match [14]. However, this strategy does not work well in practice because it is often not possible to obtain exact eigenstates for either the model or the *ab initio* system. To resolve this, we develop a general theory for generating effective quantum models that is exact when the wave functions are sampled from the manifold of low-energy states. Because this method is based on fitting the energy functional, we will show the practical application of this theory using both exact solutions and *ab initio* quantum Monte Carlo (QMC) to derive several different quantum models.

The endeavor we pursue here is to develop a multi-scale approach in which the effective interactions between quasiparticles (such as dressed electrons) are determined after an *ab initio* simulation (but not necessarily exact solution) of the continuum Schrödinger equation involving all the electrons. The method uses reduced density matrices (RDMs) of low-energy states, not necessarily eigenstates, to cast downfolding as a fitting problem. We thus call it density matrix downfolding (DMD). In this paper, our applications of DMD to physical problems employ one-body (1-RDM) and two-body (2-RDM) density matrices.

The many-body states used in DMD will typically be generated using QMC techniques [either variational Monte Carlo (VMC) or diffusion Monte Carlo (DMC)] to come close to the low energy manifold.

DMD was first applied to the benzene molecule and monolayer graphene in our previous work [15]. We showed that the downfolded Hamiltonian reproduced the experimental energy spectrum for benzene. In this paper, we will present a rigorous theory underlying the whole downfolding framework and apply the downfolding method to more systems.

The remainder of the paper is organized as follows:

- ▷ In section 2, we clarify and make precise what it means to downfold a many-electron problem to a few-electron problem. We recast the problem into minimization of a cost function that needs to be optimized to connect the many and few body problems. We further these notions both in terms of physical as well as data science descriptions, which allows us to connect to compression algorithms in the machine learning literature.
- ▷ Section 3 discusses several representative examples where we consider multiband lattice models and *ab initio* systems to downfold to a simpler lattice model.
- ▷ In section 4, we discuss future prospects of applications of the DMD method, ongoing challenges and clear avenues for methodological improvements.

2. DOWNFOLDING AS A COMPRESSION OF THE ENERGY FUNCTIONAL

In this section, we will develop a sufficient criterion for an effective Hamiltonian H_{eff} to reproduce the spectrum of a first-principles Hamiltonian H within the low-energy space. The criterion is that if a wave function $|\Psi\rangle$ is drawn from the low-energy space, then the expectation values of the effective and first principles Hamiltonian must match for that wave function. We then use the concept of *descriptors* to help parameterize the effective Hamiltonian in terms of expectation values on wave functions sampled from the low-energy space, such as hopping and double occupancy.

2.1. Theory

2.1.1. Energy Functional

Suppose we start with a quantum system with Hamiltonian H and Hilbert space \mathcal{H} .

Definition 1. Let the energy functional be $E[\Psi] = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}$ for a wave function $|\Psi\rangle \in \mathcal{H}$.

Theorem 1. $E[\Psi]$ has a critical point only where $|\Psi\rangle$ is an eigenvector of H .

Proof:

$$\begin{aligned} \frac{\delta}{\delta\Psi^*} E[\Psi] &= \frac{\delta}{\delta\Psi^*} \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{H|\Psi\rangle}{\langle\Psi|\Psi\rangle} - \langle\Psi|H|\Psi\rangle \frac{|\Psi\rangle}{|\langle\Psi|\Psi\rangle|^2} \\ &= \frac{(H - E[\Psi])|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \end{aligned} \quad (1)$$

Therefore, $\frac{\delta}{\delta\Psi^*} E[\Psi] = 0$ if and only if $(H - E[\Psi])|\Psi\rangle = 0$, i.e., $|\Psi\rangle$ is an eigenvector of H corresponding to eigenvalue $E[\Psi]$.

2.1.2. Low Energy Space

Definition 2. Let $\mathcal{LE}(H, N)$ be a subset of \mathcal{H} spanned by N vectors given by the lowest energy solutions to $H|\Psi_n\rangle = E_n|\Psi_n\rangle$.

Definition 3. H_{eff} is an operator on the Hilbert space $\mathcal{LE}(H, N)$.

Definition 4. The effective model $E_{eff}[\Psi] = \frac{\langle\Psi|H_{eff}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$ is a functional from $\mathcal{LE} \rightarrow \mathbb{R}$.

If $|\Psi\rangle \in \mathcal{LE}$ and $|\Phi\rangle \in \mathcal{H} \setminus \mathcal{LE}$, then $|\Psi\rangle \oplus |\Phi\rangle \in \mathcal{H}$. In the following, we will use the direct sum operator $\oplus 0$ to translate between the larger \mathcal{H} and the smaller \mathcal{LE} .

Lemma 1. Suppose that $|\Psi\rangle \in \mathcal{LE}$ and $|\Phi\rangle \in \mathcal{H} \setminus \mathcal{LE}$. Then $\left.\frac{\delta E[\Psi \oplus \Phi]}{\delta\Phi}\right|_{\Phi=0} = 0$.

Proof: $\langle\Psi \oplus 0|H|0 \oplus \Phi\rangle = 0$ because the two states have non-overlapping expansions in the eigenstates of H . Using that fact, we can evaluate

$$\left.\frac{\delta E[\Psi \oplus \Phi]}{\delta\Phi}\right|_{\Phi=0} = \left.\frac{(H - E[\Psi \oplus \Phi])|\Phi\rangle}{\langle\Phi|\Phi\rangle}\right|_{\Phi=0} = 0. \quad (2)$$

This is equivalent to noting that H is block diagonal in the partitioning of \mathcal{H} into \mathcal{LE} and $\mathcal{H} \setminus \mathcal{LE}$. Importantly, if $|\Psi\rangle \in \mathcal{LE}$, then $\left.\frac{\delta E[\Psi \oplus 0]}{\delta(\Psi \oplus 0)^*}\right|_{\Psi=0} = |\Psi'\rangle \oplus 0$, where $|\Psi'\rangle \in \mathcal{LE}$.

Theorem 2. Assume $E[\Psi \oplus 0] = E_{eff}[\Psi] + C$ for any $|\Psi\rangle \in \mathcal{LE}$, where C is a constant. Then $(H_{eff} + C)|\Psi\rangle \oplus 0 = H(|\Psi\rangle \oplus 0)$.

Proof: Note that

$$\left.\frac{\delta E[\Psi \oplus 0]}{\delta(\Psi \oplus 0)^*}\right|_{\Psi=0} = \frac{(H - E[\Psi \oplus 0])|\Psi \oplus 0\rangle}{\langle\Psi|\Psi\rangle} \quad (3)$$

and

$$\left.\frac{\delta E_{eff}[\Psi]}{\delta\Psi^*}\right|_{\Psi=0} = \frac{(H_{eff} - E_{eff}[\Psi])|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \quad (4)$$

Since the derivatives are equal, setting Equation 3 equal to Equation 4,

$$\begin{aligned} H|\Psi \oplus 0\rangle &= (H_{eff} + E[\Psi \oplus 0] - E_{eff}[\Psi])|\Psi\rangle \oplus 0 \\ &= (H_{eff} + C)|\Psi\rangle \oplus 0. \end{aligned} \quad (5)$$

Theorem 2 combined with Lemma 1 means that the eigenstates of H_{eff} are the same as the eigenstates of H if the derivatives of its energy functional match those of H . Such a H_{eff} always exists. Let $H_{eff} = \sum_{n=1}^N E_n|\Psi_n\rangle\langle\Psi_n|$ where $|\Psi_n\rangle$'s are eigenstates belong to $\mathcal{LE}(H, N)$. This satisfies $E[\Psi] = E_{eff}[\Psi]$ and $H_{eff}|\Psi\rangle = H|\Psi\rangle$ for any $|\Psi\rangle$ in $\mathcal{LE}(H, N)$.

We have thus reduced the problem of finding an effective Hamiltonian H_{eff} that reproduces the low-energy spectrum of H to matching the corresponding energy functionals $E[\Psi]$ and $E_{eff}[\Psi]$. Practically, this can be implemented as follows:

- (1) Generating an *ansatz* for the effective Hamiltonian in terms of operators.
- (2) Generating wave functions $|\Psi_n\rangle$'s in the low-energy subspace of the first principles Hilbert space \mathcal{H} .
- (3) Computing $E[\Psi_n]$ using the expectation value of the first principles Hamiltonian and $E_{eff}[\Psi_n]$ by taking the expectation value of the operators in the *ansatz* for H_{eff} .

This involves sampling the low-energy space, choosing the form of H_{eff} , and optimizing the parameters. An important implication of this is that it is not necessary to diagonalize either of the Hamiltonians; one must only be able to select wave functions from the low-energy space \mathcal{LE} . As we shall see, this can be substantially easier than attaining eigenstates.

Some further notes about this derivation:

- ▷ Fitting Ψ 's must come from \mathcal{LE} . It is not enough that the energy functional $E[\Psi]$ is less than some cutoff energy.
- ▷ In the case of sampling an approximate \mathcal{LE} , the error comes from non-parallelity of $E[\Psi]$ with the correct low energy manifold, up to a constant offset.
- ▷ While H_{eff} is unique, it has many potential representations and approximations.
- ▷ Our method can be applied to any manifold spanned by eigenstates.
- ▷ Model fitting is finding a compact approximation to $E_{eff}[\Psi]$. This is a high-dimensional space, so we use descriptors to do this.
- ▷ For operators that are not the Hamiltonian, it is possible to fit $\mathcal{O}_{eff}[\Psi] \simeq \mathcal{O}[\Psi]$ in a similar way. However, the eigenstates of \mathcal{O} and \mathcal{O}_{eff} will not coincide in general unless \mathcal{O} commutes with the Hamiltonian.

The theory presented above maps coarse-graining into a functional approximation problem. This is still rather intimidating, since even supposing one can generate wave functions in the low-energy space, they are still complicated objects in a very large space. An effective way to accomplish this is through the use of descriptors, $d_j[\Psi]$, which map from $\mathcal{H} \rightarrow \mathbb{R}$. Then we can approximate the energy functional as follows

$$E_{eff}[\Psi] \simeq \sum_i f_j(d_j[\Psi]), \quad (6)$$

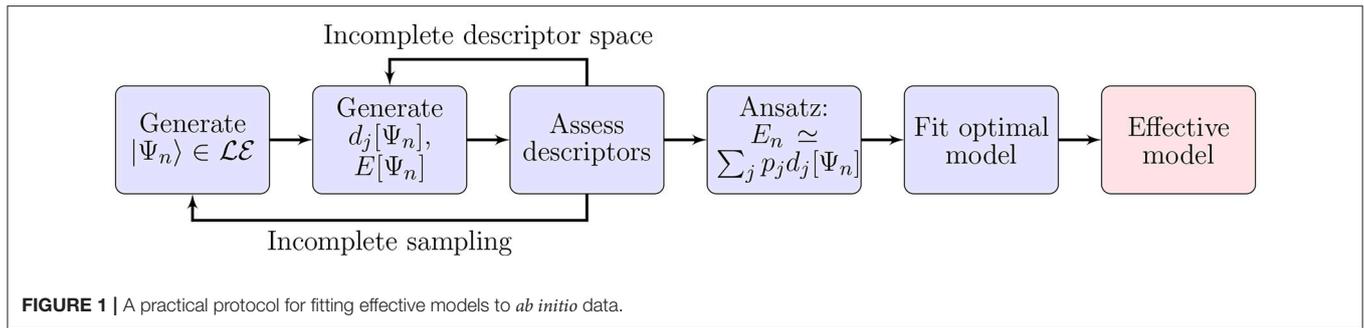
where f_j are some parameterized functions. This will allow us to use techniques from statistical learning to efficiently describe E_{eff} .

2.2. Practical Protocol

A practical protocol is presented in **Figure 1**. In this section we go through this procedure step by step.

2.2.1. Generating $|\Psi_n\rangle \in \mathcal{LE}$

Ideally one would be able to sample the entire low-energy space. Typically, however, the space will be too large and it will need to be sampled. The optimal wave functions to use depend on the models one expects to fit, which we will discuss in detail in later steps. Simple strategies that we will use in the examples below include excitations with respect to a determinant and varying spin states.



2.2.2. Generate $d_j[\Psi_n]$ and $E[\Psi_n]$

The choice of descriptor is fundamental to the success of the downfolding. In the case of a second-quantized Hamiltonian

$$H_{\text{eff}} = E_0 + \sum_{ij} t_{ij} c_i^\dagger c_j + h.c. + \sum_{ijkl} V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l, \quad (7)$$

a set of linear descriptors can be derived by simply taking the expectation value of both sides of the equation:

$$E_{\text{eff}}[\Psi_n] = E_0 + \sum_{ij} t_{ij} \langle \Psi_n | c_i^\dagger c_j + h.c. | \Psi_n \rangle + \sum_{ijkl} V_{ijkl} \langle \Psi_n | c_i^\dagger c_j^\dagger c_k c_l | \Psi_n \rangle. \quad (8)$$

Then for example, the occupation descriptor for orbital k is $d_{\text{occ}(k)}[\Psi_n] = \langle \Psi_n | c_k^\dagger c_k | \Psi_n \rangle$; the double occupation descriptor for orbital k is $d_{\text{double}(k)}[\Psi_n] = \langle \Psi_n | n_{k\uparrow} n_{k\downarrow} | \Psi_n \rangle$. The orbital that c_k represents is part of the descriptor, and in the examples we will discuss this choice. One is not limited to static orbital descriptors; they may have a more complex functional dependence on the trial function to include orbital relaxation.

2.2.3. Assess Descriptors

At this point, one has collected the data E_n and $d_j[\Psi_n]$. If two descriptors have a large correlation coefficient, then they are redundant in the data set. This could either mean that the sampling of the low-energy Hilbert space $\mathcal{L}\mathcal{E}$ was insufficient, or that they are both proxies for the same differences in states. If two data points have the same or very similar descriptor sets, but different energies, then either the descriptor set is not enough to describe the variations in the low-energy space, or the sampling has generated states that are not in the low-energy space. To resolve these possibilities, one should analyze the difference between the two wave functions.

In either case, when the model is accurate, the fits will be accurate. If descriptor values available in the reduced Hilbert space are not represented in the sampled wave functions, then intruder states can appear upon solution of the effective model. In that case, the model fitting is an extrapolation instead of an interpolation. For this reason it is desirable to have eigenstates or near-eigenstates in the sample set if possible; they are guaranteed to be on the corners of the descriptor space if the model is accurate.

2.2.4. Ansatz: $E_n \simeq \sum_j p_j d_j[\Psi_n]$

If the descriptors are chosen well, then the model can be written in linear form:

$$E[\Psi_n] = \sum_j p_j d_j[\Psi_n], \quad (9)$$

which we shorten to

$$\mathbf{E} = \mathbf{D}\mathbf{p}, \quad (10)$$

where $D_{nj} := d_j[\Psi_n]$ and $E_n := E[\Psi_n]$. If this can be done, the fitting problem is reduced to a linear regression optimization. More complex functions of the descriptors are also possible, although at the cost of making the effective model more difficult to solve and complicating the fitting procedure.

2.2.5. Fit Optimal Model

Finally, one wishes to find a set of parameters such that Equation 10 is satisfied as closely as possible. For the example given in Equation 7, this would involve optimizing the parameters E_0 , t_{ij} , and V_{ijkl} . In terms of the generalized formalism in Equation 10, this involves optimizing the parameters \mathbf{p} . It is also possible to optimize the descriptors themselves, or to choose which descriptors to use among a set. In our tests, we have successfully used LASSO [16] and matching pursuit techniques [17] to select high quality and compact model parameters. A detailed example of using the latter technique is presented in section 3.4.

We would like to emphasize several advantages of our method:

- (1) Our method provides an internal consistency check on the quality of the effective model in describing the corresponding *ab initio* system. The quality of the linear fit assesses the correctness of the model parameterization.
- (2) The low energy space is sampled by low energy states which do not have to be the eigenstates of the system. In other words, we do not need to exactly solve the *ab initio* Hamiltonian or the model Hamiltonian to know the map connecting the two Hamiltonians. The low energy non-eigenstates are computationally inexpensive to generate from first principles (e.g., QMC method);

The formulation of our method is exact in principle. However, in practical applications, there are mainly two approximations on (1) the form of the low energy Hamiltonian [Equation 7];

(2) the low energy states we used to sample the low energy manifold. We assume that the low energy Hamiltonian could be written in terms of low energy degrees of freedom [c_i 's in Equation 7] and we also only considered single-body and two-body terms. Nevertheless, the quality of the effective Hamiltonian is quantitatively measured by the quality of the linear regression fit.

3. REPRESENTATIVE EXAMPLES

Given the theoretical framework for downfolding a many orbital (or many-electron) problem to a few orbital (or few-electron) problem, we now discuss examples which elucidate the DMD method. The examples are as follows:

- ▷ Section 3.1: Three-band Hubbard \rightarrow one-band Hubbard at half filling. Demonstrates finding a basis set for the second quantized operators and uses a set of eigenstates directly sampled from the low-energy space to find a one-band model.
- ▷ Section 3.2: Hydrogen chain \rightarrow one-band Hubbard model at half filling. Demonstrates basis sets for *ab initio* systems and the possibility to use this technique to determine the quality of a model for a given physical situation.
- ▷ Section 3.3: Graphene \rightarrow one-band Hubbard model with and without σ electrons. Demonstrates using the downfolding procedure to examine the effects of screening due to core electrons.
- ▷ Section 3.4: FeSe molecule \rightarrow $3d, 4p, 4s$ system. Demonstrates the use of matching pursuit to assess the importance of terms in an effective model and to select compact effective models.

In all examples we will highlight the important ingredients associated with DMD. First and foremost is the choice of low energy space or energy window i.e., how our database of wave functions was generated. Associated with this is the choice of the one body space in terms of which the effective Hamiltonian is expressed. Finally, we discuss aspects of the functional forms or parameterizations that are expected to describe our physical problem. An important effective Hamiltonian that enters three out of our four representative examples is the one-band or single-band Hubbard model:

$$H = E_0 - t \sum_{\langle i,j \rangle, \eta} \tilde{d}_{i,\eta}^\dagger \tilde{d}_{j,\eta} + U \sum_i \tilde{n}_\uparrow^i \tilde{n}_\downarrow^i, \quad (11)$$

where t and U are downfolded (renormalized) parameters, E_0 is a constant, η is a spin index, $\tilde{d}_{i,\eta}$ is the effective one-particle operator associated with spatial orbital (or site) i and $\tilde{n}_{i,\eta} = \tilde{d}_{i,\eta}^\dagger \tilde{d}_{i,\eta}$ is the corresponding number operator. $\langle i, j \rangle$ is used to denote nearest neighbor pairs.

3.1. Three-Band Hubbard Model to One-Band Hubbard Model at Half Filling

Our first example is motivated by the high T_c superconducting cuprates [18] that have parent Mott insulators with rich phase diagrams on electron or hole doping [19, 20]. Many works have been devoted to their model Hamiltonians and corresponding parameter values [4, 5, 21–25]. A minimal model involving both

the copper and oxygen degrees of freedom is the three-orbital or three-band Hubbard model,

$$H = \epsilon_p \sum_{j \in p, \eta} n_{j,\eta} + \epsilon_d \sum_{i \in d, \eta} n_{i,\eta} + t_{pd} \sum_{\langle i \in d, j \in p \rangle, \eta} \text{sgn}(p_i, d_j) (c_{i,\eta}^\dagger c_{j,\eta} + \text{h.c.}) + U_p \sum_{j \in p} n_{j,\uparrow} n_{j,\downarrow} + U_d \sum_{i \in d} n_{i,\uparrow} n_{i,\downarrow} + V_{pd} \sum_{\langle i \in p, j \in d \rangle} n_j n_i, \quad (12)$$

where d_i, p_j refer to the $d_{x^2-y^2}$ orbitals of copper at site i and p_x or p_y oxygen at site j , respectively. $\text{sgn}(p_i, d_j)$ is the sign of the hopping t_{pd} between nearest neighbors, shown schematically in **Figure 2**. ϵ_d and ϵ_p are orbital energies, U_d and U_p are strengths of onsite Hubbard interactions, and V_{pd} is the strength of the density-density interactions between a neighboring p and d orbital. To simplify, we consider only the case where ϵ_p, U_d , and t_{pd} are non zero; t_{pd} is chosen throughout this section to be the typical value of 1.3 eV to give the reader a sense of overall energy scales. Since we work with fixed number of particles we set our reference zero energy to be $\epsilon_d = 0$, thus the charge transfer energy $\Delta \equiv \epsilon_p - \epsilon_d$ equals ϵ_p in our notation. We work in the hole notation; half filling corresponds to two spin-up and two spin-down holes on the 2×2 cell.

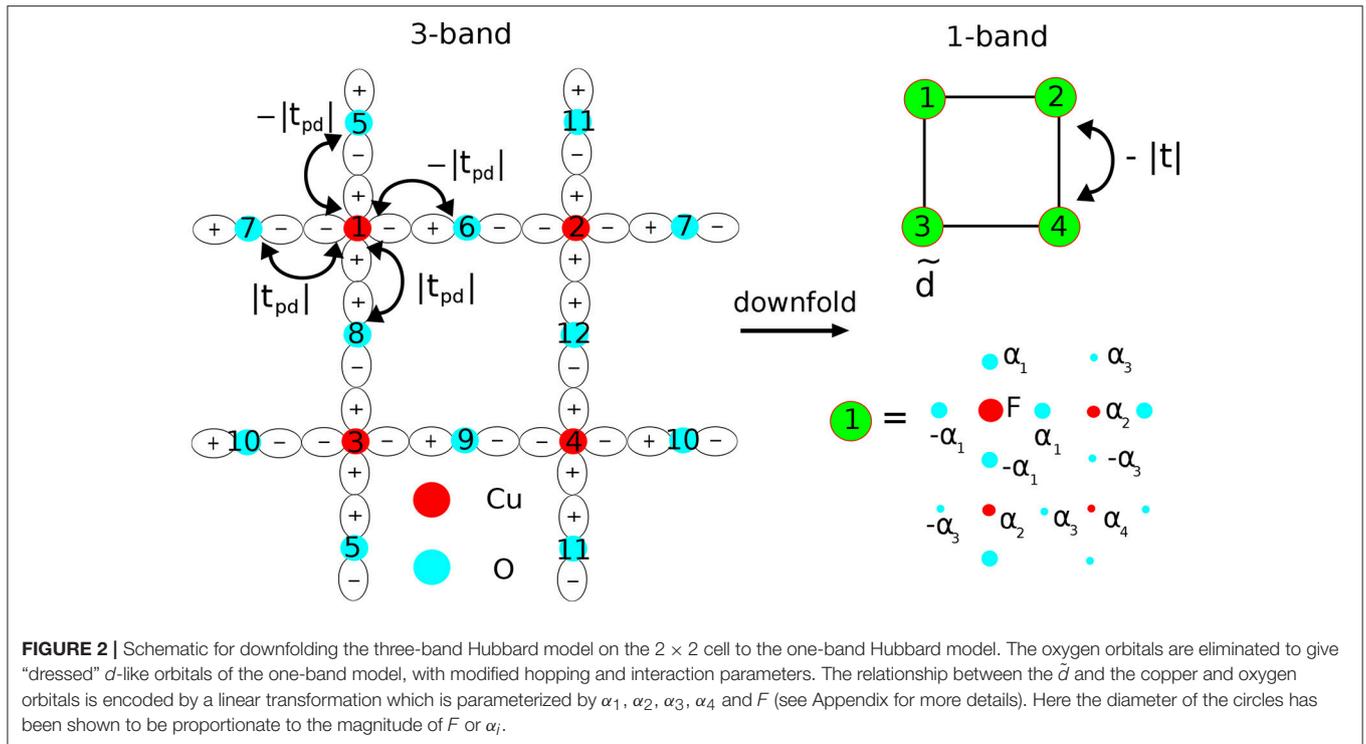
It is our objective to determine what one-band Hubbard model [Equation 11] “best” describes the three-band data. The effective *d-like* orbitals $\tilde{d}_{i,\eta}$, that enter the low energy description are mixtures of copper and oxygen orbitals; this optimal transformation also remains an unknown. Thus the model determination involves two aspects (1) what are the composite objects that give a compact description of the low energy physics? and (2) given this choice what are the effective interactions between them? (A similar problem was posed and solved by one of us in the context of spin systems [26]). In addition, the best effective Hamiltonian description depends on the energy scale of interest. All these issues will be addressed in the remainder of the section.

We first determine the optimal operators $\tilde{d}_{i,\eta}$, which are constructed from a linear superposition of the d and p orbitals,

$$\tilde{d}_{i,\eta} = \sum_j T_{ij} c_{j,\eta} \quad (13)$$

where $c_{j,\eta}$ is the hole (destruction) operator and refers to either the bare d or p orbitals and \mathbf{T} is the transformation matrix. Further generalizations of this relationship (for example, including higher body terms) are also possible, but have not been considered here. For the 2×2 unit cell \mathbf{T} is a 4×12 matrix, which we parameterize by four distinct parameters. These correspond to mixing of a copper orbital with nearest neighbor oxygens (α_1), nearest neighbor coppers (α_2), next-nearest neighbor oxygens (α_3), and next-nearest neighbor coppers (α_4) as shown schematically in **Figure 2**. The explicit form of \mathbf{T} after accounting for the symmetries of the lattice has been written out in the Appendix.

All RDMs in the three-band and one-band descriptions are also related via \mathbf{T} ; the ones that we focus on are evaluated in



eigenstate s and are given by,

$$\langle \tilde{d}_{i,\eta}^\dagger \tilde{d}_{j,\eta} \rangle_s = \sum_{mn} T_{im}^* \langle c_{m,\eta}^\dagger c_{n,\eta} \rangle_s T_{jn}, \quad (14a)$$

$$\langle \tilde{n}_{i,\uparrow} \tilde{n}_{i,\downarrow} \rangle_s = \sum_{jkmn} T_{ij}^* T_{im}^* \langle c_{j,\uparrow}^\dagger c_{m,\downarrow}^\dagger c_{n,\downarrow} c_{k,\uparrow} \rangle_s T_{in} T_{ik}. \quad (14b)$$

We determine \mathbf{T} by demanding two conditions be satisfied, (1) the effective orbitals ($\tilde{d}_{i,\eta}$) are orthogonal to each other i.e., $(\mathbf{T}\mathbf{T}^\dagger)_{mn} = \delta_{mn}$ and (2) the sum of all diagonal entries (trace) of the 1-RDM of the effective orbitals for all low energy eigenstates equals the number of electrons of a given spin i.e., $\sum_i \sum_\eta \langle \tilde{d}_{i,\eta}^\dagger \tilde{d}_{i,\eta} \rangle_s = N_\eta$. These conditions are enforced by minimizing a cost function,

$$C = \sum_s \sum_\eta \left(\sum_i \langle \tilde{d}_{i,\eta}^\dagger \tilde{d}_{i,\eta} \rangle_s - N_\eta \right)^2 + \sum_{mn} \left((\mathbf{T}\mathbf{T}^\dagger)_{mn} - \delta_{mn} \right)^2. \quad (15)$$

For the 2×2 cell, $N_\uparrow = N_\downarrow = 2$ and $i = 1, 2, 3, 4$. The number of states s was varied from three to six, depending on the energy window of interest. For a selected low energy space, we optimize Equation 15 first to determine the optimal one-body operators and then determine the parameters by a linear regression fit [Equation 9].

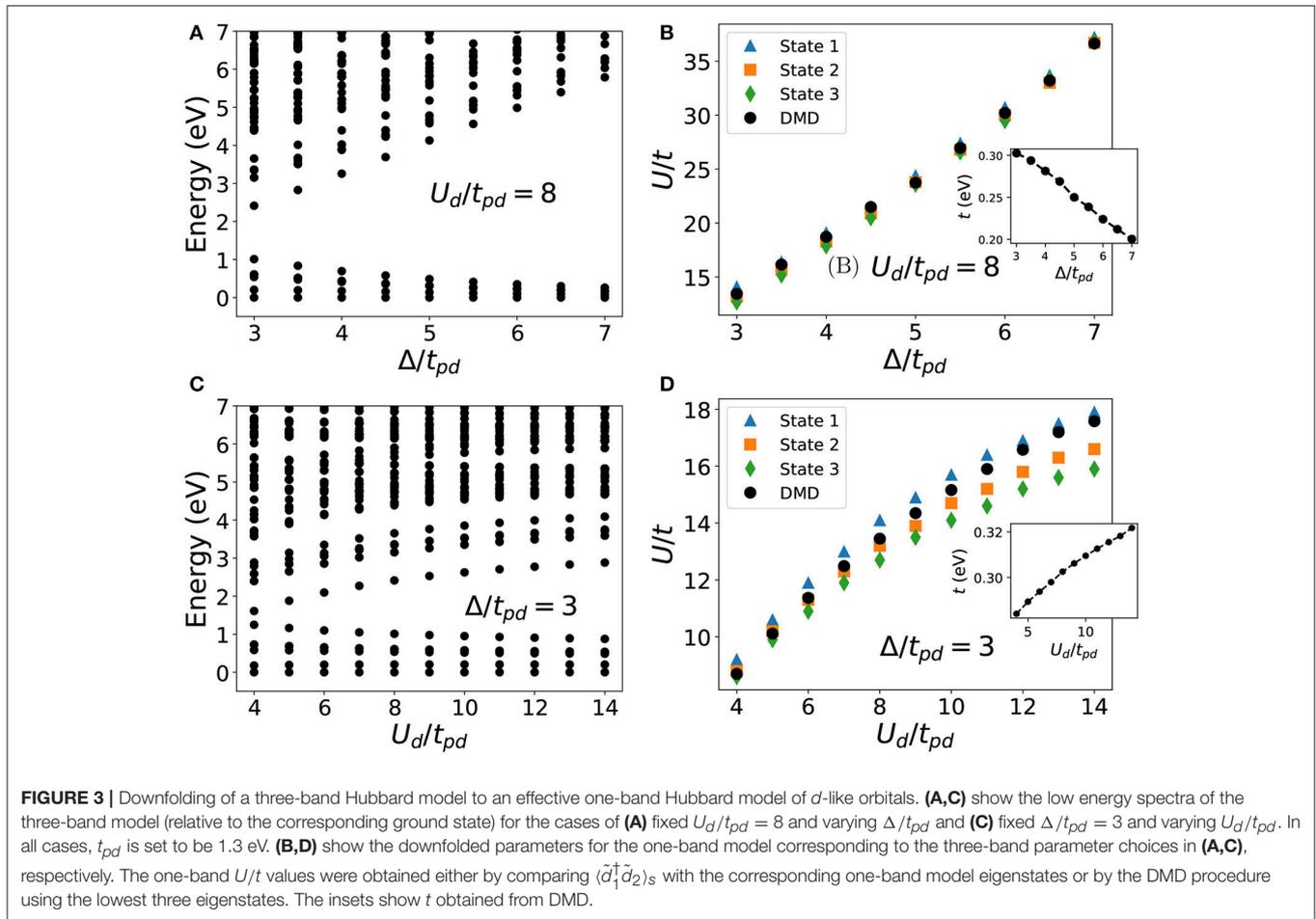
Figure 3 shows regimes of the three-band model where the lowest six eigenstates are separated from the higher energy manifold; the fourth and fifth eigenstates are degenerate. In the large U_d limit, charge fluctuations are suppressed and these six states correspond to the Hilbert space of $\binom{4}{2}$ states of the effective

spin model in its $S_z = 0$ sector. These states have primarily d -like character, an aspect we will verify in this section. The eigenstates outside this manifold involve p -like excitations which the one-band model is not designed to capture.

We chose the lowest three eigenstates of the three-band model for minimizing the cost in Equation 15. The four dimensional space of parameters of \mathbf{T} was scanned for this purpose. The corresponding trace and orthogonality conditions are simultaneously satisfied with only small deviations, confirming the validity of Equation 13. Importantly, the 1-RDM elements in the transformed basis corresponding to nearest neighbors $\langle \tilde{d}_1^\dagger \tilde{d}_2 \rangle_s$ already provide estimates for U/t of the effective model. Since the exact knowledge of the corresponding eigenstates of the one-band Hubbard model is available for arbitrary U/t by exact diagonalization, we directly look up the U/t with the same 1-RDM value. These estimates complement the one obtained by DMD which was carried out with the same three low-energy eigenstates, using their energies and the computed values of $\langle \tilde{d}_1^\dagger \tilde{d}_2 \rangle_s$ and $\langle \tilde{n}_{i,\uparrow} \tilde{n}_{i,\downarrow} \rangle_s$ from Equations 14a and 14b*. A representative example of our results for $U_d/t_{pd} = 8$ and $\Delta/t_{pd} = 3$ has been discussed in the Appendix.

Some trends in the one-band description are explored in **Figure 3** by monitoring the downfolded parameters as a function of varying Δ/t_{pd} and U_d/t_{pd} . For example, when $U_d/t_{pd} = 8$ is fixed and Δ/t_{pd} is increased, we find that the effective hopping t

*We also mimicked the situation characteristic of *ab initio* examples where no eigenstates are generally available. Several non eigenstates were generated as random linear combinations of the lowest three eigenstates and input into the DMD procedure, with similar outcomes.



decreases and U/t increases. This is physically reasonable since an increasing difference in the single particle energies of the copper and oxygen orbitals makes it energetically unfavorable for holes to hop between the two orbitals. When $\Delta/t_{pd} = 3$ is fixed and U_d/t_{pd} is increased, U/t increases. As one mechanism of avoiding the large U_d , the copper orbitals are forced to hybridize more with the oxygen ones; on the other hand, hole delocalization is suppressed in a bid to maintain mostly one hole per \tilde{d} due to the larger U/t . The net result of these effects is that t also increases.

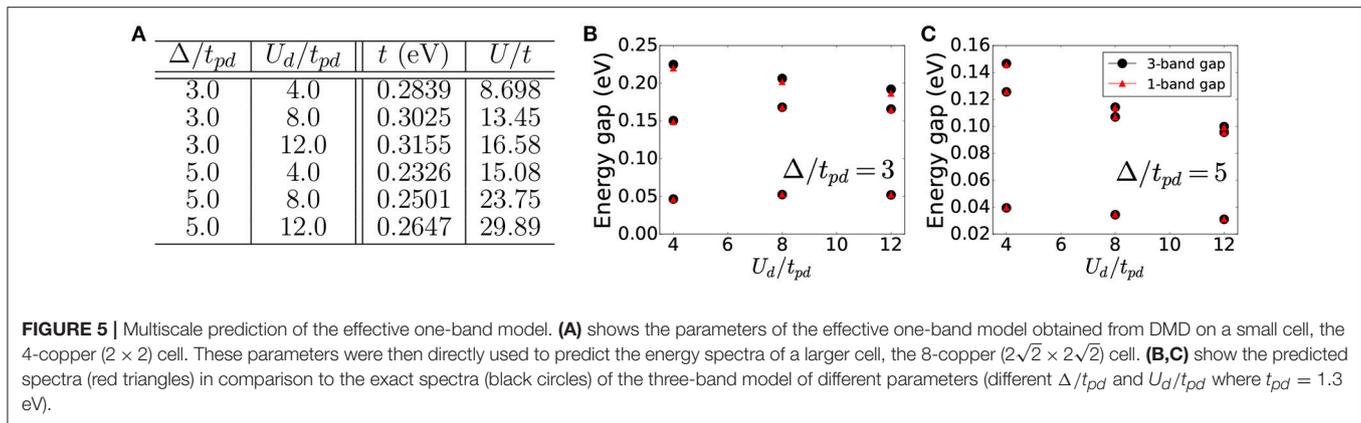
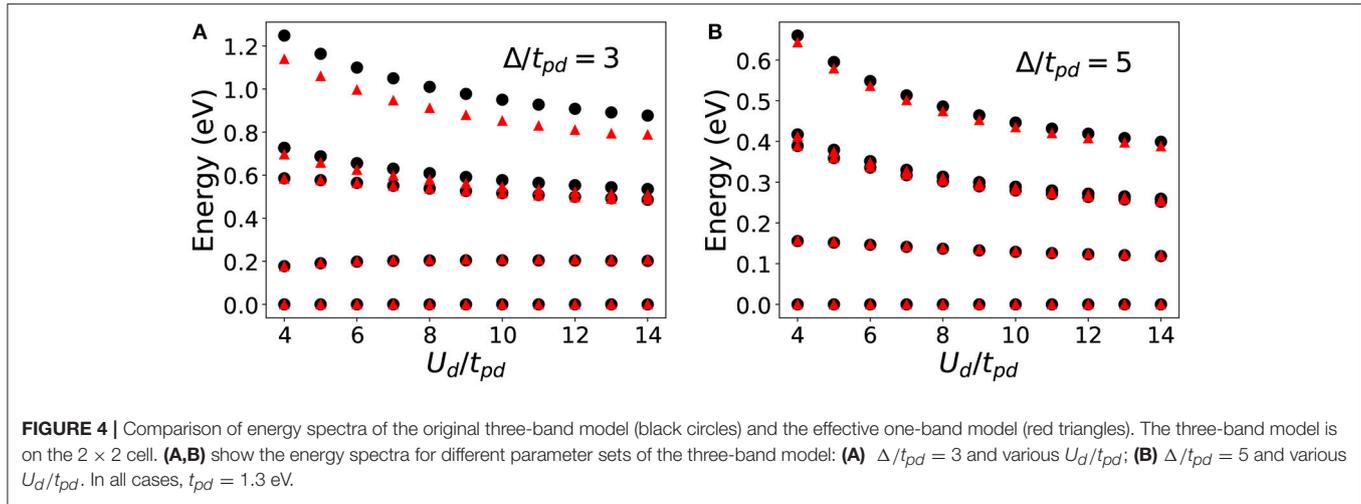
An important check for the one-band model is its ability to reproduce the low energy gaps of the three-band model; these have been compared in **Figure 4**. For the case of $\Delta/t_{pd} = 3$, we observe that for all U_d/t_{pd} the lowest three eigenstates were reproduced well. This model also reproduces the states outside of the DMD energy window, although with slightly larger errors. Similar trends are seen for the case of $\Delta/t_{pd} = 5$, with the noticeable difference being that the energy error of the highest state has reduced. This also reflects that the parameters obtained from DMD are, in general, dependent on the energy window of interest, a point which we will highlight shortly by investigating it systematically.

A promise of downfolding is the reduction of the size of the effective Hilbert space; allowing simulations of bigger unit

cells to be carried out. To show that this actually works well in practice for the three-band case, we consider the $2\sqrt{2} \times 2\sqrt{2}$ square unit cell, comprising of 8 copper and 16 oxygen orbitals. For representative test cases, we performed exact diagonalization calculations at half filling; the Hilbert space comprises of 112,911,876 basis states. Roughly 200 Lanczos iterations were carried out, enabling convergence of the lowest four energies. We compared the lowest gaps with the corresponding calculation on the one-band model on the same square geometry, with a Hilbert space size of only 4,900, using the downfolded parameters obtained from the smaller 2×2 cell.

Our results are summarized in **Figure 5**. **Figure 5A** shows six representative parameter sets of the three-band model and the corresponding downfolded one-band parameters. **Figures 5B,C** show the lowest three energy gaps for representative values of $U_d/t_{pd} = 4, 8, 12$ for $\Delta/t_{pd} = 3$ and $\Delta/t_{pd} = 5$, respectively. In all cases, the agreement between the three-band and one-band models is remarkably good. The energy gap error of the lowest gap is within 0.0004 eV (1% relative error). The largest error in the third gap is of the order of 0.005 eV (3% relative error). These results indicate the reliability of the downfolding procedure and highlight its predictive power.

Until this point, all our results focused on downfolding using only the lowest three eigenstates of the 2×2 cell. We now explore



the effect of increasing the energy window, by including higher eigenstates, using our test example of $U_d/t_{pd} = 8$ and $\Delta/t_{pd} = 3$. To do so, we now use all six low energy eigenstates for optimizing the cost function in Equation 15. We find similar (but not exactly the same) values of α_i compared to the case when only the three lowest states were used. The fact that a solution with small cost can be attained confirms our expectation that the entire low energy space of six states is consistently described by a set of \tilde{d}_i operators.

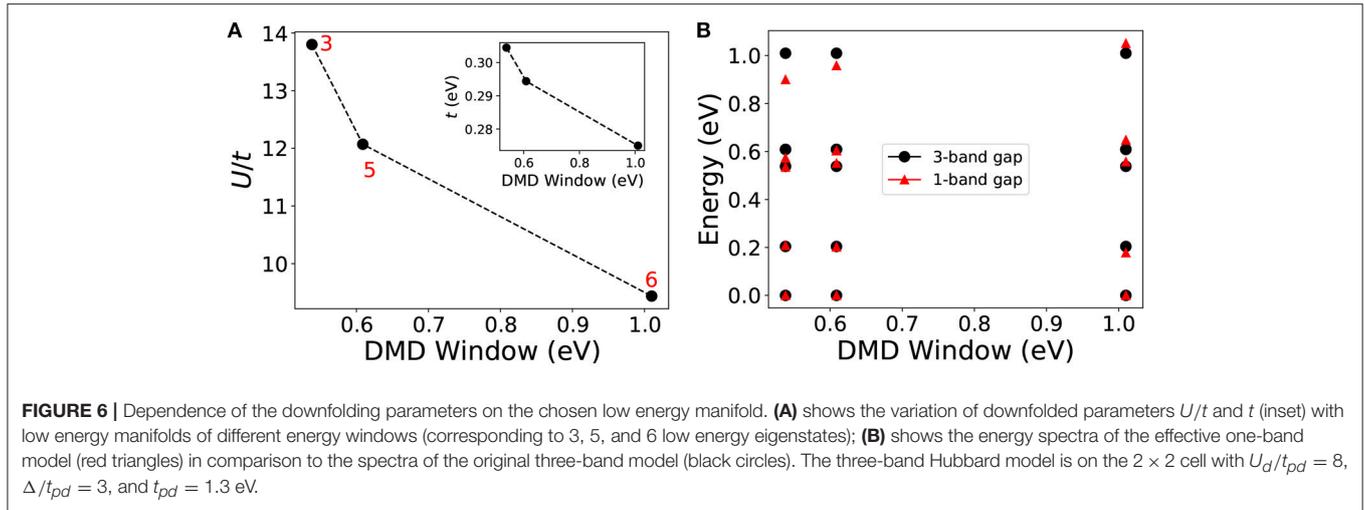
However, as **Figure 6A** shows, the estimates of U/t and t depend on how many eigenstates are used in the DMD procedure. This is because the DMD aims to provide the one-band description that best describes *all* states in a given window. If the model is not perfect within a given energy window, an energy dependent model is expected, consistent with the renormalization group perspective. For our test example, increasing the number of eigenstates from three to six changed U/t from 13.8 to 9.44 and t from 0.3045 to 0.2750 eV[†].

[†]When three states were used for optimizing the orbitals and for the DMD, we found $U/t \approx 13.45$ and $t = 0.3025$ eV. This is because \tilde{d} are slightly different in the two cases.

The features associated with the energy dependence are further confirmed in **Figure 6B**, which shows a comparison of energy gaps of the three-band and downfolded one-band model on the 2×2 cell. When only three states are used, the one-band (nearest neighbor) Hubbard model is insufficient for *accurately* describing states outside the window. When all six states are used, the DMD tries to minimize the error of the largest energy gap at the cost of errors in the smaller energy gaps. One could of course choose a different parameterization, say with additional next nearest neighbor t' , for which is may be possible to reduce this energy dependence significantly and thus have a model that describes the smaller and larger energy scales equally well.

3.2. One Dimensional Hydrogen Chain

We now move on to one of the simplest extended *ab initio* systems, a hydrogen chain in one dimension with periodic boundary conditions. The one-dimensional hydrogen chain has been used as a model for validating a variety of modern *ab initio* many-body methods [27]. We consider the case of 10 atoms with periodic boundary conditions and work in a regime where the inter-atomic distance r is in the range $1.5 - 3.0$ Å, such that the system is well described in terms of primarily *s*-like orbitals.



For a given r , we first obtain single-particle Kohn-Sham orbitals from a set of spin-unrestricted and spin-restricted DFT calculations with the PBE functional [28]. The localized orbital basis upon which the RDMs (descriptors) are evaluated is obtained by generating intrinsic atomic orbitals (IAO) [29] from the Kohn-Sham orbitals orthogonalized using the Löwdin procedure (see **Figure 7**). These are the orbitals that enter the one-band Hubbard Hamiltonian. Then, to generate a database of wavefunctions needed for the DMD, we produce a set of Slater-Jastrow wavefunctions consisting of singles and doubles excitations to the Slater determinant:

$$|s\rangle = e^J \left[a_{i\eta}^\dagger a_{k\eta} |KS\rangle \right], \quad (16a)$$

$$|d\rangle = e^J \left[a_{i\eta}^\dagger a_{j\eta'}^\dagger a_{k\eta} a_{l\eta} |KS\rangle \right], \quad (16b)$$

where $|KS\rangle$ is the Slater determinant of occupied Kohn-Sham orbitals, $\eta \neq \eta'$ are spin indices, and a_i^\dagger (a_i) is a single-electron creation (destruction) operator corresponding to a particular Kohn-Sham orbital. The k, l indices label occupied orbitals in the original Slater determinant, while i, j are virtual orbitals. e^J is a Jastrow factor optimized by minimizing the variance of the local energy.

We compute the energies (expectation values of the Hamiltonian) and the RDMs for each wave function within DMC. By computing the trace of the resulting 1-RDMs, we verify that all the electrons present in the system are represented within the localized basis of s -like orbitals. If the trace of the 1-RDM deviates from the nominal number of electrons for a particular state by more than some chosen threshold—2% in this example—it indicates that some orbitals are occupied ($2s$ - or $2p$ -like orbitals for hydrogen) that are not represented within the localized IAO basis used for computing the descriptors. Hence, these states do not exist within the $\mathcal{L}\mathcal{E}$ space, and cannot be described by a one-band s -orbital model. We exclude such states from the wave function set. The acquired data is then used in DMD to downfold to a one-band Hubbard Hamiltonian.

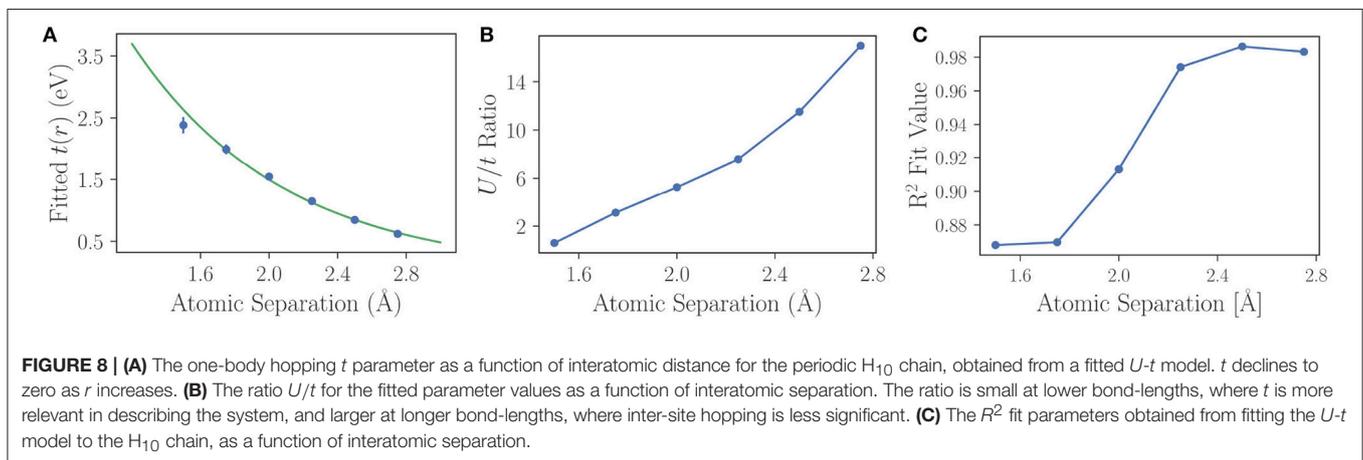
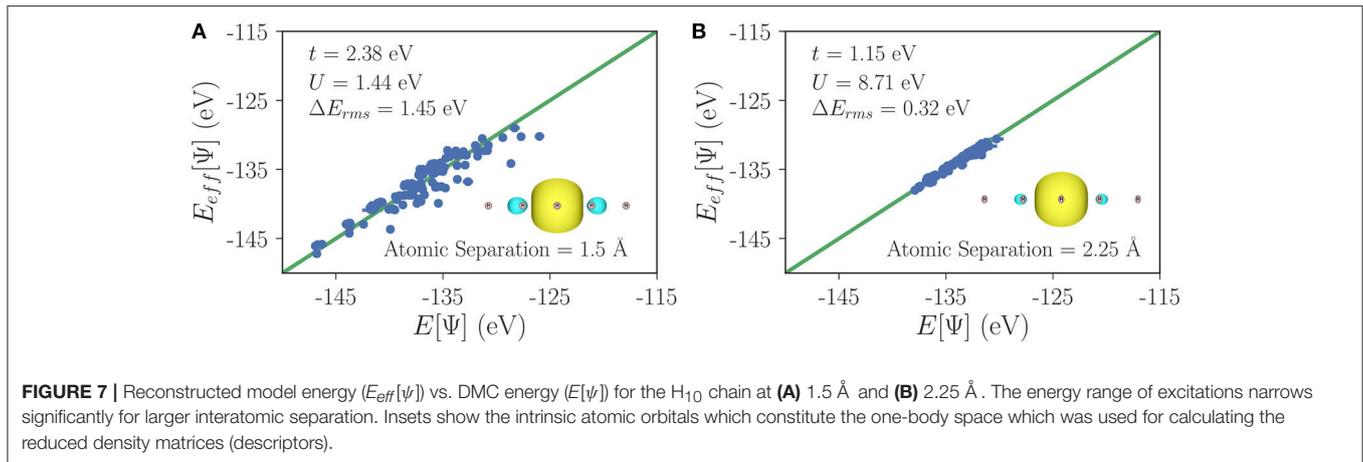
Figure 7 shows the fitting results of the energy functional $E[\Psi]$ within the sampled $\mathcal{L}\mathcal{E}$ for two representative distances (1.5 and 2.25\AA). As seen in **Figure 7**, the model $E_{\text{eff}}[\Psi]$ reproduces the *ab initio* $E[\Psi]$ up to certain error that decreases with atomic separation. That is, the fitted Hubbard model provides a more accurate description as separation distance increases, and the system becomes more atomic-like.

Figure 8 shows the fitted values of the downfolded parameters t and U/t at various distances. t decreases as the interatomic distance increases, and the value of U/t increases. The single-band Hubbard model qualitatively captures how the system approaches the atomic limit, in which t becomes zero.

The R^2 values obtained from fitting the descriptors to the *ab initio* energy (see **Figure 8C**) also show that the single-band Hubbard model is a good description of the system at large distances, but not at small distances. This is primarily because at small distances, the low energy spectrum involves the dynamics of other degrees of freedom that are not included in the effective model (e.g., $2s$ and $2p$ orbitals). Other interaction terms beyond the on-site Hubbard U , such as nearest-neighbor Coulomb interactions and Heisenberg coupling, can also become significant. Without including higher orbitals or additional many-body interaction terms, the model gives rise to an incorrect insulator state at small distances. Conversely, at larger separations ($r > 1.8\text{\AA}$), where the system is in an insulator phase [30], the model provides a better description.

3.3. Graphene and Hydrogen Honeycomb Lattice

Our third example highlights the role of the high energy degrees of freedom not present in the low energy description but which are instrumental in renormalizing the effective interactions. We demonstrate this by considering the case of graphene, and by comparing it to artificially constructed counterparts without the high energy electrons. Although many electronic properties of graphene can be adequately described by a noninteracting tight-binding model of π electrons [31], electron-electron interactions are crucial for explaining a wide range of phenomena observed



in experiments [32]. In particular, electron screening from σ bonding renormalizes the low energy plasmon frequency of the π electrons [33]. In fact a system of π electrons with bare Coulomb interactions has been shown to be an insulator instead of a semimetal [33–36]. Using DMD, we demonstrate how the screening effect of σ electrons is manifested in the low energy effective model of graphene.

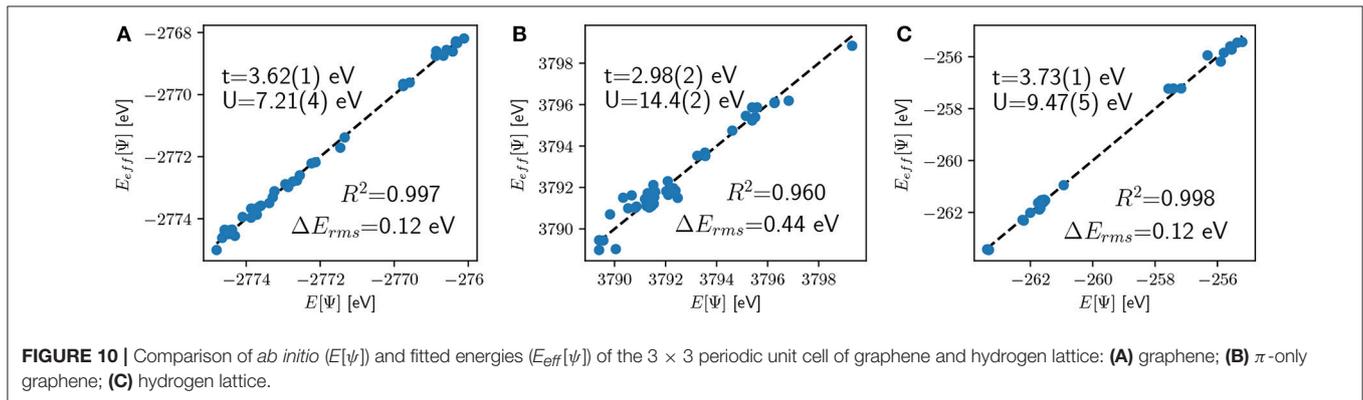
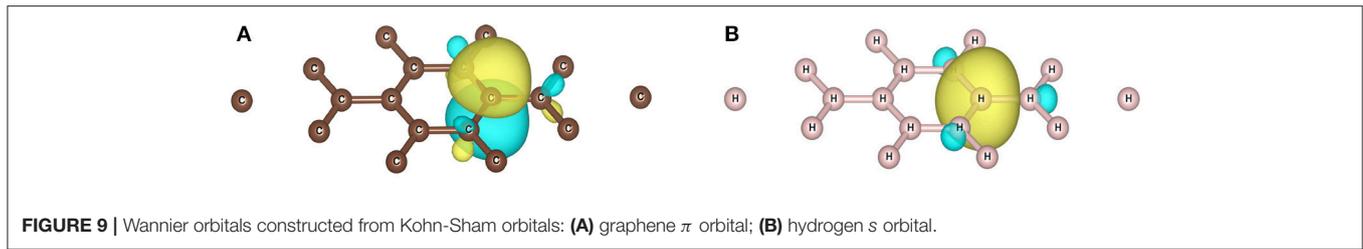
In order to disentangle the screening effect of σ electrons from the bare interactions between π electrons, we apply DMD to three different systems, graphene, π -only graphene, and a honeycomb lattice of hydrogen atoms. In π -only graphene, the σ electrons are replaced with a static constant negative charge background. The role of σ electrons is then clarified by comparing the effective model Hamiltonians of these two systems. The hydrogen system we study has the same lattice constant $a = 2.46$ Å as graphene, which has a similar Dirac cone dispersion as graphene [33].

By constructing the one-body space by Wannier localizing Kohn-Sham orbitals obtained from DFT calculations (see **Figure 9**), we verify that the low energy degrees of freedom correspond to the π orbitals in graphene and its π -only system and s orbitals in hydrogen; these enter the effective one-band Hubbard model description in Equation 11. Due

to the vanishing density of states at the Fermi level, the Coulomb interaction remains long-ranged, in contrast to usual metals where the formation of electron-hole pairs screens the interactions strongly [33]. However, for certain aspects, the long ranged part can be considered as renormalizing the onsite Coulomb interaction U at low energy [15, 37].

To estimate the one-band Hubbard parameters, we used the DMD method using a set of 50 Slater-Jastrow wave functions that correspond to the electron-hole excitations within the π channel for the graphene systems or s channel for the hydrogen system. In particular, for graphene, the Slater-Jastrow wave functions are constructed from occupied σ bands and occupied π bands, whereas for π -only graphene, Slater-Jastrow wave functions are constructed from occupied π Kohn-Sham orbitals of graphene. The *ab initio* simulations were performed on a 3×3 cell (32 carbons or hydrogens) and the energy and RDMs of these wave functions were evaluated with VMC. The error bars on our downfolded parameters are estimated using the jackknife method [38]. The results from our calculations are summarized in **Figure 10**.

We find that the one-band Hubbard model describes graphene and hydrogen very well, as is seen from the fact that R^2 is close to 1 for the fits. Our fits are shown in **Figure 10**.



For both graphene and hydrogen, U/t is smaller than the critical value of the semimetal-insulator transition $(U/t)_c \approx 3.8$ for the honeycomb lattice [39], which is consistent with both systems being semimetals. The Fermi velocity of graphene estimated from the downfolded parameters [$t = 3.62(1)$ eV and $U = 7.21(4)$ eV] using the Hartree-Fock approximation is 1.2×10^6 m/s, which is consistent with the experimental value 1.1×10^6 m/s [40]. Graphene and hydrogen have similar hopping constant t , consistent with the fact that they have similar band dispersions near the Dirac point. However, the difference in their high energy structure manifests itself as differently renormalized electron-electron interactions, explaining the difference in U . Most prominently, the π -only system has much larger U/t (~ 4.9) compared to graphene, which is large enough to push it into the insulating (antiferromagnetic) phase. Thus, downfolding shows the clear significance of σ electrons in renormalizing the effective onsite interactions of the π orbitals, making graphene a weakly interacting semimetal instead of an insulator.

3.4. FeSe Diatomic Molecule

Transition metal systems are often difficult to model because the low energy physics involves multiple orbitals and various kinds of couplings. This is seen in the proliferation of models for transition metals, which include terms like spin-spin coupling, spin-orbital coupling, hopping, Hund's coupling, and so on. Models containing all possible descriptors are unwieldy, and it is difficult to determine which degrees of freedom are needed for a minimal model to reproduce an interesting effect. Transition metal systems are challenging to describe using most electronic structure methods because of the strong electron correlations

and multiple oxidation states possible in these systems. Fixed-node DMC has been shown to be a highly accurate method on transition metal materials for improving the description of the ground state properties and energy gaps [41–44]. In this section, we applied DMD using fixed-node DMC to quantify the importance of different terms in the effective Hamiltonian in a FeSe diatomic molecule with a bond length equal to that of the iron based superconductor, FeSe [45], in order to help identify the descriptors that may be relevant in the bulk material.

We considered a low-energy space spanned by the Se $4p$, Fe $3d$, and Fe $4s$ orbitals. We sampled singles and doubles excitations from a reference Slater determinant of Kohn-Sham orbitals taken from DFT calculations with the PBE0 functional with total spin 0, 2, and 4, which were then multiplied by a Jastrow factor and further optimized using fixed-node DMC. After this procedure, 241 states were within a low energy window of 8 eV. Of these, eight states had a significant iron $4p$ component, which excludes them from the low-energy subspace. This left us with 233 states in the low-energy subspace.

We considered a set of 21 possible descriptors consisting of local operators on the iron $4s$, iron $3d$ states, and selenium $4p$ states, which is a total of 9 single-particle orbitals. We used the same IAO construction as section 3.2 to generate the basis for these operators. At the one-body level, we considered orbital energy descriptors:

$$\begin{aligned} \epsilon_s n_s, & \quad \epsilon_{\pi, \text{Se}}(n_{p_x} + n_{p_y}), & \quad \epsilon_z n_{p_z}, \\ \epsilon_{z^2} n_{d_{z^2}}, & \quad \epsilon_{\pi, \text{Fe}}(n_{d_{xz}} + n_{d_{yz}}), & \quad \epsilon_{\delta}(n_{d_{xy}} + n_{d_{x^2-y^2}}), \end{aligned} \quad (17)$$

The labels π and δ signify that these are the π and δ -bonding orbitals for each atom. We also considered the symmetry-allowed

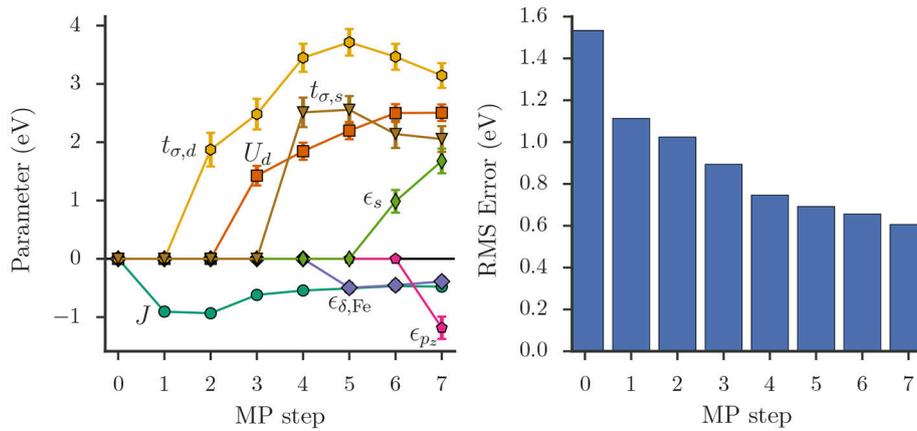


FIGURE 11 | (Left) Parameter values for each fit generated in the MP algorithm, labeled at the step where they are included in the model. A zero value indicates that parameter is not yet added to the model. The sign of J is consistent with Hund's rules, and the signs of $t_{\sigma,d}$ and $t_{\sigma,s}$ are consistent with Se being located in positive z with respect to Fe. **(Right)** RMS error of each model generated by MP as the algorithm includes parameters. The RMS error of the largest model considered was 0.61 eV.

hopping terms:

$$t_{\sigma,d} \sum_{\eta} \left(c_{d_{2,\eta}}^{\dagger} c_{p_{z,\eta}} + \text{h.c.} \right), \quad t_{\sigma,s} \sum_{\eta} \left(c_{s,\eta}^{\dagger} c_{p_{z,\eta}} + \text{h.c.} \right),$$

$$t_{\pi} \sum_{\eta} \left(c_{d_{xz,\eta}}^{\dagger} c_{p_{x,\eta}} + c_{d_{yz,\eta}}^{\dagger} c_{p_{y,\eta}} + \text{h.c.} \right). \quad (18)$$

As before, η represents the spin index. At the two-body level, we considered Hubbard interactions:

$$U_p \sum_{i \in p} n_{i,\uparrow} n_{i,\downarrow}, \quad U_{d,\delta} \sum_{i \in \{d_{xy}, d_{x^2-y^2}\}} n_{i,\uparrow} n_{i,\downarrow},$$

$$U_d \sum_{i \in d} n_{i,\uparrow} n_{i,\downarrow}, \quad U_{d,\pi} \sum_{i \in \{d_{xz}, d_{yz}\}} n_{i,\uparrow} n_{i,\downarrow}, \quad U_{d_2} n_{d_2,\uparrow} n_{d_2,\downarrow}, \quad (19)$$

where p refers to the Se-4p orbitals and d refers to the Fe-3d orbitals. Importantly, we also accounted for the Hund's coupling terms for the iron atom:

$$J \sum_{\substack{i \neq j \\ i,j \in d}} S_i \cdot S_j, \quad J_{\delta} S_{d_{xy}} \cdot S_{d_{x^2-y^2}}, \quad J_{\delta,d_2} (S_{d_{xy}} + S_{d_{x^2-y^2}}) \cdot S_{d_2},$$

$$J_{\pi} S_{d_{xz}} \cdot S_{d_{yz}}, \quad J_{\pi,d_2} (S_{d_{xz}} + S_{d_{yz}}) \cdot S_{d_2}, \quad J_{\pi,\delta} (S_{d_{xz}} + S_{d_{yz}}) \cdot (S_{d_{xy}} + S_{d_{x^2-y^2}}), \quad (20)$$

Finally, we also added a nearest neighbor Hubbard interaction:

$$V \sum_{i \in p, j \in d} n_i n_j.$$

To generate a minimal description of the system, we employed a matching pursuit (MP) method [17]. MP chooses to add descriptors based on their correlation with the residual of the linear fit. We started with a model that only consists of E_0 . The Hund's coupling descriptor [the first term in Equation 20]

has the largest correlation coefficient with the residual fit, so it is added first. The fact that the Hund's coupling is chosen first in MP is consistent with several studies in the literature which show that it is important for the magnetic properties of bulk FeSe [46–49]. Next, MP includes the descriptor that correlates most strongly with the residuals of this first minimal model, in this case the hopping between d and p σ -symmetry orbitals. We repeated this procedure until the RMS error did not improve more than 0.05 eV upon adding a new parameter. This criterion was chosen to strike a balance between the complexity of the model and the accuracy in reproducing the data set.

The following model was produced:

$$H_{\text{eff}} = \epsilon_{\delta,Fe} (n_{d_{xy}} + n_{d_{x^2-y^2}}) + \epsilon_s n_s + \epsilon_z n_{p_z}$$

$$+ t_{\sigma,d} \sum_{\eta} \left(c_{d_{2,\eta}}^{\dagger} c_{p_{z,\eta}} + \text{h.c.} \right) + t_{\sigma,s} \sum_{\eta} \left(c_{s,\eta}^{\dagger} c_{p_{z,\eta}} + \text{h.c.} \right)$$

$$+ U_d \sum_{i \in d} n_{i,\uparrow} n_{i,\downarrow} + J \sum_{\substack{i \neq j \\ i,j \in d}} S_i \cdot S_j + E_0. \quad (21)$$

E_0 is an overall energy shift, also included as a fit parameter. The parameter values and corresponding error of each model produced by MP are shown in **Figure 11**. Note that all parameters may change at each step because the entire model is refitted when an additional parameter is included in each iteration. The parameters are smoothly varying with the inclusion of new parameters, and they take the correct signs based on symmetry (where applicable). The RMS error decreases with each additional parameter, but less so as the algorithm appends additional parameters. Eventually the diminishing improvements do not merit the additional complexity of more parameters.

4. CONCLUSION AND FUTURE PROSPECTS

The density matrix downfolding (DMD) technique uses data derived from low-energy approximate solutions to a high energy Hamiltonian to systematically determine an effective Hamiltonian that describes the low-energy behavior of the system. It is based on several rather simple proofs which occupy a role similar to the variational principle; they allow us to know which effective models are closer to the correct solution than others. The method is very general and does not require a quasiparticle picture to apply, and neither does it have double-counting issues. It treats all interactions on an equal footing, so hopping parameters are naturally modified by interaction parameters and so on. While most of the applications have used the first principles QMC method to obtain the low-energy solutions, the method is completely general and can be used with any solution method that can produce high quality energy and RDMs. We have discussed several examples to present the conceptual and algorithmic aspects of DMD.

The resultant lattice model can be efficiently and accurately solved for large system sizes [50] using techniques designed and suited for small local Hilbert spaces; these include exact or selected diagonalization [51–53], density matrix renormalization group (DMRG) [54], tensor networks [55–57], dynamical mean field theory (DMFT) [58], density matrix embedding (DMET) [59], and lattice QMC methods [60–68]. These methods have also been used to obtain excited states, dynamical correlation functions and thermal properties, that have been difficult to obtain in *ab initio* approaches.

DMD, though conceptually simple, is still a method in its development stages, with room for algorithmic improvements and new applications. Advances in the field of inverse problems [69] could be incorporated into DMD to mitigate the problems associated with optimization and over-fitting. Here we briefly outline some aspects that need further research:

1. The wave function database ($|\Psi\rangle \in \mathcal{L}\mathcal{E}$): The DMD method relies crucially on the availability of a low energy space of *ab initio* wave functions. While these wave functions do not have to be eigenstates, automating their construction remains challenging and realistically requires knowledge of the physics to be described.
2. Optimal choice of basis functions. The second-quantized operators in the effective Hamiltonian correspond to a basis in the continuum. The quality of the model depends on the basis describing the changes between low-energy wave functions accurately.
3. Form of the low energy model Hamiltonian. While the exact effective Hamiltonian is unique, there may be many ways of approximating it with varying levels of compactness and accuracy.

REFERENCES

1. Wilson KG. The renormalization group: critical phenomena and the Kondo problem. *Rev Mod Phys.* (1975) **47**:773–840.

The advantage of the DMD framework is that all these can be resolved internally. Given a good sampling of $\mathcal{L}\mathcal{E}$, (2) and (3) can be resolved using regression. Given that (2) and (3) are correct or near correct, then (1) can be resolved by finding binding planes, as noted in section 2. The method thus has a degree of self consistency; it will return low errors only when 1–3 are correct.

We have shown applications to strongly correlated models (3-band), *ab initio* bulk systems hydrogen chain and graphene, and a transition metal molecule FeSe. The technique is on the verge of being applied to transition metal bulk systems; there are no major barriers to this other than a polynomially scaling computational cost and the substantial amount of work involved in parameterizing and fitting models to these systems. Looking into the future, we anticipate that this technique can help with the definition of a correlated materials genome—what effective Hamiltonian best describes a given material is highly relevant to its behavior.

AUTHOR CONTRIBUTIONS

HC, HZ, and LW conceived the initial DMD ideas and designed the project and organization of the paper. All authors contributed to the theoretical developments and various representative *ab initio* and lattice examples. All authors contributed to the analysis of the data, discussions and writing of the manuscript. LW oversaw the project; HC and HZ contributed equally to this work.

ACKNOWLEDGMENTS

We thank David Ceperley, Richard Martin, Cyrus Umrigar, Garnet Chan, Shiwei Zhang, Steven White, Lubos Mitas, So Hirata, Bryan Clark, Norm Tubman, Miles Stoudenmire, and Victor Chua for extremely useful and insightful discussions. This work was funded by the grant DOE FG02-12ER46875 (SciDAC) and NSF Grant No. DMR 1206242. HZ acknowledges support from Argonne Leadership Computing Facility, a U.S. Department of Energy, Office of Science User Facility under Contract DE-AC02-06CH11357. HC acknowledges support from the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award DE-FG02-08ER46544 for his work at the Institute for Quantum Matter (IQM). KW acknowledges support from the National Science Foundation under the Graduate Research Fellowship Program, Fellowship No. DGE-1144245. This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and the National Center for Supercomputing Applications.

2. Hubbard J. Electron correlations in narrow energy bands. *Proc R Soc Lond A Mat Phys Eng Sci.* (1963) **276**:238–57.
3. Kanamori J. Electron correlation and ferromagnetism of transition metals. *Prog Theor Phys.* (1963) **30**:275–89.

4. Chao KA, Spalek J, Oles AM. Kinetic exchange interaction in a narrow S-band. *J Phys C* (1977) **10**:L271.
5. Pavarini E, Dasgupta I, Saha-Dasgupta T, Jepsen O, Andersen OK. Band-structure trend in hole-doped cuprates and correlation with T_{cmax} . *Phys Rev Lett.* (2001) **87**:047003. doi: 10.1103/PhysRevLett.87.047003
6. Andersen OK, Saha-Dasgupta T. Muffin-tin orbitals of arbitrary order. *Phys Rev B* (2000) **62**:R16219–22. doi: 10.1103/PhysRevB.62.R16219
7. Aryasetiawan F, Imada M, Georges A, Kotliar G, Biermann S, Lichtenstein AI. Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys Rev B* (2004) **70**:195104. doi: 10.1103/PhysRevB.70.195104
8. Jeschke HO, Salvat-Pujol F, Valentí R. First-principles determination of Heisenberg Hamiltonian parameters for the spin- $\frac{1}{2}$ kagome antiferromagnet $ZnCu_3(OH)_6Cl_2$. *Phys Rev B* (2013) **88**:075106. doi: 10.1103/PhysRevB.88.075106
9. Haule K. Exact double counting in combining the dynamical mean field theory and the density functional theory. *Phys Rev Lett.* (2015) **115**:196403. doi: 10.1103/PhysRevLett.115.196403
10. Ten-no S. Stochastic determination of effective Hamiltonian for the full configuration interaction solution of quasi-degenerate electronic states. *J Chem Phys.* (2013) **138**:164126. doi: 10.1063/1.4802766
11. Zhou SQ, Ceperley DM. Construction of localized wave functions for a disordered optical lattice and analysis of the resulting Hubbard model parameters. *Phys Rev A* (2010) **81**:013402. doi: 10.1103/PhysRevA.81.013402
12. White SR. Numerical canonical transformation approach to quantum many-body problems. *J Chem Phys.* (2002) **117**:7472–82. doi: 10.1063/1.1508370
13. Yanai T, Chan GKL. Canonical transformation theory for multireference problems. *J Chem Phys.* (2006) **124**:194106. doi: 10.1063/1.2196410
14. Wagner LK. Types of single particle symmetry breaking in transition metal oxides due to electron correlation. *J Chem Phys.* (2013) **138**:094106. doi: 10.1063/1.4793531
15. Changlani HJ, Zheng H, Wagner LK. Density-matrix based determination of low-energy model Hamiltonians from ab initio wavefunctions. *J Chem Phys.* (2015) **143**:102814. doi: 10.1063/1.4927664
16. Tibshirani R. Regression shrinkage and selection via the Lasso. *J R Stat Soc Ser B* (1996) **58**:267–88.
17. Mallat SG, Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans Signal Process.* (1993) **41**:3397–415.
18. Bednorz JG, Müller KA. Possible high T_c superconductivity in the Ba-La-Cu-O system. *Z Phys B Cond Matt.* (1986) **64**:189–93.
19. Dagotto E. Correlated electrons in high-temperature superconductors. *Rev Mod Phys.* (1994) **66**:763–840.
20. Lee PA, Nagaosa N, Wen XG. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev Mod Phys.* (2006) **78**:17–85. doi: 10.1103/RevModPhys.78.17
21. Emery VJ. Theory of high- T_c superconductivity in oxides. *Phys Rev Lett.* (1987) **58**:2794–7.
22. Zhang FC, Rice TM. Effective Hamiltonian for the superconducting Cu oxides. *Phys Rev B* (1988) **37**:3759–61.
23. Hybertsen MS, Schlüter M, Christensen NE. Calculation of Coulomb-interaction parameters for La_2CuO_4 using a constrained-density-functional approach. *Phys Rev B* (1989) **39**:9028–41.
24. Hybertsen MS, Stechel EB, Schluter M, Jennison DR. Renormalization from density-functional theory to strong-coupling models for electronic states in Cu-O materials. *Phys Rev B* (1990) **41**:11068–72.
25. Kent PRC, Saha-Dasgupta T, Jepsen O, Andersen OK, Macridin A, Maier TA, et al. Combined density functional and dynamical cluster quantum Monte Carlo calculations of the three-band Hubbard model for hole-doped cuprate superconductors. *Phys Rev B* (2008) **78**:035132. doi: 10.1103/PhysRevB.78.035132
26. Changlani HJ, Ghosh S, Pujari S, Henley CL. Emergent spin excitations in a bethe lattice at percolation. *Phys Rev Lett.* (2013) **111**:157201. doi: 10.1103/PhysRevLett.111.157201
27. Motta M, Ceperley DM, Chan GKL, Gomez JA, Gull E, Guo S, et al. Towards the solution of the many-electron problem in real materials: equation of state of the hydrogen chain with state-of-the-art many-body methods. *Phys Rev X* (2017) **7**:031059. doi: 10.1103/PhysRevX.7.031059
28. Perdew JP, Burke K, Ernzerhof M. Generalized gradient approximation made simple. *Phys Rev Lett.* (1996) **77**:3865–8.
29. Knizia G. Intrinsic atomic orbitals: an unbiased bridge between quantum theory and chemical concepts. *J Chem Theory Comput.* (2013) **9**:4834–43. doi: 10.1021/ct400687b
30. Stella L, Attaccalite C, Sorella S, Rubio A. Strong electronic correlation in the hydrogen chain: a variational Monte Carlo study. *Phys Rev B* (2011) **84**:245117. doi: 10.1103/PhysRevB.84.245117
31. Castro Neto AH, Guinea F, Peres NMR, Novoselov KS, Geim AK. The electronic properties of graphene. *Rev Mod Phys.* (2009) **81**:109. doi: 10.1103/RevModPhys.81.109
32. Kotov VN, Uchoa B, Pereira VM, Guinea F, Castro Neto AH. Electron-electron interactions in graphene: current status and perspectives. *Rev Mod Phys.* (2012) **84**:1067–125. doi: 10.1103/RevModPhys.84.1067
33. Zheng H, Gan Y, Abbamonte P, Wagner LK. Importance of σ bonding electrons for the accurate description of electron correlation in graphene. *Phys Rev Lett.* (2017) **119**:166402. doi: 10.1103/PhysRevLett.119.166402
34. Drut JE, Lähde TA. Is graphene in vacuum an insulator? *Phys Rev Lett.* (2009) **102**:026802. doi: 10.1103/PhysRevLett.102.026802
35. Drut JE, Lähde TA. Critical exponents of the semimetal-insulator transition in graphene: a Monte Carlo study. *Phys Rev B* (2009) **79**:241405. doi: 10.1103/PhysRevB.79.241405
36. Smith D, von Smekal L. Monte Carlo simulation of the tight-binding model of graphene with partially screened Coulomb interactions. *Phys Rev B* (2014) **89**:195429. doi: 10.1103/PhysRevB.89.195429
37. Schüler M, Rösner M, Wehling TO, Lichtenstein AI, Katsnelson MI. Optimal hubbard models for materials with nonlocal coulomb interactions: graphene, silicene, and benzene. *Phys Rev Lett.* (2013) **111**:036601. doi: 10.1103/PhysRevLett.111.036601
38. Efron B, Stein C. The jackknife estimate of variance. *Ann Stat.* (1981) **9**:586–96.
39. Sorella S, Otsuka Y, Yunoki S. Absence of a spin liquid phase in the hubbard model on the honeycomb lattice. *Sci Rep.* (2012) **2**:1650158. doi: 10.1142/S0217979216501587
40. Siegel DA, Park C-H, Hwang C, Deslippe J, Fedorov AV, Louie SG, et al. Many-body interactions in quasi-freestanding graphene. *Proc Natl Acad Sci USA.* (2011) **108**:11365–9. doi: 10.1073/pnas.1100242108
41. Foyevtsova K, Krogel JT, Kim J, Kent PRC, Dagotto E, Reboledo FA. *Ab initio* quantum Monte Carlo calculations of spin superexchange in cuprates: the benchmarking case of Ca_2CuO_3 . *Phys Rev X* (2014) **4**:031003. doi: 10.1103/PhysRevX.4.031003
42. Wagner LK, Abbamonte P. Effect of electron correlation on the electronic structure and spin-lattice coupling of high- T_c cuprates: quantum Monte Carlo calculations. *Phys Rev B* (2014) **90**:125129. doi: 10.1103/PhysRevB.90.125129
43. Zheng H, Wagner LK. Computation of the correlated metal-insulator transition in vanadium dioxide from first principles. *Phys Rev Lett.* (2015) **114**:176401. doi: 10.1103/PhysRevLett.114.176401
44. Wagner LK, Ceperley DM. Discovering correlated fermions using quantum Monte Carlo. *arXiv:160201344* (2016).
45. Kumar RS, Zhang Y, Sinogeikin S, Xiao Y, Kumar S, Chow P, et al. Crystal and electronic structure of FeSe at high pressure and low temperature. *J Phys Chem B* (2010) **114**:12597–606. doi: 10.1021/jp1060446
46. De'Medici L. Hund's coupling and its key role in tuning multiorbital correlations. *Phys Rev B Cond Matt Mater Phys.* (2011) **83**:205112. doi: 10.1103/PhysRevB.83.205112
47. de Medici L, Mravlje J, Georges A. Janus-faced influence of Hund's rule coupling in strongly correlated materials. *Phys Rev Lett.* (2011) **107**:256401. doi: 10.1103/PhysRevLett.107.256401
48. Georges A, Medici Ld, Mravlje J. Strong electronic correlations from Hund's coupling. *Annu Rev Cond Matt Phys.* (2013) **4**:137–78. doi: 10.1146/annurev-conmatphys-020911-125045
49. Busemeyer B, Dagrada M, Sorella S, Casula M, Wagner LK. Competing collinear magnetic structures in superconducting FeSe by first-principles quantum Monte Carlo calculations. *Phys Rev B* (2016) **94**:035108. doi: 10.1103/PhysRevB.94.035108
50. LeBlanc JPF, Antipov AE, Becca F, Bulik IW, Chan GKL, Chung CM, et al. Solutions of the two-dimensional hubbard model: benchmarks and results

- from a wide range of numerical algorithms. *Phys Rev X* (2015) 5:041041. doi: 10.1103/PhysRevX.5.041041
51. Raedt HD, Frick M. *Stochastic diagonalization*. *Phys Rep.* (1993) 231:107–49.
 52. Tubman NM, Lee J, Takeshita TY, Head-Gordon M, Whaley KB. A deterministic alternative to the full configuration interaction quantum Monte Carlo method. *J Chem Phys.* (2016) 145:044112. doi: 10.1063/1.4955109
 53. Holmes AA, Tubman NM, Umrigar CJ. Heat-bath configuration interaction: an efficient selected configuration interaction algorithm inspired by heat-bath sampling. *J Chem Theory Comput.* (2016) 12:3674–80. doi: 10.1021/acs.jctc.6b00407
 54. White SR. Density matrix formulation for quantum renormalization groups. *Phys Rev Lett.* (1992) 69:2863–6.
 55. Verstraete F, Cirac JJ. Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions. *arXiv:cond-mat/0407066* (2004).
 56. Changlani HJ, Kinder JM, Umrigar CJ, Chan GKL. Approximating strongly correlated wave functions with correlator product states. *Phys Rev B* (2009) 80:245116. doi: 10.1103/PhysRevB.80.245116
 57. Neuscammen E, Changlani H, Kinder J, Chan GKL. Nonstochastic algorithms for Jastrow-Slater and correlator product state wave functions. *Phys Rev B* (2011) 84:205132. doi: 10.1103/PhysRevB.84.205132
 58. Kotliar G, Savrasov SY, Haule K, Oudovenko VS, Parcollet O, Marianetti CA. Electronic structure calculations with dynamical mean-field theory. *Rev Mod Phys.* (2006) 78:865–951. doi: 10.1103/RevModPhys.78.865
 59. Knizia G, Chan GKL. Density matrix embedding: a simple alternative to dynamical mean-field theory. *Phys Rev Lett.* (2012) 109:186404. doi: 10.1103/PhysRevLett.109.186404
 60. Scalapino DJ. Some results from quantum monte carlo studies of the 2D Hubbard model. *J Low Temper Phys.* (1994) 95:169–76.
 61. Trivedi N, Ceperley DM. Ground-state correlations of quantum antiferromagnets: a Green-function Monte Carlo study. *Phys Rev B* (1990) 41:4552–69.
 62. Zhang S, Krakauer H. Quantum Monte Carlo method using phase-free random walks with slater determinants. *Phys Rev Lett.* (2003) 90:136401. doi: 10.1103/PhysRevLett.90.136401
 63. Syljuåsen OF, Sandvik AW. Quantum Monte Carlo with directed loops. *Phys Rev E* (2002) 66:046701. doi: 10.1103/PhysRevE.66.046701
 64. Boninsegni M, Prokof'ev NV, Svistunov BV. Worm algorithm and diagrammatic Monte Carlo: a new approach to continuous-space path integral Monte Carlo simulations. *Phys Rev E* (2006) 74:036701. doi: 10.1103/PhysRevE.74.036701
 65. Booth GH, Thom AJW, Alavi A. Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space. *J Chem Phys.* (2009) 131:054106. doi: 10.1063/1.3193710
 66. Petruzielo FR, Holmes AA, Changlani HJ, Nightingale MP, Umrigar CJ. Semistochastic projector Monte Carlo method. *Phys Rev Lett.* (2012) 109:230201. doi: 10.1103/PhysRevLett.109.230201
 67. Holmes AA, Changlani HJ, Umrigar CJ. Efficient heat-bath sampling in fock space. *J Chem Theory Comput.* (2016) 12:1561–71. doi: 10.1021/acs.jctc.5b01170
 68. Booth GH, Grüneis A, Kresse G, Alavi A. Towards an exact description of electronic wavefunctions in real solids. *Nature* (2013) 493:365–70. doi: 10.1038/nature11770
 69. Nguyen HC, Zecchina R, Berg J. Inverse statistical problems: from the inverse Ising problem to data science. *Adv Phys.* (2017) 66:197–261. doi: 10.1080/00018732.2017.1341604

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation, though no other collaboration, with the authors BB, KW and LW.

Copyright © 2018 Zheng, Changlani, Williams, Busemeyer and Wagner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

In section 3.1 we discussed parameterizing the transformation \mathbf{T} as a 4×12 matrix for the 2×2 unit cell, in terms of α_1 , α_2 , α_3 and α_4 . Using the numbering of the orbitals corresponding to **Figure 2**, the explicit form of \mathbf{T} is,

$$\mathbf{T} = \begin{pmatrix} F & \alpha_2 & \alpha_2 & \alpha_4 & \alpha_1 & \alpha_1 & -\alpha_1 & -\alpha_1 & \alpha_3 & -\alpha_3 & \alpha_3 & -\alpha_3 \\ \alpha_2 & F & \alpha_4 & \alpha_2 & \alpha_3 & -\alpha_1 & \alpha_1 & -\alpha_3 & -\alpha_3 & \alpha_3 & \alpha_3 & -\alpha_1 \\ \alpha_2 & \alpha_4 & F & \alpha_2 & -\alpha_1 & \alpha_3 & -\alpha_3 & \alpha_1 & \alpha_1 & -\alpha_1 & -\alpha_3 & \alpha_3 \\ \alpha_4 & \alpha_2 & \alpha_2 & F & -\alpha_3 & -\alpha_3 & \alpha_3 & \alpha_3 & -\alpha_1 & \alpha_1 & -\alpha_1 & \alpha_1 \end{pmatrix} \quad (22)$$

where we have defined $F \equiv \sqrt{1 - 4\alpha_1^2 - 2\alpha_2^2 - 4\alpha_3^2 - \alpha_4^2}$.

A concrete and representative example of our results, shown in section 3.1 for the 2×2 cell, is explained for the case of $U_d/t_{pd} = 8$ and $\Delta/t_{pd} = 3$. The first task was to obtain the optimal transformation \mathbf{T} for which the lowest three eigenstates ($s = 1, 2, 3$) of the three-band model were used for computing the cost in Equation 15. The minimum of the cost was determined by

a brute force scan in the four dimensional space of α 's. Using a linear grid spacing of 0.002, we found $\alpha_1 = 0.216$, $\alpha_2 = 0.042$, $\alpha_3 = 0.018$, and $\alpha_4 = 0.016$. The two terms in the cost i.e., the trace and orthogonality conditions are individually satisfied to a relative error of $<0.5\%$.

$\langle c_i^\dagger c_j \rangle_s$ and $\langle c_{j,\uparrow}^\dagger c_{m,\downarrow}^\dagger c_{n,\downarrow} c_{k,\uparrow} \rangle_s$ were computed from the exact knowledge of the three-band model eigenstates and hence $\langle \tilde{d}_{i,\eta}^\dagger \tilde{d}_{j,\eta} \rangle_s$ and $\langle \tilde{n}_{i,\uparrow} \tilde{n}_{i,\downarrow} \rangle_s$ are obtained once the optimal \mathbf{T} has been determined. As mentioned in the main text, the value of $\langle \tilde{d}_{1,\eta}^\dagger \tilde{d}_{2,\eta} \rangle_s$ provides estimates for U/t of the effective model by direct comparison of its value to that in the corresponding eigenstate in the one-band model. For our test example, the absolute values of $\langle \tilde{d}_{1,\uparrow}^\dagger \tilde{d}_{2,\uparrow} \rangle_s$ in states $s = 1, 2, 3$ are approximately 0.159, 0.142, and 0.084, respectively which correspond to $(U/t)_1 \approx 14.1$, $(U/t)_2 \approx 13.2$, $(U/t)_3 \approx 12.7$. Performing DMD with the three eigenenergies and their calculated RDMs gave $t = 0.3025$ eV and $U/t = 13.45$; the latter in the correct range of the other estimates.