Variational Discrete Action Theory

Zhengqian Cheng and Chris A. Marianetti

Phys. Rev. Lett. **126**, 206402 — Published 17 May 2021

DOI: 10.1103/PhysRevLett.126.206402

Variational Discrete Action Theory

Zhengqian Cheng and Chris A. Marianetti
Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027
(Dated: April 21, 2021)

Here we propose the Variational Discrete Action Theory (VDAT) to study the ground state properties of quantum many-body Hamiltonians. VDAT is a variational theory based on the sequential product density matrix (SPD) ansatz, characterized by an integer $\mathcal{N}$, which monotonically approaches the exact solution with increasing $\mathcal{N}$. To evaluate the SPD, we introduce a discrete action and a corresponding integer time Green's function. We use VDAT to *exactly* evaluate the SPD in two canonical models of interacting electrons: the Anderson impurity model (AIM) and the $d = \infty$ Hubbard model. For the latter, we evaluate $\mathcal{N} = 2 - 4$, where $\mathcal{N} = 2$ recovers the Gutzwiller approximation (GA), and we show that $\mathcal{N} = 3$, which exactly evaluates the Gutzwiller-Baeriswyl wave function, provides a truly minimal yet precise description of Mott physics with a cost similar to the GA. VDAT is a flexible theory for studying quantum Hamiltonians, competing both with state-of-the-art methods and simple, efficient approaches all within a single framework.

Computing the ground state properties of quantum many-body Hamiltonians is a fundamental task in physics. A common strategy to approximately solve a Hamiltonian is the use of variational wave functions, which allows one to find the best solution within some fraction of the Hilbert space. A generic approach is to start from some reference wave function and apply some projector, as in the well known Jastrow[1] and Gutzwiller[2–4] variational wave functions. A key limitation to such approaches is that they often cannot be intelligently improved, meaning that it is difficult to increase the searchable region of Hilbert space efficiently. One approach to address this limitation is tensor network methods[5–7], where some control parameter increases accuracy at some computational cost; but these approaches have only proven well suited for low dimensional systems.

Here we propose a new class of variational density matrices: the sequential product density matrix (SPD). The SPD is motivated by the Trotter-Suzuki[8] decomposition, and is characterized by an integer $\mathcal{N}$. The SPD provides a paradigm for variational approaches in that the precision can be systematically improved by increasing $\mathcal{N}$, and it can be applied beyond low dimensions. In practice, such an ansatz is not useful unless one has a systematic and efficient approach for evaluating it. Our key development is the introduction of the discrete action theory (DAT) and the corresponding *integer time* Green's function (ITGF), which may be used for evaluating an SPD. The DAT has a perfect parallel to the standard many-body Green's function formalism, though with non-trivial differences. Many of the key ideas from traditional many-body theory can immediately be generalized to the DAT, such as the Dynamical Mean-Field Theory (DMFT)[9]. Using the DAT for evaluating the SPD, we can then perform the variational minimization to obtain the ground state, and we refer to this entire approach as the Variational Discrete Action Theory (VDAT). There is companion manuscript to this paper which provides extensive derivations and minimal examples to document the foundations of VDAT[10].

Given a Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$, where $\hat{H}_0$ is non-interacting and $\hat{V}$ is interacting, we motivate the SPD by considering the following variational wave function:

$$\exp(\gamma_1 \hat{H}_0) \exp(g_1 \hat{V}) ... \exp(\gamma_N \hat{H}_0) \exp(g_N \hat{V}) |\varphi_0\rangle, \quad (1)$$

where γ_i, g_i are variational parameters and $|\varphi_0\rangle$ is the ground state wave function of $\hat{H}_0$. Equation 1 can be viewed as a variational application of the Trotter-Suzuki decomposition[8, 11, 12], where the $N \to \infty$ limit will cover the exact ground state wave function. The essence of this ansatz was first proposed several decades ago by Dzierzawa *et. al* [13], motivated by the generalization of the Baeriswyl wave function[14, 15] by Otsuka[16]; and all of this work was motivated by improving upon the well known Gutzwiller wave function[2]. More recently, a unitary version of this wave function was proposed in the context of quantum computing by Farhi et. al[17], and further extended by Wecker *et. al* [18] and Grimsley *et. al*[19]. Our SPD further generalizes the idea behind Eq. 1.

Given a Hamiltonian with L spin orbitals, the SPD is given as

$$\hat{\varrho} = \exp(\boldsymbol{\gamma}_1 \cdot \hat{\boldsymbol{n}}) \hat{P}_1 \ldots \exp(\boldsymbol{\gamma}_{\mathcal{N}} \cdot \hat{\boldsymbol{n}}) \hat{P}_{\mathcal{N}} = \hat{\mathcal{P}}_1 \ldots \hat{\mathcal{P}}_{\mathcal{N}}, \quad (2)$$

where $\hat{P}_\tau$ is a generic interacting projector, and $\tau = 1, \ldots, \mathcal{N}$ is the integer time label; $\exp(\boldsymbol{\gamma}_\tau \cdot \hat{\boldsymbol{n}})$ is the noninteracting projector, where $\boldsymbol{\gamma}_\tau \cdot \hat{\boldsymbol{n}} \equiv \sum_{i=1}^{L} \sum_{j=1}^{L} [\boldsymbol{\gamma}_\tau]_{ij} [\hat{\boldsymbol{n}}]_{ij}$ and $[\hat{\boldsymbol{n}}]_{ij} = \hat{a}_i^\dagger \hat{a}_j$; and $\hat{\mathcal{P}}_\tau = \exp(\boldsymbol{\gamma}_\tau \cdot \hat{\boldsymbol{n}}) \hat{P}_\tau$. When using the SPD as a variational density matrix, it must be restricted to a Hermitian and semi-definite form; and there are two variants for a given $\mathcal{N}$[10]. It should be noted that the most general non-interacting projector would include the terms $\hat{a}_i^\dagger \hat{a}_j^\dagger$ and $\hat{a}_i \hat{a}_j$, but we presently omit them for brevity. We also define a non-interacting SPD as $\hat{\varrho}_0 = \exp(\boldsymbol{\gamma}_1 \cdot \hat{\boldsymbol{n}}) \ldots \exp(\boldsymbol{\gamma}_{\mathcal{N}} \cdot \hat{\boldsymbol{n}})$, which will be the starting point for the perturbative expansion of the SPD.

The variational parameters of the SPD are the $\boldsymbol{\gamma}_\tau$ and the parameters within $\hat{P}_\tau$. A common choice for the interacting projector is $\hat{P}_\tau = \exp(\hat{\mathcal{V}}_\tau) = \exp\left(\sum_i g_{\tau,i}\hat{V}_i\right)$ where $\sum_i \hat{V}_i = \hat{V}$ is some decomposition of the interacting portion of the Hamiltonian, though there are many possible choices (e.g. as in the Jastrow wave function for the Hubbard model[1, 20–22]). The SPD brings several generalizations over Eq. 1. First, the SPD explicitly includes all possible variational freedom at the single particle level, and formally allows for a generic interacting projector. Second, the SPD form allows for a systematic evaluation using the ITGF formalism introduced in this paper. It is useful to note that $\mathcal{N} = 1$ recovers the well known Hartree-Fock approximation; $\mathcal{N} = 2$ recovers the Gutzwiller, Baeriswyl, and Jastrow wave functions, in addition to unitary and variational coupled cluster methods[23–26]; and $\mathcal{N} = 3$ recovers the Gutzwiller-Baeriswyl[16] and Baeriswyl-Gutzwiller[13] wave functions (see [10] for a detailed discussion).

We now introduce the DAT to evaluate the Hamiltonian under the SPD at a given set of variational parameters. We begin with the ITGF formalism, where the integer time evolution in the integer time interaction representation is given as

$$\hat{O}_I(\tau) = \hat{U}_I(\tau)\hat{O}\hat{U}_I(\tau)^{-1}, \tag{3}$$

$$\hat{U}_I(\tau) = \exp(\boldsymbol{\gamma}_1 \cdot \hat{\boldsymbol{n}}) \ldots \exp(\boldsymbol{\gamma}_\tau \cdot \hat{\boldsymbol{n}}), \tag{4}$$

where $\tau = 1, \ldots, \mathcal{N}$. Taylor series expanding the interacting projector, the expectation value of some operator $\hat{O}$ under the SPD is given as

$$\langle \hat{O} \rangle_{\hat{\varrho}} = \frac{\sum_{n=0}^{\infty} \frac{1}{n!} \langle \mathrm{T}(\sum_{\tau=1}^{\mathcal{N}} \hat{\mathcal{V}}_{\tau,I}(\tau))^n \hat{O}_I(\mathcal{N}) \rangle_{\hat{\varrho}_0}}{\sum_{n=0}^{\infty} \frac{1}{n!} \langle \mathrm{T}(\sum_{\tau=1}^{\mathcal{N}} \hat{\mathcal{V}}_{\tau,I}(\tau))^n \rangle_{\hat{\varrho}_0}}, \tag{5}$$

where the quantum average is defined as $\langle \hat{O} \rangle_{\hat{\rho}} = \mathrm{Tr}(\hat{\rho}\hat{O})/\mathrm{Tr}(\hat{\rho})$; the integer time ordering operator T first sorts the operators according to ascending integer time and then according to the position in the original ordering of operators and finally the result is presented from left to right; additionally, the resulting sign must be tracked when permuting operators. It should be noted that our time convention is opposite to the usual definition[27]. Each term in Eq. 5 can be evaluated via the non-interacting ITGF

$$[\boldsymbol{g}_0]_{k\tau,k'\tau'} = \langle \mathrm{T} \hat{a}^\dagger_{k,I}(\tau)\, a_{k',I}(\tau') \rangle_{\hat{\varrho}_0}, \tag{6}$$

using the integer time Wick's theorem[10]. In general, Eq. 5 will require the evaluation of an infinite number of terms, but if the interacting projector is restricted to a local subspace, or if the system is finite, we can resum the expansion into a finite number of terms.

For a test case, we consider the Anderson impurity model (AIM) on a ring[10], which has recently been extensively studied using density matrix renormalization group (DMRG)[28]. The AIM consists of a non-interacting bandwidth of W, a hybridization V, and an impurity interaction U. The interacting projectors can be chosen as local to the impurity, and the exponential can be rewritten as a sum of Hubbard operators within the impurity as

$$\hat{P}_\tau = \exp(\mu_\tau \sum_\sigma \hat{n}_\sigma + u_\tau \hat{n}_\uparrow \hat{n}_\downarrow) = \sum_\Gamma P_{\tau,\Gamma} \hat{X}_\Gamma, \tag{7}$$

where $\hat{X}_\Gamma = |\Gamma\rangle\langle\Gamma|$ is a Hubbard operator and $P_{\tau,0} = 1$, $P_{\tau,\sigma} = \exp(\mu_\tau)$, and $P_{\tau,2} = \exp(2\mu_\tau + u_\tau)$; and the subscripts $0, \sigma, 2$ correspond to empty, singly occupied, and double occupied local states, respectively. For this interacting projector, Eq. 5 will have a finite number of terms and thus can be evaluated exactly.

The computational cost of evaluating the total energy for a given SPD in the AIM is dictated by two factors. The first cost is associated with constructing the non-interacting ITGF for the entire system, which scales at worst as $\mathcal{N}^2 L^3$. Second, there is the cost of evaluating the total energy using the integer time Wick's theorem, which scales exponentially with the number of integer time steps $\mathcal{N}$[10]. For the particular cases of $\mathcal{N} \leq 4$ and $L \approx 1000$, the computational cost is always dominated by L via the first factor, and therefore in this scenario a single evaluation of VDAT has a relatively minimal computational cost.

Having all the machinery necessary to evaluate the total energy under the SPD, we can then minimize over the variational parameters in order to determine the ground state of the Hamiltonian. Given the specific parameterization of the SPD we have chosen (see reference [10], Section VII B 1), there will be $\lceil (\mathcal{N}-1)/2 \rceil$ interacting variational parameters, while there will be $3\lfloor (\mathcal{N}-1)/2 \rfloor$ non-interacting variational parameters. Therefore, the total number of iterations for minimizing the energy under the SPD will be a constant that is independent of L, and VDAT for $1 < \mathcal{N} \leq 4$ should have a similar scaling to $\mathcal{N} = 1$ (i.e. Hartree-Fock).

We now evaluate the ground state energy $E(U, V)$ and the local impurity spin correlation[29] $\langle \hat{S}^z_f \hat{S}^z_f \rangle = \frac{1}{4}\langle (\hat{n}_{f\uparrow} - \hat{n}_{f\downarrow})^2 \rangle$ (see Figure 1, panel a), where the latter probes the double occupancy and density on the impurity site, and compare to numerically exact DMRG calculations[28]. For $\mathcal{N} = 2$, which recovers the Gutzwiller wave function, the result is only a rough approximation to the DMRG results. Alternatively, $\mathcal{N} = 3$ only shows very small deviations from the DMRG results, and $\mathcal{N} = 4$ further diminishes the differences. Clearly, the SPD converges extremely rapidly with respect to $\mathcal{N}$. What is especially remarkable is that $\mathcal{N} = 4$ has a computational scaling similar to Hartree-Fock, yet has a precision approaching

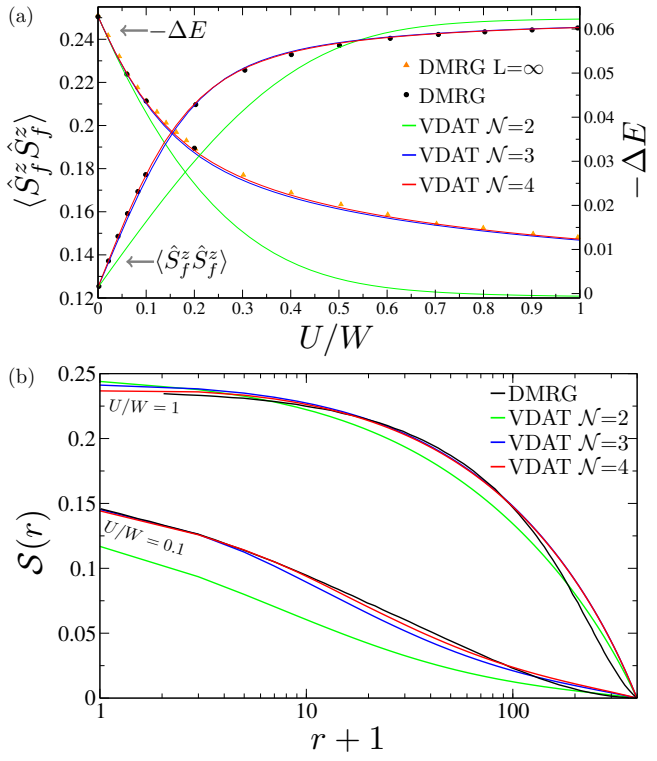

Figure 1. A comparison of VDAT ($\mathcal{N} = 2, 3, 4$) and published DMRG results[28] for the Anderson impurity model on a ring with $V/W = 0.1$. (Panel a) A plot of energy difference $-\Delta E = E(U, 0) - E(U, V)$ (right axis) and $\langle \hat{S}^z_f \hat{S}^z_f \rangle$ (left axis) vs. U/W, with $L = 1397$ unless otherwise noted. (Panel b) The unscreened spin $\mathcal{S}(r)$ vs. the distance from the impurity site with $L = 797$.

DMRG. We can also compute the unscreened spin[28, 30]

$$\mathcal{S}(R) = \langle \hat{S}^z_f (\hat{S}^z_f + \hat{S}^z_{0,c} + \sum_{r=1}^{R} (\hat{S}^z_{r,c} + \hat{S}^z_{L-r,c})) \rangle, \tag{8}$$

which is a far more challenging observable given that it involves a long range correlation between the impurity and the bath (see Figure 1, panel *b*). Once again, $\mathcal{N} = 2$ has reasonable but relatively inaccurate results, while $\mathcal{N} = 3, 4$ are much more accurate.

The preceding approach of using the integer time Wick's theorem to sum all diagrams would be intractable for a general interacting system. This motivates us to push forward our discrete action theory, and generalize the traditional tools of many-body physics. Consider the interacting ITGF under an SPD defined as

$$[\boldsymbol{g}]_{k\tau,k'\tau'} = \langle \mathrm{T} \hat{a}^\dagger_k(\tau)\, \hat{a}_{k'}(\tau') \rangle_{\hat{\varrho}}, \tag{9}$$

where $\tau = 1, \ldots, \mathcal{N}$ and $k = 1, \ldots, L$ and $\hat{O}(\tau)$ is an operator in the integer time Heisenberg representation defined as

$$\hat{O}(\tau) = \hat{U}_\tau \hat{O} \hat{U}^{-1}_\tau, \qquad \hat{U}_\tau = \hat{\mathcal{P}}_1 \ldots \hat{\mathcal{P}}_\tau. \tag{10}$$

Furthermore, when constructing interaction energies and computing the gradient of the total energy with respect to the variational parameters, we will need general integer time correlation functions $\langle \mathrm{T} \hat{O}_1(\tau_1) \ldots \hat{O}_M(\tau_M) \rangle_{\hat{\varrho}}$; which can be rewritten in the integer time Schrodinger representation as

$$\frac{\langle \mathrm{T} \hat{\mathcal{A}} \hat{O}_{1,S}(\tau_1) \ldots \hat{O}_{M,S}(\tau_M) \rangle_{\hat{1}}}{\langle \mathrm{T} \hat{\mathcal{A}} \rangle_{\hat{1}}}, \quad \text{where} \quad \hat{\mathcal{A}} = \hat{\mathcal{A}}_0 \hat{P}, \tag{11}$$

$$\hat{\mathcal{A}}_0 = \exp\big(\sum_{\tau=1}^{\mathcal{N}} \boldsymbol{\gamma}_\tau \cdot \hat{\boldsymbol{n}}_S(\tau) \big), \quad \hat{P} = \prod_{\tau=1}^{\mathcal{N}} \hat{P}_{\tau,S}(\tau), \tag{12}$$

and $\hat{O}_S(\tau)$ is an operator in the integer time Schrodinger representation where $\hat{O}_S(\tau) = \hat{O}$ after applying the time ordering operator. We refer to $\hat{\mathcal{A}}$ as the *discrete action*, given that it encodes all possible integer time correlations under the SPD. Moreover, we can generalize the form of $\hat{\mathcal{A}}$ such that it can describe integer time correlations beyond the SPD, and an important generalization allows for an $\hat{\mathcal{A}}_0$ which has off-diagonal integer time components as

$$\hat{\mathcal{A}}_0 = \exp\big(\sum_{k\tau k'\tau'} [\boldsymbol{v}]_{k\tau,k'\tau'} \hat{a}^\dagger_{k,S}(\tau) \hat{a}_{k',S}(\tau') \big), \tag{13}$$

where $\boldsymbol{v}$ is a general matrix of dimension $L\mathcal{N} \times L\mathcal{N}$[10]. We refer to this more general form as a canonical discrete action (CDA), which will be critical to exactly evaluating the SPD in $d = \infty$.

It is useful to define the discrete generating function, which encodes all information of the discrete action into a scalar function

$$Z(\boldsymbol{g}_0) \equiv \langle \hat{\mathcal{A}} \rangle_{\hat{1}} / \langle \hat{\mathcal{A}}_0 \rangle_{\hat{1}}. \tag{14}$$

For example, using the discrete generating function and the Lie group properties of the non-interacting many-body density matrix[10], we can derive the discrete Dyson equation

$$(\boldsymbol{g}^{-1} - \mathbf{1}) = (\boldsymbol{g}_0^{-1} - \mathbf{1}) \exp(-\boldsymbol{\Sigma})^T, \tag{15}$$

where the integer time self-energy $\boldsymbol{\Sigma}$ and exponential integer time self-energy $\boldsymbol{S}$ are obtained from the generating function as

$$\exp(\boldsymbol{\Sigma})^T = \boldsymbol{S}^{-1} = \mathbf{1} + \frac{\mathbf{1}}{Z\mathbf{1} - \frac{\partial Z}{\partial \boldsymbol{g}_0^T} \boldsymbol{g}_0} \frac{\partial Z}{\partial \boldsymbol{g}_0^T}. \tag{16}$$

This discrete Dyson equation plays a central role in our formalism, much like the traditional Dyson equation. In the limit of large $\mathcal{N}$, the discrete Dyson equation reverts to the usual Dyson equation assuming that the SPD is chosen as the Trotter-Suzuki decomposition[10]. While we have illustrated the single particle ITGF above, any n-particle integer time correlation function can be determined from the generating function[10].

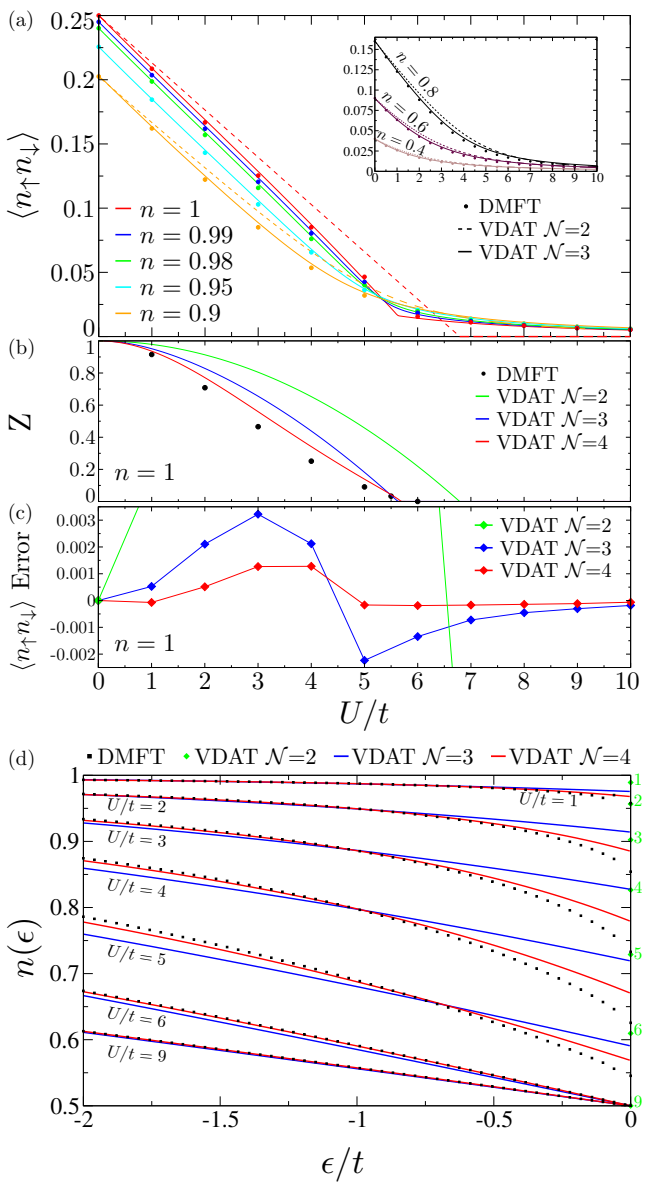

Figure 2. Results for the $d = \infty$ Hubbard model on the Bethe lattice. (Panel *a*) The double occupancy vs. U/t for various densities using VDAT($\mathcal{N} = 3$) (lines) and DMFT (points); and selected $\mathcal{N} = 2$ results (dashed lines). (Inset) Lower densities are evaluated. (Panel *b*) Quasiparticle weight vs U/t at $n = 1$ for VDAT($\mathcal{N} = 2, 3, 4$) and DMFT. (Panel *c*) The difference in the double occupancy between VDAT($\mathcal{N} = 2, 3, 4$) and DMFT vs. U/t for $n = 1$. (Panel *d*) The density distribution vs. energy at $n = 1$ for various U/t using VDAT($\mathcal{N} = 2, 3, 4$) and DMFT; the $\mathcal{N} = 2$ results are illustrated with a single green point and label for brevity.

We now have the proper tools to study the single-band Hubbard model, and we use an SPD with an interacting projector $\hat{P}_\tau = \exp(\mu_\tau \sum_{i\sigma} \hat{n}_{i\sigma} + \sum_i u_\tau \hat{n}_{i\uparrow}\hat{n}_{i\downarrow})$, while the non-interacting projector uses a diagonal $\boldsymbol{\gamma}_\tau$ in the basis that diagonalizes the non-interacting Hamiltonian. In order to evaluate the discrete generating function, we introduce the self-consistent canonical discrete action approximation (SCDA), which is the integer time analogue of DMFT[9]. The SCDA maps the SPD to a collection of CDA's, with one CDA corresponding to each site in the lattice, and the non-interacting part of the CDA is determined self-consistently while the interacting part is taken from the SPD. The essence of the SCDA is the assumption that the integer time self-energy is local

$$\boldsymbol{\Sigma}_{ij}(\boldsymbol{g}) = \delta_{ij}\boldsymbol{\Sigma}_{loc}(\boldsymbol{g}_{loc}). \tag{17}$$

Analogous to DMFT, which assumes a local self-energy and becomes exact in $d = \infty$, the SCDA exactly evaluates the SPD in $d = \infty$. For example, the SCDA for $\mathcal{N} = 2$ recovers the result that the Gutzwiller approximation exactly evaluates the Gutzwiller wave function in $d = \infty$[31]. Additionally, the SCDA for $\mathcal{N} = 3$ exactly evaluates generalizations of the Gutzwiller-Baeriswyl and Baeriswyl-Gutzwiller wave functions in $d = \infty$, which had not yet been achieved. For $\mathcal{N} > 3$, the SCDA exactly evaluates an infinite number of variational wave functions in $d = \infty$ that have not yet been considered.

The SCDA algorithm exactly parallels the DMFT algorithm. We can begin with a guess for the non-interacting ITGF $\boldsymbol{\mathcal{G}} = \int d\epsilon D(\epsilon)\boldsymbol{g}_0(\epsilon)$ for the CDA, where $D(\epsilon)$ is the density-of-states. We can then compute the generating function of the CDA, which yields $\boldsymbol{S}_{loc}$. We then use this exponential integer time self-energy to update the interacting ITGF for each energy orbital as

$$\boldsymbol{g}(\epsilon) = \frac{\mathbf{1}}{\boldsymbol{g}_0(\epsilon) + (\mathbf{1} - \boldsymbol{g}_0(\epsilon))\boldsymbol{S}_{loc}}\boldsymbol{g}_0(\epsilon). \tag{18}$$

Then we obtain the new interacting local ITGF as $\boldsymbol{g}_{loc} = \int d\epsilon D(\epsilon)\boldsymbol{g}(\epsilon)$. Finally, the new non-interacting ITGF of the CDA is

$$\boldsymbol{\mathcal{G}} = \boldsymbol{S}_{loc}\frac{\mathbf{1}}{\mathbf{1} + \boldsymbol{g}_{loc}(\boldsymbol{S}_{loc} - \mathbf{1})}\boldsymbol{g}_{loc}. \tag{19}$$

This procedure must be iterated until self-consistency is achieved, which yields a single evaluation of the SPD for a given set of variational parameters. The above procedure is applicable to any Hubbard-like model, but it will yield an *exact* evaluation of the SPD for infinite dimensions[10].

The number of variational parameters for the interacting projector will be the same as the AIM, while for the non-interacting projector we restrict to at most four variational parameters for each integer time[10]. The main computational complexity of solving the Hubbard model as compared to the AIM is that we must perform a self-consistency condition, though this can normally be achieved in small number of iterations. We now address the $d = \infty$ Hubbard model on the Bethe lattice. It should be emphasized that the computed ground state energy within VDAT is a rigorous upper bound for the

exact ground state energy, and we can compare to the numerically exact dynamical mean-field theory results obtained using the numerical renormalization group (NRG) method as the impurity solver[32].

We begin by examining the double occupancy, which is the derivative of the ground state energy with respect to U and hence a sensitive probe of the total ground state energy (see Figure 2, panel a; also see Supplementary Material[33] for plots of ground state energy). First we present VDAT results for $\mathcal{N} = 3$ and selected results for $\mathcal{N} = 2$, where the latter is the well known Gutzwiller approximation. For half-filling, shown in red, we see that VDAT $\mathcal{N} = 3$ is very close to the DMFT solutions (points), reliably capturing the Mott metal-insulator transition, and illustrating drastic improvement beyond $\mathcal{N} = 2$. Furthermore, we can see that VDAT $\mathcal{N} = 3$ clearly captures the sensitive changes with small doping, illustrated for the densities of 1, 0.99, 0.98, 0.95, and 0.9. We can also proceed to much larger dopings (see inset), where VDAT $\mathcal{N} = 3$ once again reliably describes the DMFT solution. All VDAT results discussed thus far have been for $\mathcal{N} \leq 3$, and it is interesting to consider $\mathcal{N} = 4$ to better understand the convergence of the VDAT. Therefore, we examine the error in the double occupancy at half filling for $\mathcal{N} = 2 - 4$ (see Figure 2, panel c). We see that $\mathcal{N} = 4$ has a smaller error for all values of U/t, as it must, and the error for large U/t is nearly zero. Another interesting quantity is the density distribution $n(\epsilon) = \langle \hat{a}^\dagger_{\epsilon\sigma} \hat{a}_{\epsilon\sigma} \rangle$ (see Figure 2, panel d), given that the kinetic energy is $2\int_{-\infty}^{\infty} d\epsilon D(\epsilon) n(\epsilon)\epsilon$. It is well known that $\mathcal{N} = 2$ (i.e. Gutzwiller) produces a constant density distribution below the Fermi energy (the horizontal line is denoted with a single green point in panel d), and we show that $\mathcal{N} > 2$ produces a non-trivial energy dependence. Another key aspect of $n(\epsilon)$ is the discontinuity at the Fermi energy, which dictates the quasiparticle weight $Z = n(0^-) - n(0^+)$ (see Figure 2, panel b). While $\mathcal{N} = 2$ recovers the usual Gutzwiller result, $\mathcal{N} > 2$ precisely captures the Mott transition and yields reasonable agreement for $\mathcal{N} = 4$. Furthermore, one can approximately extract the real frequency spectrum from the extended density distribution[33].

In conclusion, we have proven that VDAT with $\mathcal{N} = 3$ already yields efficient and precise results for the Anderson Impurity model and for the $d = \infty$ Hubbard model. Furthermore, it is straightforward to address the multi-orbital Hubbard model[10], which is under way. Given that VDAT recovers the Hartree-Fock and Gutzwiller wave functions, it is clear that VDAT can be combined with DFT in the same spirit as DFT+U[34] and DFT+Gutzwiller[35]; and therefore DFT+VDAT would be a prime candidate as an efficient first-principles approach to strongly correlated materials, which should rival DFT+DMFT[36]. There are many possible directions for future development. Both diagrammatic and auxiliary field quantum Monte-Carlo could be generalized to our formalism. While our present work on the SPD has used real variational parameters, we can apply VDAT using an SPD with unitary projectors[10], which could have utility in quantum computing[17–19] and unitary coupled cluster theory[24–26]. VDAT will be a key tool for parameterizing energy functionals in the context of the off-shell effective energy theory[37].

This work was supported by the grant DE-SC0016507 funded by the U.S. Department of Energy, Office of Science. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

[1] R. Jastrow, Phys. Rev. **98**, 1479 (1955).
[2] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
[3] M. C. Gutzwiller, Physical Review **134**, 923 (1964).
[4] M. C. Gutzwiller, Physical Review **137**, 1726 (1965).
[5] I. P. Mcculloch, Journal Of Statistical Mechanics-theory And Experiment **10**, 10014 (2007).
[6] U. Schollwock, Rev. Mod. Phys. **77**, 259 (2005).
[7] U. Schollwock, Annals Of Physics **326**, 96 (2011).
[8] M. Suzuki, Physica A-statistical Mechanics And Its Applications **194**, 432 (1993).
[9] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).
[10] Z. Cheng and C. A. Marianetti, Phys. Rev. B [jointly submitted with current manuscript] **0**, 0 (2020).
[11] H. F. Trotter, Proceedings of the American Mathematical Society **10**, 545 (1959).
[12] M. Suzuki, Communications In Mathematical Physics **51**, 183 (1976).
[13] M. Dzierzawa, D. Baeriswyl, and M. Distasio, Phys. Rev. B **51**, 1993 (1995).
[14] D. Baeriswyl, in *Nonlinearity in Condensed Matter*, edited by A. R. Bishop, D. K. Campbell, and S. E. Trullinger (Springer-Verlag, Berlin, 1987) 1st ed., pp. 183–193.
[15] D. Baeriswyl, Foundations Of Physics **30**, 2033 (2000).
[16] H. Otsuka, J. Phys. Soc. Jpn. **61**, 1645 (1992).
[17] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028 (2014).
[18] D. Wecker, M. B. Hastings, and M. Troyer, Physical Review A **92**, 042303 (2015).
[19] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, Nature Communications **10**, 3007 (2019).
[20] P. Fazekas and K. Penc, International Journal of Modern Physics B **2**, 1021 (1988).
[21] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **59**, 3669 (1990).
[22] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, Phys. Rev. Lett. **94**, 026406 (2005).
[23] R. J. Bartlett and J. Noga, Chemical Physics Letters **150**, 29 (1988).
[24] R. J. Bartlett, S. A. Kucharski, and J. Noga, Chemical Physics Letters **155**, 133 (1989).
[25] W. Kutzelnigg, Theoretica Chimica Acta **80**, 349 (1991).

[26] A. G. Taube and R. J. Bartlett, International Journal Of Quantum Chemistry **106**, 3393 (2006).
[27] G. D. Mahan, *Many-Particle Physics (Physics of Solids and Liquids)* (Springer, 2000).
[28] G. Barcza, F. Gebhard, T. Linneweber, and O. Legeza, Phys. Rev. B **99**, 165130 (2019).
[29] Expectation values such as $\langle \hat{O} \rangle$ imply $\langle \hat{O} \rangle_{\rho^\star}$, where $\rho^\star$ is the ground state density matrix.
[30] A. Holzner, I. P. McCulloch, U. Schollwock, J. von Delft, and F. Heidrich-Meisner, Phys. Rev. B **80**, 205114 (2009).
[31] W. Metzner and D. Vollhardt, Phys. Rev. B **37**, 7382 (1988).
[32] R. Zitko and T. Pruschke, Phys. Rev. B **79**, 085106 (2009).
[33] See Supplemental Material at [URL will be inserted by publisher] for plots of the ground state energy of the $d = \infty$ Hubbard model, in addition to the approximate local spectra; Refs. [38–42] are cited.
[34] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, Journal Of Physics-condensed Matter **9**, 767 (1997).
[35] X. Y. Deng, L. Wang, X. Dai, and Z. Fang, Phys. Rev. B **79**, 075114 (2009).
[36] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Rev. Mod. Phys. **78**, 865 (2006).
[37] Z. Cheng and C. A. Marianetti, Phys. Rev. B **101**, 081105 (2020).
[38] M. Jarrell and J. E. Gubernatis, Physics Reports-review Section Of Physics Letters **269**, 133 (1996).
[39] K. Haule, C. H. Yee, and K. Kim, Phys. Rev. B **81**, 195107 (2010).
[40] J. Karp, M. Bramberger, M. Grundner, U. Schollwock, A. J. Millis, and M. Zingl, Physical Review Letters **125**, 166401 (2020).
[41] Y. Wang, C. J. Kang, H. Miao, and G. Kotliar, Phys. Rev. B **102**, 161118 (2020).
[42] F. Lechermann, Physical Review X **10**, 041002 (2020).