Using local operator fluctuations to identify wave function improvements

Kiel T. Williams and Lucas K. Wagner*
Department of Physics, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, USA
(Received 29 January 2016; revised manuscript received 12 May 2016; published 25 July 2016)

A method is developed that allows analysis of quantum Monte Carlo simulations to identify errors in trial wave functions. The purpose of this method is to allow for the systematic improvement of variational wave functions by identifying degrees of freedom that are not well described by an initial trial state. We provide proof of concept implementations of this method by identifying the need for a Jastrow correlation factor and implementing a selected multideterminant wave function algorithm for small dimers that systematically decreases the variational energy. Selection of the two-particle excitations is done using the quantum Monte Carlo method within the presence of a Jastrow correlation factor and without the need to explicitly construct the determinants. We also show how this technique can be used to design compact wave functions for transition metal systems. This method may provide a route to analyze and systematically improve descriptions of complex quantum systems in a scalable way.

DOI: 10.1103/PhysRevE.94.013303

I. INTRODUCTION

First-principles quantum Monte Carlo calculations [1] for solids are a promising way to go beyond density functional theory (DFT). These methods directly simulate electron-electron correlations and can obtain very high accuracy on challenging materials [2–5] using current state of the art techniques such as the fixed node diffusion Monte Carlo (DMC) method. Despite this success, the DMC method's accuracy is limited by the fixed node approximation, which allows for polynomial scaling of the computational cost with system size, but results in a DMC energy that is only an upper bound to the true ground-state energy. In practical calculations, improvement of the accuracy and efficiency of the fixed node diffusion Monte Carlo method is reliant on improving trial wave functions that determine the fixed nodal surface.

In order for a trial wave function to be appropriate for quantum Monte Carlo calculations, it should be compact and efficient to calculate. For application to bulk materials, it must also be size extensive; that is, the total energy must scale with the system size. By far the most common trial wave function is the Slater-Jastrow wave function [6,7], which is simple and extensive, and initial guesses are easily obtainable from DFT codes. While truncated determinant expansions can be effective in describing small molecules [8], they cannot be used in bulk materials because they are not size extensive. Backflow wave functions [9], while they have proven effective in homogeneous [10] and inhomogeneous [11] systems, may not capture all the correlated physics in a system. It is thus of great interest, given a Slater-Jastrow wave function, whether there is a compact wave function that describes the most important improvements relative to the ground state.

In this article we present some initial steps towards a method that uses fluctuations of the local energy $\hat{H}\Psi(\mathbf{R})/\Psi(\mathbf{R})$ not to optimize a given parametrization, but to identify directions in Hilbert space that can improve trial wave functions. We first provide a summary of the imaginary time projector $\exp(-\tau\hat{H})$ and its use in improving wave functions and introduce the notation that will be used in the article. Then we show a proof of concept for multi-Slater-Jastrow wave functions, in which this method is used to select determinants in the wave function. Finally, we show how the local energy fluctuations can be used to determine *a priori* what terms to add to a variational wave function for a transition metal system TiO. These results set the stage for data mining of many-body wave functions to determine how they should be improved.

II. THEORY

In this work we use ideas that have been known for a long time for optimizing parameters in wave functions [12–15], but we follow more the work of Holzmann *et al.* [16] in that we would like to use the Feynman-Kac formulas to discover which parametrization to add to a given wave function. The quantum variational principle states that for any appropriately normalized trial wave function $\Psi_T(\mathbf{R}, P)$, where $\mathbf{R}$ is the many-body electron coordinate and P is a set of parameter values, the expectation value of the Hamiltonian of the system in state Ψ_T equals or exceeds the ground-state energy of the Hamiltonian:

$$E_T \geqslant E_g, \tag{1}$$

where

$$E_T(P) = \langle\Psi_T|\hat{H}|\Psi_T\rangle. \tag{2}$$

We then minimize $E_T(P)$ with respect to the parameter set P. Once this is done, we must alter the parametrization of the trial wave function to obtain further improvement in the energy estimate. Our ultimate goal will be not to optimize the parameters within a fixed set P, but to identify new parameters that must be added to P to improve the qualitative structure of the particular trial state.

Iteratively applying the projection operator to a trial function Ψ_T produces a sequence of new wave functions:

$$|\Psi_{H'}(\tau)\rangle = e^{-\tau\hat{H}'}|\Psi_T\rangle, \tag{3}$$

where $\hat{H}' = \hat{H} - E_{\text{ref}}$, with $E_{\text{ref}} = \langle\hat{H}\rangle$. This converges to the exact ground-state wave function $|\Psi_{\text{g.s.}}\rangle$ in the infinite limit

$$\lim_{\tau\to\infty}|\Psi_{H'}(\tau)\rangle = |\Psi_{\text{g.s.}}\rangle. \tag{4}$$

*lkwagner@illinois.edu

Performing this operation directly corresponds to a projector Monte Carlo method, such as the diffusion Monte Carlo method. The challenge in doing this is that compact representations of the operator $\exp(-\tau \hat{H}')$ are generally not known and so the imaginary time dynamics must operate in very high dimensions. Our objective here will be to find a compact representation of the short-time projector operator.

We begin by considering an arbitrary set of linear operators $\{\hat{A}_i\}$. Applying these operators to the trial state produces a new state $|\Psi_A\rangle$:

$$|\Psi_A\rangle = \left(1 + \sum_i a_i \hat{A}_i\right)|\Psi_T\rangle. \tag{5}$$

Applying this set of operators again to $|\Psi_A\rangle$ and iterating generates a new sequence of wave functions. For brevity, we define

$$|\Psi_{A_i}\rangle \equiv \hat{A}_i|\Psi_T\rangle. \tag{6}$$

We force the minimal set of operators $\{\hat{A}_i\}$ to mimic the projection operator by minimizing the square deviation of Ψ_A from $\Psi_{H'}$:

$$\int [\Psi_A(\mathbf{R}) - \Psi_{H'}(\mathbf{R})]^2 d\mathbf{R}. \tag{7}$$

This minimization procedure provides an estimate of the set of associated $\{a_i\}$ operator amplitudes. We define the local operators $A_k(\mathbf{R}) \equiv \frac{\hat{A}_k\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}$ and a local energy $E_L(\mathbf{R}) = \frac{\hat{H}'\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}$. By expanding the projection operator to first order and minimizing the square deviation, we find that

$$a_k = -\tau \int \frac{\hat{H}'\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\frac{\hat{A}_k\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}\Psi_T^2(\mathbf{R})d\mathbf{R}, \tag{8}$$

$$\frac{da_k}{d\tau} = -\langle [E_L(\mathbf{R}) - \langle \hat{H}\rangle]A_k(\mathbf{R})\rangle, \tag{9}$$

where we have assumed that elements of the set $\{\Psi_{A_i}\}$ are orthonormal such that the overlap matrix $S_{ik} = \langle \Psi_{A_i}|\Psi_{A_k}\rangle$ is approximately diagonal. Figure 1 depicts this scheme pictorially, with the exact and mimicked projection operators represented by the black and tangential red arrows, respectively. We see then that the mimicked projection operator evaluated for $\tau = 0$ can be viewed as a linearized approximation to the exact path to the ground state through Hilbert space. In this way, our approximation to the projection operator identifies the most significant elements of Hilbert space absent from an initial trial state.

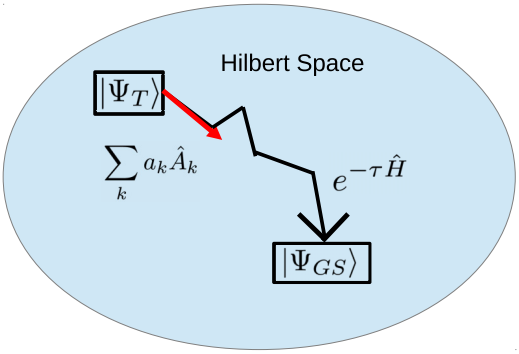

FIG. 1. Visual representation of the path through Hilbert space from the initial to the exact wave function taken by the exact (black) and mimicked (red) projection operators, respectively.

The derivation of our method is similar in spirit to the stochastic reconfiguration of Sorella [12,17–20]. The energy fluctuation potential method also shares some similarities with our technique in its focus on the correlation between the local behavior of the energy and some chosen operator [14,21,22]. A set of operators $\hat{A}_i$ is a good set if only a few terms in Eq. (8) are nonzero, while a set with many small values in Eq. (8) is not an efficient descriptor of the wave function improvement.

III. QUANTUM MONTE CARLO METHODOLOGY

We first compute the single-particle Hartree-Fock (HF) orbitals for a molecular system. We obtain all orbitals using the GAMESS computational package [23,24]. Core electrons were replaced by the corresponding Burkatzki-Filippi-Dolg pseudopotential [25] with triple-ζ basis sets.

We perform variational Monte Carlo with the QWALK computational package [26]. We begin with a trial wave function of the Slater-Jastrow form

$$\Psi = \exp(U)\mathrm{Det}[\phi_i(r_j)]. \tag{10}$$

We use the linear method of Umrigar and coworkers [13,27,28] to optimize the Jastrow U. The form of the Jastrow correlation factor U is a function of the electron and ionic coordinates

$$U = \sum_{ijI} u(r_{iI}.r_{jI},r_{ij}), \tag{11}$$

where i and j indices represent electronic coordinates and I represents ionic coordinates. The functions u are given by

$$\begin{aligned} u(r_{iI},r_{jI},r_{ij}) = &\sum_k c_k^{ei} a_k(r_{iI}) \\ &+ \sum_m c_m^{ee} b_k(r_{ij}) + \sum_{klm} c_{klm}^{eei}(a_k(r_{iI})a_l(r_{jI}) \\ &+ a_k(r_{jI})a_l(r_{iI}))b_k(r_{ij}), \end{aligned} \tag{12}$$

where the a_k and b_k functions have the general form

$$a_k(r) = \frac{1 - z(r/r_{\mathrm{cut}})}{1 + \beta z(r/r_{\mathrm{cut}})} \tag{13}$$

and $z(x)$ is a polynomial chosen to smoothly go to zero at $r = r_{\mathrm{cut}}$ [29]. This form of the Jastrow factor explicitly incorporates three-body interactions between two electrons and an ion.

IV. DETERMINANT SELECTION

The set of double excitation operators given by

$$\hat{A}_{ij,kl} \equiv c_{\uparrow k}^\dagger c_{\downarrow l}^\dagger c_{\uparrow i} c_{\downarrow j}, \tag{14}$$

where $c_{\sigma k}^\dagger$ ($c_{\sigma k}$) is the one-body creation (annihilation) operator in the σ spin channel, offers one possible choice of linear operators A_k in Eq. (8). If i,j are occupied orbitals and k,l are unoccupied orbitals, then applying $\hat{A}_{ij,kl}$ to a Slater determinant generates an excited-state determinant in which

the lower-energy i and j orbitals are now vacant and the higher-energy k and l orbitals are occupied. The elements of the two-body reduced density matrix (2RDM) are given by the expectation values of these two-body creation or destruction operators. We thus make the analogy with local energy to define a local density matrix element, given a wave function $|\Psi_T\rangle$,

$$\rho_{ijkl}(\mathbf{R}) = \frac{\hat{A}_{ij,kl}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}, \tag{15}$$

or explicitly

$$\rho_{ijkl}(\mathbf{R}) = \sum_{a\neq b}\int \phi_i^*(r_a')\phi_j^*(r_b') \times\phi_k(r_a)\phi_l(r_b)\Psi_T(R_{ab}'')\Psi_T^{-1}(R)dr_a'dr_b', \tag{16}$$

where $R = (r_1, r_2, \ldots, r_N)$, $R_{ab}'' = (r_1, r_2, \ldots, r_a', \ldots, r_b', \ldots, r_N)$ refers to the set of coordinates generated by changing the positions of two electrons, and we have omitted overall normalization. We evaluate this two-body integral in a quantum Monte Carlo calculation by sampling the coordinates r_a' and r_b' from the sum over orbitals $f(r) = \sum_i \phi_i^2(r)$ and the many-body electron coordinate $\mathbf{R}$ from $\Psi^2(\mathbf{R})$ [30]. With this, the expression given in Eq. (16) can be rearranged to give

$$\rho_{ijkl}(\mathbf{R}) = \frac{1}{N_i N_j N_k N_l} \times \sum_{a\neq b}\left\langle \frac{\frac{\Psi(\mathbf{R}_{ab}'')}{\Psi(\mathbf{R})}\phi_i^*(r_a')\phi_j^*(r_b')\phi_k(r_a)\phi_l(r_b)}{f(r_a')f(r_b')}\right\rangle_{f(r_a'),f(r_b')}, \tag{17}$$

where the normalization factor is given by

$$N_i = \sqrt{\left\langle \frac{\phi_i^2(r_a')}{f(r_a')}\right\rangle_{f(r_a')}}. \tag{18}$$

The two-particle operators in Eq. (14) are used to evaluate Eq. (8) and generate a list of important determinants missing from the initial wave function. Hence, we can select the determinants most important to the exact ground state without the need to first evaluate those determinants. The entire process of wave function generation is summarized as follows.

(i) Obtain single-particle orbitals from a HF calculation.

(ii) Optimize the single-determinant Slater-Jastrow factor

$$\langle \mathbf{R}|\Psi\rangle = e^{U(r,r')}\mathrm{Det}[\phi_i(\mathbf{r}_j)]. \tag{19}$$

(iii) Rank 2RDM elements by covariance of $\langle c_{\uparrow k}^\dagger c_{\downarrow l}^\dagger c_{\uparrow i} c_{\downarrow j}\rangle$ with E_L.

(iv) Add determinants identified as significant to the expansion

$$|\Psi_{\mathrm{new}}\rangle = |\Psi_{\mathrm{old}}\rangle + \sum_i a_i e^U [c_{\uparrow k}^\dagger c_{\downarrow l}^\dagger c_{\uparrow i} c_{\downarrow j}]|\Psi_{\mathrm{HF}}\rangle. \tag{20}$$

(v) Optimize coefficients $\{a_i\}$ of $|\Psi_{\mathrm{new}}\rangle$ using the linear method.

This process generates a determinantal expansion whose length is controlled by the user, up to the full size of the active space.

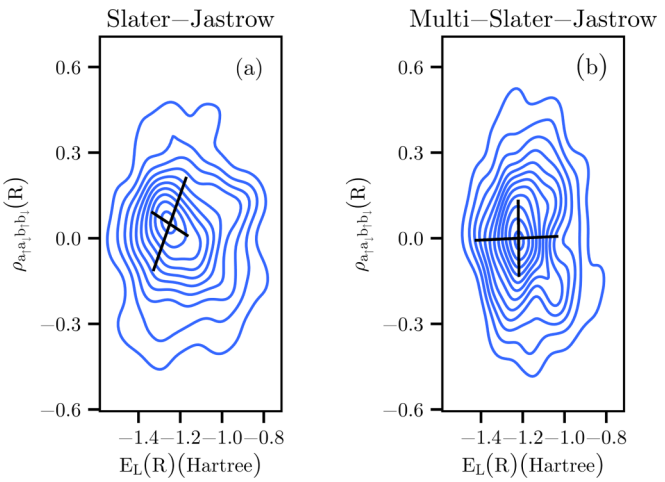

FIG. 2. Amplitude of two-body $b \to a$ bonding-to-antibonding excitation $\rho_{a_\uparrow a_\downarrow b_\uparrow b_\downarrow}(\mathbf{R})$ versus local energy $E_L(\mathbf{R})$ for two different trial wave functions, with corresponding principal components of the distribution indicated. The Slater-Jastrow wave function used to generate (a) did not include the CSF corresponding to this two-body excitation, while the wave function used to generate (b) does. The principal components rotate upon the addition of this CSF.

A. The H_2 molecule

For the case of H_2, we restrict our active Hilbert space to the set of bonding and antibonding σ-symmetry orbitals. Figure 2 shows the contours of the sampled amplitude $\rho_{a_\uparrow a_\downarrow b_\uparrow b_\downarrow}(\mathbf{R})$ of the local operator associated with a two-body $b \to a$ bonding-to-antibonding excitation in an isolated hydrogen dimer versus the sampled local energy $E_L(\mathbf{R})$ for each of two trial states:

$$\begin{aligned}\Psi_{\mathrm{SJ}} &= e^U\phi_{b\uparrow}(r_1)\phi_{b\downarrow}(r_2),\\ \Psi_{\mathrm{MSJ}} &= e^U[c_1\phi_{b\uparrow}(r_1)\phi_{b\downarrow} + c_2\phi_{a\uparrow}(r_1)\phi_{a\downarrow}],\end{aligned} \tag{21}$$

where Ψ_{SJ} and Ψ_{MSJ} are the Slater-Jastrow and multi-Slater-Jastrow wave functions containing the bonding ϕ_b and antibonding ϕ_a single-particle orbitals, respectively.

The line segments on each panel in Fig. 2 indicate the principal components of the resulting distribution. These components are given by the eigenvectors of the covariance matrix of the local energy distribution taken with respect to the local operator $\rho_{a_\uparrow a_\downarrow b_\uparrow b_\downarrow}(\mathbf{R})$:

$$\begin{pmatrix}\sigma_{\rho,\rho} & \sigma_{\rho,E_L}\\ \sigma_{E_L,\rho} & \sigma_{E_L,E_L}\end{pmatrix}$$

in this two-dimensional representation. The rotation of the principal components relative to the axes in Fig. 2(a) shows that the covariance matrix contains nonvanishing off-diagonal elements. It follows that $\rho_{a_\uparrow a_\downarrow b_\uparrow b_\downarrow}(\mathbf{R})$ and $E_L(\mathbf{R})$ are correlated for this single-determinant trial state. After the addition of the associated $b \to a$ determinant to the wave function in Fig. 2(b), the principal components rotate to align with the axes, indicating that the covariance matrix has become diagonal. This implies that the covariance between the local energy $E_L(\mathbf{R})$ and local operator $\rho_{a_\uparrow a_\downarrow b_\uparrow b_\downarrow}(\mathbf{R})$ has vanished and the two variables now have zero covariance. That is, a key element absent from the initial trial state has been identified

and added based on the covariance, pushing the wave function closer to the exact ground state.

B. Dimer molecules

As a further proof of concept, we apply the covariance method to select determinants for a set of stretched molecules: H_2 (0.88 Å bond length), N_2 (1.7 Å bond length), O_2 (1.6 Å bond length), and F_2 (1.5 Å bond length). By stretching the molecules, the electron correlations are enhanced, increasing the strength of the covariance signal. We obtain single-particle orbitals for each system from a restricted open-shell Hartree-Fock calculation using GAMESS. This method doubly fills molecular orbitals (MOs) to the greatest extent possible and places remaining unpaired electrons into singly filled MOs. We limit our active space to a set of bonding and antibonding MOs with cylindrical symmetry and either σ or π symmetry. Other states exist within the full orbital space, but their inclusion yields only small improvement to the final wave function and system energy. Because different methods of determinant selection produce significantly different rates of energy convergence [31], the covariance-based method we have described can yield interesting results even at the level of a multi-Slater-Jastrow *Ansatz*. Our chief objective in this section is to show that the covariance technique can select the most significant determinants for a particular molecule before performing a variational optimization of the wave function.

We consider only two-particle excitations featuring one particle in each spin channel. We compare these results to those obtained with the usual configuration interaction method with single and double excitation (CISD). This is natural for molecules such as N_2 with a ground-state singlet spin configuration, though it can lead to the exclusion of significant excitations in molecules like O_2, which contain unpaired electrons. Figure 3 compares the normalized weight of each configuration-state function (CSF) in conventional CISD, the optimized weight of each CSF in a multi-Slater-Jastrow *Ansatz*, and the local energy covariance for each relevant CSF in each material, respectively. We see that the determinant orderings predicted by both traditional CISD and our method based on local energy covariance are equivalent for each system across the dominant particle excitations. This indicates that the path to the ground state through Hilbert space obtained by successively applying the projection operator is approximately equivalent to that produced by the usual configuration interaction (CI) procedure in this case.

From Eq. (8) we see that the covariance signal in a 2RDM element should fall identically to zero once the corresponding excitation has been added to the trial state. In practice, we observe that the signal in an added excitation falls significantly once it has been added to the trial wave function, but it does not vanish entirely. This is a consequence of the Jastrow factor U in the trial state, which we assume commutes with the creation and annihilation operators introduced above:

$$c^\dagger_{\uparrow k} c^\dagger_{\downarrow l} c_{\uparrow i} c_{\downarrow j} (e^U |D\rangle) \approx e^U (c^\dagger_{\uparrow k} c^\dagger_{\downarrow l} c_{\uparrow i} c_{\downarrow j} |D\rangle), \tag{22}$$

where $|D\rangle$ is a determinant trial state. Because the Jastrow factor does not commute exactly with the creation and destruction operators, a small contribution to the covariance signal is neglected. Practically speaking, this approximation did not seem to affect the performance of the technique.

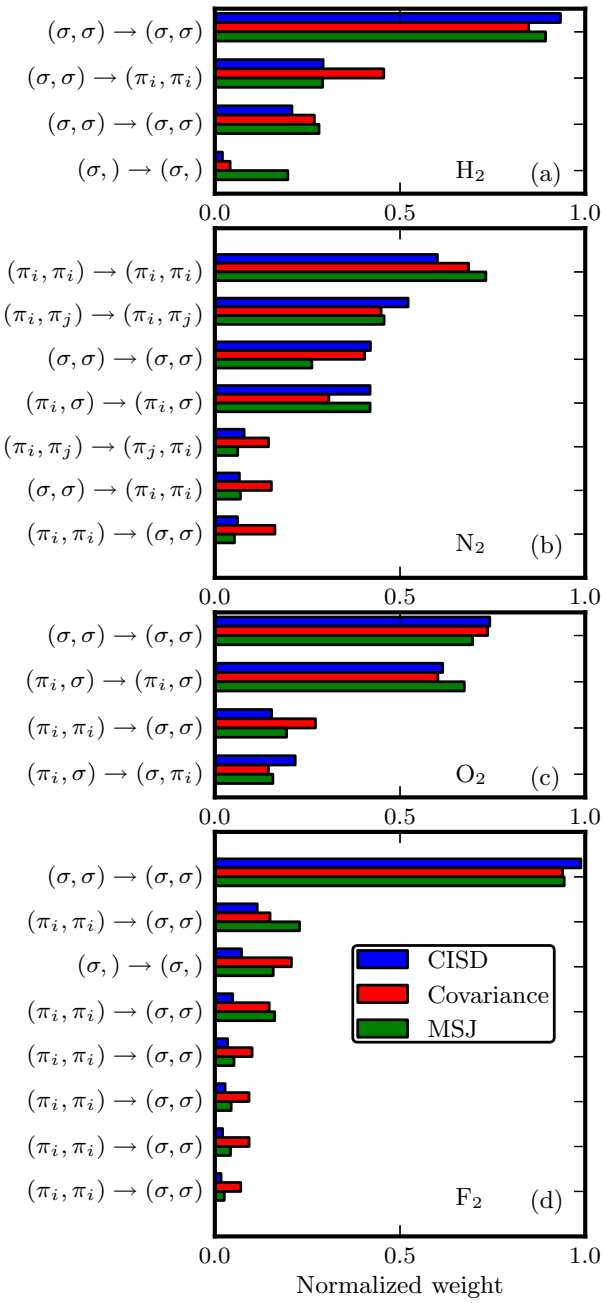

FIG. 3. Comparison of normalized signal strength for different estimators of relative CSF importance for stretched dimers of H_2, N_2, O_2, and F_2, respectively. The indicated bars are the determinant coefficients taken from a CISD calculation, the signal drawn from the $\langle (E_L - \langle H \rangle) A_k \rangle$ estimator, and the determinant coefficients taken from an optimized multi-Slater-Jastrow wave function, respectively. The CSFs are arranged such that the optimized final CSF weight declines monotonically from top to bottom. Each indicated excitation is a one- or two-particle excitation that includes both itself and any symmetry-related partners. For example, $(\pi_i,\pi_i) \to (\pi_i,\pi_i)$ is a two-particle excitation that excites a bonding π-orbital electron to an antibonding π^* orbital of the same angular momentum (x or y) in each spin channel. On the other hand, $(\pi_i,\pi_j) \to (\pi_j,\pi_i)$ involves a two-body exchange.

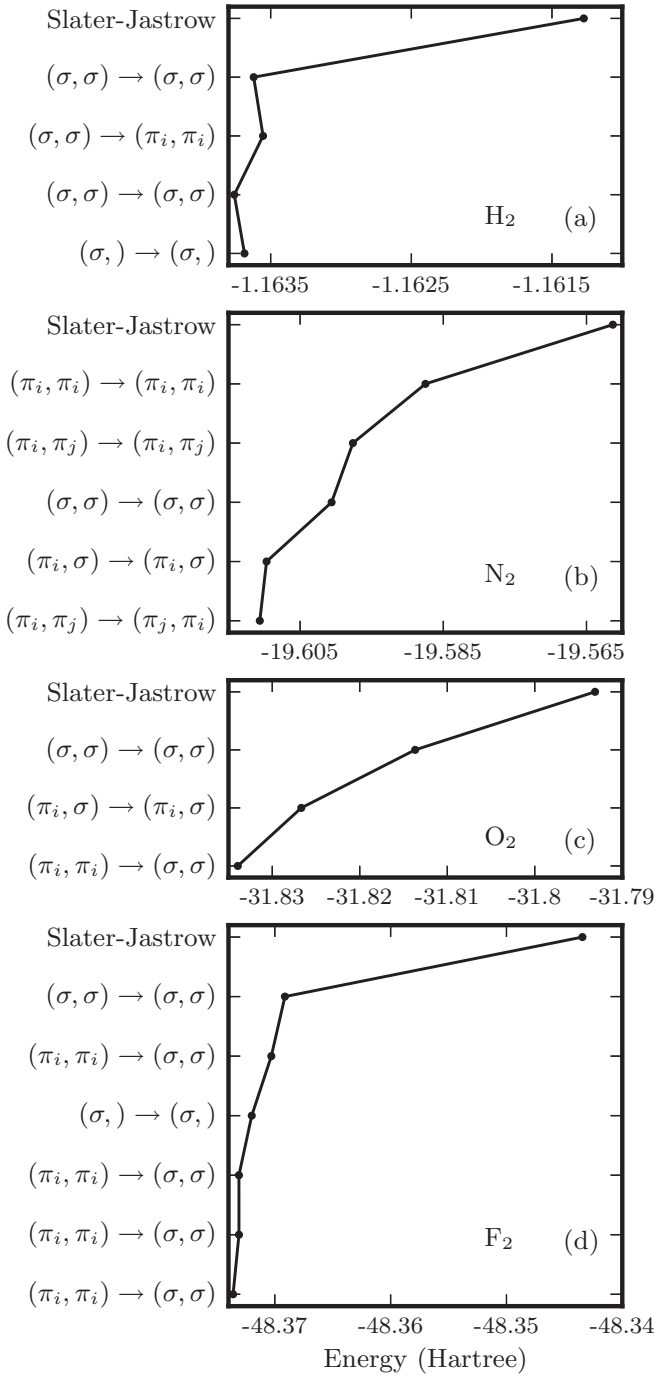

FIG. 4. Added spin-up or spin-down CSF excitations vs associated variational Monte Carlo energy in a multi-Slater-Jastrow wave function for the CSF ordering suggested by conventional CISD for each considered model system.

We also find the rate of energy convergence for the predicted CSF ordering in each model molecular system. Figure 4 shows the variational Monte Carlo energy of an optimized multi-Slater-Jastrow wave function as a function of the CSFs included in the trial state. The CSFs are ordered here according to the weight given by a conventional CISD calculation. We see that the energy converges rapidly with respect to the number of CSFs included in the wave function. This explicitly illustrates

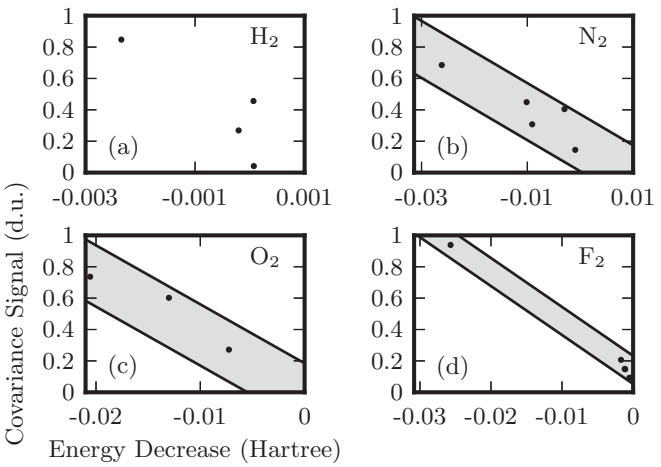

FIG. 5. Normalized covariance signal of CSFs versus the decrease in energy obtained from adding a CSF to the trial state. Significant negative correlations exist between the two values. The shading is provided as a visual guide.

that the CISD method and our covariance-based technique can drive the initial trial state asymptotically close to the exact ground state.

Finally, we also assess the degree to which the covariance in a 2RDM element predicts the energy gain obtained from adding the associated determinant to the trial state. Figure 5 compares the decrease in total system energy obtained from each additional CSF with the corresponding covariance signal. We observe that the energy gain and the covariance signal are negatively correlated with one another. This correlation indicates that the covariance in a 2RDM element can be used as a proxy for estimating the energy change from adding a determinant to the trial state.

As a method of determinant selection for these systems, this technique is less efficient than using CI to determine the weights and the results are similar. We therefore would not recommend this technique as a selection method for small molecules. However, the point of this section is that the energy fluctuations can be data mined to find the correct directions in Hilbert space to improve trial wave functions. In the case of stretched dimers, it is well known that the most important improvement over Slater-Jastrow consists of multiple determinants and the energy fluctuation technique selects the correct ones.

C. Using the 1RDM to perform selection in a simple model

Thus far, we have relied upon the covariance of elements of the 2RDM with the local energy to construct wave functions. However, for large systems, it may be computationally inconvenient to compute the 2RDM. In these cases, it may be possible to instead construct wave functions with the aid of the 1RDM, which is available at a much lower numerical cost. We can understand selection using the 1RDM within the context of a simple model Hamiltonian.

In this example, we begin by considering a two-dimensional Hilbert space consisting of the states $|D_1\rangle$ and $|D_2\rangle$. We define

the creation (destruction) operator $c_1^\dagger$ (c_1) such that

$$\langle D_1|c_1^\dagger c_1|D_1\rangle = 1,$$
$$\langle D_2|c_1^\dagger c_1|D_2\rangle = 0,$$
$$\langle D_2|c_1^\dagger c_1|D_1\rangle = 0. \quad (23)$$

That is, the orbital 1 is occupied in state $|D_1\rangle$ and unoccupied in state $|D_2\rangle$. Taking the probability amplitudes to be real valued, any state $|\Psi\rangle$ in this Hilbert space can be written in the form $\cos\theta|D_1\rangle + \sin\theta|D_2\rangle$ for a real parameter θ.

We consider a Hamiltonian $\hat{H}$ given by

$$\hat{H} = \epsilon_1|D_1\rangle\langle D_1| + \epsilon_2|D_2\rangle\langle D_2| - \Delta(|D_1\rangle\langle D_2| + |D_2\rangle\langle D_1|). \quad (24)$$

We take $\epsilon_1 = 0$ and $\epsilon_2 = 1$ to simplify the subsequent calculations. For $\Delta = 1$, the eigenvectors are $\theta_1 = 0.553$ (ground state) and $\theta_2 = 2.124$ (excited state).

In this system we can analytically compute the correlation in Eq. (8), now taking operator $\hat{A}_k$ as the number operator $c_1^\dagger c_1$ associated with $|D_1\rangle$. The result is given by

$$\langle\Psi|(\hat{H} - \langle H\rangle)c_1^\dagger c_1|\Psi\rangle = \Delta\sin 2\theta\left(\cos^2\theta - \tfrac{1}{2}\right) - \sin^2\theta\cos^2\theta. \quad (25)$$

If $\Delta \neq 0$, then there are four roots of this function in the range $[0,\pi]$, two at the high symmetry points $\theta = 0$ and $\theta = \frac{\pi}{2}$ and two at the eigenvectors (Fig. 6). So if one evaluates the one-particle correlation with the Hamiltonian using a single-determinant wave function in the single-particle orbital basis of the determinant, then there is zero signal, regardless of the value of Δ. However, if the reference wave function is not a single determinant (such as a Slater-Jastrow wave function), then the signal can be nonzero for important orbitals in the expansion. For example, in the stretched N_2 dimer, the elements corresponding to the bonding and antibonding orbitals have a covariance with the local energy of approximately 0.001 hartree, while other orbitals have much smaller signals. This allows us to select which one-particle states may be important in the determinant expansion without computing the more costly 2RDM.

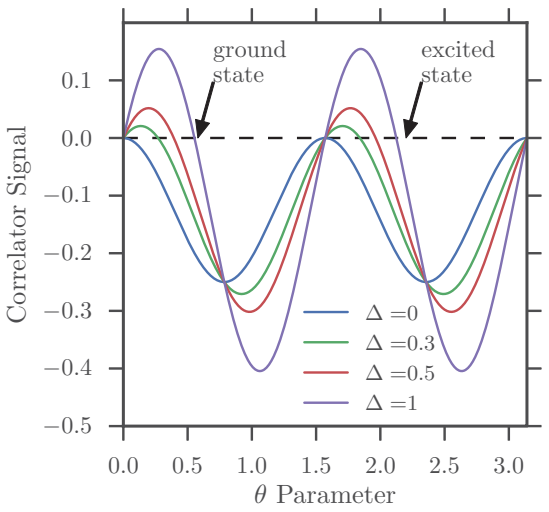

FIG. 6. Correlator signal $\langle\Psi|(\hat{H} - \langle H\rangle)c_1^\dagger c_1|\Psi\rangle$ as a function of the parameter θ appearing in the trial state $\cos\theta|D_1\rangle + \sin\theta|D_2\rangle$. In this example, we have chosen $\epsilon_1 = 0$ and $\epsilon_2 = 1$ and allowed Δ to assume several values between 0 and 1.

V. COMPARING REAL AND ORBITAL SPACES: THE TiO MOLECULE

We now proceed to use the technique to selectively improve wave function parametrizations in a more challenging case. As an example of a system where we do not know *a priori* the most important degrees of freedom, we consider a transition metal molecule TiO. The dynamic correlation present in transition metal systems is larger than in s-p systems like the dimers considered above, so the Jastrow factor could be expected to play a larger role [32,33].

In Fig. 7 the covariances of the 1RDM and the real-space electron-electron correlation function $g(\mathbf{r})$ distance are shown. The covariance signal for the 1RDM is very small, much smaller than for N_2, although we do obtain larger signals for the p and d states as one would expect. Indeed, we also find very little covariance with 2RDM matrix elements within the statistical noise. On the other hand, for our starting wave function, labeled J12, with 12 three-body parameters per atom, there is a large spin-dependent covariance with $g(\mathbf{r})$. So, from these considerations, one might expect that adding determinants would be inefficient, while improving the Jastrow factor, in particular spin-dependent terms, would be more fruitful. That is, the dynamic correlation is more poorly described in our starting wave function than the static correlation.

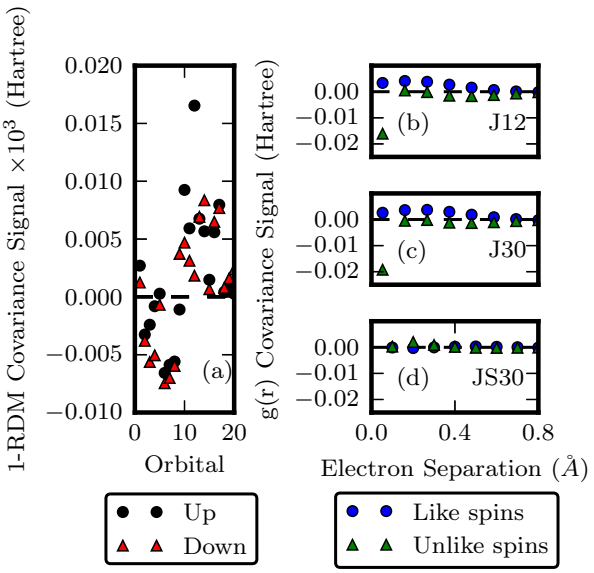

FIG. 7. (a) Covariance of the 1RDM with the local energy E_L for TiO in the J12 wave function. (b) Covariance of the pair distribution $g(\mathbf{r})$ with the local energy E_L for the J12, J30, and JS30 wave functions in both spin channels (right): J12, Slater-Jastrow state with 12 parameters per atom in the three-body part of the Jastrow factor; J30, J12, but with 30 three-body terms instead of 12; JS30, J30, but with a spin-dependent two-body portion of the Jastrow factor.

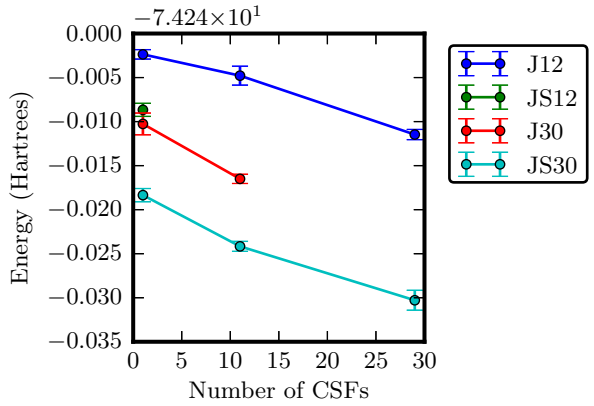

FIG. 8. Total TiO VMC energy vs the number of included CSFs using different Jastrow factors. Note that the decline in energy is quite modest with respect to the number of included CSFs, but falls dramatically when spin dependence is incorporated in the two-body portion of the Jastrow factor.

Since the determinant selection of TiO via energy covariance was not efficient, we used a CI calculation with up to sextuple excitations into eight virtual states to select CSFs and then formed a set of multi-Slater-Jastrow wave functions. If the covariance analysis was correct, then we would expect the spin-dependent terms in the Jastrow to be most effective in lowering the energy, followed by either the extra three-body terms or multiple determinants. As can be seen in Fig. 8, this supposition is correct: With only four parameters, the spin-dependent terms lower the energy by nearly 10 mhartree, while 30 determinants or a similar number of three-body parameters are necessary to achieve that decrease in energy.

This example illustrates some strengths and weaknesses of the covariance-based selection. If the set $\{A_i\}$ is selected in a basis that does not describe the needed improvement efficiently, in this case the determinant basis, then it is not the best tool. On the other hand, if several different basis sets are used, then the best basis can be used to improve the wave function. In this case, we learned that a spin-dependent Jastrow factor can improve the energy significantly for magnetic molecules, while the determinant basis is not an efficient way to improve the wave function for this molecule. The cost for performing these calculations was about a factor of 2 larger than a variational Monte Carlo (VMC) calculation and much smaller than the energy optimization technique.

VI. CONCLUSION

We have presented an outline of a technique to select not just terms in a many-body *Ansatz*, but which type of *Ansatz* with which to proceed. For example, the selection method can quickly determine whether a determinant-type basis is appropriate by evaluating the 1RDM covariance with the local energy. Similarly, if an explicitly correlated approach such as a Jastrow one is more appropriate, then the covariance of the local energy with the electron-electron distance $g(r)$ is large. The computational cost of this assessment is quite low: $g(r)$ is essentially zero cost over a VMC energy evaluation and the 1RDM is approximately a factor of 2 additional, regardless of system size. This is much less expensive than attempting energy minimization on multiple *Ansätze*.

As proof of concept, we demonstrated that the selection technique both selects the correct directions in Hilbert within a defined *Ansatz* space and also can select between alternate viewpoints of the electron correlation problem. We demonstrated the former by selecting determinants for stretched dimer molecules and the latter by differentiating between short-range dynamic correlation best described by a Jastrow factor and long-range static correlation best described by multiple determinants in the transition metal oxygen system TiO. Using standard wave functions for this problem, the dynamic correlation in TiO is more important. This work forms the base for an algorithm in which the local energy can be analyzed directly in the many-body space using feature extraction techniques to describe the most efficient basis in which to improve many-body wave functions.

ACKNOWLEDGMENTS

This material is based upon work supported by NSF Grant No. DMR-1206242 (L.K.W.) and the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1144245 (K.T.W.). We also acknowledge computer resources from the Campus Cluster program at Illinois. Useful conversations with David Ceperley are gratefully acknowledged.

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[2] L. K. Wagner, Int. J. Quantum Chem. **114**, 94 (2014).
[3] L. K. Wagner, Phys. Rev. B **92**, 161116 (2015).
[4] K. Foyevtsova, J. T. Krogel, J. Kim, P. R. C. Kent, E. Dagotto, and F. A. Reboredo, Phys. Rev. X **4**, 031003 (2014).
[5] H. Zheng and L. K. Wagner, Phys. Rev. Lett. **114**, 176401 (2015).
[6] R. Jastrow, Phys. Rev. **98**, 1479 (1955).
[7] D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977).
[8] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, J. Chem. Phys. **136**, 124116 (2012).
[9] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).
[10] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **58**, 6800 (1998).
[11] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006).
[12] E. Neuscamman, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **85**, 045103 (2012).
[13] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).
[14] C. Filippi and S. Fahy, J. Chem. Phys. **112**, 3523 (2000).
[15] D. Prendergast, D. Bevan, and S. Fahy, Phys. Rev. B **66**, 155104 (2002).

[16] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003).
[17] S. Sorella, Phys. Rev. Lett. **80**, 4558 (1998).
[18] S. Sorella and L. Capriotti, Phys. Rev. B **61**, 2599 (2000).
[19] S. Sorella, Phys. Rev. B **64**, 024512 (2001).
[20] S. Sorella, M. Casula, and D. Rocca, J. Chem. Phys. **127**, 014105 (2007).
[21] F. Schautz and C. Filippi, J. Chem. Phys. **120**, 10931 (2004).
[22] A. Scemama and C. Filippi, Phys. Rev. B **73**, 241101 (2006).
[23] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[24] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by G. Scuseria, K. Kim, C. Dykstra, and G. Frenking (Elsevier, Amsterdam, 2005), pp. 1167–1189.
[25] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007).
[26] L. K. Wagner, M. Bajdich, and L. Mitas, J. Comput. Phys. **228**, 3390 (2009).
[27] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).
[28] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **99**, 179902 (2007).
[29] K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).
[30] L. K. Wagner, J. Chem. Phys. **138**, 094106 (2013).
[31] R. C. Clay and M. A. Morales, J. Chem. Phys. **142**, 234103 (2015).
[32] L. Visscher, T. Saue, W. C. Nieuwpoort, K. Faegri, and O. Gropen, J. Chem. Phys. **99**, 6704 (1993).
[33] B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark, J. Phys. Chem. A **109**, 6575 (2005).