JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

General Correlated Geminal Ansatz for Electronic Structure Calculations: Exploiting Pfaffians in Place of Determinants

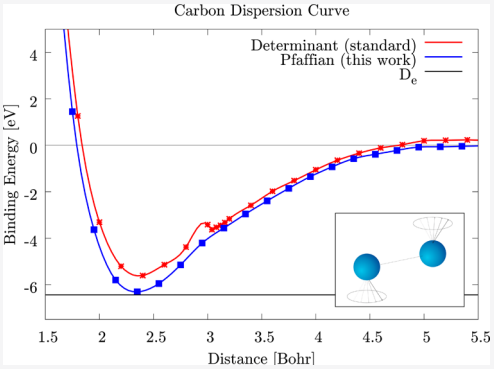
Claudio Genovese,* Tomonori Shirakawa, Kousuke Nakano, and Sandro Sorella*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 6114−6131

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: We propose here a single Pfaffian correlated variational ansatz that dramatically improves the accuracy with respect to the single determinant one, while remaining at a similar computational cost. A much larger correlation energy is indeed determined by the most general two electron pairing function, including both singlet and triplet channels, combined with a many-body Jastrow factor, including all possible spin−spin, spin−density, and density−density terms. The main technical ingredient to exploit this accuracy is the use of the Pfaffian for antisymmetrizing a highly correlated pairing function, thus recovering the Fermi statistics for electrons with an affordable computational cost. Moreover, the application of the diffusion Monte Carlo, within the fixed node approximation, allows us to obtain very accurate binding energies for the first preliminary calculations reported in this study: C_2, N_2, and O_2 and the benzene molecule. This is promising and remarkable, considering that they represent extremely difficult molecules even for computationally demanding multideterminant approaches, and opens therefore the way for realistic and accurate electronic simulations with an algorithm scaling at most as the fourth power of the number of electrons.



1. INTRODUCTION

The accurate determination of the many-electron wave function (WF) has always been a challenging task starting from the early stage of quantum mechanics.[1] So far, several attempts have been made toward this direction ranging from CCSD(T)[2] to tensor network[3] and density matrix renormalization group (DMRG),[4] up to the very recent breakthrough with the use of machine learning methodologies.[5] All these schemes pay the price of being computationally demanding, with a computational complexity ranging from a large degree polynomial of the number of electrons to exponential complexity.

Quantum Monte Carlo (QMC) techniques for electronic structure calculations have proven to be very successful in describing the electronic correlation encoded in a many-body WF.[6−8] In particular, the variational Monte Carlo (VMC)[9−11] samples the real space electronic configurations of the considered system with a probability distribution given by the WF square, thus providing the efficient evaluation not only of the total energy but also of the expectation values of most commonly used many-body operators. Within VMC, it is possible to improve the description of the ground state (GS) WF by minimization of the total energy expectation value. The WF obtained can be used as it is or further improved by the diffusion Monte Carlo calculation (DMC) method.[10−13] This technique is a projection algorithm performed statistically using the information on the sign contained in the given WF, dubbed here as a guiding function. In this way, we can considerably improve the description of the GS, projecting on the lowest possible energy WF with the same signs of the guiding WF. In the ideal case of a guiding function that, for every configuration, has the same sign of the GS, the above-described DMC algorithm provides the exact solution.[10,11]

In the framework of QMC, different *ansatzs* are used to approximate the true GS WF, with the purpose to achieve an affordable compromise between the accuracy of the calculation and its computational cost. Though a good representation of the GS can sometimes be achieved with a simple and "cheap" WF, in most cases, the use of a very complicated and computationally demanding *ansatz* is necessary to get a correct answer.

Slater determinants (SDs) are the simplest fermionic WF used for QMC. They provide a single particle picture of the quantum many-body problem, preserving the Pauli principle, that is, the Fermi statistics for electrons. They can be obtained directly from mean-field calculations. Unfortunately in many

Received: February 18, 2020
Published: August 17, 2020

situations of interest, it is not possible to give a good description of the system in terms of a single SD.[14−17] In QMC, there are two possible strategies to overcome this problem: the use of a linear combination of different SDs[17−21] or different *ansatzs* with larger variational freedom.[22,23] The multideterminant WFs can be systematically improved and in principle can describe exactly every GS with a large enough number of SDs. Unfortunately the number of SDs that has to be taken into account scales exponentially with the number of electrons preventing the calculations on large systems.[24]

The use of pairing function replaces the single particle description of the SD with a richer one in terms of electron pairs. The corresponding WF is a natural extension of the single SD *ansatz* and represents a direct and efficient implementation of the Anderson resonating valence bond (RVB)[25] theory of many-electron WFs. In particular, it provides a direct description of the singlet and triplet correlations that are absent in the SD. Depending on the definition of the pairing function, qualitatively different WFs can be obtained. They will be described in Section 2, where we will focus also on the technical details required for the calculation. We will introduce the symmetric antisymmetrized geminal power (AGP)[22,26,27] and the broken symmetry antisymmetrized geminal power (AGPu),[28] but we will mainly focus on the most general AGP. In the previous literature,[23,29] it has been indicated as Pfaffian WF, and people have been referring to the AGPs as AGP, but since the AGP (or Pfaffian WF) literally realizes the most general antisymmetrized geminal power, we dub this case with the shortest acronym, that is, AGP. It will be shown that this approach becomes very efficient in combination with an explicit correlation term, known as a Jastrow factor (JF),[14,30,31] that promotes or penalizes the bonds according to the electronic correlation. As it will be shown later, we have introduced a quite general JF, depending both on spin and electron charges. When it is applied to an AGP without definite spin, it allows its almost complete restoration, mimicking in this way a spin projection operation that, though approximate, is much cheaper than other approaches.[32,33] Even if the pairing functions cannot be improved systematically, these WFs have a much larger variational freedom than the SDs, nevertheless remaining with a similar computational cost.

If on one hand, for the multideterminant WF, the calculation can be computationally very expensive, on the other hand, for the pairing functions, the optimization of a large number of nonlinear variational parameters can become a serious limitation if not handled efficiently. Indeed, in order to exploit the full potential of these *ansatzes*, it has been fundamental to use the most recent techniques for the calculation of the derivatives and optimization strategies.

In a previous attempt, the AGP WF was used by exploiting only a very small fraction of the large variational freedom of the *ansatz*.[23,29] The results were not encouraging, and the energies obtained with this *ansatz* did not improve the ones of the AGPs that, in turn, has a lower computational cost. Despite the Pfaffian was no longer used in the electronic system to our knowledge, the experience with lattice models has shown that the AGP WF is able to improve considerably the description of magnetic and correlated systems.[34] Moreover, the introduction of a powerful JF and the recent results obtained in combination with the AGPs[14,35,36] encouraged us to look for the unexpressed potential of the full AGP WF.

In this paper, we will compare the results obtained with AGP WF with available state of the art VMC and DMC calculations. In particular, we benchmark our WF on the diatomic molecules with corresponding high spin atoms in the first row of the periodic table, i.e., carbon, nitrogen, and oxygen, and on the benzene. The first ones are systems that, despite their apparent simplicity, represent useful benchmarks for many highly correlated methods.[37−40] We will show that, with the use of our best WFs, even with a very compact basis set, we are able to achieve an accuracy comparable with the state of the art multideterminant WFs at a computational cost similar to the one of a single SD. Not only the total energies and the dissociation energies are extremely accurate but we also analyzed the magnetic proprieties of these molecules, unveiling the unexpected rich physics behind these systems. Finally we consider the benzene molecule, a system that represents the prototypical example of the RVB theory and thus a fundamental test case for our approach.

2. WFS AND PROCEDURES

For all calculations we present in this paper, we used the TurboRVB package for QMC calculations.[41,42] The WFs used for this work are factorized as the product of a fermionic mean field and an explicit bosonic correlation factor. Being $\Psi(\mathbf{X})$ the WF of a given configuration $\mathbf{X} = (\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, ..., \mathbf{r}_N\sigma_N)$ of N electrons of spins σ_i and positions $\mathbf{r}_i$, we can write $\Psi(\mathbf{X})$ as

$$\Psi(\mathbf{X}) = J(\mathbf{X}) \times \Phi(\mathbf{X}) \tag{1}$$

where $\Phi(\mathbf{X})$ takes into account the fermionic nature of the electrons, while $J(\mathbf{X})$ is the JF: an exponential modulation of the WF that substantially improves the electronic correlation description for all types of WFs studied here. The fermionic term of the WF, dubbed as $\Phi(\mathbf{X})$ in eq 1, is the most important part, directly encoding the behavior of the electrons while imposing the antisymmetrization under particle exchange. In the following section, we will describe the basis set used, the definition of the AGPs, AGPu, and AGP after a brief introduction to the SD. Finally we will discuss the JF correlator.

2.1. Basis Set. We expand our ansatz in an atom-centered basis set of Gaussian orbitals for the calculation of the JF and a hybrid basis set for the fermionic part of the WF, as it will be discussed below. The Gaussian orbitals basis set is indicated as $\{\phi_{I,\nu}(\mathbf{r})\}$, with each element being the νth orbital centered on the Ith atom at the position $\mathbf{R}_I$. The elements in the basis set have the form

$$\phi_{I,\nu}(\mathbf{r}) \propto e^{-Z_\nu|\mathbf{r}-\mathbf{R}_I|^2}[Y_{l_\nu,m_\nu} \pm Y_{l_\nu,-m_\nu}] \tag{2}$$

where Z_ν is a numerical coefficient that describes how diffused the atomic orbital is around the atom, while Y_{l_ν,m_ν} is the spherical harmonic function with angular quantum numbers l_ν and m_ν, corresponding to the orbital type ν which is always assumed to be real. This basis set has been used without further contractions for the description of the JF. Instead, for the fermionic part of our WF, we have used hybrid atomic orbitals (HOs)[26,27] to expand them over a richer set of Gaussian orbitals and, by means of the contraction, considering only an affordable number of variational parameters. The HOs, indeed, are obtained as linear combinations of all the elements of the Gaussian basis set used for a given atom, labeled by I

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

$$\bar{\phi}_{I,\omega}(\mathbf{r}) = \sum_{\nu} \mu_{\omega,\nu}\phi_{I,\nu}(\mathbf{r}) \tag{3}$$

The above hybrid orbitals allow us to take into account the modification of the standard Slater orbitals corresponding to isolated atoms, to the case when they are instead placed in a complex environment. Therefore, we have chosen to use a number of hybrid orbitals equal to the single particle ones occupied in the absence of electron−electron interactions and including also all the ones corresponding to the same shell of degenerate one particle levels. The corresponding orbitals are the ones that should physically play a role in the considered electronic systems. Hence, in all of the first row molecules, we have considered the full hybridization of five atomic orbitals, coming from two s-wave and three p-wave ones, that can be corrected by several components with much higher angular momenta. This is because the full spherical symmetry is no longer satisfied even in a simple homonuclear molecule. For the sake of compactness, we indicate here all basis elements as $\{\phi_k(\mathbf{r})\}$ combining the indices ω and I, and I and ν in a single index k for a lighter notation. Every time we refer to the AGPs, AGPu, and AGP, it is meant to be a basis of HOs.

The exponents Z_ν have been chosen from the ccpVDZ or ccpVTZ basis set according to this criterium: the contraction are removed and all exponents with $Z_\nu > 150$ a.u^{-2}. are eliminated. This is possible because contracted orbitals containing very large exponents are necessary only with a pure Gaussian basis in order to satisfy the electron-ion cusp conditions. They are instead appropriately considered by the one-body term of our JF, as described in Section 2.4. The exponents chosen are then further optimized at molecular equilibrium distance and kept fixed in the corresponding atomic calculation (where the optimization of the exponents has an almost negligible effect) and the dispersion energy curves.

2.2. Slater Determinant. Here, we will provide a preparation description of the SD that is important both for the initialization of the pairing function and for comparing our results with the existing literature.[17] From theoretical and computational point of view, the simplest fermionic WF is the SD, called Jastrow SD (JSD) in the presence of a JF. The SD is built from the vacuum state by populating a number of orthogonal single particle molecular orbitals (MOs) equal to the number of electrons in the system. Henceforth, we omit the spin indices, by assuming that each spin component corresponds to a different SD. In our basis, the MO are in the form

$$\Phi_\alpha^{\mathrm{mol}}(\mathbf{r}) = \sum_{k=1} P_{\alpha,k}\phi_k(\mathbf{r}) \tag{4}$$

The MOs can be obtained directly from a density functional theory (DFT) or Hartree−Fock calculation, but they can also be further optimized with VMC.[43] It is well known that the antisymmetric product of these MO leads to the determinant of the matrix in which every MO is evaluated for each electron position

$$\Phi_{\mathrm{SD}}(\mathbf{X}) = \det\begin{pmatrix} \Phi_1^{\mathrm{mol}}(\mathbf{r}_1) & \Phi_1^{\mathrm{mol}}(\mathbf{r}_2) & \cdots & \Phi_1^{\mathrm{mol}}(\mathbf{r}_N) \\ \Phi_2^{\mathrm{mol}}(\mathbf{r}_1) & \Phi_2^{\mathrm{mol}}(\mathbf{r}_2) & \cdots & \Phi_2^{\mathrm{mol}}(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_N^{\mathrm{mol}}(\mathbf{r}_1) & \Phi_N^{\mathrm{mol}}(\mathbf{r}_2) & \cdots & \Phi_N^{\mathrm{mol}}(\mathbf{r}_N) \end{pmatrix} \tag{5}$$

For weakly correlated systems, the JSD can often give reasonably good results with an affordable computational cost and a limited number of variational parameters. It is also a common choice to use a linear combination of SDs to improve the description of the WF, with *ansatzs* that take different names depending on the type and number of SDs considered. In this paper, we will compare directly the results of our WFs to the ones obtained with one of the most successful multideterminant WFs, the full valence complete active space (FVCAS) WF.

2.3. Pairing Function. The use of the pairing function in correlated WFs allows an electronic description that goes beyond the single particle picture of the SD. The building block of this WF has the following general form

$$f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \sum_{k,l} \lambda_{k,l}^{\sigma_1\sigma_2}\phi_{k\sigma_1}(\mathbf{r}_1)\phi_{l\sigma_2}(\mathbf{r}_2) \tag{6}$$

where all the elements of the matrix λ represent most of the WF variational parameters. They depend on the orbitals considered and on the spin σ_1, σ_2 of the so called geminal function f. In principle, when we break the spin symmetry, the basis sets used for ↑ and ↓ electrons can be different, otherwise the chosen basis does not depend on the spin component. In order to set up a consistent many-body WF starting from the geminal, several choices are possible depending on the criteria adopted for the definition of the geminal. To highlight the different possibilities, we can recast eq 6 in a way in which the spin dependency is more explicit

$$\begin{aligned} f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)f_+(\mathbf{r}_1, \mathbf{r}_2) \\ &+ \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)f_-(\mathbf{r}_1, \mathbf{r}_2) \\ &+ |\uparrow\uparrow\rangle f_\uparrow(r_1, r_2) + |\downarrow\downarrow\rangle f_\downarrow(\mathbf{r}_1, \mathbf{r}_2) \end{aligned} \tag{7}$$

where

$$\begin{aligned} f_\pm(\mathbf{r}_1, \mathbf{r}_2) &= f(\mathbf{r}_1\uparrow, \mathbf{r}_2\downarrow) \pm f(\mathbf{r}_1\downarrow, \mathbf{r}_2\uparrow), \\ f_\sigma(\mathbf{r}_1, \mathbf{r}_2) &= f(\mathbf{r}_1\sigma, \mathbf{r}_2\sigma) \qquad \text{with } \sigma = \uparrow, \downarrow \end{aligned} \tag{8}$$

In order to satisfy the Pauli principle, we have $f_\pm(\mathbf{r}_1,\mathbf{r}_2) = \pm f_\pm(\mathbf{r}_2,\mathbf{r}_1)$ and $f_\sigma(\mathbf{r}_1,\mathbf{r}_2) = -f_\sigma(\mathbf{r}_2,\mathbf{r}_1)$ for $\sigma = \uparrow, \downarrow$. Our WF is then obtained by antisymmetrizing the product over all electron pairs considered that, by definition, occupy the same pairing function. For simplicity, we will enumerate the spin up electrons from 1 to $N_\uparrow$ and the spin down ones from $N_\uparrow + 1$ to N.

As suggested by the name AGP, our goal is to define a WF that is literally the antisymmetrized product of the geminals and the unpaired orbitals (if present), namely

$$\begin{aligned} \Phi(\mathbf{X}) = \sum_\alpha \mathrm{Sgn}(\alpha)(&f(\mathbf{r}_{1_\alpha}\sigma_{1_\alpha}, \mathbf{r}_{2_\alpha}\sigma_{2_\alpha})f(\mathbf{r}_{3_\alpha}\sigma_{3_\alpha}, \mathbf{r}_{4_\alpha}\sigma_{4_\alpha})\cdots \\ &f(\mathbf{r}_{p-1_\alpha}\sigma_{p-1_\alpha}, \mathbf{r}_{p_\alpha}\sigma_{p_\alpha})\Theta_1(r_{p+1_\alpha})\cdots\Theta_{N-p}(\mathbf{r}_{N_\alpha})) \end{aligned} \tag{9}$$

where α is one of the possible way of distributing the N electrons between the $p/2$ pairs and the $N - p$ unpaired orbitals Θ, and $\mathrm{Sgn}(\alpha)$ is the sign of the corresponding permutation of the particles that is required to ensure the fermionic behavior. In particular, different choices of the pairing function, obtained by excluding one or more terms in the eq 8, lead to different ways to compute eq 9. These choices

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

also impact quantitatively and qualitatively on the kind of physics that we can describe by means of this type of WF. Therefore, we will distinguish among three distinct cases: if we consider only the singlet term in eq 8 we obtain the AGPs, if we include the singlet and the $S_z = 0$ triplet term we have the AGPu, while the most general case is just the definition adopted here for the AGP.

2.3.1. AGPs. Let us consider for the moment the unpolarized case $N_\uparrow = N_\downarrow$, the extension to the polarized cases will be straightforward and will be discussed later on. When no triplet correlations are allowed, we build our WFs using only singlet pairs, and the pairing function in eq 7 contains only the symmetric element f_+

$$f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)f_+(\mathbf{r}_1, \mathbf{r}_2) \tag{10}$$

In this case, we project a perfect singlet that we denote as AGPs. The λ matrix elements in eq 10 are nonzero only for $\sigma_1 \neq \sigma_2$, and they are symmetric for spin exchange. In order to calculate the AGPs, we can write all the possible combinations of pairs of opposite spin electrons in a matrix defined as

$$F = \begin{pmatrix} f(\mathbf{r}_1\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & f(\mathbf{r}_1\uparrow, \mathbf{r}_{N_\uparrow+2}\downarrow) & \cdots & f(\mathbf{r}_1\uparrow, \mathbf{r}_N\downarrow) \\ f(\mathbf{r}_2\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & f(\mathbf{r}_2\uparrow, \mathbf{r}_{N_\uparrow+2}\downarrow) & \cdots & f(\mathbf{r}_2\uparrow, \mathbf{r}_N\downarrow) \\ \vdots & \vdots & \ddots & \vdots \\ f(\mathbf{r}_{N_\uparrow}\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & f(\mathbf{r}_{N_\uparrow}\uparrow, \mathbf{r}_{N_\uparrow+2}\downarrow) & \cdots & f(\mathbf{r}_{N_\uparrow}\uparrow, \mathbf{r}_N\downarrow) \end{pmatrix} \tag{11}$$

In this way, each row of the matrix corresponds to an electron of spin ↑ and each column an electron of spin ↓. The definition of the matrix F in this form is convenient because it allows the antisymmetrization requested by the eq 9 in a simple and efficient way. Indeed, it can be demonstrated[26] that the correct antisymmetrization of the pairs considered in this case is given by

$$\Phi_{\mathrm{AGPs}}(\mathbf{X}) = \det F \tag{12}$$

This is somehow intuitive because we want to sum all the possible products of $N/2$ matrix elements of F, where in all these factors, a column element or a row element is present only once, exhausting all the possible configurations of the system considered with an appropriate ± sign that, in this case, is just given by the one corresponding to the determinant of F.

When the system is polarized and $N_\uparrow \neq N_\downarrow$, we cannot build the solution using only the singlet terms because the matrix F written as in eq 11 is a rectangular matrix and its determinant cannot be computed. Supposing for simplicity that $N_\uparrow > N_\downarrow$, in this case, we have to add a number $N_\uparrow - N_\downarrow$ of unpaired spin-up MOs $\{\Theta_i(\mathbf{r})\}$ not only for fulfilling the polarization required but also, most importantly, to turn the matrix F to a perfectly defined square matrix

$$F = \begin{pmatrix} f(\mathbf{r}_1\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & \cdots & \Theta_1(\mathbf{r}_1) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_1) \\ f(\mathbf{r}_2\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & \cdots & \Theta_1(\mathbf{r}_2) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_2) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ f(\mathbf{r}_{N_\uparrow}\uparrow, \mathbf{r}_{N_\uparrow+1}\downarrow) & \cdots & \Theta_1(\mathbf{r}_{N_\uparrow}) & \cdots & \Theta_{N_\uparrow-N_\downarrow}(\mathbf{r}_{N_\uparrow}) \end{pmatrix} \tag{13}$$

Also in this case, a consistent antisymmetric WF can be again calculated as the determinant[26] of the matrix F exactly in the same way as the singlet pairing in eq 12.

2.3.2. AGPu. For the AGPu, only the parallel spin term of the triplet component are omitted. This means that the spin symmetry is broken, and a magnetic order parameter can be directed along the z-quantization axis. This WF is called the broken symmetry AGP (AGPu), and the difference from the previous AGPs is the presence of the antisymmetric f_- component in the definition of the pairing function in eq 7, that for this case is

$$f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)f_+(\mathbf{r}_1, \mathbf{r}_2) + \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)f_-(\mathbf{r}_1, \mathbf{r}_2) \tag{14}$$

In order to define this pairing function, we break the spin symmetry in the opposite electron spin case with $\sigma_1 \neq \sigma_2$, by keeping equal to zero the $\sigma_1 = \sigma_2$ components of eq 6. With exactly the same procedure used in the case of the AGPs, depending on the polarization, we can build the same matrix F of eqs 11 or 13 that is no longer symmetric now. Even in this case, the correct antisymmetrized sum of these pairs is given by the determinant.[26] Thus, analogously to eq 12, we obtain

$$\Phi_{\mathrm{AGPu}}(\mathbf{X}) = \det F \tag{15}$$

that implements the simplest broken symmetry *ansatz* based on the pairing function.

2.3.3. AGP. The AGP (also known in the literature as Pfaffian WF[23]) is in our opinion the most important pairing function, being the most general one and encoding new variational freedoms into the AGPs and the AGPu. We will show that it represents the most powerful description of the chemical bond within the paradigm developed in this work. This WF represents also the most general mean-field state, that is, the GS of a mean-field Hamiltonian containing BCS anomalous terms projected on a given number N of particles and total spin projection $S_{\mathrm{tot}}{}^z = \sum_{i=1}^N \sigma_i$ along the z-quantization axis. In this case, the definition of the pairing function is exactly the one in eq 7, containing all terms including the parallel spin terms of the triplet. This means that now, when we build the AGP, we have to include also the parallel spin electron pairs in the WF. In this way, the AGP can also describe a magnetic order parameter in any direction of the space, and thus, it is also possible to rotate the spin component of the WF in any direction. This will allow us to break the symmetry along the spin-quantization axis and then rotate it. As we will explain later, this plays a crucial role when we use this WF in combination with our JF because it allows us to preserve the total S_z of the molecules and include spin fluctuations.

Of course, we cannot create a WF using only pairs if the number of electrons in the system is odd, so, for the moment, let us assume N is even. The extension to the odd number of electrons is trivial and will be discussed immediately after. We will dub as W the matrix containing all possible pairs

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

$$W = \begin{pmatrix} 0 & f(\mathbf{r}_1\uparrow, \mathbf{r}_2\uparrow) & \cdots & f(\mathbf{r}_1\uparrow, \mathbf{r}_N\downarrow) \\ f(\mathbf{r}_2\uparrow, \mathbf{r}_1\uparrow) & 0 & \cdots & f(\mathbf{r}_2\uparrow, \mathbf{r}_N\downarrow) \\ \vdots & \vdots & \ddots & \vdots \\ f(\mathbf{r}_{N-1}\downarrow, \mathbf{r}_1\uparrow) & f(\mathbf{r}_{N-1}\downarrow, \mathbf{r}_2\uparrow) & \cdots & f(\mathbf{r}_{N-1}\downarrow, \mathbf{r}_N\downarrow) \\ f(\mathbf{r}_N\downarrow, \mathbf{r}_1\uparrow) & f(\mathbf{r}_N\downarrow, \mathbf{r}_2\uparrow) & \cdots & 0 \end{pmatrix} \tag{16}$$

where the matrix is antisymmetric for the fermionic commutation rules, and thus, the elements of the diagonal are set to zero. We can recast the W highlighting its different spin sectors as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} \end{pmatrix} \tag{17}$$

where $W_{\uparrow\uparrow}$ and $W_{\downarrow\downarrow}$ are respectively $N_\uparrow \times N_\uparrow$ and $N_\downarrow \times N_\downarrow$ antisymmetric matrices that take into account the parallel spin terms of the triplet, while $W_{\uparrow\downarrow}$ is a $N_\uparrow \times N_\downarrow$ matrix such that $W_{\uparrow\downarrow} = -W_{\downarrow\uparrow}^{\mathrm{T}}$, describing the remaining triplet and singlet contributions. In the case of AGPs and AGPu, we can also build a similar matrix where the matrices $W_{\uparrow\uparrow}$ and $W_{\downarrow\downarrow}$ are identically zero.

Analogously to the case of the AGPs and AGPu, we have to identify a way to calculate the antisymmetric product of all pairs considered. In this case, it is easy to identify the antisymmetrization procedure defined in eq 9 as the Pfaffian of the matrix W. After introducing this algebraic operation, the reason will be straightforward to the reader.

The Pfaffian is an algebraic operation acting on antisymmetric square matrices with an even number of rows and columns. Being N even, the matrix W satisfies these hypothesis. The usual definition of the Pfaffian requires the introduction of the concept of partition of the matrix W

$$A(\alpha) = \mathrm{sign}(\alpha) \prod_{k=1}^{N/2} W_{i_k, j_k} \tag{18}$$

where all i_k and j_k are different, $i_k < j_k$ for each k and $i_1 < i_2 < \ldots < i_N$. The sign(α) is given by the permutation that orders the vector of the indices $\{i_1, j_1, i_2, j_2, \ldots, i_M, j_M\}$. In this way, all indices are considered only once. The Pfaffian is then defined as

$$\mathrm{Pf}(W) = \sum_\alpha A(\alpha) \tag{19}$$

where the sum over α is extended over all possible partitions. However, an alternative definition[44] of the Pfaffian can better clarify the correspondence to the eq 9. It can indeed be defined alternatively as

$$\mathrm{Pf}(W) = [(N/2)!2^{N/2}]^{-1} \sum_P \mathrm{sign}(P) \prod_{k_p=1}^{N/2} W_{i_{k_p}, j_{k_p}} \tag{20}$$

where P now represents a generic permutation of the possible row and column indices of the matrix without any constraints, and the sign(P) is the parity of the permutation. In this definition, it is easy to recognize the antisymmetrized sum corresponding to the eq 9. Let us introduce now a further property of the Pfaffian that will be useful in the following section. We will indicate with 0 a $m \times m$ matrix containing only 0 and B a generic $m \times m$ matrix, we have

$$\mathrm{Pf}\begin{bmatrix} 0 & B \\ -B^T & 0 \end{bmatrix} = (-1)^{m(m-1)/2} \det(B) \tag{21}$$

For odd number of electrons, it is necessary to use a spin-dependent unpaired orbital $\Theta_\sigma(\mathbf{r})$ so that we can accommodate the remaining electron that is not considered by the product of the pairs. The unpaired orbital introduces a supplementary row and column to the matrix W. Being $\Theta_\uparrow = (\Theta_\uparrow(\mathbf{r}_1), \Theta_\uparrow(\mathbf{r}_2), \ldots, \Theta_\uparrow(\mathbf{r}_{N\uparrow}))$ the vector containing the values of the unpaired orbital $\Theta_\uparrow$ at the $\uparrow$ electron positions and $\Theta_\downarrow = (\Theta_\downarrow(\mathbf{r}_{N_\uparrow+1}), \Theta_\downarrow(\mathbf{r}_{N\uparrow+2}), \ldots, \Theta_\downarrow(\mathbf{r}_N))$ the one calculated for the $\downarrow$ electron ones, we modify the matrix in eq 17 as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} & \Theta_\uparrow \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} & \Theta_\downarrow \\ -\Theta_\uparrow^T & -\Theta_\downarrow^T & 0 \end{pmatrix} \tag{22}$$

Also in this case, the permutation sum implied by the Pfaffian leads to the correct antisymmetrization required from eq 9. The matrix W satisfies the hypothesis of the calculation having an even leading matrix dimension $\bar{N} = N + 1$. We can further notice that no assumption has been made on the polarization of the system, and so no unpaired orbital is required except for a single one in the case of odd N.

It is however possible in principle to introduce further pairs of unpaired orbitals, if, for example, we want to describe AGPs or AGPu with a full AGP WF. We define $\Theta_{i\sigma}(\mathbf{r})$ as the set of the considered m unpaired orbitals and $\Theta_{i_\uparrow} = (\Theta_{i,\uparrow}(\mathbf{r}_1), \Theta_{i,\uparrow}(\mathbf{r}_2), \ldots, \Theta_{i,\uparrow}(\mathbf{r}_{N_\uparrow}))$ the vector containing the values of the unpaired orbital $\Theta_{i,\uparrow}$ for the $\uparrow$ electron positions and $\Theta_{i_\downarrow} = (\Theta_{i,\downarrow}(\mathbf{r}_{N_\uparrow+1}), \Theta_{i,\downarrow}(\mathbf{r}_{N_\uparrow+2}), \ldots, \Theta_{i,\downarrow}(\mathbf{r}_N))$ the one calculated for the $\downarrow$ electron ones. We can modify the matrix in eq 17 as

$$W = \begin{pmatrix} W_{\uparrow\uparrow} & W_{\uparrow\downarrow} & \Theta_{1\uparrow} & \cdots & \Theta_{m\uparrow} \\ W_{\downarrow\uparrow} & W_{\downarrow\downarrow} & \Theta_{1\downarrow} & \cdots & \Theta_{m\downarrow} \\ -\Theta_{1\uparrow}^T & -\Theta_{1\downarrow}^T & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\Theta_{m\uparrow}^T & -\Theta_{m\downarrow}^T & 0 & \cdots & 0 \end{pmatrix} \tag{23}$$

that is a $\bar{N} \times \bar{N}$ matrix where $\bar{N} = N + m$. We can again antisymmetrize this product using the definition of the Pfaffian provided in eq 20. A careful reader could have noticed that, by applying the Pfaffian definition, we are antisymmetrizing not only over the electron indices but also over the orbital indices of the unpaired orbitals. This antisymmetrization, however, contains the one over the physical electrons, and leads therefore to a physically allowed electronic WF.

Moreover, we can notice that, by using the previous definition, we can identify the AGPs and the AGPu as subcases of the general AGP. Indeed, by using the expressions of the pairing function and the unpaired orbitals of the AGPs and AGPu, we obtain $W_{\uparrow\uparrow} = 0$, $W_{\downarrow\downarrow} = 0$, $\Theta_{i_\downarrow} = 0$, and $\bar{N} = 2N_\uparrow$. By merging eqs 13 and 17, we can define

$$W = \begin{pmatrix} 0 & F \\ -F^T & 0 \end{pmatrix} \tag{24}$$

and this means that, by applying eq 21, we immediately obtain

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

$$\mathrm{Pf}(W) = \pm\det(F) \tag{25}$$

where the sign only depends on the number of electrons and is constant, thus irrelevant. This shows in a straightforward way that the AGPs and AGPu defined in the previous subsection are nothing but particular cases of the most general AGP.

2.4. Jastrow Factor. Within QMC, it is easy to improve the quality of the WF by multiplying the WF with an exponential JF. This last one enriches the description of the GS by encoding explicitly the electronic correlation, while speeding up the convergence to the complete basis set limit.[7] Indeed, with an appropriate choice, the JF can satisfy exactly the electron−electron and electron-ion cusp conditions of the many-body WF, consequences of the Coulomb $1/r$ singularity at a short distance. In this paper, we introduce a new kind of JF that contains a richer dependence on the spin and that plays a fundamental role when used in combination with the AGP WF. The JF is defined as

$$J(\mathbf{X}) = \mathrm{e}^{U_{ei}+U_{ee}} \tag{26}$$

where U_{ei} is a single body term that deals explicitly with the electron−ion interaction and U_{ee} is a many-body term that properly accounts for the electronic correlation. The single body term is

$$U_{ei} = \sum_{i=1}^{N} u_{ei}(\mathbf{r}_i) \tag{27}$$

with u_{ei} being

$$u_{ei}(\mathbf{r}_i) = \sum_{I=1}^{\#\mathrm{ions}} -Z_I \frac{1-\exp(b_{ei}|\mathbf{r}_i - R_I|)}{b_{ei}} + g_I(\mathbf{r}_i) \tag{28}$$

In eq 28, Z_I is the atomic number of the atom I and b_{ei} is a variational parameter, while $g_I(\mathbf{r}_i)$ encodes the most general nonhomogeneous electron−ion one-body term, that is, depending explicitly on all nuclear and electron coordinates and not only on their relative distances, that is defined as

$$g_I(\mathbf{r}_i) = \sum_{\nu} \xi_{I,\nu}\phi_{I,\nu}(\mathbf{r}_i) \tag{29}$$

where the summation is extended over all Gaussian orbitals in the JF basis set centered on the Ith atom. The electron−electron term instead is written as

$$U_{ee} = \sum_{i<j} u_{ee}(\mathbf{r}_i\sigma_i, \mathbf{r}_j\sigma_j) \tag{30}$$

where the sum is extended over the pairs of different electrons and where

$$u_{ee}(\mathbf{r}_i\sigma_i, \mathbf{r}_j\sigma_j) = k_{\sigma_i,\sigma_j}\frac{|\mathbf{r}_i - \mathbf{r}_j|}{1 + b^{ee}_{\sigma_i,\sigma_j}|\mathbf{r}_i - \mathbf{r}_j|} + g_{ee}(\mathbf{r}_i\sigma_i, \mathbf{r}_j\sigma_j) \tag{31}$$

with the 2×2 matrix ${b_{\sigma,\sigma'}}^{ee}$ described by one ${b_{\sigma,\sigma'}}^{ee} = b^{ee}$ or two variational parameters for $\sigma_i = \sigma_j$ when $k_{\sigma_i,\sigma_j} = 1/4$ and ${b_{\sigma,\sigma'}}^{ee} = b_{\parallel}^{ee}$ for $\sigma_i \neq \sigma_j$ when $k_{\sigma_i,\sigma_j} = 1/2$ and ${b_{\sigma,\sigma'}}^{ee} = b_{\perp}^{ee}$. The conventional expression for the JF can be obtained by removing all spin dependencies in the previous expressions and considering only the variational parameters corresponding to the opposite spin case $k_{\sigma_i,\sigma_j} = 1/2$ and ${b_{\sigma,\sigma'}}^{ee} = b^{ee}$.

In our expression, the first term in eq 31 deals explicitly with the electron−electron cusp conditions, the second term in eq 31 instead is a bosonic pairing function in the form

$$g_{ee}(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) = \sum_{k,l} \zeta_{k,l}^{\sigma_1\sigma_2}\phi_{k\sigma_1}(\mathbf{r}_1)\phi_{l\sigma_2}(\mathbf{r}_2) \tag{32}$$

with the elements of the matrix ζ defining further variational parameters. Notice that both g_I and g_{ee} do not affect the cusp conditions because they are expanded over cuspless Gaussian orbitals. The g_{ee} term has the same form of eq 6, but since the fermionic behavior is already encoded in the fermionic part of the WF, this term is symmetric under particle exchange. The use of a pairing function in the JF enriches the description of the charge and spin correlations of the system, noticeably improving the quality of the global WF. It is a common practice to adopt a simplified or even absent spin dependency in the function u of eq 31. This is often accurate for systems where the magnetic properties are not relevant. We will refer to it below with the prefix Js in the WF, in contrast with the prefix J used for the full spin-dependent JF.

A perfect singlet remains as such after the multiplication of a spin-independent Jastrow, and so our spin-dependent JF is not appropriate if we do not want to break the spin symmetry. It is, instead, necessary if we want to recover, at least approximately, the singlet from a spin contaminated broken symmetry ansatz. A general spin-dependent u, as defined in eq 31, is therefore of fundamental importance for the AGPu or the AGP ansatzs.

Let us start with a simple example. We consider two atoms with opposite spins and break the spin symmetry by orienting the spins of the atoms along the z-quantization axis. In this case, the JF is not able to change the classical antiferromagnetic spin state because it acts as an irrelevant constant when applied to it. It is instead more physical to orient the spin moment of the atoms in a direction perpendicular to the quantization axis chosen for the JF. In this way, the JF can act on the electrons and the spins while the magnetic moment is free to fluctuate and recover its genuine quantum character. As previously mentioned, with the AGP it is possible to rotate the spin of the WF in every direction and orient the magnetic moment in any direction of the space. This works particularly well in combination with our JF that can suppress the unfavored triplet configurations with parallel spins generated by the rotation. This optimal spin orientation of the atoms, that is perpendicular to the JF one, is rigorously valid within the well-known spin-wave theory of a quantum antiferromagnet.[34] In this case, the JF defined with a spin-quantization axis perpendicular to the magnetic moment of the atoms allows the description of the quantum fluctuations and the corresponding zero point energy (ZPE), even for a finite (as is our case) number of atoms.[34]

2.5. Procedure. The first step to calculate and optimize our WF is to identify a reasonable starting point. We chose to start from a DFT calculation because of its flexibility. We have used LDA calculations for spin symmetric systems, while for the ones with opposite spin antiferromagnetic moments we have broken the symmetry with a LSDA calculation, by adopting an appropriate initialization of the WF. The SD obtained from a DFT calculation is mapped without loss of information into AGPs or AGPu and then in a second analogous step, we convert the AGPs and AGPu into a full AGP.

For the first conversion, let us consider eq 6. If we compute it in the basis set of the MOs obtained from the DFT, we have

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

$$f(\mathbf{r}_1 \uparrow, \mathbf{r}_2 \downarrow) = \sum_{\alpha} \bar{\lambda}^{\uparrow\downarrow}_{\alpha,\alpha} \Phi^{\mathrm{mol}}_{\alpha\uparrow}(\mathbf{r}_1)\Phi^{\mathrm{mol}}_{\alpha\downarrow}(\mathbf{r}_2) \tag{33}$$

namely only the diagonal terms in the matrix $\bar{\lambda}$ are present. Moreover, we can also remove the spin dependency if there is no symmetry breaking. For the polarized case, the unpaired orbitals are the last occupied MOs. By substituting the definition of the MOs with their expansion over a localized atomic basis set, as given in eq 4, we can recast the above equation exactly in the same form shown in eq 6 with a matrix λ

$$\lambda_{k,l} = \sum_{k} P^{\dagger}_{k,\alpha} \bar{\lambda}^{\uparrow\downarrow}_{\alpha,\alpha} P_{\alpha,l} \tag{34}$$

When we convert AGPs or AGPu into an AGP WF, we already have an initialization for the sectors of the pairs with different spins $W_{\uparrow\downarrow}$ and $W_{\downarrow\uparrow}$ that can be obtained directly from the AGPs or AGPu pairing functions. The main challenge is to find a reasonable initialization for the two sectors $W_{\downarrow\downarrow}$ and $W_{\uparrow\uparrow}$ that are not described by the AGPs or AGPu.

There are two different procedures that we can follow: the first one is used for polarized systems, and the second one instead is preferred in the case of broken spin symmetry and in the presence of antiferromagnetism, namely, molecules well described by opposite atomic magnetic moments. If there is no antiferromagnetism and the polarization is such that $|S^z_{\mathrm{tot}}| < 1$, the $W_{\downarrow\downarrow}$ and $W_{\uparrow\uparrow}$ are instead identically zero. This holds true not only for $S_{\mathrm{tot}}{}^z = 0$ but also for $S_{\mathrm{tot}}{}^z = \pm 1/2$, where the single unpaired MO used in eq 22 acquires also a spin dependency, not present in the AGPs and AGPu cases.

Obviously the atoms, and also the O_2 molecule, do not have antiferromagnetism, but, on the other side, they have a net polarization. We can build the $W_{\uparrow\uparrow}$ block of the matrix using the two unpaired orbitals Θ_1 and Θ_2 for the definition of the parallel spin matrices of the AGPs or AGPu in the following way

$$f(\mathbf{r}_1 \uparrow, \mathbf{r}_2 \uparrow) = \bar{\lambda}(\Theta_1(\mathbf{r}_1)\Theta_2(\mathbf{r}_2) - \Theta_1(\mathbf{r}_2)\Theta_2(\mathbf{r}_1)) \tag{35}$$

where the presence of the minus sign guarantees the pairing function to be antisymmetric under particle exchange, while the $\bar{\lambda}$ is an arbitrary scaling factor that has no influence on the final value of the WF. Once we map the unpaired orbitals in the desired basis set, we obtain the variational parameters of the matrix λ for the $\uparrow\uparrow$ sector.

In the presence of opposite atomic magnetic moments, it is possible to rotate the spin component of the pairing function to initialize the $W_{\downarrow\downarrow}$ and $W_{\uparrow\uparrow}$ sectors. As we mentioned earlier, a further effect of this operation is to direct the atomic magnetic moments in a direction perpendicular to the spin-quantization axis. It is worth mentioning that, within our method, the spin orientation with respect to the molecular axis is irrelevant because in a nonrelativistic Hamiltonian, the spin−orbit coupling is not present. In this case, we have chosen to work with the atomic magnetic moments perpendicular to the z—axis, hence we applied a rotation of $\pi/2$ around the $\hat{y}$ direction. This operation maps

$$\begin{aligned} |\uparrow\rangle &\rightarrow \tfrac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) \text{ and} \\ |\downarrow\rangle &\rightarrow \tfrac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) \end{aligned} \tag{36}$$

If we apply this transformation to the pairing function from eq 14, we obtain

$$\begin{aligned} f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2) &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) f_+(\mathbf{r}_1, \mathbf{r}_2) \\ &+ |\uparrow\uparrow\rangle f_-(\mathbf{r}_1, \mathbf{r}_2) - |\downarrow\downarrow\rangle f_-(\mathbf{r}_1, \mathbf{r}_2) \end{aligned} \tag{37}$$

This transformation provides a meaningful initialization to our AGP WF that now has to be optimized to reach the best possible description of the GS. Indeed, within VMC, it is only thanks to the optimization that we can improve the description of the GS. So far, we have only converted the DFT WF from one *ansatz* to another, but the key for the success of this procedure is the optimization of all possible variational parameters. It is indeed crucial to optimize not only the ones corresponding to the matrix λ and the JF parameters but also the coefficients of the hybrid orbitals μ and the exponents of the Gaussian basis set Z_ν. This is realized computationally in a very efficient way using a coding technique called adjoint algorithmic differentiation[45] that allows calculations of total energy derivatives with respect to all variational parameters involved in a given algorithm that computes only the energy. This is remarkably done by paying a very small slowing down of a factor ≈2−3 with respect to the latter algorithm. We have also used a state of the art optimization scheme[46,47] for a correct search of the energy minimum. Remarkably, even when there is some possible dependency among the many variational parameters considered in our ansatz, the stochastic reconfiguration technique remains stable and efficient, thanks to an appropriate regularization of the stochastic matrix S.[10] Once we calculate the variational minimum, the best description of the GS is then obtained with the DMC calculation.

Even considering that the number of variational parameters involved in the calculation may be quite large, the optimization has a very small impact over the total computational cost, that is indeed mostly given by the DMC for all cases reported in this work. In Table 1, we compare the computational cost of the DMC calculations for the different WFs considered. We notice that the JAGP is even less expensive than the JSD and JsAGPs WFs. The JAGP and JAGPu are so efficient because, in this case, the variance of the energy is considerably smaller, as we can see from Table 1. This implies that JAGP and JAGPu require a smaller number of DMC iterations to reach the desired accuracy because they have lower variance compared with the JsAGPs and the JSD, thanks to the spin-dependent JF.

Table 1. DMC Computational Time Required to Obtain an Accuracy of 0.1 mH and Energy Variance on the Oxygen Dimer with an Intel Xeon Architecture Using a Recent LRDMC Algorithm[48] With a Lattice Spacing Equal to 0.05 bohr, the Smallest Used in This Work[a]

WF	CPU time	variance $[H^2]$
JSD	2806	2.909
JsAGPs	2526	2.819
JAGPu	14,523	2.455
JAGP	1857.79	2.125

[a]In these systems, the cost for doing about 10,000 iterations for the VMC optimization of our WFs is less than 30 h. The CPU time reported in the table corresponds to the total one (time spent by a single core times the number of cores) for obtaining the required accuracy, for example, with 256 cores parallel computation, the JAGP calculation can be obtained with about 7 h of walltime.

For large number N of electrons, the DMC calculation should scale as N^4 for fixed total energy accuracy, and the main

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

question, that we have not studied here, is whether the optimization remains computationally negligible because the number of variational parameter scales as N^2. In this respect, we have experienced that an optimization technique performed with a slow but very stable method, that is, with a large number of "cheap" optimization steps, each one determined by a relatively small number of samples (even much smaller than the number of parameters) is very promising for future large N applications.

Finally we introduced a technique to deal with particularly unstable AGP WFs. Indeed, it is possible, after a very large number of optimization steps (>10,000), that some eigenvalues of the matrix λ become too small as compared with the largest reference eigenvalue. This creates some instabilities in the inversion of the matrix W required for the QMC fast updates. For this reason, by an appropriate use of the PFAPACK library,[49] we identified a procedure to map the diagonalization of a full skew-symmetric matrix λ to the one corresponding to a real tridiagonal symmetric matrix. After this mapping, we can use the most powerful and stable LAPACK routines for diagonalization. Indeed, most linear algebra packages cannot deal with antisymmetric matrices, and a general diagonalization tool was not available for this case. The introduction of this procedure, described in detail in the Appendix, allows us to describe the matrix λ in terms of eigenvalues and orthogonal orbitals playing the role of eigenvectors of an antisymmetric matrix. We will refer to them in the Appendix as MOs because it may be considered their formal definition, within the formulation introduced in this work. With this meaningful decomposition, we can finally regularize the matrix λ by replacing the too small eigenvalues with reasonable lower bounds and continue, if necessary, with the optimization of the variational parameters.

2.6. S^2 Operator. The basic concept of QMC relies on the real space configurations sampling of a general electronic system. All observables can be indeed calculated in the basis where the electron positions and their spins are defined. In particular, for the systems considered, it is interesting to estimate the spin observables in order to understand their magnetic properties and the quality of the corresponding WFs. If during the simulation the value of S_z is fixed, when we break the symmetry the value of the S^2 is instead the result of the interplay between the JF and the AGP or the AGPu. The efficient computation of the expectation value of the S^2 operator has already been described in ref 50 for the JsAGPu and will be shown now for the JAGP.

Here, we show how to evaluate S^2 in a region of the space with a fast and computationally cheap approach based on the fast update algebra of the AGP and the spin-dependent JF. Let us consider the expectation value of the S^2 operator over a generic WF Ψ by direct application of its definition. Here, we use the completeness of the spatial configurations

$$\sum_{\mathbf{X}} |\mathbf{X}\rangle\langle\mathbf{X}| = \mathbf{I} \tag{38}$$

where the summation symbol implies a $3N$-multidimensional integral over the electron coordinates. Assuming a fixed polarization, we can write the explicit expression of the total spin square as

$$\begin{aligned}\langle S^2\rangle_\Psi &= \frac{\langle\psi|(S^z)^2|\psi\rangle}{\langle\psi|\psi\rangle} + \frac{1}{2}\sum_{\mathbf{X}}\sum_{i,j}^{N}\frac{\langle\psi|(S_i^+S_j^- + S_i^-S_j^+)|\mathbf{X}\rangle\langle\mathbf{X}|\psi\rangle}{\langle\psi|\psi\rangle}\\ &= \sum_{i=1}^{N_\uparrow}\sum_{j=N_\uparrow+1}^{N}\sum_{\mathbf{X}}\frac{\langle\psi|\mathbf{X}\rangle\langle\mathbf{X}|S_i^+S_j^-|\psi\rangle}{\langle\psi|\psi\rangle} + \frac{1}{4}(N_\uparrow - N_\downarrow)^2 + \frac{1}{2}(N_\uparrow + N_\downarrow)\\ &= \sum_{i=1}^{N_\uparrow}\sum_{j=N_\uparrow+1}^{N}\sum_{\mathbf{X}} p(\mathbf{X})\frac{\langle\mathbf{X}|S_i^+S_j^-|\psi\rangle}{\langle\mathbf{X}|\psi\rangle} + \frac{1}{4}(N_\uparrow - N_\downarrow)^2 + \frac{1}{2}N\end{aligned} \tag{39}$$

where the operators S_i in the above equation act on the spin component corresponding to the electron position $\mathbf{r}_i$ of the configuration $\mathbf{X}$. We can notice that

$$p(\mathbf{X}) = \frac{|\langle\psi|\mathbf{X}\rangle|^2}{\sum_{\mathbf{X}}|\langle\psi|\mathbf{X}\rangle|^2} \tag{40}$$

and that, by using QMC sampling, we generate configurations according to the probability density $p(\mathbf{X})$. Thus, we can evaluate the above multidimensional integral by directly sampling the estimator $S^2(\mathbf{X})$ that multiplies $p(\mathbf{X})$ in eq 39, as follows

$$S^2(\mathbf{X}) = \sum_{i=1}^{N_\uparrow}\sum_{j=N_\uparrow+1}^{N}\frac{\langle\mathbf{X}|S_i^+S_j^-|\psi\rangle}{\langle\mathbf{X}|\psi\rangle} + \frac{1}{4}(N_\uparrow - N_\downarrow)^2 + \frac{1}{2}N \tag{41}$$

The content of the former equation can be evaluated efficiently as we will explain in the following section. Indeed, the application of the operator $S_i^+S_j^-$ to the configuration $\mathbf{X}$ generates only a configuration $\mathbf{X}_{ij} = \{(\mathbf{r}_1\uparrow), ..., (\mathbf{r}_i\downarrow), ..., (\mathbf{r}_j\uparrow), ..., (\mathbf{r}_N\downarrow)\}$. Considering $\mathbf{X}$ our sampled configuration and using the previously given definition of $\mathbf{X}_{ij}$, we can recast eq 41 as

$$S^2(\mathbf{X}) = \frac{1}{4}(N_\uparrow - N_\downarrow)^2 + \frac{1}{2}N + \sum_{i=1}^{N_\uparrow}\sum_{j=N_\uparrow+1}^{N}\frac{\langle\mathbf{X}_{ij}|\psi\rangle}{\langle\mathbf{X}|\psi\rangle} \tag{42}$$

The only hard challenge of eq 42 is the calculation of the $N_\uparrow \times N_\downarrow$ ratios

$$r_{ij} = \frac{\langle\mathbf{X}_{ij}|\psi\rangle}{\langle\mathbf{X}|\psi\rangle} \tag{43}$$

for $i = 1, 2, ..., N_\uparrow$ and $j = N_\uparrow + 1, N_\uparrow + 2, ..., N$, that in our case read

$$r_{ij} = \frac{J(\mathbf{X}_{ij})\Psi_{\mathrm{AGP}}(\mathbf{X}_{ij})}{J(\mathbf{X})\Psi_{\mathrm{AGP}}(\mathbf{X})} = \mathbf{r}_{ij}^{\mathrm{AGP}}\mathbf{r}_{ij}^{\mathrm{JF}} \tag{44}$$

The configurations $\mathbf{X}$ and $\mathbf{X}_{ij}$ differ for a spin flip of the electrons i and j, but we can also consider $\mathbf{X}_{ij}$ as the configuration in which the electron i evolved to the position previously occupied by j and vice versa. We can then calculate the ratios in eq 44 using a fast algebra to update two positions for the AGP and for the JF with a direct evaluation based on

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

the Sherman−Morrison algebra and some simple manipulations, as discussed in detail later on.

It is also possible to calculate the value $S^2(\Lambda)$ of the S^2 operator in a sub-region of the space Λ. For this quantity, we can obtain the similar expression to eq 42

$$S_\Lambda{}^2(\mathbf{X}) = \frac{1}{4}(N_\uparrow^\Lambda - N_\downarrow^\Lambda)^2 + \frac{1}{2}N^\Lambda + \sum_{i=\{\Lambda,\uparrow\}}\sum_{j=\{\Lambda,\downarrow\}}\frac{\langle \mathbf{X}_{ij}|\psi\rangle}{\langle \mathbf{X}|\psi\rangle} \tag{45}$$

where N_σ^Λ ($\sigma = \uparrow, \downarrow$) is the number of σ-electrons in the region Λ, $N^\Lambda = N_\uparrow^\Lambda + N_\downarrow^\Lambda$. The summation symbol over $i \in \{\Lambda, \sigma\}$ indicates the sum for all σ-electron whose coordinate is in the region Λ. Therefore, we can use the same method as described below.

2.6.1. AGP Contribution. To calculate the AGP contribution to r_{ij}, we were able to find a slim and fast algebra making an extensive use of the Pfaffian properties.[51] It was fundamental to find an efficient algebra to calculate the whole matrix of the ratios r with a computational cost that is $O(N^3)$, by using mostly BLAS3 operations, thus avoiding that this computation could become the bottleneck of the whole procedure. In this way, we could ensure the evaluation cost of S^2 to be comparable with the one of a typical QMC cycle over all N electrons that is at most $O(N^3)$. Before describing the fast updating rules for the position of two electrons with a single move, we need to introduce some quantities fundamental for the calculation.

Let us denote W^{-1} as the inverse of W. This inverse W^{-1} can be computed from scratch for each configuration used to sample the spin square. The electron coordinates $\mathbf{r}_i$ are given for $i = 1, ..., N$, but since the corresponding spin can change with respect to the original choice ($\uparrow$ for $i \leq N_\uparrow$, and $\downarrow$ for $i > N_\uparrow$) because of the spin flips mentioned in the previous subsection, we will consider explicitly the values of the spin here.

We then define the matrix θ as

$$\theta_{ij} = f(\mathbf{r}_i\uparrow, \mathbf{r}_j\downarrow) + f(\mathbf{r}_i\downarrow, \mathbf{r}_j\uparrow) - f(\mathbf{r}_i\uparrow, \mathbf{r}_j\uparrow) - f(\mathbf{r}_i\downarrow, \mathbf{r}_j\downarrow) \tag{46}$$

For the spin $\uparrow$ electrons, we can define the vectors

$$v_k^\uparrow = \begin{pmatrix} f(\mathbf{r}_1\uparrow, \mathbf{r}_k\uparrow) - f(\mathbf{r}_1\uparrow, \mathbf{r}_k\downarrow) \\ f(\mathbf{r}_2\uparrow, \mathbf{r}_k\uparrow) - f(r_2\uparrow, \mathbf{r}_k\downarrow) \\ \vdots \\ f(\mathbf{r}_N\downarrow, \mathbf{r}_k\uparrow) - f(\mathbf{r}_N\downarrow, \mathbf{r}_k\downarrow) \end{pmatrix} \tag{47}$$

while for the spin $\downarrow$, we have instead

$$v_l^\downarrow = \begin{pmatrix} f(\mathbf{r}_1\uparrow, \mathbf{r}_l\downarrow) - f(\mathbf{r}_1\uparrow, \mathbf{r}_l\uparrow) \\ f(\mathbf{r}_2\uparrow, \mathbf{r}_l\downarrow) - f(\mathbf{r}_2\uparrow, \mathbf{r}_l\uparrow) \\ \vdots \\ f(\mathbf{r}_N\downarrow, \mathbf{r}_l\downarrow) - f(\mathbf{r}_N\downarrow, \mathbf{r}_l\uparrow) \end{pmatrix} \tag{48}$$

We can use these vectors to build the $N \times N$ matrix

$$V = (v_1^\uparrow v_2^\uparrow \cdots v_{N_\uparrow}^\uparrow v_{N_\uparrow+1}^\downarrow \cdots v_N^\downarrow) = (V^\uparrow V^\downarrow) \tag{49}$$

that allows us to define

$$U = (U^\uparrow U^\downarrow) = W^{-1}V = (W^{-1}V^\uparrow W^{-1}V^\downarrow) \tag{50}$$

and finally

$$D = (V^\uparrow)^T U^\downarrow \tag{51}$$

Now, we have all ingredients that we need for our fast updating algebra, and upon application of Sherman−Morrison algebra, we arrive at the ratio

$$r_{ij}^{\mathrm{AGP}} = \mathrm{Pf}[W(\mathbf{X}_{ij})]/\mathrm{Pf}[W(\mathbf{X})] = (1 + U_{ii})(1 + U_{jj}) - U_{ij}U_{ji} - (\theta_{ij} + D_{ij})W_{ij}^{-1} \tag{52}$$

We can notice that the preliminary calculation of the auxiliary matrices θ, V, U, and D, including the inversion of W, amounts to a total of $O(N^3)$ operations, while the calculation of the ratios is $O(N^2)$ once the matrices have been computed.

2.6.2. JF Contribution. In the JF that we introduced in the previous section, only the two-body term of eq 30 has a spin dependency, and thus, only this part contributes to the ratio. By simple substitution, it is easy to prove that

$$r_{ij}^{\mathrm{JF}} = \exp(D_i - D_j + u_{ee}(\mathbf{r}_i\uparrow, \mathbf{r}_j\downarrow) + u_{ee}(\mathbf{r}_i\downarrow, \mathbf{r}_j\uparrow) - u_{ee}(\mathbf{r}_i\uparrow, \mathbf{r}_j\uparrow) - u_{ee}(\mathbf{r}_i\downarrow, \mathbf{r}_j\downarrow)) \tag{53}$$

where we have defined

$$D_k = \sum_l u_{ee}(\mathbf{r}_l\sigma_l, \mathbf{r}_k\downarrow) - u_{ee}(\mathbf{r}_l\sigma_l, \mathbf{r}_k\uparrow) \tag{54}$$

The whole operation has a $O(N^2)$ computational cost and so does not limit the calculation in terms of performances.

3. RESULTS AND DISCUSSION

We apply this new approach for two types of systems: the first row high spin atoms (carbon, nitrogen, and oxygen) and their diatomic molecules and the benzene molecule. The first ones still represent useful benchmarks for the quantum chemistry approach, and a reasonable description of their properties and binding energies requires very expensive multireference methods. It is therefore very interesting to test our approach to find whether we are able to obtain a good description with a single Pfaffian *ansatz*. Benzene molecule on the other side is the most famous and important example of the RVB theory, so it represents a fundamental benchmark test for a method inspired by this theory. Here, we compare our results with exact available solutions, JSD WFs and with the JFVCAS multideterminant expansions for QMC. We will also show that our WF satisfies the size consistency both at VMC and DMC levels, a primary requirement if we want to use this approach for more challenging chemical studies.

3.1. Carbon. Carbon dimer is probably the most interesting example discussed in this study. A full understanding of the behavior of the carbon−carbon interaction is still missing, and the bond order of this molecule is still under debate.[52] The role of the spin fluctuations in this molecule has already been discussed,[53] but we believe that it is very instructive and represents the most important achievement of the JAGP WF. Indeed, it is only thanks to the spin fluctuations that we can have a correct description of its dimer bond.

The carbon atoms have spin triplet electronic configurations, and their mutual interaction leads to a singlet molecule. As we can see from Figure 1 and Table 3, the JAGP not only

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

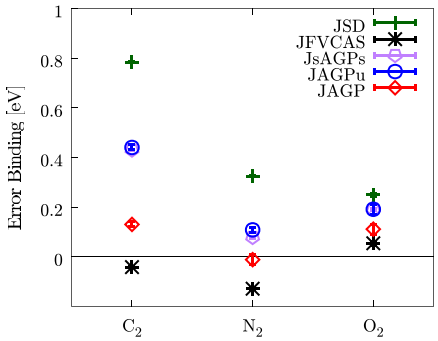


Figure 1. Comparison of the different DMC energies for different WFs. The results are shown for the three dimers described in this paper. The JFVCAS and the JSD results are taken from the literature.[17]

Table 2. Spin Measures with Different WFs for the Carbon Atom and Dimer at VMC Level

	S^2		$2\mu_B$
	atom	molecule	moment$\|z$
JsAGPs	2.00	0.00	0.0005(4)
JAGPu	2.00534(3)	0.1743(5)	0.5833(4)
JsAGP	2.00418(5)	0.2880(4)	0.7194(4)
JAGP	2.00542(1)	0.0327(1)	0.0013(5)
exact	2.00	0.00	

Table 3. Carbon Energies[a]

	carbon		
source	atom energy [H]	molecule energy [H]	binding energy [eV]
JSD	−37.81705(6)[b]	−75.8088(5)[b]	4.75(1)[b]
JFVCAS	−37.82607(5)[b]	−75.8862(2)[b]	6.369(6)[b]
JsAGPs	−37.8243(1)	−75.8611(2)	5.78(1)
JAGPu	−37.8263(1)	−75.8706(2)	5.93(1)
JAGP	−37.827965(3)	−75.88650(4)	6.274(3)
JSD (DMC)	−37.82966(4)[b]	−75.8672(1)[b]	5.656(3)[b]
JFVCAS (DMC)	−37.83620(1)[b]	−75.9106(1)[b]	6.482(3)[b]
JsAGPs (DMC)	−37.8364(1)	−75.8938(2)	6.01(1)
JAGPu (DMC)	−37.8364(1)	−75.8935(2)	6.00(1)
JAGP (DMC)	−37.8363(1)	−75.9045(2)	6.31(1)
estimated exact	−37.8450[c]	−75.9265[d]	6.44(2)[d,e]

[a]The JsAGPs, JAGPu, and JAGP results are calculated with an optimized ccpVTZ basis set. [b]Reference 17. [c]Reference 54. [d]Reference 55. [e]A more recent estimate yields 6.39 eV (Cyrus Umrigar, private communication).

improves the results of the JSD WF but remarkably also the description given by the JsAGPs and JAGPu. The huge difference between the multideterminant expansion JFVCAS and the JSD binding energies helps to quantify the effect of the multideterminantal nature of this molecule, and what makes this even more surprising is that the quality of the results obtained with a single JAGP WF, with a computational cost comparable to a SD, is already very close to the exact value.

As already mentioned before, the explanation for the impressive improvement of the binding energy from JsAGPs and JAGPu to JAGP resides on the description of the strong spin fluctuations in this molecule. The JAGP gives a very accurate picture of its magnetic properties as we can see from Table 2, giving results very close to $S^2 = 2$ for the atom and $S^2 = 0$ for the molecule. Conversely, by using the JsAGP (the AGP without spin-dependent JF) and the JAGPu, we cannot recover the singlet from the broken symmetry initialization. Interestingly, as expected, the molecule does not have any magnetic moment on the z-direction because it is an almost perfect singlet. The atomic spins, localized around each atom, point in opposite directions in order to form the singlet molecular state. Because there is no magnetic moment along z, we can measure its magnetic moment only by separately evaluating the S^2 in the two semi-infinite regions, each one containing a single atom, separated by a plane perpendicular to the molecular axis and at the same distance from the two atoms. In Figure 2, we show that, even at bond distance, there

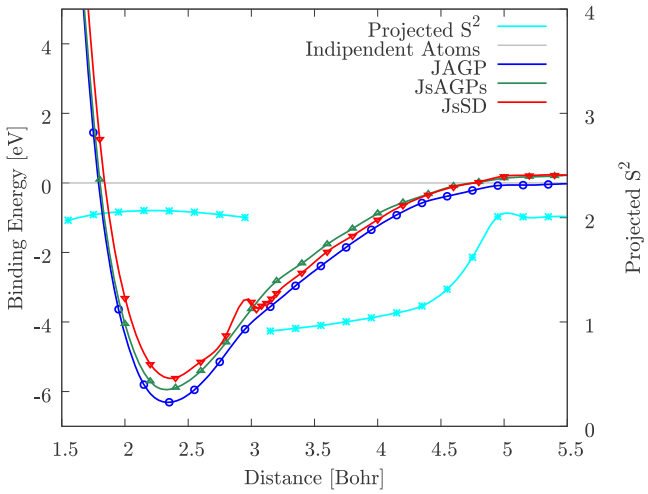


Figure 2. DMC energy dispersion of the carbon dimer: only the JAGP allows the system to be size consistent at large distance, which means that it is able to recover the energy and the expectation value of the S^2 operator of two isolated atoms at bond distance; however, the carbon atoms maintain a large value of S^2. The sharp change of the projected S^2 value at around 3 a.u. is probably due to an avoided crossing of two energy levels belonging to the same irreducible representation, in agreement with DMRG.[59] Within LSDA, this effect is reproduced by a discontinuous change in the occupation of the π orbitals in the corresponding SD. Lines are guides to the eye.

is a very strong magnetic moment around the atoms and, in this way, we can explain the strong effect of the ZPE of the spin fluctuations described by the JAGP.

Moreover, Figure 2 shows that only with the JAGP WF, we have a size consistent solution with the molecule that recovers the energy of two independent atoms at large distance. This feature is fundamental if we want to use this WF to describe chemical reactions and perform large-scale simulations, with a size consistent behavior at large distances. The importance of the variational optimization of the WF is particularly evident in this small molecule. With the standard approach, by applying DMC to a SD taken by DFT (here obtained with Purdue and Zunger LDA[56]), a level crossing in the occupation of the π MOs occurs at around 3 bohr distance, above which the π bonding orbitals are only partially occupied. This implies clear artifacts in the DMC energies. We have verified that this level crossing is reproduced with a standard DFT-LDA calculation by Gaussian 16 A.03 revision[57] and an almost converged basis set (the standard cc-pVQZ). The level crossing has also been observed in ref 58. In our variational optimization instead, we have verified that it is important to start at large distance with

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

the WF predicted by LSDA, otherwise a sizably higher energy is obtained. This effect is reflected also by the sharp change of the projected S^2 at around 3 bohr distance (see Figure 2), that could be compatible with an avoided crossing between two energy levels belonging to the same $^1\Sigma_g^+$ representation.[59]

3.2. Nitrogen. Nitrogen is in some sense similar to the carbon case: also its dimer is indeed a singlet formed by two large spin 3/2 atoms.

As we can notice from Figure 1 and Table 4, at the DMC level, the JsAGPu and JAGP are both exact within chemical accuracy. All our calculations compare with the exact result better than the JFVCAS solution. Surprisingly, at the VMC level, the binding energies calculated with JAGP, JsAGPs, and JAGPu are also very good.

Table 4. Nitrogen Energies[a]

	nitrogen		
source	atom energy [H]	molecule energy [H]	binding energy [eV]
JSD	−54.5628(1)[b]	−109.4520(5)[b]	8.88(1)[b]
JFVCAS		−109.4851(3)[b]	9.78(1)[b]
JsAGPs	−54.55794(6)	−109.4781(7)	9.856(3)
JAGPu	−54.55998(5)	−109.48155(7)	9.840(3)
JAGP	−54.56633(5)	−109.49226(7)	9.785(3)
JSD (DMC)	−54.57587(4)[b]	−109.5039(1)[b]	9.583(3)[b]
JFVCAS (DMC)		−109.5206(1)[b]	10.037(3)[b]
JsAGPs (DMC)	−54.5765(1)	−109.5164(2)	9.88(1)
JAGPu (DMC)	−54.5767(3)	−109.5140(2)	9.81(1)
JAGP (DMC)	−54.57709(9)	−109.5192(1)	9.933(6)
Fermi net	−54.58882(6)[c]	−109.5388(1)[c]	9.828(5)[c]
estimated exact	−54.5892[d]	−109.5427[e]	9.908(3)[e]

[a]The JsAGPs, JAGPu, and JAGP results are calculated with an optimized ccpVTZ basis set. [b]Reference 17. [c]Reference 5. [d]Reference 54. [e]Reference 55.

In this case, however, the difference between JAGP and JsAGPs/JAGPu is much smaller than in the previous case and should be related to a less important role of the spin fluctuations and also to a smaller magnetic moment of the atoms at equilibrium distance. By repeating the reasoning done for the carbon dimer, we can quantify the magnetic moment from the S^2 value in the semi-infinite region separated by a plane perpendicular to the axis of the molecule and equidistant from the atoms. As shown in Figure 3, at bond distance, the S^2 of the atom is much smaller than the one of an independent atom, and therefore, even if the nitrogen atom has a large spin, when it is forming a dimer it does not give rise to a strong antiferromagnetism.

Also in this case, it is important to notice that the JAGP solution is size consistent both in energy and spin. Despite the very good description at bond distance provided by the JsAGPs, we notice from Figure 3 that it is not perfectly size consistent. Within our approach, a fully consistent picture and a very accurate dispersion are possible only by means of the JAGP ansatz, that is able to work properly also in the strong correlation regime at large interatomic distance.

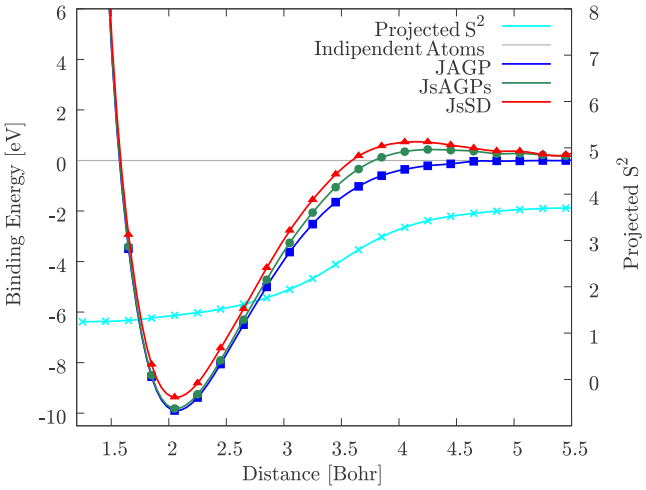


Figure 3. DMC energy dispersion of the nitrogen dimer: only the JAGP appears to be perfectly size consistent, thus recovering the energy and the expectation value of the S^2 operator of two isolated atoms at large interatomic distance. At bond distance, however, the nitrogen atoms have a smaller value of S^2, in contrast to what observed for the carbon dimer. Lines are guides to the eye.

3.3. Oxygen. The oxygen is very different from the previous cases but nevertheless very interesting for different reasons. The oxygen dimer consists of two triplet atoms, but this time the molecule is a triplet. There are small atomic magnetic moments in the GS of the oxygen molecule, but the role of the magnetic interaction remains important, as shown by the application of the JAGP ansatz. In this case, it looks that the interaction of parallel spin electrons is particularly important, and this can be described by the JAGP ansatz more accurately than the corresponding JsAGPs and JAGPu ones, as discussed in the previous sections. Thus, we expect to recover with the JAGP some correlation that we miss when we simplify the ansatz by using the unpaired orbitals in the JsAGPs and in the JAGPu WFs.

As shown in Figure 1 and Table 5, at the DMC level, the energies obtained with the JAGP WF are extremely good even for the oxygen dimer. In this case, the correct description of

Table 5. Oxygen Energies[a]

	oxygen		
source	atom energy [H]	molecule energy [H]	binding energy [eV]
JSD	−75.0352(1)[b]	−150.2248(5)[b]	4.20(1)[b]
JFVCAS		−150.2436(2)[b]	4.713(8)[b]
JsAGPs	−75.0268(3)	−150.2372(6)	5.00(3)
JAGPu	−75.0339(3)	−150.2503(5)	4.97(3)
JAGP	−75.0346(2)	−150.2572(4)	5.11(2)
JSD (DMC)	−75.05187(7)[b]	−150.2872(2)[b]	4.992(7)[b]
JFVCAS (DMC)		−150.29437(9)[b]	5.187(5)[b]
JsAGPs (DMC)	−75.0518(3)	−150.2894(3)	5.06(2)
JAGPu (DMC)	−75.0519(3)	−150.2902(4)	5.06(2)
JAGP (DMC)	−75.05289(7)	−150.2942(1)	5.127(5)
estimated exact	−75.0673[c]	−150.3724[d]	5.241[d]

[a]The JsAGPs, JAGPu, and JAGP results calculated with an optimized ccpVTZ basis set. [b]Reference 17. [c]Reference 54. [d]Reference 55.

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

the triplet pairing correlations, possible within the JAGP *ansatz*, appears to be fundamental. Indeed, the final result is so accurate that the binding energy is comparable to the one obtained with the multideterminant JFVCAS WF. It is even more surprising that the absolute energies of the atom and molecule are very close where not even better than the ones provided by the multideterminant expansion both at VMC and DMC levels. We have to point out, however, that within JFVCAS method, it is not possible to improve the JSD atom[17] and that the binding energy slightly better than the JAGP one derives from the poorer quality of the atom rather than a better description of the molecule.

The problem of the size consistency for the oxygen dimer is absolutely nontrivial and even more complicated than the previous cases. Starting from bond distance, we have a molecule of spin one and fixed projection $S_z = 1$, but we have to recover the behavior of two independent atoms. This means that, by keeping the projection $S_z = 1$ constant, while separating the atoms far apart, we have to recover the correct atoms of spin one and thus we need to have one atom with the spin oriented in a direction perpendicular to the z-axis. This is impossible for the JsAGPs and the JAGPu but allowed by the JAGP, a remarkable and absolutely nontrivial feature of this WF. As we can see from Figure 4, at large distance, only with

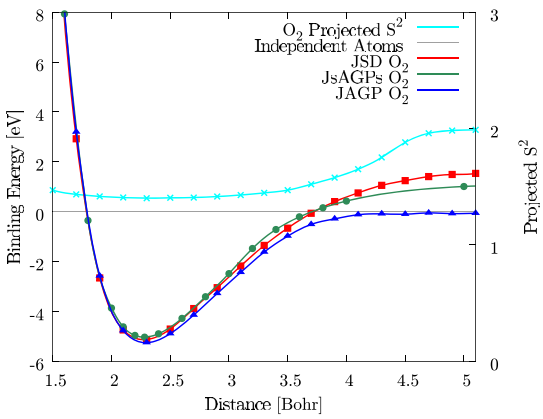


Figure 4. DMC energy dispersion of the oxygen dimer with the JAGP, JsAGP,[60] and JSD (with the SD obtained from DFT calculations): at large distance, only the JAGP WF is size consistent. In the plot, also the expectation value of the projected S^2 operator on the atoms for the JAGP that recovers the value of two isolated atoms at large distance. Lines are guides to the eye.

the JAGP the system recovers the energy and the spins of the independent atoms, showing that, by means of our advanced optimization tools, it is possible to dramatically change the WF up to the point of rotating completely the spin of an atom.

3.4. Benzene. The benzene molecule represents one of the most successful examples of the RVB theory with the carbon−carbon bonds resonating among several valence bond configurations, for example, Kekulé and Dewar. QMC methods are able to provide a very good description of this important molecule,[61,62] and thus, it is interesting to check whether, with our new approach, we can obtain a very accurate result. In particular, in Table 6, we compare the results obtained by JSD, JAGP, JAGPu, and JsAGPs WF, showing that all results obtained with a pairing function (from JsAGPs to JAGP) provide a very good estimate of the absolute energies, noticeably improving the results of the JSD. Moreover, the corresponding atomization energies are extremely accurate at

Table 6. Benzene Energies

	benzene		
	C atom[a]	molecule	atomization energy
source	energy[H]	energy[H]	energy[eV]
JSD	−37.8074(1)	−232.0261(3)	59.37(1)
JsAGPs	−37.82383(4)	−232.0805(3)	58.166(8)
JAGPu	−37.82651(5)	−232.0900(3)	57.986(8)
JAGP	−37.82921(4)	−232.1060(2)	57.982(7)
JSD(DMC)	−37.8299(1)	−232.1879(6)	60.09(2)
JsAGPs(DMC)	−37.8368(1)	−232.1947(6)	59.16(2)
JAGPu(DMC)	−37.8367(1)	−232.1943(6)	59.16(2)
JAGP(DMC)	−37.83751(9)	−232.1998(5)	59.18(2)
estimated exact	−37.8450[b]	−232.250(1)	59.32(2)[c]

[a]Calculated with the same basis set used for the benzene molecule. [b]Reference 54. [c]Reference 63.

the DMC level, whereas the JSD largely overestimates it. It is finally interesting to notice that, even if there is a sizeable gain in terms of absolute energy with our best ansatz, that is, the JAGP, it is not clear why this systematic improvement does not sizeably affect the atomization energy, likewise this could be almost converged to the exact value. This might be in principle explained because, at present, the accuracy of the state of the art "estimated exact" calculation is probably not enough to establish an energy difference ≪0.1 eV. For instance, the ZPE has been estimated by DFT[63] and some work is certainly necessary to clarify this issue, for example, by calculating the ZPE directly with QMC.

We remark here that the JsAGPs description of the benzene molecule is already very accurate and it is not improved by the JAGP. This is probably due to the lack of any sizeable spin moment around any atom composing this molecule. Indeed, the S^2 value calculated for the JAGP and JAGPu solutions are 0.032(1) and 0.0123(7), respectively, proving that any local magnetic moment is almost completely melted during the optimization, despite its nonzero initialization. We conclude therefore that in the benzene molecule, the spin fluctuations are not relevant and the use of the Pfaffian leads only to a marginal improvement of the total energy while the molecule is correctly described by a perfect singlet RVB ansatz given by the JsAGPs, in agreement with the classical RVB picture by L. Pauling.[64]

4. CONCLUSIONS

In this work we have proposed a new WF for QMC calculations given by the most general fermionic pairing function *ansatzs* in combination with a spin JF that provides a very rich description of the electronic correlation by means of a bosonic pairing function complementary to the fermionic one. With a computational cost comparable to a SD, we were able to improve not only the results achieved with a simple JSD but also with JsAGPs and JAGPu, reaching a level of accuracy comparable to the one obtained with the multideterminant JFVCAS WF. The powerful optimization techniques are probably the keys to explain the remarkable improvement we obtained with this WF, compared to previous attempts.[23,29] In particular, we have shown that the JAGP *ansatz* provides a very accurate description of high spin atoms and their dimers and that it is size consistent. This should increase the number of possible applications, providing a reasonably accurate and computationally feasible tool for studying chemical reactions.

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

Table 7. Carbon Energy Dispersion (Hartree)[a]

	numerical technique					
distance	JAGP (DMC)	DMRG	HCI	UCCSD-T_{frozen}	UCCSD-T_{full}	FCI
2.0787	−75.86652(3)	−75.76125[c]	−75.76701[d]	−75.76085	−75.78683	−75.7624[e]
2.2677	−75.90207(3)	−75.79924[c]	−75.80461[d]	−75.78450	−75.80878	−75.7987[e]
2.3480	−75.90456(3)	−75.80269[c]	−75.80786[b,d]	−75.78370	−75.80754	−75.8025[e]
2.4566	−75.90008(3)	−75.79937[c]	−75.80444[d]	−75.77928	−75.80247	−75.7993[e]
2.6456	−75.87825(4)	−75.77937[c]	−75.78460[d]	−75.76465	−75.78664	−75.7798[e]
3.0235	−75.81700(8)	−75.72405[c]	−75.72895[d]	−75.71762	−75.73765	−75.7243[e]
3.7794	−75.73649(8)	−75.64560[c]	−75.65043[b,d]	−75.62162	−75.63996	−75.6454[e]

[a]The JAGP results were obtained with the optimized ccpVDZ basis set (as explained in Section 2.1), the DMRG results with the ccpVQZ basis, the HCI with ccpV5Z basis set, the FCI with ccpVQZ basis set, whereas the UCCSD-T ones, both full- and frozen-core, are shown for ccpV5Z basis sets. [b]Interpolated. [c]Reference 59. [d]Reference 66. [e]Reference 67.

Table 8. Nitrogen Energy Dispersion (Hartree)[a]

	numerical technique				
distance	JAGP (DMC)	DMRG	MRCC	UCCSD-T (DZ)	UCCSD-T (5Z)
2.118	−109.51694(5)	−109.27833[c]	−109.27683[c]	−109.27652	−109.41303[b]
2.4	−109.46459(6)	−109.23838[c]	−109.23687[c]	−109.23202	−109.35926
2.7	−109.37935(6)	−109.16029[c]	−109.15895[c]	−109.14731	−109.26936
3.0	−109.29961(6)	−109.08619[c]	−109.08442[c]	−109.06570	−109.18331
3.6	−109.19745(6)	−108.99489[c]	−108.99272[c]	−108.97982	−109.08833
4.2	−109.16376(7)		−108.96471[c]	−108.96002	−109.06204

[a]The JAGP results were obtained with the optimized ccpVDZ basis set (as explained in Section 2.1), the DMRG and MRCC results with the ccpVDZ basis, whereas the corresponding UCCSD-T ones are shown also for a much larger basis (ccpV5Z), resulting in much better agreement with the present DMC results. [b]Interpolated. [c]Reference 68.

The triplet correlations have proven to be necessary to take into account correctly the ZPE of the spin fluctuations that we can now correctly describe thanks to a physical and accurate setup obtained by orienting the atomic magnetic moments of the AGP in the direction perpendicular to the spin-quantization axis chosen for the JF. For this reason, we have obtained a very good description of the carbon and nitrogen dimers, remarkably even when the first molecule was found to be very poorly described by the JsAGPs and the JAGPu. Moreover, it is only thanks to the presence of the triplet correlations that we were able to improve the description of the oxygen dimer as a strongly correlated triplet molecule with a highly entangled spin interaction among the atoms. Comparison with other methods different from QMC is shown in App. A. Our QMC variational energy is much better than state of the art quantum chemistry methods that seem to be affected by strong basis set errors even when considering only energy differences. For instance, the total energy difference ΔE at R = 4.2 a.u. and R = 2.11 a.u. in Table 8 should be close to the estimated exact binding energy (i.e. ≃9.91 eV from ref 55) at most weakly corrected by the residual dispersive interaction. Both DMRG and MRCI clearly miss more than 1 eV with the DZ basis, that is, $\Delta E \simeq 8.49$ eV. In order to show more clearly that the discrepancy between our DMC results and DMRG and MRCI is actually an artifact of the small basis, we have carried out UCCSD(T) calculation both for small (ΔE = 8.6 eV) and large (ΔE = 9.55 eV) basis set, and, as expected, our calculation (ΔE = 9.63) is much more in agreement with the most accurate large basis set calculation. In any event, our binding energy for N_2 (9.933 ± 0.006 eV) is surprisingly more accurate than the best state of the art calculation with CCSD(T) (9.73 eV from the Computational Chemistry Comparison and Benchmark DataBase[65]), implying that, most likely, our results should be considered to be the state art for the full dispersion curve of these small molecules.

Finally we demonstrated that for the benzene dimer, the JAGP is able to provide a very accurate atomization energy, though it is not clear in this case whether the triplet correlations are crucial for a highly accurate calculation. However, it is important to highlight that the accuracy in the binding energy is always much better than the accuracy in the total energy and that therefore there exists always a remarkable cancellation of errors in the total energy differences. This feature indeed is fundamental for a compact *ansatz* like the JAGP, and it challenges other very expensive highly correlated methods, even when these are able to achieve almost exact total energies as it was the case for the Fermi net approach to the nitrogen dimer.

The relatively low computational cost of QMC combined with powerful optimization techniques, allowing a reasonably large number of variational parameters, makes this approach ideal for studying systems even much larger than the ones considered in this work. Indeed, we believe that the paradigm presented in this paper could represent in the future a very powerful tool to investigate the electronic structure of interesting chemical compounds and physical systems where the spin interaction may play an important role, that in turn may be a number much larger than previously believed, as we have presented here the C_2 molecule as the very first and remarkable example of an antiferromagnetic chemical bond.

■ APPENDIX

Energy Dispersion Comparison

Comparison between the energy dispersion calculated with DMC JAGP WF and unrestricted single reference coupled cluster (UCCSD-T) with ccpVDZ and ccpV5Z basis sets. We further compared the carbon energy dispersion with DMRG,

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

heat-bath configuration interaction (HCI) and full configuration interaction (FCI) from the literature,[59,66,67] and the nitrogen dispersion with multireference coupled cluster (MRCC) and DMRG.[68] UCCSD(T) calculations were performed using Gaussian 16 A.03 revision with the counterpoise correction, with the frozen-core approximation and the full-core correlation.[57] Table 7 and Figure 5 show that there

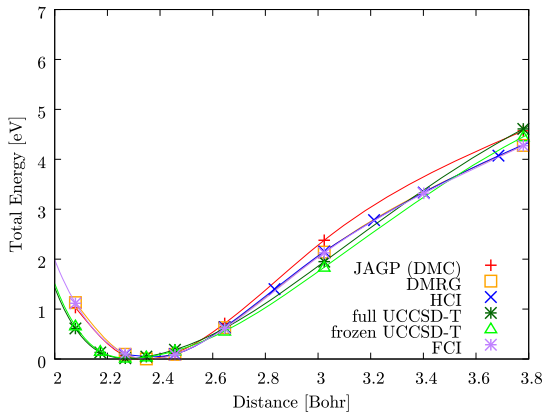


Figure 5. Energy dispersion of the carbon dimer calculated with JAGP (DMC), UCCSD-T (ccpV5Z), DMRG,[59] HCI,[66] and FCI.[67] Lines are guides to the eye.

are significant discrepancies between different methods in the carbon dimer dispersion curve at large distances. However, one has to consider that even in a quadruple zeta basis, the binding energy D_e = 6.22 eV[67] is about 6 mH lower than the estimated exact one and therefore if we reference all curves at the bond length minimum energy, as reported in the mentioned figure, a method that is supposed to be weakly dependent on the basis, as our DMC, should be slightly higher in energy at large distance, provided it remains close to the exact dispersion energy curve. Moreover, there may be sizeable corrections due to the frozen-core approximation employed by DMRG, HCI, and FCI. We have indeed verified that they are non-negligible in the UCCSD-T calculation, implying that core−valence interaction can lead to a further nonparallelity error of about 3 mH (see Figure 5). Core−valence interaction is considered in our DMC calculations simply because, within this technique, it is not possible to employ the frozen-core approximation. Nevertheless, it is clear that our results may have some error, but it is remarkable that if we use the corresponding energy values for computing the ZPE of the dimer, we find excellent agreement with the experimental value, given by 0.1146 eV.[69] Indeed, the ZPE calculated values, using a standard fit with a quartic polynomial close to the equilibrium distance, are 0.1153(6), 0.108, 0.106, 0.112, 0.114, and 1133(3) eV for our DMC, UCCSD-T full-core, UCCSD-T frozen-core, DMRG, HCI, and FCI, respectively. In summary, by taking into account all possible sources of error, we believe that our results are in reasonable agreement with the expected "exact result" converged in the complete basis set limit and with full core−valence interaction taken into account. Indeed, we believe that only a more direct comparison with experiments or a full-core FCI/DMRG or HCI extrapolated to the complete basis set limit can further improve the accuracy of the dispersion curve.

Diagonalization of a Skew Symmetric Generally Complex Matrix λ

Here, we discuss a general procedure to transform a generic complex antisymmetric matrix into a canonical Youla form that represents the equivalent of the standard diagonalization of Hermitian matrices. This is obtained by means of an appropriate unitary matrix U defined by an orthonormal set of states that we will address in the following MOs.

Given a $\bar{N} \times \bar{N}$ antisymmetric matrix λ, our goal is to identify a set of p paired states $\{(\phi_j^1, \phi_j^2)\}$ of orthonormal MOs, such that $p \leq \bar{N}$ is even and

$$\lambda\phi_j^1 = a_j\phi_j^2 \tag{55}$$

$$\lambda\phi_j^2 = -a_j\phi_j^1 \tag{56}$$

where the LHS of the above equations indicate standard matrix vector products, with shorthand notations adopted also in the remaining part of this Appendix. In this basis, we can write any skew-symmetric matrix λ in the canonical Youla form

$$\lambda_{\mathrm{MO}} = \begin{pmatrix} 0 & a_1 & 0 & 0 & \cdots & 0 \\ -a_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & a_2 & \cdots & 0 \\ 0 & 0 & -a_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \tag{57}$$

using only $p/2$ strictly positive parameters a_j. These ones play the same role of the eigenvalues for an ordinary Hermitian matrix and henceforth we will use this name for them, even if the matrix λ_{MO} is not diagonal but represents the simplest nonvanishing skew-symmetric matrix.

The transformation of the original matrix λ to the corresponding canonical Youla form by means of an appropriate unitary transformation $\lambda = U^*\lambda_{\mathrm{MO}}U^\dagger$ provides us also a very simple way to regularize the matrix λ, as discussed in the main text. In the case of odd $\bar{N}$, it will be shown later that there exists always an eigenvector of λ with a vanishing eigenvalue, but the decomposition remains possible as λ_{MO} will contain at least one vanishing row and corresponding column. Here, we define that an eigenvector is singular if it corresponds to a vanishing eigenvalue as in the odd $\bar{N}$ case.

It would be ideal for this calculation to use a very robust and stable diagonalization routine to maintain machine accuracy for the MOs. Unfortunately these routines are not commonly available for antisymmetric matrices and thus several mathematical transformations are necessary to map our task to a sequence of more commonly used or at least easily available algorithms.

A generic $\bar{N} \times \bar{N}$ antisymmetric matrix is written in the following way

$$\lambda = \begin{pmatrix} 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,\bar{N}} \\ -a_{1,2} & 0 & a_{2,3} & \cdots & a_{2,\bar{N}} \\ -a_{1,3} & -a_{2,3} & 0 & \cdots & a_{3,\bar{N}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{1,\bar{N}} & -a_{2,\bar{N}} & -a_{3,\bar{N}} & \cdots & 0 \end{pmatrix} \tag{58}$$

The first step is to transform λ in a tridiagonal antisymmetric real matrix. This operation is implemented in the subroutine

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

zsktrd (dsktrd) contained in the PFAPACK library.[49] The use of the Householder algorithm allows us to decompose the generic matrix λ as

$$\lambda = U_1^* \lambda_{\mathrm{Tr}} U_1^\dagger \tag{59}$$

where U_1 is the unitary transformation matrix output of the algorithm, while λ_{Tr} is a tridiagonal real antisymmetric matrix written in the standard tridiagonal form

$$\lambda_{\mathrm{Tr}} = \begin{pmatrix} 0 & b_1 & 0 & \cdots & 0 \\ -b_1 & 0 & b_2 & \cdots & 0 \\ 0 & -b_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \tag{60}$$

Thus, we can multiply the matrix λ_{Tr} for the imaginary unit i, yielding a more conventional tridiagonal Hermitian matrix λ_{iH}, defined by purely imaginary matrix elements.

We highlight that it is possible to map the matrix λ_{iH} into a real Hermitian matrix via a unitary transformation and use the appropriate LAPACK routine for its fast diagonalization. This procedure is well known and will be discussed later.

At this point, we can use the spectral theorem for Hermitian matrices to decompose the matrix $\lambda_{iH} = \psi \lambda_{\mathrm{diag}} \psi^\dagger$, where λ_{diag} is a diagonal matrix containing in its diagonal part the real eigenvalues a_i of λ_{iH}, and ψ is the unitary matrix, where each column is given by the eigenvector in the principle complex, corresponding to each eigenvalue, in the chosen order. This decomposition implies:

$$\lambda = -iU_1^* \psi \lambda_{\mathrm{diag}} \psi^\dagger U_1^\dagger \tag{61}$$

However, because the matrix ψ is generally complex and $\psi^\dagger \neq \psi^T$, some manipulation is necessary if we want to satisfy the skew-symmetry property of λ in an easy and transparent way.

If we consider one eigenvector $\overline{\psi}_j$ associated to an eigenvalue $a_j > 0$, we have

$$\lambda_{iH} \overline{\psi}_j = i\lambda_{\mathrm{Tr}} \overline{\psi}_j = a_j \overline{\psi}_j \tag{62}$$

the complex conjugate of this expression is

$$-i\lambda_{\mathrm{Tr}} \overline{\psi}_j^* = -a_j \overline{\psi}_j^* \tag{63}$$

where we have used that both λ_{Tr} and the eigenvalues a_j are real. This means that if $\overline{\psi}_j$ is an eigenvector of λ_{iH} relative to the eigenvalue a_j, then $\overline{\psi}_j^*$ is an eigenvector corresponding to the eigenvalue $-a_j$ and thus orthogonal to $\overline{\psi}_j$ because of the orthogonality between eigenvectors of an Hermitian matrix corresponding to different eigenvalues $\pm a_j$. We can thus easily verify, by using the relations given in eqs 62 and 63, the following simple equations

$$\lambda_{iH}(\overline{\psi}_j + \overline{\psi}_j^*) = a_j(\overline{\psi}_j - \overline{\psi}_j^*) \tag{64}$$

$$\lambda_{iH}(\overline{\psi}_j - \overline{\psi}_j^*) = a_j(\overline{\psi}_j + \overline{\psi}_j^*) \tag{65}$$

[a]In this way, we can define pairs of real orthogonal vectors $\overline{\psi}_j^1 = \sqrt{2}\,\Re(\overline{\psi}_j)$ and $\overline{\psi}_j^2 = \sqrt{2}\,\Im(\overline{\psi}_j)$ such that

$$\lambda_{iH} \overline{\psi}_j^1 = ia_j \overline{\psi}_j^2 \tag{66}$$

$$\lambda_{iH} \overline{\psi}_j^2 = -ia_j \overline{\psi}_j^1 \tag{67}$$

Once we have identified all pairs corresponding to all positive eigenvalues $a_j > 0$, we can write the unitary matrix $\overline{\psi}$, that is now real, by adding the remaining eigenvectors (that can be also chosen real, as shown in subsection Singular Eigenvectors) with vanishing eigenvalues in the remaining rightmost columns. In this way, we can finally define a unitary real matrix $\overline{\psi}$ yielding $\lambda_{\mathrm{Tr}} = -i\lambda_{iH} = \overline{\psi}\lambda_{\mathrm{MO}}\overline{\psi}^T$, where λ_{MO} is defined in eq 57 and therefore by using eq 59

$$\lambda = U_1^* \overline{\psi} \lambda_{\mathrm{MO}} \overline{\psi}^T U_1^\dagger \tag{68}$$

which represents the desired decomposition because the product of two unitary matrices $U^* = U_1^* \overline{\psi}$ remains a unitary matrix and its transponce $U^\dagger$ coincides with $\overline{\psi}^T U_1^\dagger$, yielding $\lambda = U^* \lambda_{\mathrm{MO}} U^\dagger$, that is the purpose of this Appendix.

Triangular Hermitian Matrices: A Mapping from Imaginary to Real

In order to use the LAPACK routines for diagonalization, we have to map the tridiagonal fully imaginary Hermitian matrix λ_{iH}, defined only (the diagonal elements are zero to fulfill Hermitianity) by its upper diagonal elements ib_j with b_j real for $j = 1, 2, \ldots, N - 1$, into a tridiagonal real symmetric matrix λ_{R}. We can implement this mapping by applying a unitary transformation to the matrix λ_{iH}. For this purpose, we introduce the following transformation described by the matrix U_2

$$\lambda_{\mathrm{R}} = U_2^\dagger \lambda_{iH} U_2 \tag{69}$$

The matrix U_2 is a complex diagonal matrix defined as

$$U_2 = \begin{pmatrix} e^{i\phi_1} & 0 & \cdots & 0 \\ 0 & e^{i\phi_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{i\phi_{N-1}} \end{pmatrix} \tag{70}$$

The explicit calculation of the right-hand side of the eq 69 gives

$$\lambda_{\mathrm{R}} = \begin{pmatrix} 0 & ib_1 e^{i(\phi_2 - \phi_1)} & 0 & \cdots & 0 \\ -ib_1 e^{i(\phi_1 - \phi_2)} & 0 & ib_2 e^{i(\phi_3 - \phi_2)} & \cdots & 0 \\ 0 & -ib_2 e^{i(\phi_2 - \phi_3)} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix} \tag{71}$$

By setting the imaginary units $\pm i = \exp(\pm i\pi/2)$ (when not exponentiated in the previous equation), we can easily impose that all phase factors cancel in all the corresponding matrix elements of λ_{R} with the choice

$$\phi_j = -\frac{\pi}{2}(j - 1) \tag{72}$$

that therefore implies that λ_{R}, with the above definition, is a real symmetric matrix. At this point, we can diagonalize the matrix λ_{R} by means of a real unitary matrix U_{R}, that is the output of a standard LAPACK diagonalization routine of tridiagonal real matrices (e.g., dstevx for double precision arithmetic). In this way, λ_{iH} can be diagonalized as $\lambda_{iH} = U_2 U_{\mathrm{R}} \lambda_{\mathrm{diag}} U_{\mathrm{R}}^T U_2^\dagger$, where λ_{diag} is a diagonal matrix containing the corresponding eigenvalues of the LAPACK diagonalization.

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

Singular Eigenvectors

Within this formulation, it is also particularly easy to compute all the real singular eigenvectors of λ_{iH} corresponding to the possible vanishing eigenvalues. They were used in this Appendix to complete the columns of the unitary real matrix $\overline{\psi}$. From the outcome of the previous subsection, any eigenvector ϕ_k^j of λ_{iH} can be obtained by applying the diagonal matrix U_2 to a real eigenvector $\overline{\phi}_k^j$ of λ_R, namely, $\phi_k^j = \overline{\phi}_k^j \exp(-i\pi/2(k-1))$, implying that even k-components are purely imaginary and odd k-components are purely real. Then, it is obvious to realize that $\overline{\phi}_k^j$ corresponds to a singular eigenvector of λ_R, and also $\Re(\phi_k^j)$ and $\Im(\phi_k^j)$ [and therefore also $\Re(\phi_k^j) + \Im(\phi_k^j)$] correspond to singular eigenvectors or at most null vectors (not both) of λ_{iH} because this matrix is purely imaginary, and the complex conjugation of a singular eigenvector is again a singular eigenvector by eqs 62 and 63 with $a_j = 0$.

Then, it follows that all orthogonal eigenvectors $\overline{\phi}^j$ (output of dstevx) corresponding to the zero eigenvalues of the matrix λ_R can be used to define the real singular eigenvectors corresponding to the matrix λ_{iH}, that is

$$\tilde{\phi}_k^j = \Re(\phi_k^j) + \Im(\phi_k^j) = \begin{cases} (-1)^{(k-1)/2}\overline{\phi}_k^j & k \text{ odd} \\ (-1)^{k/2}\overline{\phi}_k^j & k \text{ even} \end{cases} \tag{73}$$

that are explicitly real and orthogonal to each other because $\sum_k \tilde{\phi}_k^j \tilde{\phi}_k^l = \sum_k \overline{\phi}_k^j \overline{\phi}_k^l = \delta_{l,j}$. They are also orthogonal to all the other pairs of non-singular eigenvectors because of the orthogonality property of eigenvectors of an Hermitian matrix λ_{iH} that we have already used in the previous section.

AUTHOR INFORMATION

Corresponding Authors

Claudio Genovese − *SISSA, International School for Advanced Studies, 34136 Trieste, Italy*; orcid.org/0000-0002-8296-8132; Email: claudio.genovese@sissa.it

Sandro Sorella − *SISSA, International School for Advanced Studies, 34136 Trieste, Italy*; Email: sorella@sissa.it

Authors

Tomonori Shirakawa − *Computational Materials Science Research Team, RIKEN Center for Computational Science (R-CCS), Kobe, Hyogo 650-0047, Japan*

Kousuke Nakano − *SISSA, International School for Advanced Studies, 34136 Trieste, Italy; School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa 923-1292, Japan*; orcid.org/0000-0001-7756-4355

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00165

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge PRIN-2017 for financial support and CINECA PRACE-2019 for computational resources. Parts of numerical simulations have been performed on the HOKUSAI supercomputer at RIKEN (project ID: 19011 and 19030) and K-computer at RIKEN (project ID: hp190107). K.N. is grateful for the facilities of Research Center for Advanced Computing Infrastructure at Japan Advanced Institute of Science and Technology (JAIST). K.N. also acknowledges financial support from the Simons Foundation and that from Grant-in-Aid for Scientific Research on Innovative Areas (no. 16H06439).

ADDITIONAL NOTE

[a]The same argument holds true if the eigenvalue aj corresponds to $p > 1$ degenerate eigenvectors. The mentioned orthogonality property of Hermitian matrix eigenvectors leads to the straightforward definition of p pairs of mutually orthonormal real ones used for the present decomposition with a block diagonal matrix, where each 2×2 block corresponds to one of the p degenerate eigenvectors.

REFERENCES

(1) Dirac, P. A. M. Quantum Mechanics of Many-Electron Systems. *Proc. R. Soc. London, Ser. A* **1929**, *123*, 714−733.

(2) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256−4266.

(3) Orús, R. Tensor networks for complex quantum systems. *Nat. Rev. Phys.* **2019**, *1*, 538−550.

(4) White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **1992**, *69*, 2863−2866.

(5) Pfau, D.; Spencer, J. S.; de, G.; Matthews, A. G.; Foulkes, W. M. C. Ab-Initio Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks. **2019**, arXiv:1909.02487v2.

(6) Booth, G. H.; Grüneis, A.; Kresse, G.; Alavi, A. Towards an exact description of electronic wavefunctions in real solids. *Nature* **2013**, *493*, 365−370.

(7) Motta, M.; Ceperley, D. M.; Chan, G. K.-L.; Gomez, J. A.; Gull, E.; Guo, S.; Jiménez-Hoyos, C. A.; Lan, T. N.; Li, J.; Ma, F.; Millis, A. J.; Prokof'ev, N. V.; Ray, U.; Scuseria, G. E.; Sorella, S.; Stoudenmire, E. M.; Sun, Q.; Tupitsyn, I. S.; White, S. R.; Zgid, D.; Zhang, S. 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.

(8) Motta, M.; Genovese, C.; Ma, F.; Cui, Z.-H.; Sawaya, R.; Chan, G. K.-L.; Chepiga, N.; Helms, P.; Jimenez-Hoyos, C.; Millis, A. J.; Ray, U.; Ronca, E.; Shi, H.; Sorella, S.; Stoudenmire, E. M.; White, S. R.; Zhang, S. Ground-state properties of the hydrogen chain: insulator-to-metal transition, dimerization, and magnetic phases. **2019**, arXiv:1911.01618v2.

(9) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **2001**, *73*, 33−83.

(10) Becca, F.; Sorella, S. *Quantum Monte Carlo Approaches for Correlated Systems*; Cambridge University Press, 2017.

(11) Hammond, B. L.; Lester, W. A.; Reynolds, P. J. *Monte Carlo Methods in Ab Initio Quantum Chemistry*; WORLD SCIENTIFIC, 1994.

(12) Reynolds, P. J.; Tobochnik, J.; Gould, H. Diffusion Quantum Monte Carlo. *Comput. Phys.* **1990**, *4*, 662−668.

(13) Anderson, J. B. A random-walk simulation of the Schrödinger equation: H+3. *J. Chem. Phys.* **1975**, *63*, 1499−1503.

(14) Genovese, C.; Meninno, A.; Sorella, S. Assessing the accuracy of the Jastrow antisymmetrized geminal power in the H4 model system. *J. Chem. Phys.* **2019**, *150*, 084102.

(15) Gasperich, K.; Deible, M.; Jordan, K. D. H4: A model system for assessing the performance of diffusion Monte Carlo calculations using a single Slater determinant trial function. *J. Chem. Phys.* **2017**, *147*, 074106.

(16) Braïda, B.; Toulouse, J.; Caffarel, M.; Umrigar, C. J. Quantum Monte Carlo with Jastrow-valence-bond wave functions. *J. Chem. Phys.* **2011**, *134*, 084108.
(17) Toulouse, J.; Umrigar, C. J. Full optimization of Jastrow-Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules. *J. Chem. Phys.* **2008**, *128*, 174101.
(18) Knowles, P. J.; Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **1984**, *111*, 315−321.
(19) Olsen, J.; Roos, B. O.; Jo/rgensen, P.; Jensen, H. J. r. A. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.* **1988**, *89*, 2185−2192.
(20) Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157−173.
(21) Filippi, C.; Umrigar, C. J. Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules. *J. Chem. Phys.* **1996**, *105*, 213−226.
(22) Linderberg, J. The Antisymmetrized Geminal Power, a Simple Correlated Wave Function for Chemical Bonding. *Isr. J. Chem.* **1980**, *19*, 93−98.
(23) Bajdich, M.; Mitas, L.; Drobný, G.; Wagner, L. K.; Schmidt, K. E. Pfaffian Pairing Wave Functions in Electronic-Structure Quantum Monte Carlo Simulations. *Phys. Rev. Lett.* **2006**, *96*, 130201.
(24) Morales, M. A.; McMinis, J.; Clark, B. K.; Kim, J.; Scuseria, G. E. Multideterminant Wave Functions in Quantum Monte Carlo. *J. Chem. Theory Comput.* **2012**, *8*, 2181−2188.
(25) Anderson, J. B. A random-walk simulation of the Schrödinger equation: H+3. *J. Chem. Phys.* **1975**, *63*, 1499.
(26) Casula, M.; Sorella, S. Geminal wave functions with Jastrow correlation: A first application to atoms. *J. Chem. Phys.* **2003**, *119*, 6500−6511.
(27) Casula, M.; Attaccalite, C.; Sorella, S. Correlated geminal wave function for molecules: An efficient resonating valence bond approach. *J. Chem. Phys.* **2004**, *121*, 7110−7126.
(28) Neuscamman, E. Improved Optimization for the Cluster Jastrow Antisymmetric Geminal Power and Tests on Triple-Bond Dissociations. *J. Chem. Theory Comput.* **2016**, *12*, 3149−3159.
(29) Bajdich, M.; Mitas, L.; Wagner, L. K.; Schmidt, K. E. Pfaffian pairing and backflow wavefunctions for electronic structure quantum Monte Carlo methods. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2008**, *77*, 115112.
(30) Neuscamman, E. Size Consistency Error in the Antisymmetric Geminal Power Wave Function can be Completely Removed. *Phys. Rev. Lett.* **2012**, *109*, 203001.
(31) Neuscamman, E. The Jastrow antisymmetric geminal power in Hilbert space: Theory, benchmarking, and application to a novel transition state. *J. Chem. Phys.* **2013**, *139*, 194105.
(32) Mahajan, A.; Sharma, S. Symmetry-Projected Jastrow Mean-Field Wave Function in Variational Monte Carlo. *J. Phys. Chem. A* **2019**, *123*, 3911−3921.
(33) Henderson, T. M.; Scuseria, G. E. Spin-projected generalized Hartree-Fock method as a polynomial of particle-hole excitations. *Phys. Rev. A* **2017**, *96*, 022506.
(34) Franjic, F.; Sorella, S. Spin-Wave Wave Function for Quantum Spin Models. *Prog. Theor. Phys.* **1997**, *97*, 399−406.
(35) Sorella, S.; Seki, K.; Brovko, O. O.; Shirakawa, T.; Miyakoshi, S.; Yunoki, S.; Tosatti, E. Correlation-Driven Dimerization and Topological Gap Opening in Isotropically Strained Graphene. *Phys. Rev. Lett.* **2018**, *121*, 066402.
(36) Nakano, K.; Maezono, R.; Sorella, S. All-Electron Quantum Monte Carlo with Jastrow Single Determinant Ansatz: Application to the Sodium Dimer. *J. Chem. Theory Comput.* **2019**, *15*, 4044−4055.
(37) Watts, J. D.; Bartlett, R. J. Coupled-cluster calculations on the C2 molecule and the C+2 and C−2 molecular ions. *J. Chem. Phys.* **1992**, *96*, 6073−6084.
(38) Karton, A.; Tarnopolsky, A.; Martin, J. M. L. Atomization Energies of the Carbon Clusters Cn (n = 2-10) Revisited by Means of W4 Theory as well as Density Functional, Gn, and CBS Methods. *Mol. Phys.* **2009**, *107*, 977−990.
(39) Zou, W.; Cremer, D. C2 in a Box: Determining Its Intrinsic Bond Strength for the X1 Σg + Ground State. *Chem.—Eur. J.* **2016**, *22*, 4087−4099.
(40) Nemec, N.; Towler, M. D.; Needs, R. J. Benchmark all-electron ab initio quantum Monte Carlo calculations for small molecules. *J. Chem. Phys.* **2010**, *132*, 034111.
(41) Nakano, K.; Attaccalite, C.; Barborini, M.; Capriotti, L.; Casula, M.; Coccia, E.; Dagrada, M.; Genovese, C.; Luo, Y.; Mazzola, G.; Zen, A.; Sorella, S. TurboRVB: A many-body toolkit for ab initio electronic simulations by quantum Monte Carlo. *J. Chem. Phys.* **2020**, *152*, 204121.
(42) *TurboRVB*. https://people.sissa.it/sorella/web/index.html, 2020.
(43) Marchi, M.; Azadi, S.; Casula, M.; Sorella, S. Resonating valence bond wave function with molecular orbitals: Application to first-row molecules. *J. Chem. Phys.* **2009**, *131*, 154116.
(44) Kasteleyn, P. W. Dimer Statistics and Phase Transitions. *J. Math. Phys.* **1963**, *4*, 287−293.
(45) Sorella, S.; Capriotti, L. Algorithmic differentiation and the calculation of forces by quantum Monte Carlo. *J. Chem. Phys.* **2010**, *133*, 234111.
(46) Sorella, S.; Casula, M.; Rocca, D. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.* **2007**, *127*, 014105.
(47) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions. *Phys. Rev. Lett.* **2007**, *98*, 110201.
(48) Nakano, K.; Maezono, R.; Sorella, S. Speeding up ab initio diffusion Monte Carlo simulations by a smart lattice regularization. *Phys. Rev. B* **2020**, *101*, 155106.
(49) Wimmer, M. Algorithm 923: Efficient Numerical Computation of the Pfaffian for Dense and Banded Skew-Symmetric Matrices. *ACM Trans. Math Software* **2012**, *38*, 1−17.
(50) Zen, A.; Coccia, E.; Luo, Y.; Sorella, S.; Guidoni, L. Static and Dynamical Correlation in Diradical Molecules by Quantum Monte Carlo Using the Jastrow Antisymmetrized Geminal Power Ansatz. *J. Chem. Theory Comput.* **2014**, *10*, 1048−1061.
(51) Caracciolo, S.; Sokal, A. D.; Sportiello, A. Algebraic/combinatorial proofs of Cayley-type identities for derivatives of determinants and pfaffians. *Adv. Appl. Math.* **2013**, *50*, 474−594.
(52) Shaik, S.; Danovich, D.; Wu, W.; Su, P.; Rzepa, H. S.; Hiberty, P. C. Quadruple bonding in C2 and analogous eight-valence electron species. *Nat. Chem.* **2012**, *4*, 195−200.
(53) Genovese, C.; Shirakawa, T.; Sorella, S. The nature of the quadruple chemical bond in the dicarbon molecule. **2019**, arXiv:1911.09748v3.
(54) Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; p Fischer, C. F. Ground-state correlation energies for atomic ions with 3 to 18 electrons. *Phys. Rev. A* **1993**, *47*, 3649−3670.
(55) Bytautas, L.; Ruedenberg, K. Correlation energy extrapolation by intrinsic scaling. IV. Accurate binding energies of the homonuclear diatomic molecules carbon, nitrogen, oxygen, and fluorine. *J. Chem. Phys.* **2005**, *122*, 154110.
(56) Perdew, J. P.; Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1981**, *23*, 5048.
(57) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Petersson, G.; Nakatsuji, H.; et al. *Gaussian 16*, 2016.
(58) Sharma, S.; Alavi, A. Multireference linearized coupled cluster theory for strongly correlated systems using matrix product states. *J. Chem. Phys.* **2015**, *143*, 102815.
(59) Sharma, S. A general non-Abelian density matrix renormalization group algorithm with application to the C2 dimer. *J. Chem. Phys.* **2015**, *142*, 024107.
(60) Zen, A.; Trout, B. L.; Guidoni, L. Properties of reactive oxygen species by quantum Monte Carlo. *J. Chem. Phys.* **2014**, *141*, 014305.

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131

(61) Azadi, S.; Cohen, R. E. Chemical accuracy from quantum Monte Carlo for the benzene dimer. *J. Chem. Phys.* **2015**, *143*, 104301.
(62) Sorella, S.; Casula, M.; Rocca, D. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.* **2007**, *127*, 014105.
(63) Franck, E. U. J. D. Cox, D. D. Wagman, V. A. Medvedev: CODATA - Key Values for Thermodynamics, aus der Reihe: CODATA, Series on Thermodynamic Properties. Hemisphere Publishing Corporation, New York, Washington, Philadelphia, London 1989. 271 Seiten, Preis: £ 28.00. *Ber. Bunsenges. Phys. Chem.* **1990**, *94*, 93.
(64) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Itaca, New York, 1960.
(65) Johnson, R. D., III *Computational Chemistry Comparison and Benchmark Database*, Release 20; National Institute of Standards and Technology (NIST), http://cccbdb.nist.gov (Online, accessed June 17, 2020).
(66) Holmes, A. A.; Umrigar, C. J.; Sharma, S. Excited states using semistochastic heat-bath configuration interaction. *J. Chem. Phys.* **2017**, *147*, 164111.
(67) Booth, G. H.; Cleland, D.; Thom, A. J. W.; Alavi, A. Breaking the carbon dimer: The challenges of multiple bond dissociation with full configuration interaction quantum Monte Carlo methods. *J. Chem. Phys.* **2011**, *135*, 084104.
(68) Chan, G. K.-L.; Kállay, M.; Gauss, J. State-of-the-art density matrix renormalization group and coupled cluster theory studies of the nitrogen binding curve. *J. Chem. Phys.* **2004**, *121*, 6110−6116.
(69) Irikura, K. K. Experimental Vibrational Zero-Point Energies: Diatomic Molecules. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389−397.

https://dx.doi.org/10.1021/acs.jctc.0c00165
J. Chem. Theory Comput. 2020, 16, 6114−6131