Generalized valence bond wave functions in quantum Monte Carlo

Amos G. Anderson[a] and William A. Goddard III[b]
Materials and Process Simulation Center, Division of Chemistry and Chemical Engineering, California Institute of Technology, (MC 139-74) Pasadena, California 91125, USA

(Received 31 August 2009; accepted 10 March 2010; published online 26 April 2010)

We present a technique for using quantum Monte Carlo (QMC) to obtain high quality energy differences. We use generalized valence bond (GVB) wave functions, for an intuitive approach to capturing the important sources of static correlation, without needing to optimize the orbitals with QMC. Using our modifications to Walker branching and Jastrows, we can then reliably use diffusion quantum Monte Carlo to add in all the dynamic correlation. This simple approach is easily accurate to within a few tenths of a kcal/mol for a variety of problems, which we demonstrate for the adiabatic singlet-triplet splitting in methylene, the vertical and adiabatic singlet-triplet splitting in ethylene, 2+2 cycloaddition, and Be_2 bond breaking. © *2010 American Institute of Physics*. [doi:10.1063/1.3377091]

I. INTRODUCTION

The quantum Monte Carlo (QMC) algorithm[1,2] is rapidly advancing as a tool competitive with the highest accuracy *ab initio* electronic structure methods. It has already been used with remarkable success to calculate energies and other properties for a wide variety of molecules and periodic systems across the periodic table. Although it will probably never replace cheaper methods such as density functional theory (DFT), given advances in computing power, it will surely begin to serve as a complementary method, brought in for calibration or to resolve disagreements.

QMC methods can capture the correlation energy in two ways. First, when using variational QMC (VMC) sampling, we can introduce arbitrary functions of interparticle coordinates, called Jastrow functions, to explicitly model interparticle interactions. Second, and even better, provided a best guess for the wave function nodes, we can include all the dynamic correlation energy through the diffusion QMC (DMC) algorithm. The nodal assumption results in an error called the fixed-node approximation, which is not negligible. Fortunately, some of the same techniques used to capture the static correlation energy for self-consistent field (SCF) wave functions can be used to capture the fixed-node energy, and thus multiconfiguration SCF (MCSCF) techniques can be considered to be quite complementary to DMC.

Fully accurate SCF techniques can be expensive, typically scaling quite poorly with molecule size, motivating a search for methods which do not require more work than necessary. We often need at least chemical accuracy of 1 kcal/mol, which is too demanding for many methods. Explored in this paper is an approach where we use generalized valence bond (GVB) wave functions[3–5] to correct for the fixed-node error in DMC, without using QMC to optimize the orbitals. By working with valence bond orbitals, GVB has the advantage over more general approaches of being chemically intuitive and of scaling well with molecule size, while efficiently correcting for the important sources of static correlation.

To demonstrate the usefulness of the GVB approach, as well as to validate our overall methodology, we present a study of a few molecules for which experimental data or reliable calculations are available, testing excitation energies and bond breaking. The methylene singlet-triplet adiabatic splitting is one of the few processes for which experimental data are accurate to within tenths of a kcal/mol. Thus it is a good test to see exactly how close to the exact answer we can get. Among the most studied processes is surely the ethylene singlet-triplet splitting (both adiabatic and vertical), with quite a few experimental and computational studies. Both of these processes have even been the subject of several other QMC studies, providing an excellent basis on which to compare our results with those of more standardized approaches. We look at the cycloaddition of two ethylene molecules to make a cyclobutane molecule and show how our approach is successful at modeling multiple bond changes at once. Finally, we study Be_2, a classic challenge for quantum chemistry techniques.

II. METHOD

In this section, we present the QMC approach we use in our QMcBEAVER (Ref. 6) code, which is available online for free. First, we discuss GVB as the SCF part of the wave function, and, second, our modifications to the Jastrow functions recommended by Drummond *et al.*[7] Third, we talk about our strategy in optimizing this kind of wave function, starting from the approach of Toulouse and Umrigar.[8] Fourth, we diverge from Umrigar's recommended DMC algorithm[1] to use the reconfiguration method for Walker branching provided by Assaraf *et al.*,[9] with more of our customizations. Finally, we summarize our approach.

[a]Electronic mail: amosa@alumni.caltech.edu.
[b]Electronic mail: wag@wag.caltech.edu.

A. Generalized valence bond wave functions

A GVB wave function[3–5] starts with a localized restricted Hartree–Fock (RHF) wave function and replaces an orbital (e.g., a single bond) with two singlet paired orbitals in a geminal called a perfect pair,

$$\Psi_{\mathrm{GVB}} = \mathcal{A}[\{\mathrm{core}\}\{\varphi_u\varphi_v\}\{\alpha\beta - \beta\alpha\}], \quad (1)$$

where $\mathcal{A}$ is the antisymmeterizer, or determinant, operator. Although we allow φ_u and φ_v to overlap each other, they are orthogonal to all the other orbitals in the wave function. This can be thought of as permitting each electron to have its own orbital. We can rotate these intuitive orbitals into the more computationally useful, but fully equivalent, natural orbital form,

$$\Psi_{\mathrm{GVB}} = \mathcal{A}[\{\mathrm{core}\}\{\sigma_u\phi_u^2 - \sigma_v\phi_v^2\}\{\alpha\beta\}], \quad (2)$$

where $\sigma_u^2+\sigma_v^2=1$. We typically interpret ϕ_u as a "bonding" orbital and ϕ_v as an "antibonding" orbital. Where a perfect pair is used to represent a single bond, the benefit is to add left-right correlation to the bond, allowing the electrons to get away from each other a little bit, and this is the simplest wave function that permits H_2 to dissociate to 2H. In the same way, we can add left-right correlation to double or triple bonds. When it comes to lone pairs, the perfect pairing scheme can be used to add in an important orbital left out by RHF (such as 1 b_1 in $^1A_1$ methylene) to incorporate some angular correlation, or, in other cases, to add in-out correlation to a lone pair.

Although GVB is a subset of MCSCF calculations, which have been well studied in the QMC context,[10,11] the main advantage to GVB over MCSCF is that it is the only variety that is able to avoid integral transformations.[12] Additionally, it allows a simple, modular, and balanced way of selecting the active space, since everything is localized. The researcher perhaps does not even need to look at any orbitals to do this, since reliable routines exist to generate good initial guesses[13] for a GVB wave function based on RHF orbitals.

For our QMC wave functions, we expand the geminals in each N_{GVB} pair wave function into the equivalent $2^{N_{\mathrm{GVB}}}$ determinant wave function. Although the number of determinants grows quickly, we use a simple algorithm to sort these determinants such that sequential determinants in the wave function differ by only one column (orbital). To calculate the local energy of the wave function, the algorithm only needs to perform one Sherman–Morrison update per determinant in the wave function. This is a significant performance boost where many pairs are used.

All of the cases we present here are adequately modeled with perfect pairing. However, for increased accuracy in some of our calculations, we can add restricted configuration interaction (RCI) terms[14] to the GVB reference wave function, without reoptimizing the orbitals. With these terms, the GVB-RCI geminal now takes the "excited" form $\{\sigma_u\phi_u^2 + \phi_u\phi_v - \sigma_v\phi_v^2\}$, adding some charge-transfer character in the pair. Although we could add these RCI terms to all geminals, for a total of $3^{N_{\mathrm{GVB}}}$ determinants, we excite only up to two geminals per determinant.

Recently, there has been some interest[15] in antisymmetrized geminal power (AGP) wave functions, so it is useful to compare the two methods. As Rassolov discusses,[16] both of these wave functions start by approximating the wave function with two electron geminals. To make this functional form tractable, AGP requires the geminals to be identical, while GVB requires them to be strongly orthogonal. This means that GVB wave functions are size consistent even during bond breaking, while AGP wave functions are in general not size consistent. Furthermore, because the GVB strong orthogonality is associated with chemical locality (e.g., bonds, lone pairs), whereas AGP geminals are forced to delocalize across the entire wave function, GVB and AGP can be considered to be opposite approaches,[17] even though in their respective unrestricted forms there is no formal distinction on a computational level.[3] We use Be_2 to compare their energies in Sec. III D.

B. Length scaled Jastrows

We implemented the two and three particle Jastrow functions recommended by Drummond *et al.*[7] because we like the cutoffs, flexibility, and simplicity. However, we found that their length scale parameter L was too difficult to optimize for the algorithms we use, so we use the following modifications instead. For two particle interactions, we use the functional form

$$u_{ij}[x \leftarrow r_{ij}S] = (x-1)^3\left(\sum_{k=0}^{M} a_k x^k\right), \quad \text{if} \quad 0 \le x \le 1, \quad (3)$$

$$= 0, \quad \text{if} \quad x > 1,$$

where r_{ij} is the distance between the two particles (electrons or nuclei) i and j, S is the length scale parameter ($x=r_{ij}S$), and a_1 is constrained to satisfy the cusp conditions. The $(x-1)^C$ prefactor is used to force the $C-1$ lowest order derivatives to go to zero at the cutoff. We found that $C<3$ inhibits the optimization of S using our routines, and that $C>3$ does not make much difference. Our three customizations are that the function uses the scaled coordinate x instead of r, we optimize $1/L$ instead of L, and we only use $C=3$. These do not change the variational flexibility of the function, but they do make the a_k parameters less dependent on S, easing their optimization. This makes a total of $M+1$ independent parameters, and in all calculations presented here, we use $M=8$. Optimizing the a_k parameters was still delicate during concurrent optimization with S, so we eventually turn off the optimization of S for some final fine tuning, as discussed in Sec. II C. We make analogous modifications to their electron-electron-nuclear Jastrows, using $M=3$, for our software.

Our tests did not indicate that differentiating between spin for electron-nuclear Jastrows significantly changed the energy, so we use the same Jastrow for all electrons. For hydrocarbons, then, we use four two particle Jastrow classes: carbon-electron, hydrogen-electron, opposite-electron, and parallel-electron. Adding the eight parameters for Jastrow's polynomial and the one length scale parameter, there are nine

parameters for each two particle Jastrow, for a total of 36 parameters for two particle Jastrows in most of our calculations.

Similarly, we ignore spin distinctions in our three particle Jastrows, leaving us with only one three particle Jastrow per element represented in the molecule. Although there are 4^3 terms of the form $x_i^a x_j^b x_{ij}^c$ in the polynomial for three particle Jastrows, there are several necessary constraints including symmetry and cusp conditions. Thus, the number of independent parameters is reduced considerably to only 27 parameters, including the length scale, per Jastrow class. As a further simplification, we found three particle Jastrows centered on hydrogen atoms to be unhelpful. This makes physical sense given that these Jastrows are primarily useful for modeling the interaction of two 1s electrons with the nucleus, and on average only one electron will be near a hydrogen nucleus.

C. Wave function optimization

To optimize our wave functions, we use the method recommended by Toulouse and Umrigar,[8] with the following modifications. To make a wave function, we copy into our input file the best Jastrows we have from among similar systems, noting that it is more important that we match the basis set than the type of SCF wave function. If we found that two CI coefficients were the same (or additive inverses) to within a relative difference of 10^{-5}, we constrained them to maintain the relationship. Furthermore, even though QMC is unaffected by the normalization of the wave function, we do not take this opportunity to eliminate a degree of freedom in the CI coefficients. Starting at around 20 000 samples per optimization step, we double the number of samples collected per iteration, with a maximum of 500 000 samples, if the variational energy does not go below the statistical error between successive iterations. Umrigar makes use of an a_{diag} factor to stabilize the eigenvector from the solver. Just as he does, we obtain this factor on the basis of a short correlated sampling run in between optimization steps. Our correlated sampling runs are produced using the best optimized wave function from the previous iteration as the guiding trial function, and including seven wave functions produced with preselected a_{diag} factors, logarithmically spaced between 10^{-7} and 10^3. The larger a_{diag} is, the less the wave function will change as compared with the previous iteration. The wave function for the next step is chosen from the seven by selecting the one i with the lowest quantity,

$$0.95(E_i - E_0) + 0.05(\sigma_i^2 - \sigma_0^2)/\sigma_0^2, \tag{4}$$

thus optimizing for lowering the energy compared to the guiding trial function ($i=0$), while penalizing a wave function with too large of a sample variance. With this scheme, if an optimization step goes bad, the step can effectively be ignored by choosing $a_{\mathrm{diag}}=1000$.

There are two occasional problems with this procedure applied to our Jastrow functions. First, despite our improvements, the length scale parameter remains a source of instability. Thus once we observe the length scale to be changing by less than a few percent, we turn off its optimization, allowing us to fine tune the other Jastrow parameters. Second, the algorithm occasionally leads to local minima. Some of our wave functions, for example the $^3B_1$ methylene wave function, initially optimized to an absurd parallel spin Jastrow, which was only discovered upon examining a plot of the Jastrow itself. In these cases, neither the energy nor the variance is necessarily suspicious, since after all, we do not know how deep the global minima go. The problem is that some of the local minima we found raised the VMC energy by about a few kcal/mol. For DMC, this is not a problem upon time step extrapolation, but we are not doing time step extrapolation as will be discussed in Sec. II E. For this reason, and since the CI coefficients might be affected by poor Jastrows, we currently have no choice but to carefully monitor our Jastrows during optimization.

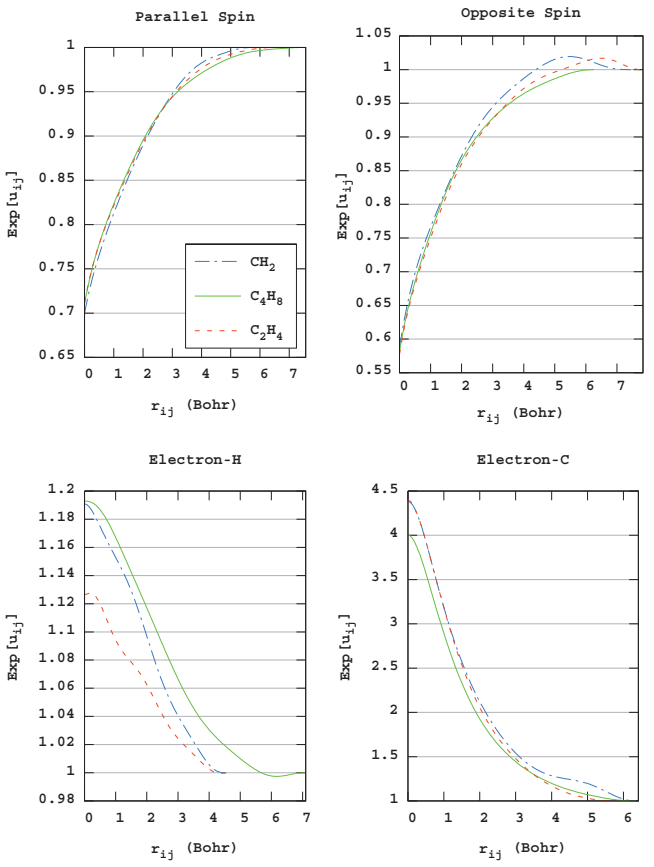

FIG. 1. Typical ground state Jastrow functions used in this study for the aug-tz basis set. Based on our experience, we do not believe that any Jastrows would look significantly different than these. In our optimization, we ignored minor flaws in the Jastrows, such as wiggles in the electron-H Jastrows, or the brief crossing of the $\exp(u_{ij})=1$ line in opposite spin Jastrows.

Once satisfactorily optimized, all Jastrows within each class looked qualitatively very similar when three particle Jastrows are not used. A few examples are plotted in Fig. 1, exponentiated. With this in mind, we were easily able to identify bad Jastrows as ones which cross the $\exp(u_{ij})=1$ line, which were not monotonic, or which took on extremely high or low values. In some exceptional cases, the global minima were only obtained by first optimizing all Jastrows except the troublesome one, constraining it to a good Jastrow from another system. Once that converges, we optimize the

troublesome Jastrow (and possibly its length scale) holding all the others fixed. We repeat this cycle until all of the Jastrows are sufficiently close to the global minima that concurrent optimization of all the Jastrows can lock it in. There were not many cases like this, but we estimate that this casts an uncertainty of perhaps a tenth of a kcal/mol on our results.

D. Walker stabilization

There are a variety of ways to design the branching process such that the number of walkers is always constant, and we use the algorithm designed by Assaraf and co-workers.[9] This is made possible with a reconfiguration step, where low weight walkers are replaced with duplicates of high weight Walkers. This is done by calculating the average walker weight W_{avg}, and using W_{avg} to bifurcate the list of walkers. We delete a total of

$$N_{\text{replacements}} \propto \sum_{i \in \{w_i < W_{\text{avg}}\}} \left| \frac{w_i}{W_{\text{avg}}} - 1 \right| \tag{5}$$

walkers, where a walker with weight w_i is selected with probability proportional to $|w_i/W_{\text{avg}}-1|$. The same probability proportionality relation is used to select enough high weight walkers for duplication, so that the total number of walkers is restored. After this, the weights of all walkers are set to W_{avg}, so that the total weight of the walkers is also unchanged. This method adds significant stabilization to the ordinary DMC process since any instabilities affecting one walker are instantaneously disbursed to the others.

We add further stabilization to the method, partly because of the added instability of our all-electron move iterations.[18] This is done simply by selectively ignoring in the duplication and elimination candidate lists walkers which fail our criteria. That is, we keep W_{avg}, the probabilities, and $N_{\text{replacements}}$ the same as prescribed. The only difference is that the actual length of either of the two lists might be different than the original algorithm predicted. The penalty for this is that, in rare cases, the algorithm will be unable to maintain the same number of walkers it started with. We modify our elimination lists to ensure that walkers with $w_i < 10^{-5}$ are guaranteed to be replaced this iteration, since they are a complete waste of computational effort. Defining "age" as the number of iterations since the walker last moved and dW as the multiplicative factor by which the weight changed this iteration, our acceptable duplication criteria are:

1. age < 4,
2. $dW^{(\text{age}+1)} < 5$,
3. and if the walker has not been duplicated this iteration.

Persistent walkers, those stuck in one location, can be a problem in a Monte Carlo calculation. Our improvement is to ensure through Criteria (1) that at least these walkers never become duplicated. Duplication will also be prohibited by Criteria (2) if a slow walker is in a location where its weight grows too fast. The reason is that we have found that some walkers can become stuck close to a wave function node, which is a singularity in the local energy, where they often spawn more quickly than they can move away. Lastly, with Criteria (3), we do not allow a walker to duplicate more than once per iteration, a fail-safe to slow the damage that one walker might cause. Although these techniques do not satisfy Markov chain criteria, we found them necessary for weight control, while only adding perhaps less than a tenth of a kcal/mol error.

E. Further details

To make our wave functions, we used both JAGUAR (Ref. 19) and GAMESS-U.S.,[20] and we obtained our basis sets from the EMSL website.[21,22] For this study, we have chosen two basis sets, which we label aug-tz and tz. Our aug-tz basis set is aug-cc-pwCVTZ, which is Peterson and Dunning's new[23] weighted basis functions, which were optimized with the inclusion of some core-core correlation energy for better overall performance. This basis set puts 25 basis functions from [4s3p2d] on H, 69 basis functions from [7s6p4d2f] on Be, and 69 basis functions from [7s6p4d2f] on C. We also use cc-pCVTZ, labeled here as tz, which uses their recent[24] scheme for adding core-valence correlation. This basis puts 15 basis functions from [3s2p1d] on H and 49 basis functions from [6s5p3d1f] on C. All Hartree–Fock, coupled-cluster,[25] GVB,[3] and MCSCF results were obtained in GAMESS using the same geometry as the corresponding QMC calculation. We included all determinants in the CI expansion, except where noted. All of our DFT calculations were done using JAGUAR with high precision settings. We include results using the LDA, PBE, B3LYP, and the M06-2X (Ref. 26) functionals, using the same geometries as the corresponding QMC wave functions. We used Jaguar to make our GVB wave functions, since it has a good mechanism for generating initial guesses.[13] But any wave functions that we used were handed over to GAMESS for final convergence since JAGUAR restricts us to seven f basis functions, and we want to use all ten Cartesian functions. Our wave functions are all-electron; none use any type of pseudopotential.

Our QMC calculations are done using the QMcBEAVER (Ref. 6) software developed in our group. The C++ source code is available online under the GNU public license. Starting with a script generated input file based on an SCF calculation and similar Jastrows, we use our own efficient algorithm[27] to initialize the walkers. We evaluate the local energy in all-electron updates, using the cusp replacement algorithm of Ma *et al.*[28] We use VMC to optimize all CI coefficients and Jastrows by the method recommended by Toulouse and Umrigar,[8] with our modifications as outlined in Sec. II C. Using the resulting optimized parameters, we run DMC based on Umrigar's seminal algorithm,[1] with our modifications as described here. Our calculations are run on four CPU cores, for a total of 400 walkers, using a different parallelization[29] technique than is typical for QMC calculations. All energies reported have been fully decorrelated using our efficient algorithm,[30] which automatically finds the smallest decorrelated block size.

Based on the results shown in Figs. 2 and 3, and other comparisons not included here, we can see that the majority of the time step error cancels off for each time step. This indicates that the dominant source of error is not the time

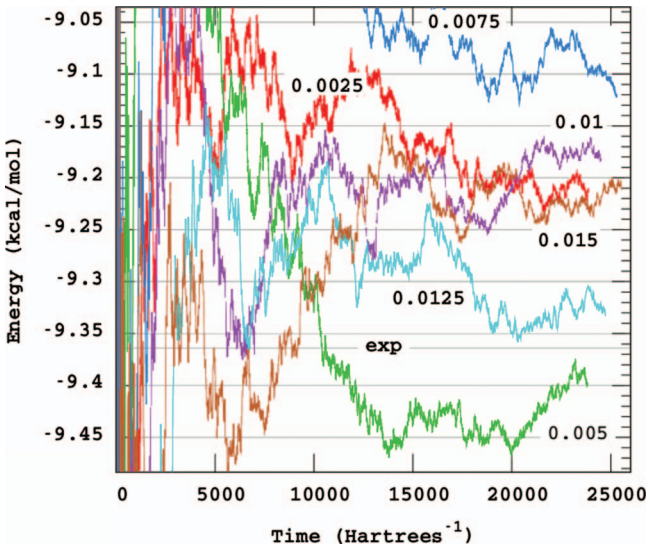

FIG. 2. Cancellation of time step error between triplet to singlet energies in methylene, using three pair delocalized GVB wave function. For this plot, individual calculations were stopped when they reached a statistical error of exactly 0.065 kcal/mol, corresponding to an error of 0.092 kcal/mol for the difference. We plot the differences here against the amount of simulated time, iterations × time step × average move acceptance probability.

step error itself, but an instability of the order of a few tenths of a kcal/mol. With this in mind, the consensus result appears to be about 9.2(1) kcal/mol. Since the computational cost of the calculation scales linearly with the number of samples, we are motivated to choose just one time step as large as reasonable to maximize the independence of each sample. We can also see that after running for about 15 000 a.u.$^{-1}$, most of the calculations converged to within the 0.092 kcal/mol statistical error. Based on this observation, we typically choose a time step of 0.0075 and run for 20 000 a.u.$^{-1}$, which corresponds to 2.7×10^6 iterations. Looking ahead at Table I, our converged result is 9.239 ± 0.088 kcal/mol for this case, in agreement with our qualitative assessment of Fig. 2.

The length of wallclock time for each calculation varied with many factors, but ranged from about 40 h on methylene to about 100 h on ethylene to about 400 h on cyclobutane. However, for these same calculations, each processor only required about 15–40 megabytes of random access memory (RAM) each, since low memory requirements are one of the benefits of all-electron updates. It is illustrative to compare these performance numbers with coupled-cluster methods, which not only scale poorly in computation time with larger molecules, but scale poorly in memory requirements as well. Even if a researcher is willing to wait for completion, memory is certainly a finite resource, and RAM will remain a bottleneck resource for the foreseeable future. In contrast, even though QMC scales somewhat poorly in computation time at $\mathcal{O}(N^3)$, assuming dense matrices, with a large prefactor, where N is the number of electrons, the memory requirements are negligible. This is a favorable situation since machines are rapidly getting faster, and it is even possible to run QMC on a graphical processing unit[31] for remarkable speedups.

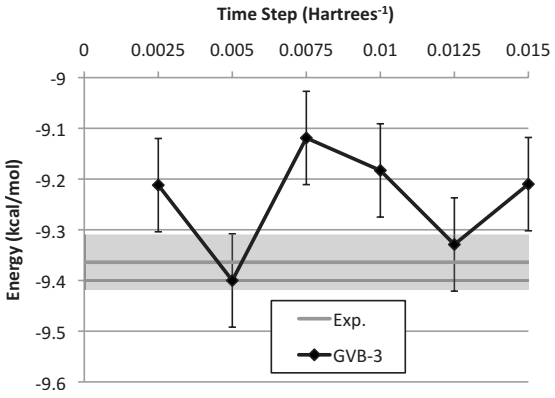

FIG. 3. The final energies from Fig. 2, plotted against their time step. All appear to be producing approximately the same result, so we choose a 0.0075 time step for the rest of our calculations, in lieu of extrapolating to zero time step.

III. RESULTS

We present our results for several related molecules for which good experimental or computational results are available to use as a reference. We wish to examine the effectiveness of adding GVB pairs to our wave functions, as well as the importance of different basis sets. In this section, we examine methylene, ethylene, cyclobutane, and beryllium dimer. The level of theory that we consider most reliable is adding GVB perfect pairs to all non-CH bonds, using our augmented basis set, and not using three particle Jastrows. This leads to a level of theory that is appropriate for studying large molecules. In each table, we highlighted this result, even where we present more accurate results.

A. Methylene

The singlet-triplet splitting Δ_e in methylene is among the most studied problems in quantum chemistry. The earliest GVB calculations[32,33] led to a $\Delta_e=11.5$ kcal/mol but the first direct spectroscopic experiments[34] led to $\Delta_0=19.4$ kcal/mol. Harding and Goddard later calculated[35] that GVB-CI led to $\Delta_e=11.1$ kcal/mol, while reinterpreting the experiments to exclude hot bands to conclude that the experimental data are $\Delta_0=9.0$ kcal/mol. The most recent spectroscopic results[36] are $\Delta_0=8.998\pm0.014$ kcal/mol. The issue here is the zero point energy (ZPE) correction δ_{ZPE} to convert Δ_0 to Δ_e. To make a comparison between theory and experiment, we need to correct for the differential ZPE of the $^1A_1$ and $^3B_1$ states, δ_{zpe}. The estimated $\delta_{\mathrm{ZPE}}=0.22$,[36] 0.20,[37] and 0.35 kcal/mol,[38] leading to leading to an experimental $\Delta_e=9.215\pm0.022$, 9.192 ± 0.014, and 9.364 ± 0.053, respectively. We consider the last number to be the most reliable and use it for comparing with the QMC results, which leads to 9.4 ± 0.10 kcal/mol.

This remains a very useful benchmark for QMC. The 2s and 2p atomic orbitals on carbon are nearly degenerate, necessitating the inclusion of all four into any carbon containing molecule. Any $^3B_1$ wave function does this, while one orbital is left out at the RHF level for $^1A_1$. Thus the simplest reasonable description of the $^1A_1$ state is to add the missing orbital by perfect pairing it with the lone pair as an angular correlation term. It is important to also recognize that triplet

TABLE I. Methylene excitations: $^1A_1 \leftarrow {}^3B_1$ and $^1B_1 \leftarrow {}^3B_1$. For $^3B_1$, $[R_{CH}, \Theta_{HCH}]=[1.0753$ Å, 133.93] from experiment (Ref. 36), for $^1A_1$ [1.107 Å, 102.4] from experiment (Ref. 53), and for $^1B_1$ [1.0723 Å, 142.44] from theory (Ref. 54). By "B," we are indicating our basis, by "O" we are indicating, where it matters, whether our GVB pairs are localized or delocalized, and by "J" we are indicating whether we are using two or three particle Jastrows. We highlighted the results which meets our selection criteria.

SCF	B	O	J	Δ_e kcal/mol	$^1A_1$ a.u.	$^3B_1$ a.u.
GVB-3	aug-tz	L	2	9.071(80)	−39.121 669(91)	−39.136 124(89)
GVB-3	aug-tz	D	2	9.239(88)	−39.120 847(81)	−39.135 57(11)
GVB-3	aug-tz	D	3	9.340(71)	−39.124 461(79)	−39.139 345(82)
Exp[a]				9.364(53)		
RCI-3	aug-tz	L	2	9.37(11)	−39.121 76(13)	−39.136 70(12)
CAS-3[b]	VB1		3	9.4(1)		
GVB-1	aug-tz		2	9.40(10)	−39.121 49(12)	−39.136 48(12)
CAS-1[b]	VB1		3	9.5(1)		
RCI-3	aug-tz	D	2	9.519(95)	−39.121 37(11)	−39.136 54(11)
GVB-3	tz	D	2	9.53(10)	−39.120 65(11)	−39.135 84(12)
GVB-3	tz	D	3	9.557(74)	−39.123 975(83)	−39.139 205(83)
GVB-1	aug-tz		3	9.560(76)	−39.124 248(91)	−39.139 483(80)
GVB-1	tz		2	9.65(11)	−39.120 93(12)	−39.136 31(12)
GVB-1	tz		3	9.673(73)	−39.123 838(84)	−39.139 253(81)
CAS-3	aug-tz		2	9.792(92)	−39.123 53(10)	−39.139 13(10)
RHF	aug-tz		2	13.80(10)	−39.114 49(12)	−39.136 48(12)
RHF	aug-tz		3	13.844(73)	−39.117 421(85)	−39.139 483(80)
SCF	B	O	J	Δ_e kcal/mol	$^1B_1$ a.u.	$^3B_1$ a.u.
PES[c,d]				31.897		
GVB-1	aug-tz		2	32.06(11)	−39.085 39(12)	−39.136 48(12)
GVB-1	aug-tz		3	32.114(71)	−39.088 306(80)	−39.139 483(80)
MRCI[c]				32.807		

[a]Experimental $T_e=T_0+\Delta ZPE$, where $T_0=3147\pm5$ cm^{-1} (Ref. 36) and $\Delta ZPE=128\pm18$ cm^{-1} (Ref. 38).
[b]Optimizing basis functions and orbitals, from Ref. 39.
[c]From Ref. 37.
[d]From Ref. 54.

paired electrons are much better correlated, due to orbital orthogonality, than closed shell counterparts in a singlet wave function. Consistency requires at least that the number of orbitals on each side of a comparison is the same, providing another reason for the perfect pairing.

We present our results in Table I. Our GVB-1 calculations represent RHF for the triplet state, and one GVB perfect pair for the singlet, indicating our policy of using the label from the comparison with the highest number of pairs. This is the level of theory consistent with our other calculations, by excluding CH bonds. The GVB-3 level adds correlation to the bonds, for a total active space of six orbitals, and there are two ways to do this. GVB-3 is supposed to use localized bonding and antibonding orbitals, but we also include a version with the same four orbitals delocalized, even though the SCF GVB splittings are 0.02 kcal/mol different. The RCI-3 level of theory excites up to two perfect pairs into their corresponding open shell singlet, without optimizing the orbitals. Finally, by CAS-3, we mean the complete active space (CAS) in the six orbitals, optimizing the orbitals in SCF. We also note that in contrast with other theoretical studies of this system, we do not incorporate any other energy corrections to our measurements. It is gratifying to note how good our result of 9.40 ± 0.10 kcal/mol is compared with other recent DMC calculations from Zimmerman *et al.*,[39] who also obtain 9.4 ± 0.1 kcal/mol with wave functions are significantly more superior than our own since they optimize their orbitals and basis functions.

For methylene, we run each calculation shown in Table I twice so that we can average some of the instabilities out, a luxury we do not employ for our other molecules. Additionally, one of these two runs for our GVB-3/aug-tz was run for much longer, since we were surprised that the localized orbital results are further from experiment. This error is compensated for at the RCI-3 level. All of our results are within 0.4 kcal/mol of the experimental estimate, with the exception of our RHF calculation which does quite poorly at an error of about 4 kcal/mol. Additionally, we include our estimation of the lowest singlet-singlet vertical excitation, even though there is little consensus for what the accurate energy should be. Adding augmented basis functions improves our estimates by 0.1 to 0.2 kcal/mol, while 50% more basis functions increased computational time by only 10%–30%. There is no reason not to use the augmented version of the chosen basis set class. Looking at our timing data, we see that if we stopped our calculations at an error of 0.1 kcal/mol, our three particle wave functions would have finished 30%–40%

TABLE II. The VMC energies for $^1A_1 \leftarrow {}^3B_1$ corresponding to the DMC calculations from Table I. The labels have the same meaning. We include the rms fluctuations for each calculation. We highlighted the result which meets our selection criteria.

SCF	B	O	J	Δ_e (kcal/mol)	$^1A_1$ (a.u.)	σ_A (a.u.)	$^3B_1$ (a.u.)	σ_B (a.u.)
GVB-3	aug-tz	D	3	8.214(66)	−39.103 026(71)	0.0829	−39.116 115(77)	0.0813
GVB-3	aug-tz	L	2	8.294(100)	−39.079 68(11)	0.1115	−39.092 89(11)	0.1113
GVB-3	aug-tz	D	2	8.333(98)	−39.076 41(11)	0.1148	−39.089 68(11)	0.1155
GVB-3	tz	D	2	8.533(96)	−39.076 11(10)	0.1091	−39.089 71(11)	0.1103
RCI-3	aug-tz	D	2	9.329(94)	−39.079 41(11)	0.1123	−39.094 28(10)	0.1125
Exp				9.364(53)				
RCI-3	aug-tz	L	2	9.489(91)	−39.080 52(10)	0.1120	−39.095 64(10)	0.1120
GVB-3	tz	D	3	9.769(67)	−39.101 465(78)	0.0797	−39.117 033(73)	0.0822
CAS-3	aug-tz		2	10.183(94)	−39.083 11(11)	0.1098	−39.099 33(11)	0.1104
GVB-1	aug-tz		2	10.23(10)	−39.071 47(11)	0.1124	−39.087 77(11)	0.1133

quicker, demonstrating their value in variance reduction. This comparison encourages their use, but this conclusion changes for cyclobutane.

It is clear that beyond the statistical error, there are some additional sources of error. As mentioned previously in reference to Fig. 2, there is some error due to instability in the convergence, which we attempted to minimize for methylene by running each calculation twice. But, more importantly, there appears to be some error due to incomplete optimization of wave function parameters. For example, using our tz basis set, the addition of three particle Jastrows does not appreciably change the energy difference, a result which makes sense given our assumption that the time step error cancels out. This is not the case for our aug-tz basis set, which changes by at least 0.1 kcal/mol with the addition of three particle Jastrows. Our CAS-3 results are also puzzling, given how far it stands from the better results from theoretically worse wave functions. Our CAS-3 result is 0.4 kcal/mol higher than the result from Zimmerman *et al.*, highlighting two possible sources of error. First, there is the possibility that, despite our efforts, our wave function had become stuck in a local minima of CI coefficients. Second, that optimizing CI coefficients without optimizing the orbitals themselves produces an error that grows with the length of the CI expansion. If this is the case, then this favors the use of GVB wave functions based on our results, whose orbitals do not need to be optimized.

When running QMC calculations, the DMC energy is more interesting than its VMC energy counterpart. The DMC energy includes all correlation energy except for that left out by the fixed node approximation, whereas the VMC energy only includes as much correlation energy as the wave function describes. Although we believe that the value in a method lies in the DMC energies themselves, the VMC energies are useful for examining the wave function. We present a few VMC results for methylene in Table II. The most obvious result is that a three particle wave function is far superior than a two particle wave function, both in producing lower energies and errors for individual calculations, even though they do not appear to consistently improve energy differences. We note that the ordering of the individual energies is consistent with the theoretical quality of the wave function, even if the energy differences are not.

We wanted to discover the effect of optimizing different parts of the wave function. We pursued this by choosing some standard state for each atom for the Jastrows, and then selectively optimizing parts of the wave function, and comparing these results to the comparable result from Table I. Our results, shown in Table III, tell us that optimizing the CI coefficients was worth 0.5 kcal/mol, and that optimizing the electron-nuclear Jastrows was worth another 0.4 kcal/mol. Of course, in the limit of zero time step Jastrows should not matter, so much of the error here can be called time step error. But it appears to be crucial that we optimize the CI coefficients, even if we are not optimizing the orbitals. Returning to the question of the CAS discrepancy, we tried optimizing all the Jastrows, while keeping the original CI coefficients. This produced 15.80 ± 0.11 kcal/mol as the DMC energy splitting, the worst of all the results we obtained. Since this represents the comparison of the SCF functions directly without worrying about whether the VMC optimization is falling into local minima, we can see that, given our methodology, a better SCF wave function does not always improve accuracy unless everything is optimized.[39] On the other hand, in separate investigations,[40] we found that CAS wave functions were the simplest accurate representations we tried.

In Table IV, we present our results for several single-

TABLE III. The effect of optimizing different parts of our aug-tz GVB-3 delocalized methylene wave functions, with two particle Jastrows, and all calculations run at 0.0075 time step. The starting point for these calculations is the CI coefficients from GVB, electron-carbon and electron-electron Jastrows from optimized carbon GVB-1 atom, and electron-hydrogen Jastrows from an optimized GVB-1 H_2 molecule. Each row corresponds to a different set of parameters which were optimized, where EN stands for electron-nuclear and EE for electron-electron.

Optimization	Δ_e (kcal/mol)	$^1A_1$ (a.u.)	$^3B_1$ (a.u.)
EN and CI	9.20(11)	−39.121 84(13)	−39.136 50(12)
Fully optimized	9.239(88)	−39.120 847(81)	−39.135 57(11)
EE and CI	9.51(11)	−39.121 33(13)	−39.136 48(12)
CI	9.58(11)	−39.121 61(13)	−39.136 87(12)
EE and EN	9.97(11)	−39.118 72(12)	−39.134 60(12)
EE	10.05(11)	−39.118 01(12)	−39.134 03(12)
No optimization	10.08(11)	−39.118 06(13)	−39.134 12(11)

TABLE IV. Methylene excitations using single determinant wave functions. Using our aug-tz basis set, we obtained orbitals from RHF or DFT, and added two particle Jastrows.

	Δ_e (kcal/mol)	$^1A_1$ (a.u.)	$^3B_1$ (a.u.)
CAS-3	13.76(10)	−39.114 99(13)	−39.136 93(10)
RHF	13.80(10)	−39.114 49(12)	−39.136 48(12)
B3LYP	14.05(11)	−39.115 39(12)	−39.137 78(12)
LDA	14.39(15)	−39.114 81(18)	−39.137 73(16)
PBE	14.64(17)	−39.115 27(24)	−39.138 60(12)

determinant representations of the trial wave function. Here we can see that none of these wave functions are capable of addressing the missing angular correlation, and that perhaps DFT is worse than RHF. This demonstrates that the fixed node approximation is associated with static correlation, which DFT cannot do much about.

B. Ethylene

There has been continued interest in calculating various excitation energies for ethylene in QMC, from the ground state singlet ^1A_g, also known as the N state, to the first excited triplet $^3B_{1u}$, the T state, or singlet $^1B_{1u}$, the V state. Experimentally measured energies for the N-T splitting will tend to be artificially low since the molecule twists immediately upon excitement to the triplet state, and, indeed, measured values span a range of 4.32 eV (Ref. 41) to 4.6 eV.[42] Calculations have been in better agreement, with results ranging from recent QMC calculations[43,44] both producing 4.50 ± 0.02 and 4.51 eV for a CCSD(T)/CBS (Ref. 45) calculation, up to about 4.6 eV for MRCI (Ref. 46) and auxiliary field Monte Carlo.[47] In the many comparisons made with experimentally based results, researchers typically do not account for the zero point energy, which is difficult to calculate for the vertical triplet state, so we do not either.

DMC was wrong by several kcal/mol from the correct energy splitting when we used RHF wave functions. Part of the problem, as discussed for methylene, is that the RHF level of theory is inconsistent between the N and T states. Thus the simplest level of theory for which we obtained correct results was the GVB-1 level, which perfect-pairs the π^* orbital to the π orbital for the N state. This indicates that, for ethylene, the most important source of fixed-node error is the left-right correlation in the double bond. The $\pi\pi^*$ electrons in the triplet RHF wave function are already correlated at the GVB-1 level since they occupy orthogonal orbitals, and both states use the same nine orbitals, satisfying consistency. The next level of theory is GVB-2, which adds left-right correlation to the CC single bond for both states. Finally, for GVB-6, we add correlation to all four CH bonds. RCI and CAS have the same meaning as we described for methylene.

Among our consistent results for aug-tz, shown in Table V, we can see agreement to within 0.26 kcal/mol with each other and with the other DMC results of 103.5 ± 0.5 kcal/mol, with the exception of our RCI-6 calculation. This is a clear indication that the GVB level of theory is sufficient to capture this process, and that going beyond our result of 103.45 ± 0.17 kcal/mol is unnecessary. Here, we can see that, unlike our methylene CAS wave functions, our ethylene CAS wave functions are correct. On the other hand, our RCI calculation seems to have a problem whereas our methylene RCI calculations were good. This is again evidence of local minima among the optimized parameters, or more evidence of the need to optimize orbitals as well. Presumably, we could pay as much attention to these wave functions as we did for our methylene CAS wave function and perhaps improve the result, but this would represent an unfair selection bias to our overall methodology. Either way, this speaks well for GVB, which does not appear to have any problems.

Examining our inconsistent results, below the horizontal line, we can see that left-right correlation in the double bond (found by comparing GVB-1 with RHF) is worth 1.80 or 2.96 kcal/mol. Our GVB-(1,1) case, an inconsistent wave function which correlates the CC single bond for the T state, but only the double bond for the N state, does produce an excellent energy difference, showing that consistency is not always critical. Once the double bond's correlation is included, the QMC results have reached convergence, suggesting that the remaining correlation energy from the SCF perspective is almost entirely dynamical. Looking at the SCF results, the RHF splitting was 83 kcal/mol, GVB calculations all produced about 100 kcal/mol, RCI calculations produced 108 kcal/mol, and our CASSCF(12,12) calculation produced 110 kcal/mol. We can clearly see the advantage of QMC, even when inconsistent, over other approaches which converge slowly to the accurate result.

Originally, we used MP2 and the tz basis set to obtain our ethylene geometry, and we include those results in Table VI. Concerned about the disagreement of about 2 kcal/mol between these results and the other DMC results, we decided to switch and use exactly the same geometry as Schautz,[44] and our results did agree. We include these results to illustrate a few key lessons. First, we point out that most of the difference came from the energy of the T state with its steep energy slope, demonstrating that poor geometries can result in errors as large as those in the method itself. Second, notice that previously our RCI-6 calculation was not an outlier, as it is with the new geometry. One difference was that, previously, we used a determinant cutoff of 0.001 for our RCI-6 and our CAS-6 wave functions, whereas for the new geometry, we raised the cutoff to 0.01 so that they would run faster (about two to three times for the N state). This change in truncation appears to have helped the CAS-6 calculation relative to consensus, but hurt the RCI-6 calculation. This forced us to lower the truncation cutoff back to 0.001 for RCI-6, and results at both levels are shown in Table V. Third, in rerunning the calculation, we used the optimized Jastrows in the new wave functions and reoptimized everything. This appears to have helped improve consistency, which can be seen by comparing the spread in Δ_e for aug-tz. If RCI-6 and CAS-6 are this sensitive to determinant cutoffs, then this is yet another reason to prefer wave functions with fewer CI coefficients.

For the N-V vertical splitting, at the bottom of Table V, our energies are 8 kcal/mol higher than the best values re-

TABLE V. Vertical ethylene: $^3B_{1u} \leftarrow {}^1A_g$ and $^1B_{1u} \leftarrow {}^1A_1$. For $^3B_{1u}$. For all calculations, we used R_{CC} =1.339 Å, R_{CH}=1.086 Å, and Θ_{HCH}=117.6, in order for our results to be directly comparable with Schautz (Ref. 44). The entries below the horizontal line are inconsistent with the number of GVB pairs indicated in parentheses for each state individually. We highlighted the result which meets our selection criteria.

SCF	B	J	Δ_e kcal/mol	$^3B_{1u}$ a.u.	^1A_g a.u.
Exp[a]			100.54		
GVB-1	tz	2	103.13(16)	−78.398 72(18)	−78.563 07(18)
GVB-6	tz	2	103.38(27)	−78.397 81(40)	−78.562 56(16)
GVB-2	aug-tz	2	103.45(17)	−78.397 59(18)	−78.562 45(22)
DMC[b]	Partridge	3	103.5(3)		
DMC[c,d]		2	103.5(5)		
CAS-6[e]	aug-tz	2	103.51(26)	−78.402 08(38)	−78.567 03(17)
GVB-6	aug-tz	2	103.56(14)	−78.397 42(16)	−78.562 46(16)
CAS-2	aug-tz	2	103.68(41)	−78.397 81(17)	−78.563 03(63)
GVB-1	aug-tz	2	103.71(42)	−78.397 24(18)	−78.562 51(64)
GVB-2	tz	3	103.91(39)	−78.403 20(60)	−78.568 79(14)
GVB-2	tz	2	103.98(16)	−78.396 76(18)	−78.562 46(18)
CCSD(T)[f]	CBS		104.1		
RCI-6[g]	aug-tz	2	104.29(14)	−78.398 97(16)	−78.565 16(16)
RCI-6[e]	aug-tz	2	105.14(38)	−78.397 27(15)	−78.564 83(59)
Exp[h]			106.1		
RHF	tz	2	100.17(31)	−78.398 72(18)	−78.558 36(47)
RHF	aug-tz	2	101.91(38)	−78.397 24(18)	−78.559 64(57)
GVB-(1,1)	aug-tz	2	103.49(42)	−78.397 59(18)	−78.562 51(64)
SCF	B	J	Δ_e kcal/mol	$^1B_{1u}$ a.u.	^1A_g a.u.
Exp[i]			177.57		
DMC[c,d]			182.9(5)		
CAS-2	aug-tz	2	190.81(41)	−78.258 96(19)	−78.563 03(63)
"CAS 6-6"[c]			192.6(5)		
CAS-6	aug-tz	2	199.65(15)	−78.248 87(16)	−78.567 03(17)

[a]Energy-loss spectra, from Ref. 41.
[b]Single determinant from CASSCF(4,8), using pseudopotentials, from Ref. 43.
[c]Using pseudopotentials and their custom basis set, from Ref. 44.
[d]These DMC results use VMC optimized orbitals.
[e]Only determinants with coefficients >0.01 were included.
[f]Computed value from Ref. 45.
[g]Only determinants with coefficients >0.001 were included.
[h]Optical spectra, from Ref. 42.
[i]Adsorption spectra, from Ref. 55.

ported by Schautz and Filippi,[44] for which they even optimized orbitals within their QMC treatment. This illustrates that for some molecules, it is important to include dynamic correlation during[44,48] orbital optimization. We use the same geometry, but our results are only directly comparable when neither of us optimizes orbitals. Our CAS-2 N-V splitting, based on a CASSCF(4,4) calculation, is about 2 kcal/mol better than their "CAS 6-6." The difference could be due to pseudopotentials, or because we did not need to truncate our CI expansion, like they did for coefficients below 0.01. However, our CAS-6 calculation, based on a CASSCF(12,12) wave function with determinants truncated at 0.01, is 7 kcal/mol worse than theirs, perhaps due to a failure on our part to fully optimize this wave function. This is more evidence that simpler wave functions can work well.

The N-T vertical splitting is difficult to study experimentally, since the triplet state is far from its D_{2d} minimum. In Table VII, we examine the adiabatic splitting, the geometry for which we obtained by optimizing the structure with MP2 using the tz basis set. We use the same N state QMC energies as before, but include them again in this table for completeness. Although there does not appear to be sufficient experimental data to make a good comparison, we do have the recent high quality CCSD(T)/CBS result[45] of 68.8 kcal/mol to compare against, and with which our recommended result of 69.20 ± 0.17 kcal/mol only differs by 0.4 kcal/mol. El Akramine and co-workers[43] also recently studied this transition using QMC to obtain 69.6 ± 0.3 kcal/mol, and our results are in agreement with theirs, even given the differences in our wave functions.

C. 2+2 cycloaddition

The ethylene+ethylene reacting to make cyclobutane is the textbook example of a concerted reaction forbidden by the Woodward–Hoffman rules. We are only doing a two

TABLE VI. Vertical ethylene: $^3B_{1u} \leftarrow {}^1A_g$ and $^1B_{1u} \leftarrow {}^1A_1$. For $^3B_{1u}$. For these calculations, we used MP2 optimized R_{CC}=1.331 046 Å, R_{CH}=1.080 564 Å, and Θ_{HCH}=121.35. We highlighted the result which meets our selection criteria.

SCF	B	J	Δ_e kcal/mol	$^3B_{1u}$ a.u.	^1A_g a.u.
GVB-2	aug-tz	2	105.05(16)	−78.394 80(18)	−78.562 20(18)
GVB-6	aug-tz	2	105.14(14)	−78.395 56(16)	−78.563 11(16)
GVB-6	tz	2	105.38(14)	−78.394 64(16)	−78.562 57(16)
RCI-6[a]	aug-tz	2	105.55(14)	−78.395 58(16)	−78.563 79(15)
GVB-2	tz	2	105.64(16)	−78.394 01(18)	−78.562 36(18)
CAS-2	aug-tz	2	105.82(16)	−78.395 02(18)	−78.563 66(19)
GVB-1	tz	2	105.99(16)	−78.394 00(18)	−78.562 90(18)
GVB-2	tz	3	106.14(13)	−78.399 84(15)	−78.568 99(14)
RCI-2	tz	2	106.16(16)	−78.394 60(18)	−78.563 77(18)
CAS-6[a]	aug-tz	2	106.54(14)	−78.397 33(15)	−78.567 11(15)
SCF	B	J	Δ_e kcal/mol	$^1B_{1u}$ a.u.	^1A_g a.u.
CAS-2	aug-tz	2	192.27(17)	−78.257 26(19)	−78.563 66(19)
CAS-2	tz	2	193.83(18)	−78.254 89(22)	−78.563 77(18)

[a]Only determinants with coefficients >0.001 were used.

point calculation, one for an isolated ethylene molecule, and one for an isolated cyclobutane molecule, bypassing any questions related to allowed reaction paths. This is one of the simplest reactions that DFT gets incorrect, disagreeing with experiment by 5–10 kcal/mol, even with some of the more recent functionals, so we consider this to be an ideal showcase for QMC. Our cyclobutane geometry was obtained by optimizing the D_{2d} structure with MP2 using the tz basis set.

Below the dashed line in Table VIII we show our single determinant results using RHF, and also using the orbitals from an M06-2X DFT calculation. We were disappointed to be unable to get any single-determinant DMC calculation to do any better than DFT, with errors of 3–4 kcal/mol. This process breaks and then makes two bonds, suggesting that at least 2 GVB pairs should be used. However, since the CC bonds in cyclobutane are equivalent, we cannot justify using fewer than four GVB pairs on either side of the reaction. Indeed, upon adding left-right correlation to the bonds, our recommended answer 22.05 ± 0.32 kcal/mol agrees perfectly with our experimental estimate to within our statistical error. We should mention here that this near perfect agreement should be considered coincidental, since there is perhaps as much error in the ZPE and geometry as in the calculation.

Looking back to our tz ethylene calculations, we estimated that the static correlation in the double bond was worth 2.96 kcal/mol. Seeing here that our single-determinant calculation is in error by about 5.4 kcal, we conclude that most of this error comes from ethylene. With this in mind, we could have accepted decent results by only correlating the double bond in ethylene, which is our GVB-(1,0) result from

TABLE VII. Adiabatic ethylene: $^3B_{1u} \leftarrow {}^1A_g$. For $^3B_{1u}$, we use R_{CC}=1.449 148 Å, R_{CH}=1.080 469 Å, and Θ_{HCH}=121.5, and we use the same geometry as previously for $^1A_1$. The entries below the horizontal line are unbalanced in terms of the number of orbitals. We highlighted the result which meets our selection criteria.

SCF	B	J	Δ_e kcal/mol	$^3B_{1u}$ a.u.	^1A_g a.u.
Exp[a]			61.(3)		
CCSD(T)[a]	CBS		68.8		
GVB-1	aug-tz	2	69.14(42)	−78.452 33(18)	−78.562 51(64)
GVB-2	aug-tz	2	69.20(17)	−78.452 17(17)	−78.562 45(22)
DMC[a,b]	Partridge	3	69.6(3)		
GVB-2	tz	2	69.79(16)	−78.451 24(18)	−78.562 46(18)
GVB-1	tz	2	70.13(16)	−78.451 31(17)	−78.563 07(18)
GVB-6	tz	2	70.31(14)	−78.450 51(16)	−78.562 56(16)
CASSCF-6[c]	aug-tz		72.13		
RHF	tz	2	67.17(31)	−78.451 31(17)	−78.558 36(47)
RHF	aug-tz	2	67.34(37)	−78.452 33(18)	−78.559 64(57)

[a]We "uncorrect" the experimental value from Ref. 56 of 58(3) kcal/mol and all the computed results from Ref. 45 by ΔZPE=3.2 kcal/mol, so that we can directly compare calculations.
[b]Single determinant from CASSCF(4,8), using pseudopotentials, computed value from Ref. 43.
[c]Our own CASSCF(12,12) calculation.

TABLE VIII. Cycloaddition: $2C_2H_4 \leftarrow C_4H_8$. We use the same ethylene geometry as previously, and our cyclobutane geometry is $R_{CC}=1.545\ 029$ Å, $R_{CHax}=1.089\ 404$ Å, $R_{CHeq}=1.0877$ Å, and $\Theta_{HCH}=109.18$. Below the solid horizontal line are inconsistent calculations, where the number of GVB pairs for the two states are in the parentheses. We highlighted the result which meets our selection criteria.

SCF	B	J	Δ_e (kcal/mol)	C_2H_4 (a.u.)	C_4H_8 (a.u.)
GVB-4	tz	2	21.98(28)	−78.562 46(18)	−157.159 93(27)
GVB-4	aug-tz	2	22.05(32)	−78.562 45(22)	−157.160 04(27)
Exp[a]			22.3(2)		
CCSD(T)	tz		22.54		
GVB-4	tz	3	22.65(23)	−78.568 79(14)	−157.173 67(23)
M06–2X	tz	2	25.67(30)	−78.561 21(19)	−157.163 33(30)
RHF	tz	2	27.37(61)	−78.558 36(47)	−157.160 34(28)
GVB-(1,0)	tz	2	21.45(28)	−78.563 07(18)	−157.160 34(28)
GVB-(1,4)	aug-tz	2	21.98(82)	−78.562 51(64)	−157.160 04(27)
GVB-(0,4)	aug-tz	2	25.58(74)	−78.559 64(57)	−157.160 04(27)
GVB-(0,4)	tz	2	27.12(61)	−78.558 36(47)	−157.159 93(27)

[a]Enthalpy (Refs. 57 and 58) difference of 68.97(71) kJ/mol, corrected with ΔZPE (Ref. 59) 5.84 kcal/mol.

below the horizontal line in Table VIII. Although this provides some opportunity for shortcuts in larger calculations, when possible only balanced calculations should be considered, such as those above the line.

We note that all three of our calculations were successful, disagreeing by only 0.3 kcal/mol. Therefore, the augmented basis functions were unnecessary. If we would have stopped all cyclobutane calculations once, they reached 0.2 kcal/mol error, our cyclobutane RHF/tz wave function with three particle Jastrows (not in the table) would have spent 33% more computational time than the equivalent wave function without the three particle Jastrows. Additionally, the analogous GVB-4/tz three particle Jastrow calculation would have taken 4% more time than when we left the three particle Jastrows out. The reason is because at a length scale of over $6a_0$, the three particle Jastrows can reach almost four times as many electrons in cyclobutane than in methylene. In contrast with our conclusions for methylene, the three particle Jastrows are not worth the computational effort, even if we were entirely confident in their optimization.

D. $Be_2 \rightarrow 2Be$

The Be_2 molecule is notable for being particularly challenging for quantum chemistry methods. The difficulty stems from the near degeneracy of the 2s and 2p orbitals, significantly distorting the symmetry of the Be atom, and making the RHF wave function a particularly poor choice. In the molecular orbital picture of Be_2, the four valence electrons fill the 2s and $2s^*$ bonding orbitals, resulting in an attraction that is not quite a bond. The GVB approach makes a single bond out of the 2s orbitals and adds left-right correlation, but then puts the other two electrons in an antibond orbital made from the $2p_z$ orbitals for the overall $^1\Sigma_g^+$ symmetry. Martin pointed out[49] that a CASSCF(4,4) reference is preferred over even a CASSCF(4,8) reference, who argue that $2p_{x,y}$ and $3p_{x,y}$ should either all be included or all excluded. This strengthens the credibility of a GVB wave function.

There are numerous results to compare against, and we have not attempted a literature survey. The best result comes from a recent experimental study of the system by Merritt and co-workers,[50] who fit a function to rotational spectra, obtaining 2.6581 ± 0.0057 kcal/mol. This calculation has been attempted by several QMC researchers as well. Recently Toulouse and Umrigar[51] demonstrated their orbital and basis function optimization for a full valence wave function that is possibly superior to ours, depending on the relevance Martin's warning in a DMC context.

As discussed in Sec. II A, it is interesting to compare GVB wave functions with a related type of wave function, the AGP wave function,[15,52] which has also been used for DMC calculations. The advantage of the GVB wave function is illustrated in Table IX, where our recommended GVB result of 2.44 ± 0.12 kcal/mol is much closer to experiment, whether or not we optimize our orbitals, than the AGP result 3.30 ± 0.14 kcal/mol. The differences shown here are probably larger than they should be, so further comparisons would be interesting, including wave functions[16] which bridge the gap between AGP and GVB.

We present our results for the well depth of the Beryllium dimer in Table IX, using the aug-tz basis set used in our other studies. We implemented a naive form of orbital optimization. Starting with GVB orbitals, we optimize all nonzero orbital coefficients as parameters treated the same way as all the other Jastrow and CI parameters. Unfortunately, the number of orbital parameters is quite high, so this is the only reaction that we used orbital optimization to study. We note that our unoptimized orbital results are only 0.22 and 0.11 kcal/mol from the 2.6581 ± 0.0057 kcal/mol reference.

IV. CONCLUSION

In this paper, we used QMC to study the effect of various types of wave functions on calculations for which we have high quality results to compare against, and found that our recommended GVB wave function was sufficient to obtain results accurate to a few tenths of a kcal/mol without optimizing the orbitals. Based on this, we recommend GVB wave functions with two particle Jastrows for studying large molecules. This conclusion is drawn with the exception of singlet-singlet ethylene, for which our simple wave functions

TABLE IX. $2Be \leftarrow Be_2$ at the experimental geometry of 4.63 bohr. We indicate in the Opt column whether the each result used QMC optimized orbitals. We highlighted the result which meets our selection criteria.

SCF	B	J	Opt	Δ_e (kcal/mol)	Be (a.u.)	Be_2 (a.u.)
AGP[a]	STO-DZ	3	y	0.43(19)		
CAS[b]	CVB1	3	y	1.770(15)		
GVB-2	aug-tz	2	y	2.24(11)	−14.660 410(69)	−29.324 39(11)
CAS[c]	tz	3	n	2.37(18)		
GVB-2	aug-tz	2	n	2.44(12)	−14.660 254(72)	−29.324 39(12)
GVB-2	aug-tz	3	y	2.555(85)	−14.660 549(54)	−29.325 169(84)
Ab initio[d]				2.582(23)		
Exp[e]				2.6581(57)		
Ab initio[f]				2.699(71)		
GVB-2	aug-tz	3	n	2.772(90)	−14.660 479(52)	−29.325 375(98)
RHF[b]	CVB1	3	y	2.893(33)		
JAGPn* [g]		4	y	3.30(14)		

[a]DMC on an AGP wave function, Ref. 15.
[b]Everything optimized, including basis functions, from Ref. 51.
[c]From Ref. 60.
[d]Explicitly correlated r_{12}-MR-CI, from Ref. 61.
[e]Stimulated emission pumping, from Ref. 50.
[f]Extrapolated *ab initio*, from Ref. 49.
[g]Everything optimized, including basis functions, from Ref. 52.

were unable to obtain the correct splitting, demonstrating that GVB restricted to perfect pairing is not suitable for all problems.

Furthermore, we discussed our difficulty in studying these same problems using extended CASSCF wave functions and RCI wave functions when we do not optimize orbitals. There are two issues that affected our results. First, our results have been somewhat sensitive to how we truncate the CI expansion for inclusion in our QMC wave functions, and it appears that 0.01 is not always good enough for them to perform even as well as the simpler GVB. Second, even where we applied concentrated effort in optimizing CASSCF wave functions with all determinants included, we were still unable to get as close of a result as GVB.

Finally, regardless of minor issues, it is remarkable how well QMC performs even for difficult cases, since all our consistent calculations were within chemical accuracy. We believe that given a simple GVB description with two particle Jastrows, we are able to describe a significant amount of chemistry, and given the excellent scalability of both QMC and GVB, we are confident that this high accuracy approach can be applied to ever larger molecules.

ACKNOWLEDGMENTS

We wish to thank the reviewer who pointed out that some error in our CAS wave functions is likely the result of CI coefficient optimization without orbital optimization. We also wish to thank Dan Fisher for his contributions to the software, as well as useful discussions.

[1] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).
[2] P. Reynolds, D. Ceperley, B. Alder, and W. Lester, J. Chem. Phys. **77**, 5593 (1982).
[3] F. W. Bobrowicz and W. A. Goddard III, in *Methods of Electronic Structure Theory, Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 3, p. 79.
[4] J. P. Hay, W. J. Hunt, and W. A. Goddard III, J. Am. Chem. Soc. **94**, 8293 (1972).
[5] W. A. Goddard III, T. H. Dunning, W. J. Hunt, and J. P. Hay, Acc. Chem. Res. **6**, 368 (1973).
[6] A. G. Anderson, D. Fisher, M. Feldmann, and D. R. Kent, " QMcBEAVER," 2009, http://qmcbeaver.sourceforge.net.
[7] N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 235119 (2004).
[8] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).
[9] R. Assaraf, M. Caffarel, and A. Khelif, Phys. Rev. E **61**, 4566 (2000).
[10] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).
[11] M. D. Brown, J. R. Trail, P. L. Ríos, and R. J. Needs, J. Chem. Phys. **126**, 224110 (2007).
[12] M. W. Schmidt and M. S. Gordon, Annu. Rev. Phys. Chem. **49**, 233 (1998).
[13] J.-M. Langlois, T. Yamasaki, R. P. Muller, and W. A. Goddard III, J. Phys. Chem. **98**, 13498 (1994).
[14] R. B. Murphy, R. A. Friesner, M. N. Ringnalda, and W. A. Goddard III, J. Chem. Phys. **101**, 2986 (1994).
[15] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).
[16] V. A. Rassolov, J. Chem. Phys. **117**, 5978 (2002).
[17] V. A. Rassolov and F. Xu, J. Chem. Phys. **126**, 234112 (2007).
[18] See supplementary material at http://dx.doi.org/10.1063/1.3377091 for some more discussion of our all-electron moves compared with one-electron moves.
[19] M. N. Ringnalda, J.-M. Langlois, R. B. Murphy, B. H. Greeley, C. Cortis, T. V. Russo, B. Marten, R. E. Donnelly, Jr., W. T. Pollard, Y. Cao, R. P. Muller, D. T. Mainz, J. R. Wright, G. H. Miller, W. A. Goddard III, and R. A. Friesner, JAGUAR v7.5, 2008.
[20] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., J. Comput. Chem. **14**, 1347 (1993).
[21] D. Feller, J. Comput. Chem. **17**, 1571 (1996).
[22] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, J. Chem. Inf. Model. **47**, 1045 (2007).
[23] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **117**, 10548 (2002).
[24] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).
[25] P. Piecuch, S. A. Kucharski, K. Kowalski, and M. Musial, Comp. Phys. Commun. **149**, 71 (2002).
[26] Y. Zhao and D. G. Truhlar, Acc. Chem. Res. **41**, 157 (2008).
[27] D. R. Fisher, D. R. Kent IV, M. T. Feldmann, and W. A. Goddard III, J. Comput. Chem. **29**, 2335 (2008).
[28] A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs, J. Chem. Phys. **122**, 224322 (2005).
[29] M. T. Feldmann, J. C. Cummings, D. R. Kent IV, R. P. Muller, and W. A.

Goddard III, J. Comput. Chem. **29**, 8 (2008).
[30] D. R. Kent IV, R. P. Muller, A. G. Anderson, W. A. Goddard III, and M. T. Feldmann, J. Comput. Chem. **28**, 2309 (2007).
[31] A. G. Anderson, W. A. Goddard III, and P. Schröder, Comp. Phys. Commun. **177**, 298 (2007).
[32] W. R. Wadt and W. A. Goddard III, J. Am. Chem. Soc. **96**, 5996 (1974).
[33] W. A. Goddard III, Science **227**, 917 (1985).
[34] P. F. Zittel, G. B. Ellison, S. V. ONeil, E. Herbst, W. C. Lineberger, and W. P. Reinhardt, J. Am. Chem. Soc. **98**, 3731 (1976).
[35] L. B. Harding and W. A. Goddard III, J. Chem. Phys. **67**, 1777 (1977).
[36] P. Jensen and P. R. Bunker, J. Chem. Phys. **89**, 1327 (1988).
[37] J. P. Gu, G. Hirsch, R. J. Buenker, M. Brumm, G. Osmann, P. R. Bunker, and P. Jensen, J. Mol. Struct. **517-518**, 247 (2000).
[38] T. Furtenbacher, G. Czak, B. T. Sutcliffe, A. G. Császár, and V. Szalay, J. Mol. Struct. **780–781**, 283 (2006).
[39] P. M. Zimmerman, J. Toulouse, Z. Zhang, C. B. Musgrave, and C. J. Umrigar, J. Chem. Phys. **131**, 124103 (2009).
[40] A. G. Anderson, "Quantum Monte Carlo: Faster, more reliable, and more accurate," Ph.D. thesis, California Institute of Technology, Pasadena, California, 2010.
[41] W. M. Flicker, O. A. Mosher, and A. Kuppermann, Chem. Phys. Lett. **36**, 56 (1975).
[42] D. F. Evans, J. Chem. Soc. **347**, 1735 (1960).
[43] O. E. El Akramine, A. C. Kollias, and J. W. A. Lester, J. Chem. Phys. **119**, 1483 (2003).
[44] F. Schautz and C. Filippi, J. Chem. Phys. **120**, 10931 (2004).
[45] M. T. Nguyen, M. H. Matus, W. A. Lester, and D. A. Dixon, J. Phys. Chem. A **112**, 2082 (2008).
[46] B. Gemein and S. D. Peyerimhoff, J. Phys. Chem. **100**, 19257 (1996).
[47] S. Jacobi and R. Baer, J. Chem. Phys. **120**, 43 (2004).
[48] W. T. Borden and E. R. Davidson, Acc. Chem. Res. **29**, 67 (1996).
[49] J. M. L. Martin, Chem. Phys. Lett. **303**, 399 (1999).
[50] J. M. Merritt, V. E. Bondybey, and M. C. Heaven, Science **324**, 1548 (2009).
[51] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).
[52] M. Marchi, S. Azadi, M. Casula, and S. Sorella, J. Chem. Phys. **131**, 154116 (2009).
[53] H. Petek, D. J. Nesbitt, D. C. Darwin, P. R. Ogilby, C. B. Moore, and D. A. Ramsay, J. Chem. Phys. **91**, 6566 (1989).
[54] A. Kalemos, T. Dunning, A. Mavridis, and J. Harrison, Can. J. Chem. **82**, 684 (2004).
[55] R. McDiarmid, Adv. Chem. Phys. **110**, 177 (1999).
[56] F. Qi, O. Sorkhabi, and A. G. Suits, J. Chem. Phys. **112**, 10707 (2000).
[57] L. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances*, 4th ed. (Hemisphere, New York, NY, 1989).
[58] K. Frenkel, M. Marsh, R. Wilhoit, G. Kabo, and G. Roganov, *Thermodynamics of organic compounds in the gas state* (Thermodynamics Research Center, College Station, TX, 1994).
[59] D. Feller and D. A. Dixon, J. Chem. Phys. **115**, 3484 (2001).
[60] J. A. W. Harkless and K. K. Irikura, Int. J. Quantum Chem. **106**, 2373 (2006).
[61] R. J. Gdanitz, Chem. Phys. Lett. **312**, 578 (1999).