

# The well-tempered auxiliary-field Monte Carlo

Shlomit Jacobi and Roi Baer<sup>a)</sup>

*Department of Physical Chemistry and the Lise Meitner Minerva-Center for Quantum Chemistry,  
The Hebrew University of Jerusalem, Jerusalem 91904 Israel*

(Received 7 August 2003; accepted 8 October 2003)

The auxiliary-field Monte Carlo (AFMC) is a method for computing ground-state and excited-state energies and other properties of electrons in molecules. For a given basis set, AFMC is an approximation to full-configuration interaction and the accuracy is determined predominantly by an inverse temperature “ $\beta$ ” parameter. A considerable amount of the dynamical correlation energy is recovered even at small values of  $\beta$ . Yet, nondynamical correlation energy is inefficiently treated by AFMC. This is because the statistical error grows with  $\beta$ , warranting increasing amount of Monte Carlo sampling. A recently introduced multi-determinant variant of AFMC is studied, and the method can be tuned by balancing the sizes of the determinantal space and the  $\beta$ -parameter with respect to a predefined target accuracy. The well-tempered AFMC is considerably more efficient than a naïve AFMC. As a welcome “byproduct” low lying excitation energies of the molecule are supplied as well. We demonstrate the principles on dissociating hydrogen molecule and torsion of ethylene where we calculate the (unoptimized) torsional barrier and the vertical singlet-triplet splitting. © 2004 American Institute of Physics. [DOI: 10.1063/1.1630020]

## I. INTRODUCTION

Possessing the detailed structure of excited state molecular potential energy surfaces (PESs) is a prerequisite for understanding and controlling many of their properties. Numerous important biological processes, such as photosynthesis or the vision, involve dynamics on the excited state PESs. Yet, excited state PESs are in general poorly known. The reason is that it is usually much more difficult to compute excited state than the ground state potentials. Several difficulties are encountered in the process. For example, a proper basis set should include also diffuse functions. Moreover, the dense manifold of electronic excited states sometimes severely complicates the problem because they tend to mix when insufficient degree of correlation is used.

The development of methods for excited states is an ongoing effort since the first days of quantum chemistry. An often used method is the multireference double configuration interaction (MRDCI) method,<sup>1</sup> which often yields good excited state energies. Another successful approach is the second-order perturbation theory within a complete active space (CASPT2) method which can deliver in certain cases very accurate description of electronically excited states.<sup>2</sup> This method relies on complete active space self-consistent field (CASSCF) which has a high numerical complexity and thus correlates only a small number of electrons in a restricted space of molecule orbitals. Recently methods relying on time-dependent density functional theory<sup>3</sup> (TDDFT) have been shown to provide useful excitation energies<sup>4</sup> and other observables concerning the potential energy surfaces.<sup>5</sup> While being considerably less sensitive to system size, these methods sometimes deliver highly accurate excitation energies.

Yet, the current versions of TDDFT methods are not general enough. They do describe quantitatively Rydberg states,<sup>6,7</sup> electron-transfer processes,<sup>8</sup> and anions.<sup>7</sup> Thus, present day capability concerning excited molecular state is still an open problem and there is need for new accurate methods which can act as benchmarks for the rapidly developing TDDFT.

Among the multitude of quantum chemistry approaches, Monte Carlo methods hold a special place, delivering the hope of rigorous convergence while alleviating abuse of computer memory and power in large systems.<sup>9–12</sup> The most popular approaches, variational, diffusion, and Green’s function Monte Carlo, are especially useful for the ground state energy calculations. These methods are attractive because they are essentially free of basis set errors. Yet systematic, uncontrollable, errors do come in—through the fixed nodes approximation.<sup>13</sup> Removing or alleviating the fixed node approximation has been shown to be a practical method only for small number of electrons (less than five). Even for three electrons there are still unresolved claims that biases exist in the Green’s function Monte Carlo, when the fixed node approximation is not used.<sup>14</sup> The fixed nodes also prevent correlated sampling of, for example, forces, and special methods for alleviating this need to be applied.<sup>15</sup>

This article concentrates on a relatively new Monte Carlo method for electronic structure, the auxiliary field Monte Carlo (AFMC). The great benefit of the method is that it is essentially free of fixed node approximations. AFMC was developed and used in several fields of physics in pioneering work by several authors.<sup>16–22</sup> Application of AFMC to molecular electronic structure problems was unsuccessful due to very large statistical fluctuations. This problem essentially prevented progress until the development of the shifted contour AFMC by Rom and Neuhauser.<sup>23</sup> Analysis of the method and demonstration of its utility has been published in several studies.<sup>11,24–27</sup> As mentioned above, the greatest ad-

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: roi.baer@huji.ac.il

vantage of the method, over other Monte Carlo approaches to electronic structure, is the *absence* of the uncontrolled fixed node approximation. This makes the method more expensive in individual energy calculations. However, there is an added value for the increased price: the possibility to perform correlated sampling. This has been established in several studies: calculation of spectroscopic constants,<sup>28</sup> calculation of deformation barriers,<sup>29</sup> and determination of singlet-triplet splittings.<sup>30</sup>

As an alternative to “conventional” electronic structure methods, Monte Carlo techniques have had a hard time to compete with less rigorous but deterministic methods, such as the truncated coupled-cluster approaches. The primary reason for this is that the deterministic methods enjoy fortuitous error cancellations. This is noticeable when deterministic methods are pushed to their rigorous limits.<sup>31</sup> Another reason is the *fixed node approximation*, which despite some impressive achievements<sup>32–36</sup> is limited in going beyond the single-energy calculations.

This present article discusses the relation of various sources of errors in AFMC: the finite time step, the statistical error (SE), and the finite  $\beta$  parameter (to be defined in Sec. II). We also discuss the dependence of these errors on the size of a multi-determinantal space in which AFMC is performed. One important conclusion is that it is important to balance the method by a suitable tuning of the various convergence parameters—arriving at a well-tempered AFMC. Another conclusion is that a suitable choice of the multi-determinantal space allows the full CI quality computation of excited state surfaces, within the given basis set. We demonstrate these points on the potential surfaces of hydrogen and ethylene.

## II. AFMC THEORY AND APPLICATION

### A. Review of AFMC theory

The basic utility of AFMC is its practically unrivalled ability to compute the following matrix elements related to any two determinantal wave functions  $\Phi_i$  and  $\Phi_j$ :

$$S_{\mu\nu} = \langle \Phi_\mu | e^{-\beta\hat{H}} | \Phi_\nu \rangle, \quad \mathbf{H}_{\mu\nu} = \langle \Phi_\mu | \hat{H} e^{-\beta\hat{H}} | \Phi_\nu \rangle, \quad (2.1)$$

where  $\hat{H}$  is the full Hamiltonian (within a given one-electron basis set) and  $\Phi_\mu$  and  $\Phi_\nu$  ( $\mu, \nu = 1, \dots, K$ ) are  $K$  arbitrary  $N_e$  electron determinantal wave functions containing  $N_e$  orbitals spanned by the basis set. These orbitals are usually obtained from Hartree–Fock calculation. Yet they are not limited to this choice and multi-configuration self-consistent-field theory, natural orbitals from various choice of CI or even density functional orbitals.

The basic idea of AFMC is the use of a well-known integral transform ( $V > 0$ ):

$$e^{-(1/2)V\rho^2} = \sqrt{\frac{V}{2\pi}} \int e^{iV\sigma\rho} e^{-(1/2)V\sigma^2} d\sigma \quad (2.2)$$

for converting the Boltzmann operator  $e^{-\Delta\beta\hat{H}}$  where  $\hat{H}$  includes two-electron interactions into a weighted sum of Boltzmann operators describing noninteracting electrons. The integration variable  $\sigma$  of Eq. (2.2) is called an “auxiliary

density.”<sup>37</sup> On the left-hand side the exponent includes a quadratic term in  $\rho$  while on the right-hand side the exponent includes a linear term. The connection to electronic structure, is evident:  $\rho$  is taken not as a number but as a one-particle operator  $\hat{\rho}_i \equiv \hat{\rho}_{ij}$  (see Appendix A for a review of notation). We then have

$$e^{-(1/2)\rho^T V \rho \Delta\beta} = \sqrt{\det\left(V \frac{\Delta\beta}{2\pi}\right)} \int e^{i\sigma^T V \hat{\rho} \Delta\beta} e^{-(1/2)\sigma^T V \sigma \Delta\beta} dN^2 \sigma. \quad (2.3)$$

Here we used the fact that  $V$ , the matrix of two-electron integrals, is positive definite. This identity is called the Hubbard–Stratonovich transformation. In Appendices A and B we review the basic theory, showing that the matrix element of a Boltzmann operator  $e^{-\beta\hat{H}}$  can be written as a functional average [Eq. (B1)], while the matrix elements of  $\hat{H} e^{-\beta\hat{H}}$  are analogously given by

$$\langle \Psi_i | \hat{H} e^{-\hat{H}\Delta\beta} | \Psi_j \rangle = N' \times \langle \langle \Psi_i | \hat{H}_{\hat{\rho}} e^{iJ(\sigma(\beta) - i\alpha(\beta))^T V \hat{\rho} \Delta\beta} | \Psi_j \rangle \rangle_{\sigma(\beta, V\Delta\beta)}. \quad (2.4)$$

### B. Multi-determinantal AFMC

Having an in-principle accurate method for calculating the quantities of Eq. (2.1), for a given  $\beta$ , we discuss how to use this tool, namely how the  $M$  lowest eigenenergies of  $\hat{H}$  (of given symmetry and spin) can be calculated. We assert that these are accessible by solving the generalized  $K \times K$  ( $K \geq M$ ) eigenvalue problem:

$$\mathbf{HC} = \mathbf{SCE}. \quad (2.5)$$

All matrices in Eq. (2.5) are  $K \times K$  and  $\mathbf{E}$  is diagonal, holding the approximate eigenenergies. The utility in working with a large  $K$  will become evident in the sections to come. We profit twice: increasing the efficiency of the ground state calculates and obtaining excited state information.

### C. Computer program—GSAFMC

The computational application of AFMC starts where a conventional self-consistent field (SCF) electronic structure treatment ends. We have therefore written a computer program, GSAFMC, which performs an AFMC calculation based on an output file of a GAMESS(US) program. The basic algorithm of the program is quite straightforward and was given in Ref. 11. We discuss here a few methodological details.

To facilitate the generation of Gaussian distributed auxiliary densities, we perform a Cholesky factorization of the two-electron-integral matrix  $V$ :

$$V = LL^T, \quad (2.6)$$

where  $L_{IK}$  is a lower triangular matrix in the  $I, J$  indices (please refer to Appendix A for explanation of these notations). Since the matrix  $V$  is in practice usually not positive definite due to numerical round-offs, care must be exercised when performing this factorization. In order to derive an expression with  $L$  instead of  $V$  we change the variable of inte-

gration from  $\sigma$  to  $L\sigma$ . Thus, the rotated Gaussian weight is transformed to an efficacious Gaussian weight:

$$e^{-(1/2)\sigma^T V \sigma \Delta \beta} \rightarrow e^{-(1/2)\sigma^T \sigma \Delta \beta}. \quad (2.7)$$

The operator, including the contour shift, becomes

$$e^{i(\sigma-i\alpha)^T V \hat{\rho} \Delta \beta} \rightarrow e^{i(\sigma-i\bar{\alpha})^T L^T \hat{\rho} \Delta \beta}, \quad (2.8)$$

where  $\bar{\alpha} = L^T \alpha$ . The importance of this procedure is that it replaces expensive diagonalization by a more numerically efficacious Cholesky factorization.

Another effective method is related to the expensive procedure of the two-electron repulsion energy calculation  $\langle V \rangle = \langle \Phi_L | V | \Phi_R \rangle$ , where  $\Phi_L$  is a constant matrix and  $\Phi_R$  changes at each Monte Carlo (MC) iteration. A naive calculation necessitates the use of the full  $V_{ijkl}$  matrix is numerically expensive. Utilizing the fact that  $\Phi_L$  is always the same determinant, we can save considerable memory and speed by the following scheme. We first write

$$\langle V \rangle = \sum_{\alpha\beta}^{N_e} \sum_{ijkl}^N L_{i\alpha} R_{j\alpha} (V_{ijkl} - V_{ilkj}) L_{k\beta} R_{l\beta}, \quad (2.9)$$

where the  $L_{i\alpha}$  ( $R_{i\alpha}$ ) denote the coefficient of the  $i$ th MO in the  $\alpha$ th orbital of the determinant  $\Phi_L$  ( $\Phi_R$ ). Defining  $W_{jl\alpha\beta} = \sum_{ik}^{N_b} L_{i\alpha} (V_{ijkl} - V_{ilkj}) L_{k\beta}$ , a quantity which can be calculated just once before the MC iterations begin, we find

$$\langle V \rangle = \sum_{\alpha\beta}^{N_e} \sum_{jl}^N R_{j\alpha} W_{lj\alpha\beta} R_{l\beta}. \quad (2.10)$$

Using this equation in place of Eq. (2.9) during the multitude of MC iterations is numerically efficacious.

Finally, we comment on the use of memory. When the memory needed by the  $L$ =matrix is prohibitive, it is possible to store it on disk and structure the loops of AFMC in such a way as to minimize the number of reads needed simply by performing parallel sampling, i.e., sampling several  $\sigma$ 's at once. This makes the penalty of disk access negligible.

#### D. Estimation of the "statistical error"

Next, we discuss the estimation of the SE. The procedure is based on statistical theory and we briefly review it for our purposes here. Each MC iteration yields a results  $\alpha_i$  ( $i = 1, \dots, I$ ) which is a sample of the random variable  $A$ . This random variable is distributed in an unknown fashion with an expectation value  $\langle A \rangle$  and variance  $\Sigma(A)^2$ . We want to use the samples  $\alpha_i$  to estimate  $\langle A \rangle$  and we want determine the accuracy of this estimate (the SE). We first group the results to a small number ( $J$ ) of groups  $I_j$ ,  $j = 1, \dots, J$ , of  $N_j = I/J$  members. We then define the mean of  $A$  for each group:

$$\bar{\alpha}_j = \frac{1}{N_j} \sum_{i \in I_j} \alpha_i. \quad (2.11)$$

These estimates can be viewed as samples of a new random variable  $\bar{A}$ , defined as

$$\bar{A} = \frac{1}{N_j} \sum_{j=1}^J A_j, \quad (2.12)$$

where  $A_j$  are random variables which are distributed identically to  $A$ . The central limit theorem asserts that (in the limit of large  $N_j$ ) the random variable  $\bar{A}$  is distributed according to a Gaussian distribution, with the same expectation value as that of  $A$ ,  $\langle \bar{A} \rangle = \langle A \rangle$  and variance which is smaller by a factor  $\sqrt{N_j}$ :  $\Sigma(\bar{A})^2 = \Sigma(A)^2 / \sqrt{N_j}$ .

We have  $J$  samples  $\bar{\alpha}_j$ ,  $j = 1, \dots, J$ , from this distribution and we use them to obtain a  $\gamma$ -confidence interval for  $\langle A \rangle$ . (A  $\gamma$ -confidence interval is a smallest interval  $[a, b]$  for which there is  $\gamma$  probability that  $\langle A \rangle$  is in  $[a, b]$ .) According to standard statistical theory this interval is  $[\alpha - SE, \alpha + SE]$  where

$$\bar{\alpha} = \frac{1}{J} \sum_{j=1}^J \bar{\alpha}_j \approx \langle A \rangle \quad (2.13)$$

and

$$SE = t_{(J-1), (1-\gamma)/2} \times \sqrt{\frac{1}{J(J-1)} \sum_{j=1}^J (\alpha_j - \bar{\alpha})^2}, \quad (2.14)$$

where  $t_{df,p}$  is the Student  $t$  distribution. For a confidence interval of  $\gamma = 90\%$  and  $J = 5$ ,  $t = 2.13$ , while for a confidence interval of  $\gamma = 95\%$ ,  $t = 2.77$ . Equations (2.11), (2.13), and (2.14) are the working equations. We use  $J = 5$  and confidence  $\gamma = 90\%$  in all calculations of this article.

### III. RESULTS ON MODEL SYSTEMS

In this section we report a few computational results that demonstrate the performance of multi-determinantal AFMC. We study two systems: the hydrogen molecule bond stretch and the ethene torsional motion.

#### A. Ground and excited states of H<sub>2</sub> bond stretching

We consider here the ground and valence excited states of H<sub>2</sub> described within the cc-pVDZ basis. This basis facilitates full CI calculation, so the AFMC accuracy can be assessed within this basis. The accuracy in absolute terms is not large due to the limitations of the basis set. We use a  $K = 6$  determinantal space as follows:  $|1\sigma_g^2|$ ,  $|1\sigma_g^1 1\sigma_u^1|$ ,  $|1\sigma_g^1 1\sigma_u^1|$ ,  $|1\sigma_g^1 2\sigma_g^1|$ ,  $|1\sigma_g^1 2\sigma_g^1|$ , and  $|1\sigma_u^2|$ . The orbitals are taken from a Hartree-Fock treatment. These determinants span a six-dimensional space which is composed of 4D singlet+2D triplet subspaces. The corresponding AMFC estimates for the energies of the eigenstates converge towards the full CI limit as  $\beta$  increases. This is shown in Fig. 1, where it is seen that the states are approached from above as  $\beta$  grows from 0 to 5 au. For the low energy manifold the multi-determinantal set already captures the main features of the PES, and AFMC nicely corrects it to quantitative fit. For the higher excitation energies AFMC is more effective as the multi-determinantal fares less well. However, at some regions, a quantitative fit requires a much larger  $\beta$  than we use. It is difficult to enlarge  $\beta$  because the SE increases in response (see the next section for a full discussion on this). The last singlet state is an extreme example. The reason that this last state is not well converged is because other singlet states of similar energies exist and have a near-degeneracy effect. These are not described in the small six-dimensional space

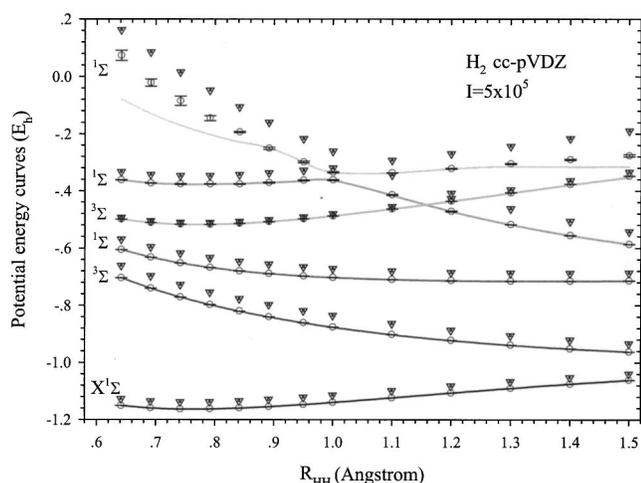


FIG. 1. Low lying excited state potential energy curves for  $H_2$  in the cc-pVDZ basis set. The lines are full-CI results; large circles are the  $K=6$  result (no AFMC); small circles with statistical error bars are  $K=6$  AFMC ( $\beta=5$  au).

we chose. If the higher energy states are required, the determinantal manifold must be enlarged for an effective treatment.

A closer look for the low-lying states reveals that AFMC is extremely effective at including the dynamical correlation which is missing from the 6D determinantal space. In Fig. 2, for example, the CI misses dynamic correlation energy of almost 1 eV and AFMC restores it very nicely.

A close look at the higher excited states (Fig. 3) reveals again that AFMC in combination with the determinantal space is extremely effective. Indeed, even the approach to the singlet-singlet avoided crossing is very effectively corrected by AFMC.

These results show that the combination of AFMC with a determinantal space that captures the salient features of nondynamical correlation is an effective method for computing both ground and excited molecular states.

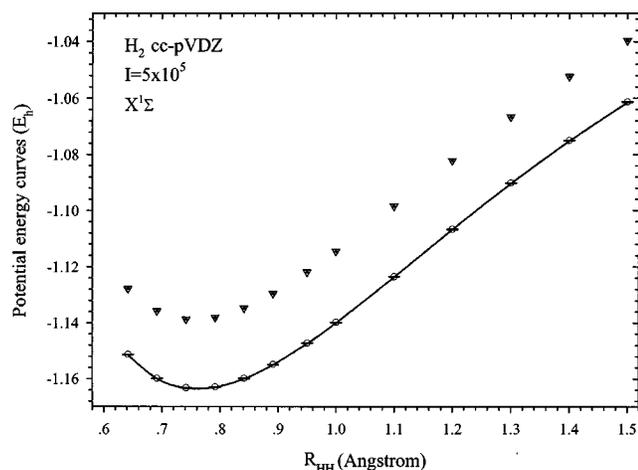


FIG. 2. Close-up on the ground state potential curve of Fig. 1.

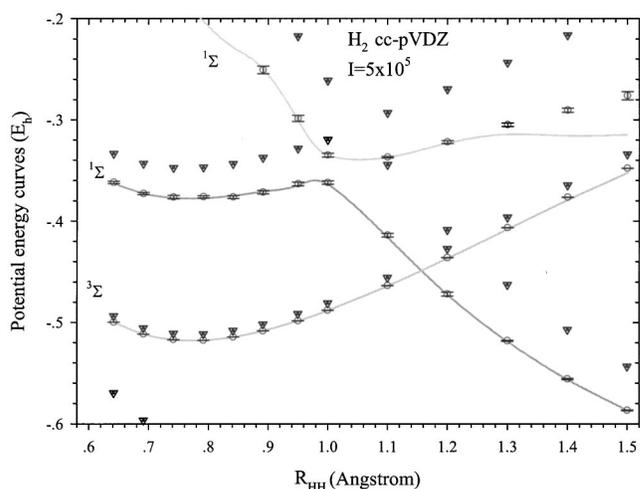


FIG. 3. Close-up on a portion of the states depicted in Fig. 1.

## B. The ethylene excited states and torsional barrier

The ethylene molecule at its ground state is planar so that the two  $CH_2$  groups, connected by a  $C=C$  double bond, share the same plane. We consider the torsional motion where one  $CH_2$  group rotates around the  $C-C$  bond. This motion is described by a torsional angle  $\theta$  which has a value  $\theta=0$  in the planar configuration ( $N$ ) and  $\theta=90^\circ$  in the fully vertical configuration ( $V$ ). During the torsional motion we freeze the bond-lengths  $R_{CC}=1.334$  Å and  $R_{CH}=1.081$  Å and the HCH angle is  $\alpha_{HCH}=117.4^\circ$ .

In the planar (stable) configuration, the HOMO orbital of this molecule is a sum of the two atomic orbitals (AOs) of character  $p$  on each carbon atom which are perpendicular to the molecular plane. We name this orbital, rather loosely, the  $\pi$  orbital. Accordingly, the LUMO is the difference between these AOs and we name it, for obvious reasons, the  $\pi^*$  orbital. Because the overlap between the two AOs decreases monotonically with growing  $\theta$ , the  $\pi-\pi^*$  orbital energy difference decreases as well. Both reach zero at the vertical ( $\theta=90^\circ$ ) configuration where their axes are perpendicular. This degeneracy in fact causes a breakdown of the Hartree-Fock approximation and a proper description of the torsional motion usually needs to be described by a multi-reference method that take into account nondynamical correlation.<sup>1,38</sup>

In the present calculation, we freeze four core electrons in their restricted Hartree-Fock core orbitals (mainly carbon  $1s$ ). Thus we correlate only the 12 valence electrons. We work in a  $K=4$  space consisting of excitations from the HOMO " $\pi$ " orbital to the LUMO " $\pi^*$ " orbital. The four determinants composed of  $|\cdots\pi^2\rangle$ ,  $|\cdots\pi\pi^*\rangle$ ,  $|\cdots\bar{\pi}\pi^*\rangle$ , and  $|\cdots(\pi^*)^2\rangle$  form the determinantal space. The orbitals are Hartree-Fock orbitals.

In Fig. 4 we show the ground state curves as  $\beta$  increases from 0 [restricted Hartree-Fock (RHF)] to 4.5 au. The nondynamical correlation, allowed for by using the four determinantal basis specified above, is clearly not enough to obtain an even qualitatively correct PES. The AFMC iterations damp the unphysical feature at the vertical configuration. The computed barrier we obtain (without relaxation of the barrier geometry) is 3.2 eV.

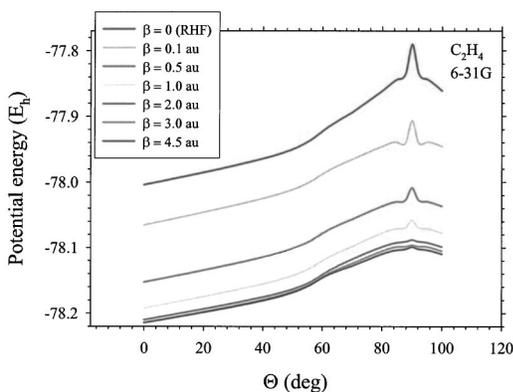


FIG. 4. The potential energy curve, as a function of the torsion angle  $\Theta$  as computed by Hartree–Fock approximation and by AFMC at increasing values of  $\beta$ .

A quantity of experimental interest is the singlet–triplet vertical excitation of ethylene (Fig. 5).<sup>39–41</sup> Our results, in 6-31G basis set, is  $4.65 \pm 0.1$  eV, virtually identical to the full CI estimation of Gemein *et al.*,<sup>1</sup> using multi-reference configuration interaction method. Similar results of 4.4 eV were obtained by Serrano-Andres *et al.*,<sup>2</sup> using CASPT2 and a much larger basis set. The results are also in agreement with the recent QMC calculation of El Akramine *et al.*,<sup>34</sup> which obtained 4.5 eV. All these calculated results are within the experimental error bars of electron impact spectroscopic measurements of van Veen.<sup>39</sup> Yet, with exclusion of Ref. 2, they are some  $\sim 5\%$ – $10\%$  above the value inferred from electron energy loss spectroscopy.<sup>40,41</sup>

#### IV. WELL-TEMPERED AFMC

For any given basis set and value of  $\beta$ , the quantities in Eq. (2.1) can be calculated by AFMC to within two controlled approximations:<sup>42,43</sup> the finite step-size  $\Delta\beta$  (where  $L$  is the number of  $\beta$  steps) and the statistical error (SE). We depict the sources of errors in Fig. 6. It is seen that the finite step-size error is always positive, i.e., it systematically converges from above. It is a controlled approximation (compu-

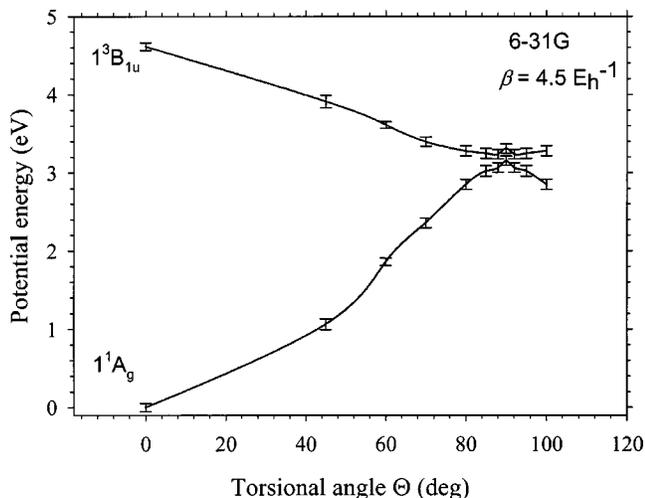


FIG. 5. The singlet and triplet states of ethylene calculated within the 6-31G basis set, using AFMC, at  $\beta = 4.5E_h^{-1}$ . Number of MC iterations: 50 000.

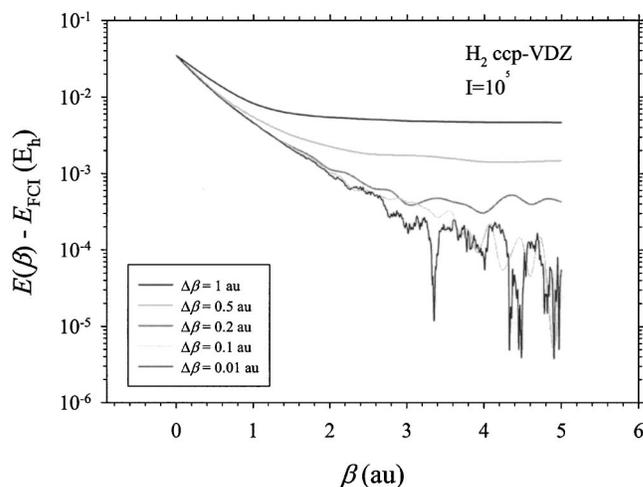


FIG. 6. Three sources of errors in AFMC, shown for a calculation of  $H_2$ . The finite step-size error converges as  $\Delta\beta \rightarrow 0$ . The finite  $\beta$  error diminishes as  $\beta \rightarrow \infty$ . The size of the statistical error (SE) is approximately that of the fluctuations. The SE at a given  $\beta$  is determined by the number of MC iterations  $I$ . For efficient applications a balance must be made between the three sources.

tational work is linear with  $L$ ). Formally, the finite step-size error converges no slower than  $O(\Delta\beta)$  (see Appendix A). In actual applications the observed convergence is much faster than this. The SE is proportional to  $I^{-1/2}$  where  $I$  is the number of iterations, so it too can be reduced indefinitely simply by more sampling.

What happens when  $\Delta\beta$  decreases? Does that cause the statistical error to grow? We study this in Fig. 7 where the SE is plotted as a function of  $\beta$  and  $\Delta\beta$ . Essentially, it is seen that the SE is independent of  $\Delta\beta$ .<sup>44</sup>

Consider now two extreme situations of applying the scheme depicted in Eqs. (2.1) and (2.5). It is well known that when  $\beta=0$ , taking  $\{\Phi_{\mu}\}_{\mu=1}^K$  to contain the largest possible set of linearly independent determinants (i.e., the “full-CI” limit) will yield the essentially exact eigenenergies within a give basis set. However, this is not a practical solution for any but the smallest systems. Consider now the other ex-

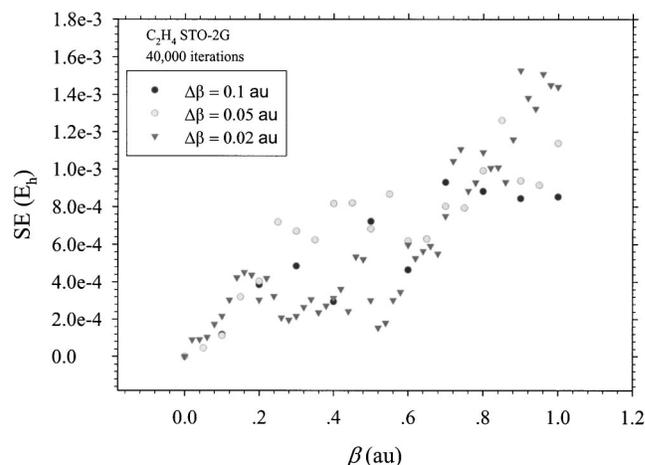


FIG. 7. SE as a function of  $\beta$  for several values of step size  $\Delta\beta$ . The SE is independent of  $\Delta\beta$  and grows as  $\beta$ .

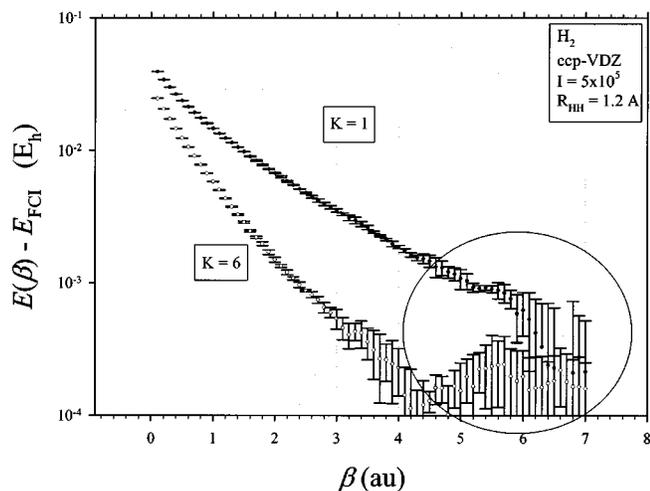


FIG. 8. The AFMC error in total energy in a stretched  $H_2$  molecule and the relative statistical error bars. Two calculations are shown, one with  $K=1$  and  $K=6$ . The ellipse indicates the region where the SE dominates.

treme: take the minimal number of determinants,  $K=M$ , and use an extremely large value of  $\beta$ . This procedure too will yield a formal solution to the problem, since the desired energies are obtained in the limit  $\beta \rightarrow \infty$ . Yet, this solution is impractical: the SE grows exponentially with  $\beta$ , as clearly seen in Fig. 7.

We shall see now that the finite  $\beta$  error is considerably reduced by taking a multi-determinantal approach. Of course, such an approach does not change the exponential scaling, since for a systematic decrease of the error, an exponential increase of the dimension  $K$  of the determinantal space must follow. Yet, for a given target accuracy there is an optimal balance between size of the determinantal space and the value of  $\beta$ . We conclude that the scheme of Eqs. (2.1) and (2.5) has a certain flexibility that enables to balance the two extremes. For any given system, a good choice of  $K$ ,  $\{\Phi_{\mu}^K\}_{\mu=1}^K$ , and  $\beta$  will yield an accurate and efficient electronic structure method.

We demonstrate this point on a simple example, the  $H_2$  molecule at stretched bond length (Fig. 8). The picture nicely depicts two sources of AFMC errors—the deterministic finite- $\beta$  error and the statistical error (SE). Two scaling cases are considered: when  $K=1$  and  $K=6$ . In both cases, an increase of  $\beta$  steadily reduces the deterministic error. Yet the statistical error grows with  $\beta$  as well. Eventually, at  $\beta = 4.5$  au the deterministic error is sufficiently reduced and the SE is sufficiently blown up for a transition to occur: the total error becomes dominated by the SE. There is no point in increasing  $\beta$  at this stage. If better accuracy is required, the increase of  $\beta$  must be accompanied in an increase in the number of iterations  $I$ , for reducing the SE.

A well-tempered AFMC is obtained when a correct choice of all convergence parameters is chosen. We show in Table I an example of how the choice of parameters as a function of target accuracy and dimension of determinantal space  $K$  affects the numerical work. The superiority of the  $K=6$  method for high accuracies is evident, as well as the large increase in computational work necessary for better accuracy. Because AFMC allows correlated sampling, a tar-

TABLE I. Computational work as a function of the convergence parameters  $K$ ,  $L$  and  $\Delta\beta$ . Results are for  $H_2$  at  $R=1.2A$ .

Target accuracy in absolute energy ( $E_h$ )	$K$	$\Delta\beta$ (au)	$\beta$ (au)	Iterations needed $I$	Relative numerical work
					$\left(K \times \frac{\beta}{\Delta\beta} \times I\right)$
0.01	1	0.5	1.5	100	300
	6	0.5	0.5	100	600
0.001	1	0.2	5	$5 \times 10^5$	$4 \times 10^6$
	6	0.2	2.3	$1 \times 10^3$	$3 \times 10^4$
0.0001	1	0.1	10	$5 \times 10^9$ (est.)	$5.10^{11}$
	6	0.1	4	$5 \times 10^5$	$1 \times 10^8$

get accuracy of 0.001  $E_h$  in the absolute energy is often enough to ensure chemical accuracy in properties (which involve energy differences).

## V. CONCLUSIONS

In this article we have examined multi-determinant AFMC and analyzed the various convergence parameters of a typical AFMC calculation. We have shown that for a given  $\beta$  parameter the time-step error and the statistical errors are controlled. The only uncontrolled parameter is thus the  $\beta$ -parameter, which should be in principle taken to infinity. This limit cannot be achieved in actual applications since it causes exponential rise in the SE. The need for large  $\beta$  is only a result of low-lying excited states. These, however, can be dealt with using a multi-determinant approach. Thus, the combination of AFMC and multi-determinantal AFMC, including well-tempered tuning of all the convergence parameters, is a requisite for good performance of the method.

Indeed, an important conclusion of the article is that correct tuning can lead to huge reduction in computational effort. We have also applied the method, using correlated sampling to accurately compute the ground and excited state *potential energy curves* of hydrogen and ethylene within a given basis set. In the latter we have shown that the method can very effectively reduce the cusp at vertical geometry which appears in Hartree–Fock and low-order determinantal expansions. We have also calculated in a correlated manner both the singlet and triple potential curves of the torsional motion.

AFMC is a general method for including correlation and, with a correct choice of determinantal space, it can be made very efficient, too. It can build into a determinantal basis the dynamical correlation which it misses. For a given accuracy target the overall scaling is gentle [ $O(N^4)$  where  $N$  is the number of electrons or basis functions even when no linear scaling methods are used<sup>45</sup>]. However, the prefactors are still quite large and, in the future, a large effort must be invested for reducing the prefactors.

## ACKNOWLEDGMENTS

We gratefully acknowledge the Israel Science Foundation established by the Israel Academy of Sciences and Humanities for their support of this research.

### APPENDIX A: DETAILS OF AFMC THEORY

Why is this useful for electronic structure? For answering this, let us write the electronic Hamiltonian in second quantized form:

$$\hat{H} = \sum_{i,j,s} T_{ij} \hat{c}_{is}^\dagger \hat{c}_{js} + \frac{1}{2} \sum_{ijkl,ss'} V_{ijkl} \hat{c}_{is}^\dagger \hat{c}_{ks}^\dagger \hat{c}_{ls'} \hat{c}_{js}, \quad (\text{A1})$$

where  $\hat{c}_{i,s}^\dagger$  ( $\hat{c}_{i,s}$ ) is a creation (annihilation) operator of an electron of spin  $s$  in molecular orbital  $\phi_i(\mathbf{r})$ ,  $i=1,\dots,N$ :

$$T_{ij} = \int \phi_i^*(\mathbf{r}) [-\nabla^2/2 + v_{\text{nuc}}(\mathbf{r})] \phi_j(\mathbf{r}) d^3r \quad (\text{A2})$$

and

$$V_{ijkl} = \int \int \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) \times v(|\mathbf{r}-\mathbf{r}'|) \phi_k^*(\mathbf{r}') \phi_l(\mathbf{r}') dr dr'. \quad (\text{A3})$$

Atomic units are used throughout. In Eq. (A2)  $v_{\text{nuc}}(\mathbf{r})$  is the sum over all nuclei of the Coulombic attraction potentials and in Eq. (A3)  $v(r) = e^2/r$  is the Coulomb repulsion between electrons. The Hamiltonian of Eq. (A1) can be written as

$$\hat{H} = \sum_{ij} t_{ij} \hat{\rho}_{ij} + \frac{1}{2} \sum \hat{\rho}_{ij} V_{ijkl} \hat{\rho}_{kl}, \quad (\text{A4})$$

where  $\hat{\rho}_{ij} = \sum_{s=\uparrow\downarrow} \hat{c}_{is}^\dagger \hat{c}_{js}$  is the single-particle-density-matrix operator and  $t_{ij} = T_{ij} - \frac{1}{2} \sum_l V_{illj}$ . It is worthwhile to introduce a shorthand notation  $I=(ij)$  ( $I$  is an index of  $N^2$  components), with which the Hamiltonian of Eq. (A4) is written as

$$\hat{H} = t^T \hat{\rho} + \frac{1}{2} \hat{\rho}^T V \hat{\rho}. \quad (\text{A5})$$

We use here the notation

$$A^T B = \sum I A_I B_I. \quad (\text{A6})$$

The real matrix  $V_{IK}$  defined in Eq. (A1) is in fact ‘‘positive definite,’’ thus for any real  $M$ -vector  $\sigma$ ,

$$\sigma^T V \sigma > 0. \quad (\text{A7})$$

We now consider the practical evaluation Boltzmann operator  $e^{-\beta \hat{H}}$ . As usual with all quantum path integral methods, the first step is to break up to a product of  $L$  time steps:

$$e^{-\beta \hat{H}} = e^{-\Delta \beta \hat{H}} \dots e^{-\Delta \beta \hat{H}} \quad (\text{A8})$$

with  $\Delta \beta = \beta/L$ . Let us now consider a single step operating with  $e^{-\Delta \beta \hat{H}}$  on a determinantal wavefunction  $\Phi = c_{i_1}^\dagger \dots c_{i_{N_e}}^\dagger |0\rangle$  (we suppress the spin indices for clarity). For this, let us further approximate

$$e^{-\Delta \beta \hat{H}} \Phi \approx e^{-(1/2) \rho^T V \hat{\rho} \Delta \beta} e^{-\Delta \beta t^T \hat{\rho}} \Phi. \quad (\text{A9})$$

The approximation becomes exact as fast as  $\Delta \beta^2$  goes to zero [since there are  $L$  steps, the total error this introduces is actually  $O(\Delta \beta)$ ]. The first operator is an exponential of a one-electron operator, and is easily applied as we now show. We first Taylor expand the expression  $e^{-s \hat{y} \hat{x}} e^{s \hat{y}}$ , obtaining

$$e^{-s \hat{y} \hat{x}} e^{s \hat{y}} = \hat{x} + [\hat{y}, \hat{x}] s + \frac{1}{2!} [\hat{y}, [\hat{y}, \hat{x}]] s^2 + \frac{1}{3!} [\hat{y}, [\hat{y}, [\hat{y}, \hat{x}]]] s^3 + \dots \quad (\text{A10})$$

We want to apply Eq. (A10) for the case  $\hat{y} = \Delta \beta t^T \hat{\rho}$  and  $\hat{x} = \hat{c}_i^\dagger$ . For this, we use the commutation relation

$$[\hat{c}_j^\dagger \hat{c}_k, \hat{c}_i^\dagger] = \delta_{ik} \hat{c}_j^\dagger, \quad (\text{A11})$$

from which

$$[t^T \hat{\rho}, \hat{c}_i^\dagger] = \sum_j t_{ij} \hat{c}_j^\dagger, \quad (\text{A12})$$

$$[t^T \hat{\rho}, [t^T \hat{\rho}, \hat{c}_i^\dagger]] = \sum_j (t^2)_{ij} \hat{c}_j^\dagger,$$

[where the definition  $(t^2)_{ij} = \sum_k t_{ik} t_{kj}$  is used]. Equations (A10)–(A12) show that

$$e^{-\Delta \beta t^T \hat{\rho}} \hat{c}_i^\dagger e^{\Delta \beta t^T \hat{\rho}} = \sum_j (e^{-\Delta \beta t^2})_{ij} \hat{c}_j^\dagger \equiv \hat{d}_i^\dagger(\Delta \beta). \quad (\text{A13})$$

It is evident that  $\hat{d}_i^\dagger$  simply creates an electron in the molecular orbital

$$\psi_i(\mathbf{r}) = \sum_j (e^{-\Delta \beta t^2})_{ij} \phi_j(\mathbf{r}), \quad (\text{A14})$$

Thus, the operation of  $e^{-\Delta \beta t^T \hat{\rho}}$  on a determinant is again a determinant, of new molecular orbitals, computed by Eq. (A14):

$$e^{-\Delta \beta t^T \hat{\rho}} |\Phi\rangle = e^{-\Delta \beta t^T \hat{\rho}} \hat{c}_{i_1}^\dagger \dots \hat{c}_{i_{N_e}}^\dagger |0\rangle = \hat{d}_{i_1}^\dagger \dots \hat{d}_{i_{N_e}}^\dagger |0\rangle = |\Psi(\Delta \beta)\rangle. \quad (\text{A15})$$

The next problem is the operation of  $e^{-\Delta \beta \hat{\rho}^T V \hat{\rho}}$  on the determinant  $\Psi(\Delta \beta)$ . Here, in general, the result is a long and unyielding sum of determinants. This sum can be written as a functional integral, using Eq. (2.3):

$$e^{-\Delta \beta (1/2) \hat{\rho}^T V \hat{\rho}} |\Psi\rangle = \sqrt{\det \left[ \frac{V \Delta \beta}{2\pi} \right]} \times \int e^{-(1/2) \sigma^T V \sigma \Delta \beta} e^{i \sigma^T V \hat{\rho} \Delta \beta} |\Psi\rangle d^{N^2} \sigma. \quad (\text{A16})$$

The operation on the determinant is now a sum of operation of single-particle evolutions with an auxiliary electron density  $\sigma$ . The integral is a weighted sum over all possible densities  $\sigma_{ij}$  and the weight is Gaussian  $e^{-(1/2) \sigma^T V \sigma \Delta \beta}$ . Thus, we may write this scheme as a weighted average, taking into account all the segments and, closing from the left with an arbitrary determinant, we have

$$\langle \Theta | e^{-\hat{H} \Delta \beta} | \Psi \rangle = \mathbb{N} \langle \langle \Theta | \hat{\rho} e^{i \int \sigma^T (\beta) V \hat{\rho} d\beta} | \Psi \rangle \rangle_{\sigma(\beta, V \Delta \beta)}, \quad (\text{A17})$$

where  $\hat{\rho}$  is the path ordering operator,  $\mathbb{N} = (1/2) \int \text{Tr} \ln[(V(\beta))d\beta/2\pi]$  (in obvious notation) and the average is weighted by  $e^{-(1/2) \int \sigma^T(\beta) V \sigma(\beta) d\beta}$  and normalized accordingly.

## APPENDIX B: THE CONTOUR SHIFT

While formally correct, the expression in Eq. (A17) is not very useful because any application suffers from large random noises. This is explained in Ref. 24, where the solution, developed by Rom and Neuhauser,<sup>23</sup> is described in detail. The result of the analysis is that for any “density shift”  $\alpha_{ij}(\beta)$ , the following modified AFMC formula is equivalent to (A17):

$$\langle \Theta | e^{-\hat{H}\Delta\beta} | \Psi \rangle = \mathbb{N}' \langle \langle \Theta | \hat{\rho} e^{i \int (\sigma(\beta) - i\alpha(\beta))^T V \hat{\rho} d\beta} | \Psi \rangle \rangle_{\sigma(\beta, V\Delta\beta)}, \quad (\text{B1})$$

where  $\mathbb{N}' = (1/2) \int [ \text{Tr} \ln(V(\beta)d\beta/2\pi) + \alpha^T V \alpha d\beta ]$ . In Ref. 24 it is shown why taking the mean field density reduces the statistical fluctuations. We find that often taking a  $\beta$ -dependent shift  $\alpha_{ij} = \langle \Psi | \hat{\rho}_{ij} | \Psi \rangle$  is very effective.

- <sup>1</sup>B. Gemein and S. D. Peyerimhoff, *J. Phys. Chem.* **100**(50), 19257 (1996).
- <sup>2</sup>L. Serrano-Andres, M. Merchán, I. Nebotgil, R. Lindh, and B. O. Roos, *J. Chem. Phys.* **98**, 3151 (1993).
- <sup>3</sup>M. Petersilka, U. J. Gossmann, and E. K. U. Gross, *Phys. Rev. Lett.* **76**(8), 1212 (1996).
- <sup>4</sup>C. Jamorski, M. E. Casida, and D. R. Salahub, *J. Chem. Phys.* **104**, 5134 (1996).
- <sup>5</sup>R. Baer, *Chem. Phys. Lett.* **264**, 75 (2002).
- <sup>6</sup>M. E. Casida, C. Jamorski, K. C. Casida, and D. R. Salahub, *J. Chem. Phys.* **108**, 4439 (1998).
- <sup>7</sup>D. J. Tozer and N. C. Handy, *J. Chem. Phys.* **109**, 10180 (1998).
- <sup>8</sup>A. Dreuw, J. Wiesman, and M. Head-Gordon, *J. Chem. Phys.* **119**, 2943 (2003).
- <sup>9</sup>W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- <sup>10</sup>D. M. Ceperley and L. Mitás, *Adv. Chem. Phys.* **93**, 1 (1996).
- <sup>11</sup>R. Baer and D. Neuhauser, in *Recent Advances in Quantum Monte Carlo Methods II*, edited by S. R. a. S. T. W. A. Lester Jr. (World Scientific, Singapore, 2002).
- <sup>12</sup>A. Luchow and J. B. Anderson, *Annu. Rev. Phys. Chem.* **51**, 501 (2000).
- <sup>13</sup>J. B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
- <sup>14</sup>S. L. Mielke, B. C. Garrett, and K. A. Peterson, *J. Chem. Phys.* **111**, 3806 (1999).
- <sup>15</sup>C. Filippi and C. J. Umrigar, *Phys. Rev. B* **61**, R16291 (2000).

- <sup>16</sup>G. Sugiyama and S. E. Koonin, *Ann. Phys. (N.Y.)* **168**, 1 (1986).
- <sup>17</sup>E. Y. Loh, Jr. and J. E. Gubernatis, in *Electronic Phase Transitions*, edited by W. Hanke and Y. V. Kopayev (North-Holland, Amsterdam, 1992), Vol. 32.
- <sup>18</sup>G. H. Lang, C. W. Johnson, S. E. Koonin, and W. E. Ormand, *Phys. Rev. C* **48**(4), 1518 (1993).
- <sup>19</sup>D. R. Hamann and S. B. Fahy, *Phys. Rev. B* **47**(4), 1717 (1993).
- <sup>20</sup>P. L. Silvestrelli, S. Baroni, and R. Car, *Phys. Rev. Lett.* **71**(8), 1148 (1993).
- <sup>21</sup>D. M. Charutz and D. Neuhauser, *J. Chem. Phys.* **102**, 4495 (1995).
- <sup>22</sup>S. W. Zhang, J. Carlson, and J. E. Gubernatis, *Phys. Rev. Lett.* **74**(18), 3652 (1995).
- <sup>23</sup>N. Rom, D. M. Charutz, and D. Neuhauser, *Chem. Phys. Lett.* **270**(3–4), 382 (1997).
- <sup>24</sup>R. Baer, M. Head-Gordon, and D. Neuhauser, *J. Chem. Phys.* **109**(15), 6219 (1998).
- <sup>25</sup>N. Rom, E. Fattal, A. K. Gupta, E. A. Carter, and D. Neuhauser, *J. Chem. Phys.* **109**(19), 8241 (1998).
- <sup>26</sup>R. Baer, *Phys. Rev. A* **62**(6), 063810 (2000).
- <sup>27</sup>R. Baer and D. Neuhauser, *J. Chem. Phys.* **112**, 1679 (2000).
- <sup>28</sup>R. Baer, *J. Chem. Phys.* **113**, 473 (2000).
- <sup>29</sup>R. Baer, *Chem. Phys. Lett.* **324**(1–3), 101 (2000).
- <sup>30</sup>R. Baer, *Chem. Phys. Lett.* **343**(5–6), 535 (2001).
- <sup>31</sup>T. Helgaker, J. Gauss, P. Jorgensen, and J. Olsen, *J. Chem. Phys.* **106**(15), 6430 (1997).
- <sup>32</sup>J. C. Grossman and L. Mitás, *Phys. Rev. Lett.* **79**(22), 4353 (1997).
- <sup>33</sup>J. C. Grossman, L. Mitás, and K. Raghavachari, *Phys. Rev. Lett.* **75**(21), 3870 (1995).
- <sup>34</sup>O. El Akramine, A. C. Kollias, and W. A. Lester, *J. Chem. Phys.* **119**, 1483 (2003).
- <sup>35</sup>J. C. Grossman, M. Rohlfing, L. Mitás, S. G. Louie, and M. L. Cohen, *Phys. Rev. Lett.* **86**(3), 472 (2001).
- <sup>36</sup>L. Mitás, *Isr. J. Chem.* **42**(2–3), 261 (2002).
- <sup>37</sup>In fact, in the present formulation it should be called an auxiliary density.
- <sup>38</sup>A. I. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon, *J. Chem. Phys.* **109**, 10669 (1998).
- <sup>39</sup>E. van Veen, *Chem. Phys. Lett.* **41**, 540 (1976).
- <sup>40</sup>D. Wilden and P. Hicks, *J. Phys. B* **12**, 1579 (1979).
- <sup>41</sup>D. E. Love and K. D. Jordan, *Chem. Phys. Lett.* **235**(5–6), 479 (1995).
- <sup>42</sup>A controlled approximation is one for which the error can be diminished by a factor  $f > 1$  by making the computer work more by a factor  $p(f)$  where  $p$  is a polynomial.
- <sup>43</sup>There are also basis-set errors and one can make additional approximations by neglecting the (or other) electrons and decide to work within a predefined active space of orbitals. We do not address these in this article.
- <sup>44</sup>We should note that the step size error itself is much smaller than the SE for the number of iterations used.
- <sup>45</sup>Formally there is a N 5 step, but this is done once and can probably be avoided by localization methods. Currently it is not the bottleneck. Also, the N 4 step can probably be reduced when localization methods are introduced. This too is not the current computational bottleneck.