

EXPLICITLY CORRELATED WAVE FUNCTIONS IN QUANTUM MONTE CARLO*

DARIO BRESSANINI, GABRIELE MOROSI, AND SILVIA TARASCO

*Dipartimento di Scienze Chimiche, Fisiche e Matematiche, Universita' dell'Insubria,
via Lucini 3, 22100 Como, Italy*

e-mails: Dario.Bressanini@iminsubria.it Gabriele.Morosi@uninsubria.it Silvia.Tarasco@uninsubria.it

(Rec. 23 January 2004)

Abstract: The use of explicitly correlated wave functions in conjunction with Variational and Diffusion Monte Carlo techniques is reviewed. The functional forms investigated are explicitly correlated exponentials, exponentials of Pade's approximants, Jastrow functions and the product of determinants and a correlation factor. Results for four and six electron systems, namely the anions Li^- and B^- , are presented.

1. INTRODUCTION

The main problem of Quantum Chemistry is the accurate description of the correlation between particles in atoms and molecules. Mean-field methods, like Hartree-Fock, give a qualitative description of atomic and molecular systems, but in order to get quantitative results the instantaneous correlation between electrons must be taken into account. The most common way to include correlation is, starting from the Hartree-Fock picture, to approximate the exact wave function using MC-SCF or CI expansions. Unfortunately, methods based on the orbital approximation converge very slowly to the non-relativistic limit. The reason is that these wave functions include the interelectronic distances only in an implicit form. Furthermore this implicit dependence is quadratic instead of linear, so the cusp conditions [1] of the exact wave functions, i.e. the behavior of the wave function when two particles collide, are reproduced only for infinite expansions.

A very efficient and effective approach to accurately describe the local behavior of the wave function when two electrons collide is the explicit inclusion of the interelectronic distances into an approximate wave function. Hylleraas [2], Pekeris [3], James and Coolidge [4], and Kolos and Wolniewicz [5-7] showed that very accurate results for two electron systems are obtained by including the interelectronic distance into the wave function. Unfortunately, it is not easy to generalize these methods to many-electron systems since the resulting integrals are extremely difficult or even impossible to evaluate analytically.

* Dedicated to the memory of Professor Jacek Rychlewski

results on a variety of two-, three- and four-electron systems, provided that a careful optimization of the nonlinear parameters is performed. Unfortunately, this type of function poorly reproduces the cusp conditions and this slows down the convergence to the exact wave function. In order to overcome this problem, a very large number of functions is required to reach high accuracy; but still, this increases the number of nonlinear parameters and makes their optimization a very demanding task. Beyond four electron systems, with at most two nuclei, the analytical approach becomes almost infeasible [12] and we are not aware of published calculations with more than four electrons using ECG functions.

When integrals cannot be evaluated analytically, one must resort to numerical methods. Grid methods cannot be used owing to the high dimensionality of the integrals, so one is left with the variational Monte Carlo (VMC) method [14], VMC is a very powerful numerical technique to estimate the energy, and all the desired properties, of a given trial wave function without any need to compute analytically the matrix elements. The advantage of using VMC is that any wave function, even of great functional complexity, can be used: Monte Carlo methods require only the evaluation of the wave function, its gradient and its Laplacian, and do not impose any restriction on the wave function form since analytical integration is not required. An early application of VMC and an example of its capability was the calculation of the Born-Oppenheimer potential between two helium atoms using Hylleraas-type electronic wave functions [15].

2. WAVE FUNCTION FORM

We recently proposed [16-19] to approximate the electronic wave function of an atomic or molecular system with N electrons with the linear expansion

$$\Psi = \sum_i c_i \phi_i \quad (1)$$

where

$$\phi_i = \hat{A} \left\{ F(\mathbf{r}) f g \Theta_{S, M_S}^N \right\}. \quad (2)$$

In this equation \hat{A} is the antisymmetrizer operator, F is a function of all the electron-nucleus distances, and g is a function of all the electron-electron distances called the correlation factor. Both functions include variational parameters. Θ_{S, M_S}^N is an eigenfunction of the spin operators \hat{S}^2 and \hat{S}_z with the correct spin multiplicity. The functions f and g , being dependent only on interparticle distances, are rotationally invariant. This means that their product can describe only S states, with zero angular momentum. To describe higher angular momentum states, it is necessary to include a function $F(\mathbf{r})$ with the correct rotational symmetry, $F(\mathbf{r})$ is a function of the Cartesian electronic coordinates (x, y, z) , but might include also the electron-nucleus distances [18], Ψ is the two-body term of a many-body expansion of the wave function and this function might be generalized by including products of the interparticle distances.

Various functional forms are possible for the functions f and g . If we use Gaussian functions we recover the ECG functional form. However, to assure a high quality wave function while retaining the compactness, it is particularly important to satisfy both the cusp conditions [1], representing the behavior of the exact wave function at the coalescence of two particles, and the asymptotic conditions [20], representing the behavior when one of the particles goes to infinity.

The first choice for the functions f and g is a simple exponential of the distance. This choice is motivated by the fact that, at particles coalescence, the wave function behaves as

$$\left. \frac{1}{\Psi} \frac{\partial \Psi}{\partial r_{ij}} \right|_{r_{ij}=0} = c, \quad (3)$$

where c is a constant, depending on the type of the colliding particles. The local solution of the above equation

$$\Psi = \exp(cr_{ij}) \quad (4)$$

suggests that an exponential of all the interparticle distances might be a good choice. Also the asymptotic behavior of the wave function is an exponential decay when a particle goes to infinity [20]. Application to various small systems showed the rapid convergence of this expansion [16, 17].

A more flexible function to describe the electron nucleus interaction is an exponential of a Pade's approximant

$$\exp[(ar + br^2)/(1 + cr)] . \quad (5)$$

This form is a good choice for the electron-nucleus part, because it is the best compromise between flexibility and small number of parameters. In fact this function goes as e^{ar} for $r \rightarrow 0$ and $e^{(b/c)r}$ for $r \rightarrow \infty$. So with different exponents it can accommodate both the coalescence at the nucleus and the decay for large r .

As to the electron-electron correlation functions, in order to satisfy the cusp conditions, for $r \rightarrow 0$, they should behave as exponentials, while for $r \rightarrow \infty$ they must go to a constant value. A Jastrow factor [21] $g(r) = \exp(dr(1 + er))$ can satisfy both conditions.

Another popular and effective approach to building compact explicitly correlated wave functions is to multiply a determinantal wave function by a correlation factor, the most commonly used being a Jastrow factor. The determinants are generated from analytic calculations, in general SCF or MC-SCF calculations, for a given basis set. Then the correlation factor is added, and its variational parameters optimized using VMC calculations. In fact, the inclusion of the Jastrow factor does not allow the analytical evaluation of the integrals, so one again has to turn to Monte Carlo techniques.

3. THE VMC ALGORITHM

The VMC algorithm is a stochastic numerical integration scheme. It allows the calculation of the expectation value of any operator \hat{O} given a trial wave function $\Psi_T(\mathbf{R})$. The estimate of any desired property $\langle \hat{O} \rangle$ of a system is written as:

$$\langle \hat{O} \rangle = \frac{\int \Psi_T(\mathbf{R}) \hat{O}(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T^2(\mathbf{R}) d\mathbf{R}} = \frac{\int \Psi_T^2(\mathbf{R}) \frac{\hat{O}(\mathbf{R}) \Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})} d\mathbf{R}}{\int \Psi_T^2(\mathbf{R}) d\mathbf{R}} \quad (6)$$

and the estimate is obtained by generating a large number of points \mathbf{R}_i (called *configurations* or *walkers*) distributed according to $\Psi_T^2(\mathbf{R})$:

$$\langle \hat{O} \rangle \cong \frac{1}{N} \sum_{i=1}^N \frac{\hat{O} \Psi_T(\mathbf{R}_i)}{\Psi_T(\mathbf{R}_i)}. \quad (7)$$

So the expectation value of the energy can be estimated by averaging the local energy $E_L(\mathbf{R}) = \hat{H} \Psi_T(\mathbf{R}) / \Psi_T(\mathbf{R})$ over an ensemble of configurations distributed as $\Psi_T^2(\mathbf{R})$. This ensemble is generated using the Metropolis algorithm [22] or the Langevin algorithm [23].

VMC being an application of the Metropolis algorithm to the variational principle, the mean energy value of the trial wave function is an upper bound to the exact value and the quality of the result depends critically on the trial wave function.

3.1. Metropolis sampling

In order to generate the distribution proportional to $\Psi_T^2(\mathbf{R})$, the most common method is Metropolis sampling. A Markov chain is generated by "box-sampling" $\Psi_T^2(\mathbf{R})$. Starting from an ensemble of walkers, a new generation is obtained by attempting to move each walker: each point is moved by $\mathbf{R} = \mathbf{R}' + \vec{\xi} \Delta$, with Δ the box size, and $\vec{\xi}$ a $3M$ -dimensional vector of uniformly distributed random numbers $\xi \in [-1, +1]$. The move is accepted with probability

$$P(\mathbf{R}' \rightarrow \mathbf{R}) = \min \left[1, \frac{\Psi_T^2(\mathbf{R})}{\Psi_T^2(\mathbf{R}')} \right]. \quad (8)$$

If the move is accepted, then the walker is updated, otherwise, the old coordinate is retained. The random walk generated by such kind of steps results in an asymptotic equilibrium distribution proportional to $\Psi_T^2(\mathbf{R})$. The expectation value of the observables of interest are obtained by averaging on the quantities calculated at each point \mathbf{R} of the Markov chain. Obviously, the greater the number of configurations generated, the smaller is the statistical error of the calculated quantity.

As is normally done in standard applications of the Metropolis method, proper care must be taken when estimating the statistical error, since the configurations generated by a Markov chain are not statistically independent, but are serially correlated. The device of dividing

the simulation into blocks of sufficient length, and computing the statistical error only over the block averages, is usually sufficient to eliminate this problem.

3.2. Langevin sampling

The sampling efficiency of the simple Metropolis algorithm can be improved when one switches to the Langevin simulation scheme. In order to obtain a distribution proportional to $f(\mathbf{R}, \tau) = \Psi_T^2(\mathbf{R})$, in the Langevin approach one simulates the Fokker-Planck differential equation:

$$-\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = -\sum_i D_i \nabla_i^2 f(\mathbf{R}, \tau) + \sum_i D_i \nabla_i \cdot f(\mathbf{R}, \tau) \mathbf{F}_i(\mathbf{R}), \quad (9)$$

where

$$\mathbf{F}_i(\mathbf{R}) = \nabla_i \ln \Psi_T(\mathbf{R})^2 \quad (10)$$

is called the quantum force and $D_i = 1/(2/m_i)$ is the diffusion coefficient for a given particle. The simulation is realized using the Langevin equation

$$r_i = r_i' + D_i \tau \mathbf{F}_i(\mathbf{R}) + \chi, \quad (11)$$

where χ is a three-dimensional Gaussian random variable with zero mean and τ/m_i variance. This equation is used to select an attempted displacement for each particle, which is accepted with probability

$$P(\mathbf{R}' \rightarrow \mathbf{R}) = \min \left[1, \frac{T(\mathbf{R}' \rightarrow \mathbf{R}, \tau) \Psi_T(\mathbf{R})^2}{T(\mathbf{R}' \rightarrow \mathbf{R}, \tau) \Psi_T(\mathbf{R}')^2} \right], \quad (12)$$

where $T(\mathbf{R}' \rightarrow \mathbf{R}, \tau)$ is the transition probability for the particles to move from \mathbf{R}' to \mathbf{R} . For the Langevin equation it is

$$T(\mathbf{R}' \rightarrow \mathbf{R}, \tau) = \prod_i \frac{1}{(4D_i \tau)^{3/2}} \exp \left(- \left[\mathbf{R}_i - \mathbf{R}'_i - D_i \tau \mathbf{F}_i(\mathbf{R}') \right]^2 / 4D_i \tau \right). \quad (13)$$

The dependence of the transition probability on the time step results in a time step bias, so that the estimated values must be extrapolated to zero time step.

3.3. Diffusion Monte Carlo

The amount of correlation energy computed by VMC depends on the quality of the trial wave function. The recovery of the remaining correlation energy can be done using the diffusion Monte Carlo (DMC) method. Exact results in principle can be obtained projecting out the exact ground state of the time-independent Schrödinger equation, $\phi_0(\mathbf{R})$, from any initial

distribution $\Psi_T(\mathbf{R}, 0)$ not orthogonal to $\phi_0(\mathbf{R})$, using the DMC method. From a formal point of view, the time dependent Schrödinger equation in imaginary time

$$-\frac{\partial \Psi(\mathbf{R}, \tau)}{\partial \tau} = (-D\nabla^2 + V)\Psi(\mathbf{R}, \tau) \quad (14)$$

is analogous to a classical diffusion equation, along with a birth/death process that mimics a reaction exhibiting first-order kinetics. In order to make feasible this kind of interpretation, the wave function must be considered a probability density. This constraint is always verified for symmetric (bosonic) ground-states, but presents difficulties when treating excited state wave functions (including fermions ground-state wave functions) that change sign as nodes are crossed: in these cases, the classical analogy breaks down.

Usually DMC simulations employ the importance sampling technique [14]. A trial wave function $\Psi_T(\mathbf{R})$ is used to bias the random walk so as to sample a different probability distribution $f(\mathbf{R}, \tau)$, where

$$f(\mathbf{R}, \tau) = \Psi_T(\mathbf{R})\Psi(\mathbf{R}, \tau) . \quad (15)$$

In this case, the time dependent Schrodinger equation in imaginary time takes the form

$$\frac{\partial f(\mathbf{R}, \tau)}{\partial \tau} = \sum_i D_i \nabla_i^2 f(\mathbf{R}, \tau) - \sum_i D_i \nabla_i \cdot f(\mathbf{R}, \tau) \mathbf{F}_i(\mathbf{R}) - [E_R - E_l(\mathbf{R})]f(\mathbf{R}, \tau) , \quad (16)$$

where the only difference with the Fokker-Planck equation, used in VMC, is the additional term $[E_R - E_l(\mathbf{R})]f(\mathbf{R}, \tau)$. This is simulated varying the population of the configurations during the simulation. In order to maintain the diffusion analogy, $f(\mathbf{R}, \tau)$ rather than $\Psi(\mathbf{R}, \tau)$ must be positive definite. The most commonly used way to overcome this problem is to employ the fixed-node approximation [24]: the nodes of $\Psi_T(\mathbf{R})$ are imposed as a boundary condition upon the evolving solution $\Psi(\mathbf{R}, \tau)$. This boundary condition ensures that $f(\mathbf{R}, \tau)$ is positive definite since $\Psi_T(\mathbf{R})$ and $\Psi(\mathbf{R}, \tau)$ have the same sign for all \mathbf{R} . If the nodes of $\Psi_T(\mathbf{R})$ coincide with those of the exact wave function $\phi_0(\mathbf{R})$, the fixed-node DMC energy will be exactly E_0 (i.e., there would be no inaccuracy), otherwise it can be shown that it is an upper bound to E_0 . In any case most of the correlation energy is recovered by FN-DMC calculations. Since this method is already well described in the literature, details can be found in the available reviews [14, 25].

The DMC energy can be evaluated by means of the mixed estimator

$$E = \frac{\int \Psi(\mathbf{R})H\Psi_T(\mathbf{R})d\mathbf{R}}{\int \Psi(\mathbf{R})\Psi_T(\mathbf{R})d\mathbf{R}} = \frac{\int \Psi(\mathbf{R})\Psi_T(\mathbf{R})\frac{H\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}d\mathbf{R}}{\int \Psi(\mathbf{R})\Psi_T(\mathbf{R})d\mathbf{R}} = \frac{\int f(\mathbf{R}, \infty)E_l(\mathbf{R})d\mathbf{R}}{\int f(\mathbf{R}, \infty)d\mathbf{R}} . \quad (17)$$

The trial wave function is used to guide the walk of the configurations in the space, to reduce the fluctuations of the population simulating the birth-death process described by the term $|E_t - E_t(\mathbf{R})|f(\mathbf{R}, \tau)$ and to compute the energy by means of the mixed estimator.

4. TRIAL WAVE FUNCTION OPTIMIZATION

For a given system, the exact wave function is not known a priori. Nevertheless, it is possible to write down a wave function that depends on some parameters and try to obtain the best parameters for that system. Moreover, if possible one would like to be able to arbitrarily improve the "goodness" of the description, in order to approach the true wave function. The general scheme of optimization starts from an initial choice of the parameters of the wave function. After the evaluation of the quantity one wants to minimize, a new set of parameters is selected modifying the old ones. If the new parameters give a "better" wave function, they are retained. Note that the term "better" is referred to the wave function that minimizes the observable that one wants to minimize. This procedure is iterative, and is repeated until convergence is reached.

5. ENERGY OPTIMIZATION

The commonly used criterion of goodness is given by the variational principle

$$E[\Psi_T] = \langle H \rangle = \frac{\int \Psi_T(\mathbf{R}) H \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T(\mathbf{R}) \Psi_T(\mathbf{R}) d\mathbf{R}} \geq E_0 \quad (18)$$

which provides an upper bound to the exact ground state energy E_0 . Thus the better the energy approaches the exact energy, the better is the wave function. Consequently, the best wave function is the exact wave function.

Let us choose a family of trial wave functions $\Psi_T(\mathbf{R}; \mathbf{c})$, where \mathbf{c} is a vector of parameters $\mathbf{c} \equiv \{c_1, c_2, \dots, c_n\}$ on which the wave function depends parametrically. The best function in the family is selected by solving the problem

$$\min_{\mathbf{c}} \langle H(\mathbf{c}) \rangle \quad (19)$$

The naive application of the variational principle to the optimization problem is affected by the statistical uncertainty inherent in every Monte Carlo calculation, because the convergence of the optimization depends on its magnitude.

The optimization may be carried out both in the framework of VMC and DMC.

5.1. Variance optimization

The statistical error in any VMC estimate of the energy obtained using N sample points is related to the fluctuation of the local energy function. Specifically,

$$E = \langle E_L \rangle \equiv \frac{1}{N} \sum_{i=1}^N E_L(\mathbf{R}_i) \pm \frac{\sigma(E_L)}{\sqrt{N}} , \quad (2.0)$$

where

$$\sigma^2(E_L) = \langle E_L^2 \rangle - \langle E_L \rangle^2 . \quad (2.1)$$

Suppose the exact wave function is known: in that case the local energy would be a constant over the whole configuration space, and the error in its estimate would be rigorously zero. Similarly, if very good wave functions could be obtained, the fluctuations of the local energy would be very small. The error thus provides a measure of the quality of the trial wave function. This leads naturally to the idea of optimizing the wave function by minimizing the error of the estimate. It is easy to see that $\langle E_L \rangle = \langle H \rangle$ and that $\langle E_L^2 \rangle = \langle H^2 \rangle$, so that

$$\langle H^2 \rangle = \langle E_L^2 \rangle \equiv \frac{1}{N} \sum_{i=1}^N E_L^2(\mathbf{R}_i) . \quad (2.2)$$

We can now rewrite Eq. 2.1 as

$$\sigma^2(H) = \langle H^2 \rangle - \langle H \rangle^2 . \quad (2.3)$$

Optimizing a wave function using this principle means solving the problem

$$\min_{\mathbf{c}} \sigma^2(H(\mathbf{c})) . \quad (2.4)$$

There are advantages in using the "sigma variational principle" instead of the more common "energy variational principle." Firstly, the minimum possible value, namely zero, is exactly known and is obtained only for the exact wave function. The minimum value of the energy, on the other hand, is not known. A second quality is that this variational principle can also be used for excited states, and not just for the first state of a given symmetry:

$\sigma^2(H)$ is a minimum for any state. This is in contrast with the energy variational principle, where for an arbitrary stationary state the energy is not necessarily required to be a local minimum.

Instead of minimizing $\sigma^2(H)$, sometimes it is preferable to optimize a related quantity, namely the second moment of the local energy with respect to a fixed (or reference) energy E_R ,

$$\mu_{E_R}^2(H) = \frac{\int \Psi(\mathbf{R})(E_L(\mathbf{R}) - E_R)^2 \Psi(\mathbf{R}) d\mathbf{R}}{\int \Psi(\mathbf{R}) \Psi(\mathbf{R}) d\mathbf{R}} . \quad (2.5)$$

This last form of minimization procedure is preferable when looking for an excited state, since by a careful choice of the reference energy one can prevent falling to the ground state.

5.2. Robust optimization procedures

The assumption of an underlying Gaussian distribution of the local energy, implicit in the standard variance minimization scheme, sometimes is not justified and generates convergence problems. Among the algorithms that can be chosen, the minimization of the mean absolute deviation is usually more reliable than the variance minimization, void of convergence problems, and able to generate trial wave functions with better variational energies [26].

6. RESULTS AND DISCUSSION

In order to show the capabilities of quantum Monte Carlo techniques to deal with explicitly correlated wave functions, we performed calculations on the negative ions of lithium and boron. The choice of atomic anions is dictated by the fact that they show very strong correlation effects, and are a challenge for standard quantum chemistry methods.

For Li^- in its ground state we adopted a single term wave function with different functional forms for the electron-nucleus and electron-electron functions. All functions were optimized minimizing the variance of the local energy. From Table 1 one can see that 85.4% of the correlation energy is recovered with a single exponential term. A substantial improvement can be gained using a Jastrow factor as the electron-electron correlation factor and another millihartree can be recovered by using the more flexible Pade' as the electron-nucleus correlation function instead of the simple exponential form. Despite the fast convergence of this approach, in practice it is not possible to use this kind of function for more than four or five electrons: the weak point of the method lies in the $N!$ cost generated by the antisymmetrizer operator. This weakness is shared with the ECG approach. One way to get around this problem is to use, as electron-electron correlation function, the same function for every electron pair. In this way the electron-electron part becomes totally symmetric and drops out of the antisymmetrizer operator. The remaining electron-nucleus part can now be expressed as a sum of few determinants, lowering the computational scaling cost from $N!$ to N^3 roughly. The energy recovered using a symmetric Jastrow, despite the lower number of variational parameters, is only slightly higher than the value computed without this constraint. This is a very important observation since it shows that using different parameters for different pairs of electrons is not necessary and prevents a more efficient evaluation of the wave function.

A single term wave function already recovers more than 95% of correlation energy. It would be possible to improve the trial wave function by adding more terms in the linear

expansion. However when used in a Diffusion Monte Carlo calculation, more than 99.9% of the correlation energy is recovered.

Table 1. Li^- results

Type	e-n	e-e	Energy (hartree)	% E_{corr}
VMC	Exp	Exp	-7.49019(8)	85.4
VMC	Exp	Jastrow	-7.49625(9)	93.8
VMC	Pade'	Jastrow	-7.49730(2)	95.2
VMC	Pade'	Symmetric Jastrow	-7.49721(3)	95.1
DMC	Pade'	Symmetric Jastrow	-7.50069(2)	99.9
HF			-7.4282299 ^a	
NR Limit			-7.5007727 ^b	

^aRef. [27], ^bRef. [28]

As a second example, we show results for the six-electron system B^- . To avoid the problem of the 720 permutations generated by the antisymmetrizer operating on a product of exponential functions, the wave functions was written as product of a determinantal part times an electron-electron correlation factor written as product of Jastrow factors for every electron pair

$$\Psi_T = D^{(\alpha)} D^{(\beta)} e^U . \quad (26)$$

Table 2. B^- results

Type	CSF	Energy (hartree)	% E_{corr}
VMC	1	-24.5908(12)	49.4
VMC	2	-24.5954(32)	52.5
DMC	2	-24.6565(2)	94.7
HF		-24.519192 ^a	
NR Limit		-24.664190	

^a Ref. [27]

The SCF wave function was calculated using the Hartree-Fock functions for the anion [27]. The d Jastrow parameter was fixed to satisfy the electron-electron cusp condition, while the remaining e parameter was optimized minimizing the energy variance. In a second calculation the configuration $1s^2 2p^4$ was added to the Hartree-Fock $1s^2 2s^2 2p^2$, the relative weights were the result of a configuration interaction calculation. The results of VMC calculations with these two trial wave functions are reported in Table 2 together with the DMC value computed using the second trial wave function. The non-relativistic limit for the anion

was computed from the estimated NR limit of the neutral atom [29] and the experimental electron affinity [30].

This form of trial wave function allows us to recover about half of the correlation energy by VMC calculations. The VMC energy computed by Moskowitz and Schmidt [31] is $-24.6243(1)$ hartree (72% of the correlation energy). They used a more sophisticated electron-electron correlation factor including 17 parameters. The product of correlated exponentials for Li^- was able to recover roughly 95% of correlation energy: even taking into account the smaller number of electrons, the comparison evidences the better efficiency of fully correlated exponential functions than the product of a determinantal function and a correlation factor. DMC again outperforms VMC, but its result depends on the nodal surface defined by the determinantal part of the trial wave function.

7. CONCLUSIONS

In this paper we have shown that the use of explicitly correlated trial wave functions, coupled with quantum Monte Carlo techniques, permits the calculation of very accurate energies. For the lithium anion, using the full antisymmetrizer operator and no constraints on the variational parameters, we recovered 95% of the correlation energy, at the variational level, and more than 99% at the DMC level. However, the computational cost of such an approach scales as $N!$ and quickly becomes infeasible as the number N of electrons grows. We showed that the worsening of the wave function, when the electron-electron parameters are constrained to be the same, is very mild, but now the computational cost scales as N^3 , the cost of a determinant evaluation. We performed calculations on a determinantal function times a Jastrow factor for B^- recovering 52.5% of the correlation energy at VMC level and 94.7% at DMC level.

Acknowledgments

This work has been done in the framework of a project financed by the University of Insubria.

References

- [1] T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [2] E. A. Hylleraas, *Z. Phys.* **54**, 347 (1929).
- [3] C. L. Pekeris, *Phys. Rev.* **112**, 1649 (1958).
- [4] H. M. James and A. S. Coolidge, *J. Chem. Phys.* **1**, 825 (1933).
- [5] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **41**, 3663 (1964).
- [6] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **43**, 2429 (1965).
- [7] W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **49**, 404 (1968).
- [8] S. F. Boys, *Proc. R. Soc. A* **258**, 402 (1960).
- [9] K. Singer, *Proc. R. Soc. A* **258**, 412 (1960).
- [10] J. Rychlewski, W. Cencek, and J. Komasa, *Chem. Phys. Lett.* **229**, 657 (1994).
- [11] J. Rychlewski, *Int. J. Quantum Chem.* **49**, 477 (1994).
- [12] J. Rychlewski, in *Adv. Quantum Chem.* (Academic, New York, 1999), Vol. 31, p. 173.
- [13] Y. Suzuki and K. Varga, *Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems.* (Springer-Verlag, Berlin, 1998).

- [14] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo methods in ab initio quantum chemistry*, 1st ed. (World Scientific, Singapore, 1994).
- [15] R. L. Coldwell and R. E. Lowther, *Int. J. Quantum Chem., Quantum Chem. Symp.* **12**, 329 (1978).
- [16] D. Bressanini, M. Mella, and G. Morosi, *Chem. Phys. Lett.* **240**, 566 (1995).
- [17] D. Bressanini, M. Mella, and G. Morosi, *Chem. Phys. Lett.* **272**, 370 (1997).
- [18] L. Bertini, D. Bressanini, M. Mella, and G. Morosi, *Int. J. Quantum Chem.* **74**, 23 (1999).
- [19] L. Bertini, M. Mella, D. Bressanini, and G. Morosi, *J. Phys. B At, Mol. Opt. Phys.* **34**, 257 (2001).
- [20] S. H. Patil, K. T. Tang, and J. P. Toennies, *J. Chem. Phys.* **111**, 7278 (1999).
- [21] R. Jastrow, *Phys. Rev.* **98**, 1479 (1955).
- [22] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. M. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087(1953).
- [23] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., *J. Chem. Phys.* **77**, 5593 (1982).
- [24] J. B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
- [25] W. A. Lester, Jr. and B. L. Hammond, in *Ann. Rev. Phys. Chem.* (Annual Reviews, Palo Alto, 1990), Vol. 41, p. 283.
- [26] D. Bressanini, G. Morosi, and M. Mella. *J. Chem. Phys.* **116**, 5345 (2002).
- [27] E. Clementi and C. Roetti, *At. Data and Nucl. Data Tables* **14**, 177 (1974).
- [28] J. Komasa, *Phys. Rev. A* **65**, 012506 (2002).
- [29] S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. F. Fischer, *Phys. Rev. A* **47**, 3649(1993).
- [30] M. Scheer, R. C. Bilodeau, and H. K. Haugen, *Phys. Rev. Lett.* **80**, 2562 (1998).
- [31] J. W. Moskowitz and K. E. Schmidt, *J. Chem. Phys.* **97**, 3382 (1992).