
Linear Expansions of Correlated Functions: Variational Monte Carlo Case Study

LUCA BERTINI,¹ DARIO BRESSANINI,² MASSIMO MELLA,¹
GABRIELE MOROSI¹

¹*Dipartimento di Chimica Fisica ed Elettrochimica, Università di Milano, via Golgi 19, 20133 Milano, Italy*

²*Dipartimento di Scienze Chimiche, Fisiche e Matematiche, Università dell'Insubria, Polo di Como, via Lucini 3, 22100 Como, Italy*

Received 27 February 1998; revised 29 January 1999; accepted 16 February 1999

ABSTRACT: The relative performance of trial wave functions expressed as linear combination of correlated exponentials has been tested on a variety of systems. The results are compared against other correlated functions commonly used in the literature to assess the capabilities of the proposed ansatz. A possible departure from the simple exponential functional form used in previous works is discussed, along with its advantages and drawbacks. We also discuss how to implement an efficient optimization procedure for this correlated basis set. © 1999 John Wiley & Sons, Inc. *Int J Quant Chem* 74: 23–33, 1999

Key words: correlated wave functions; quantum Monte Carlo; variational Monte Carlo; variance optimization; few-electron systems

Introduction

Since the early developments of quantum mechanics, it was immediately recognized that the accurate description of particles correlation is an extremely difficult problem. Mean-field methods, like the Hartree–Fock method, neglect the

instantaneous correlation between particles. They often give quite good results and provide a qualitative description for many atomic and molecular systems. However, in order to get accurate quantitative results, the simple mean-field picture must be abandoned in favor of methods that try to accurately describe the electron correlation. The most common way to include correlation into a wave function is to start from the Hartree–Fock picture, and to approximate the exact wave func-

Correspondence to: G. Morosi.

tion using multiconfiguration–self-consistent field (MC–SCF) or configuration interaction (CI) expansions. Unfortunately, methods based on the orbital approximation converge very slowly to the exact results. An alternative approach is the explicit inclusion of the interelectronic coordinates into an approximate wave function; this is an old and well-known method to build very accurate solutions of the Schrodinger equation. Hylleraas [1] and James and Coolidge [2] computed very good results for two-electron systems by including the interelectronic distance r_{12} into the wave function and allowing the trial wave function to reproduce the cusps of the exact one. These methods, however, are difficult to generalize for systems with more than two electrons since the resulting integrals are extremely difficult or impossible to evaluate analytically.

A popular and effective approach to building compact explicitly correlated wave functions is to multiply an SCF wave function by a correlation factor, the most commonly used being a Jastrow factor [3]. The inclusion of the Jastrow factor does not allow the analytical evaluation of the integrals, so one has to turn to a numerical technique. Although the inclusion of a Jastrow factor has been shown to largely improve the simple SCF function, using a more sophisticated determinantal part, such as MC–SCF or truncated CI expansions, multiplied by a Jastrow factor [4, 5], does not always lead to large improvements in the quality of the wave function. As a consequence the research so far has focused mainly on developing better correlation factors [6, 7].

In this study we explore the possibility of expanding the entire wave function, rather than the correlation factor only, as a linear combination of explicitly correlated functions, a field recently reviewed by Rychlewski [8]. The aim is toward the development of a good correlated basis set. This should allow a correct description of the cusp conditions and should keep the number of terms needed to obtain the desired accuracy small, alleviating the problem of the optimization of the nonlinear parameters.

Trial Wave Function Form

We approximate the electronic wave function of an atomic or molecular system with N electrons

and M nuclei with the linear expansion

$$\Psi = \sum_{i=1}^L c_i \phi_i, \quad (1)$$

where

$$\phi_i = \hat{A} \left\{ \left[\hat{O} f_i(\mathbf{R}) F(\mathbf{r}, \mathbf{k}_i) \right] \Theta_{S, M_S}^N \right\}. \quad (2)$$

In this equation \hat{A} is the antisymmetrizer operator, \hat{O} is an operator used to fix the space symmetry, when needed, F is a function of all the electron–electron and electron–nucleus distances, collected in the \mathbf{r} vector, while \mathbf{k}_i is the i th parameter vector; Θ_{S, M_S}^N is an eigenfunction of the spin operators \hat{S}^2 and \hat{S}_z of the correct spin multiplicity; $f_i(\mathbf{R})$ is a function of the Cartesian electronic (x, y, z) and nuclear (X, Y, Z) coordinates:

$$f_i(\mathbf{R}) = \prod_{j=1}^N \prod_{k=1}^M (x_j - X_k)^{\alpha_{ijk}} (y_j - Y_k)^{\beta_{ijk}} \times (z_j - Z_k)^{\gamma_{ijk}} \mathbf{r}_{jk}^{\delta_{ijk}}, \quad (3)$$

where α , β , γ , and δ are appropriate integers greater than or equal to zero. The function F , depending only on the interparticle distances, is rotationally invariant. This means that it can describe only S states, with zero angular momentum. To describe higher angular momentum states, it is necessary to include a function $f_i(\mathbf{R})$ with the correct rotational symmetry.

A linear expansion of the wave function on a basis of correlated functions, with the choice of a Gaussian function as F , first suggested by Boys [9] and Singer [10] in 1960, is currently used by various researchers [11–15]. They have shown that these explicitly correlated Gaussian functions can give very accurate 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 [16], i.e., the behavior of the wave function when two particles collide, and this has the unpleasant effect of slowing down the convergence. As a result, a very large number of functions is needed to reach high accuracy, increasing the number of nonlinear parameters, and making their optimization a very demanding task.

The only motivation for using a Gaussian function is the possibility to compute analytically all the integrals needed to minimize the expectation value of the energy [17]. The variational Monte

Carlo (VMC) approach [18, 19] is a very powerful technique that estimates the energy, and all the desired properties, of a given trial wave function without any need to analytically compute the matrix elements. For this reason it poses no restrictions on the functional form of the trial wave function, requiring only the evaluation of the wave function value, its gradient, and its Laplacian, and these are easily computed. Using the VMC algorithm, essentially a stochastic numerical integration scheme, the expectation value of the energy for any trial wave function form can be estimated by averaging the local energy $H\Psi/\Psi$ during a random walk in the configuration space using a Metropolis algorithm [20] or a Langevin algorithm [21]. In recent years, the VMC method has been successfully used for this task with a variety of explicitly correlated trial wave functions. So as function F in Eq. (2) we are free to choose a more suitable functional form than an explicitly correlated Gaussian, getting faster convergence.

In previous studies [22–24] we proposed to expand the solution of the electronic Schrodinger equation as a linear expansion of explicitly correlated exponentials, that is, we assumed $F(\mathbf{r} \cdot \mathbf{k}) = \exp(\mathbf{r} \cdot \mathbf{k})$, an exponential of a linear combination of all the interparticle distances. We showed that the presence of the interparticle distances in the exponent allows a good description of the cusp conditions and reduces the number of terms needed to obtain the desired accuracy by an order of magnitude, in comparison with correlated Gaussians expansions. The choice of the exponential was motivated by the fact that, at particles coalescence, the wave function behaves as

$$\frac{1}{\Psi} \frac{\partial \Psi}{\partial r_{if}} \Big|_{r_{ij}=0} = c, \quad (4)$$

where c is a constant, depending on the type of the colliding particles. The local solution of Eq.(4)

$$\Psi = \exp(cr_{if}) \quad (5)$$

suggests that an exponential of all the interparticle distances might be a good choice.

This expansion was applied to some test systems and was shown to converge rapidly. Let us briefly review the results we obtained.

For the hydrogen molecule in its ground state at the equilibrium geometry, the convergence of the expansion is very fast [22]. One-term and two-term functions already recover 87 and 98% of the corre-

lation energy, while the 20-term wave function has an error in the energy of the order of 10^{-5} hartree.

For the ${}^2\Sigma_u^+$ ground state of the He_2^+ molecular ion at a bond distance of 2.0625 bohrs we optimized a series of functions of the form

$$\Psi_L = \sum_j^L c_j \hat{A} \left[(1 - \hat{i}) e^{-\mathbf{k}_j \cdot \mathbf{r}} (\alpha\alpha\beta - \alpha\beta\alpha) \right]. \quad (6)$$

A 1-term wave function is already able to recover about 80% of the correlation energy, while a 12-term function gives better results than an extensive full CI calculation [22]. A successive use of these wave functions in a diffusion Monte Carlo simulation allowed to estimate the exact ground-state energy with high accuracy.

This basis was also applied to nonadiabatic systems: For the hydrogen molecular ion [23] a single term of the expansion

$$\Psi_L = \sum_i^L c_i (1 + \hat{P}_{AB}) e^{k_{i1}r_A + k_{i2}r_B + k_{i3}r_{AB} + k_{i4}r_{AB}^2} \quad (7)$$

recovers 99.84% and a 2-term function already recovers 99.97% of the exact energy, while the 10-term wave function has an error in the energy of the order of 10^{-6} hartree. The linear expansion in Eq. (7) works exceptionally well, giving a very fast convergence.

These results are even more striking if we compare them with linear expansions of explicitly correlated Gaussians: two terms of our expansion are already better than Gaussian expansions with more than 200 terms [25, 26]. The improved quality of our basis can be explained by the fact that Gaussian functions poorly reproduce the cusps, while the exponentials we use can account for them.

The positronium molecule Ps_2 , formed by two electrons and two positrons and sometimes called “dipositronium,” was studied [23, 24] optimizing a nonadiabatic trial wave function of the form

$$\Psi_L = \sum_i^L c_i \hat{O} e^{k_{i1}r_{12} + k_{i2}r_{13} + k_{i3}r_{14} + k_{i4}r_{23} + k_{i5}r_{24} + k_{i6}r_{34}}, \quad (8)$$

where the numerical indices indicate the four particles. The energies are quickly convergent as the number of terms increases, as evidenced from the comparison with long expansions of Hylleraas-like functions [27] and correlated Gaussians [26]. A 12-term wave function is within 0.00014 hartree from the estimated exact energy. To recover the remaining energy, a diffusion Monte Carlo calcula-

tion was performed; the exact energy was computed, putting to an end the long debate on the ground-state energy of this system.

Having checked the very good performance of this basis with the previously mentioned systems, we decided to study different atomic and molecular systems, in order to extend the comparison of its performances against other functional forms. In this work we also explore the possibility of using, instead of the simple exponential, different functions F as basis set, in particular an electrons–electron Jastrow term:

$$J(r_{ij}) = e^{ar_{ij}/(1+br_{ij})}. \quad (9)$$

Generation of the Basis

Departing from the usual determinantal wave function form can be very fruitful, as our previous work has shown, but it is computationally much more demanding. The bottleneck to the extension of this method to many-electron systems is the calculation of the $N^\alpha!N^\beta!$ terms generated by the antisymmetrizer operator, N^α and N^β being the number of alpha and beta electrons. For this reason, special care must be taken of the design of an efficient way of generating and optimizing the trial wave function. These steps need to be done in the most effective, fast, and efficient way available. We describe here the procedure we have developed.

It is now well established that the best way to optimize a trial wave function using VMC is not to minimize the energy, but instead to minimize the variance of the local energy, as described in detail by Umrigar [6], since this has been proved to be numerically much more stable.

The optimization of the first term of the expansion is usually performed starting from a trial wave function with all the electron–electron parameters set to zero, and with the electron–nucleus ones coming from some standard Slater orbital basis or from a small basis set optimized at the SCF level. In some cases, when the system may be approximated as a sum of two fragments, the starting wave function may be constructed by combining the wave functions of the two parts. Such a procedure was used, for example, in the optimization of the LiH molecule. However, usually the initial parameters are not very important,

and even “reasonable” randomly chosen parameters quickly converge to the optimum values.

The generation and optimization of additional terms is a crucial point if one aims at a procedure that systematically improves the trial wave function. In principle, one could generate an L -term wave function by simultaneously choosing all the parameters randomly, or by an educated guess, and by performing a minimization of the variance of the local energy. This procedure, however, is doomed to failure, due to the enormous number of local minima present in the nonlinear parameters space.

Instead, we build the $(L + 1)$ -term wave function by adding an extra term to an optimized L -term wave function. We have found this procedure to be very efficient, and able to produce high-quality basis functions. In practice, after having optimized a L -term function Ψ , we perform a VMC simulation and generate an ensemble of walkers (on the order of thousands) with Ψ^2 distribution. At this point a new term is included in the basis, selecting it among a set of terms, all with randomly chosen parameters. The term that gives the largest decrease in the variance of the local energy, estimated using the previously generated fixed ensemble of walkers, is included in the basis, and subsequently optimized using a nonlinear minimization method. After the inclusion of the additional term into the basis, we reoptimize all the terms, one by one, as this leads to a much better wave function, at least when the number of terms is small.

The evaluation of the wave function is the most expensive part of the optimization process. A considerable speed up can be obtained when using a basis with a large number of terms by noting the trivial fact that, during the optimization of a single term, all the others do not change, so a precomputation of their values and of the local energy contributions for each term and each walker gives a substantial saving in CPU time. After the optimization of a given term is completed, its contributions are stored and the next term is optimized in the same way. This process must be repeated until the decrease of the variance is negligible. This usually takes a few “sweeps” (3–10) of all the terms.

The nonlinear parameters of each term to be tested as candidate for inclusion in the basis are randomly chosen within a given range: Usually between 0 and $-k$, for the electron–nucleus terms,

and between $-h$ and h for the electron–electron terms, for some reasonable positive k and h . The linear parameters as well could be randomly chosen in a given range; however, due to the nonorthogonality of the basis, there is little clue on what should be the best range, or even a reasonable one. A better idea is to choose the linear parameter for the candidate term with some guidance.

We seek the conditions on Ψ that minimize the variance of the local energy

$$\begin{aligned}\sigma^2(H) &= \langle H^2 \rangle - \langle H \rangle^2 \\ &= \int \Psi H^2 \Psi d\mathbf{R} - \left(\int \Psi H \Psi d\mathbf{R} \right)^2\end{aligned}\quad (10)$$

with the constraint that the wave function is normalized: $\int \Psi^2 d\mathbf{R} = 1$. Using standard calculus of variations, we find the differential equation that must be satisfied by Ψ . Calling λ the Lagrange's undetermined multiplier and taking the functional derivative, one easily gets

$$H^2\Psi - 2\langle H \rangle H\Psi - \lambda\Psi = 0. \quad (11)$$

If we now expand Ψ on a basis

$$\Psi_L = \sum_{i=1}^L c_i \phi_i \quad (12)$$

substituting Eq. (12) into Eq. (11) and allowing the linear parameters c_i to vary, but not the basis functions, we obtain a matrix equation than must be solved to find the best linear parameters. Unfortunately, the presence of $\langle H \rangle$ in Eq. (11) makes the equation a nonlinear one: It can be solved, at best, by an iterative procedure [28]. However, if we choose to minimize the second moment of the Hamiltonian with respect to some fixed reference energy E_R ,

$$\begin{aligned}\mu^2(H) &= \int \Psi (H - E_R)^2 \Psi d\mathbf{R} \\ &= \sigma^2(H) + (\langle H \rangle - E_R)^2\end{aligned}\quad (13)$$

we obtain the linear equation in Ψ

$$H^2\Psi - 2E_R H\Psi - \lambda\Psi = 0, \quad (14)$$

which, in turn, expanding Ψ in a basis, gives the linear system

$$\sum_j c_j (G_{ij} - 2E_R H_{ij} + \lambda S_{ij}) = 0, \quad i = 1 \dots L, \quad (15)$$

whose nontrivial solution is given by the secular determinant

$$\det|G_{ij} - 2E_R H_{ij} + \lambda S_{ij}| = 0, \quad (16)$$

where we have defined

$$\begin{aligned}G_{ij} &= \int \phi_i H^2 \phi_j d\mathbf{R} = \int (H\phi_j)(H\phi_i) d\mathbf{R}, \\ H_{ij} &= \int \phi_i H \phi_j d\mathbf{R}. \\ S_{ij} &= \int \phi_i \phi_j d\mathbf{R}.\end{aligned}\quad (17)$$

Of course, the above integrals are evaluated using the fixed ensemble of Monte Carlo walkers. So, the linear parameter of the tentative term, given all the nonlinear parameters randomly chosen, can be selected by solving Eqs. (15) and (16). In this way we choose the best possible linear parameter for a given set of nonlinear parameters.

Results and Discussion

TWO-ELECTRON SYSTEMS

The usual testing ground for any new proposed trial wave function form is the helium atom: Despite its apparent simplicity with only two electrons, it is already very difficult to treat, if one is aiming at high accuracy. For this system, and for any two electron atom, the integrals of the exponential ansatz

$$\Psi = \sum_{i=1}^L c_i (1 + \hat{P}_{12}) e^{k_1 r_1 + k_2 r_2 + k_3 r_{12}} \quad (18)$$

are analytically computable, so were able to optimize a rather large number of terms minimizing the energy, a procedure that is not usually possible in VMC calculations. The best energies obtained for 1–19 terms are reported in Table I, while the corresponding wave functions' parameters for these functions and all the other ones in the following can be obtained from the authors. The generation of these basis functions was done according to the scheme previously described. The expansion has an amazingly fast convergence. A two-term wave function recovers already 98.9% of the correlation energy, while our best function has an error in the energy of the order of 10^{-7} hartree.

Considering the relative small number of terms, this functional form gives a remarkable compact and accurate description of the helium atom. This behavior is partially explained by the high number of nonlinear parameters that must be optimized,

TABLE I
Helium atom, analytic calculation.

Terms	Energy(hartree)
1	-2.899534375
2	-2.903269243
3	-2.903640610
4	-2.903687268
5	-2.903713090
6	-2.903717300
7	-2.903720078
8	-2.903722779
9	-2.903723087
10	-2.903723798
11	-2.903723965
12	-2.903724020
13	-2.903724072
14	-2.903724095
15	-2.903724176
16	-2.903724231
17	-2.903724237
18	-2.903724238
19	-2.903724249
HF limit ^a	-2.86168
Exact ^b	-2.903724377

^a Ref. [29].

^b Ref. [30].

so the compactness of the trial wave function is paid with a more difficult optimization problem.

As already mentioned, a quite commonly used correlation factor is the Jastrow factor, so it is interesting to compare its performance when used in place of the simple electron–electron exponential. It is not possible to analytically compute the integrals needed to optimize the Jastrow factor, so the optimizations were done using the VMC method, as previously explained. To provide a fair comparison, the one-term exponentials were reoptimized using the VMC method. Both energies and parameters of the one-term wave functions for H^- , He, and Li^+ are reported in Table II. From this simple comparison, it is clear that the Jastrow factor describes the electron–electron part of the wave function better than a simple exponential. This result is not unexpected: The Jastrow factor is a generalization of the exponential form, and setting $b = 0$, we recover the previous functional form.

A more complete comparison was performed using the hydrogen molecule as testing ground. We have optimized a series of functions of the ground state at the equilibrium geometry using the electron–electron Jastrow, to compare with our previous results with the exponentials [22]. The superiority of the Jastrow in describing the electron–electron correlation is apparent from Table III (CE denoting the correlation energy), although this effect becomes less pronounced as the number of terms increases, the flexibility of the larger expansion compensating for the better behavior, of the single term.

TABLE II
One term functions for H^- , He and Li^+ . Energies in hartrees.

	H^-	He	Li^+
Exponential	-0.52285(13) $k_1 = -1.088$ $k_2 = -0.465$ $k_3 = 0.165$	-2.89950(9) $k_1 = -2.227$ $k_2 = -1.507$ $k_3 = 0.254$	-7.27474(4) $k_1 = -3.288$ $k_2 = -2.457$ $k_3 = 0.262$
Jastrow	-0.52420(5) $k_1 = -1.080$ $k_2 = -0.528$ $a = 0.454$ $b = 0.248$	-2.90143(10) $k_1 = -2.200$ $k_2 = -1.428$ $a = 0.452$ $b = 0.439$	-7.27625(14) $k_1 = -3.297$ $k_2 = -2.394$ $a = 0.465$ $b = 0.686$
HF Limit ^a		-2.86168	-7.23641
Exact ^b	-0.52775	-2.90372	-7.27991

^a Ref. [29].

^b Ref. [30].

TABLE III
H₂ ground state, $R = 1.4011$ bohrs.^a

Terms	Exponential	CE (%)	Jastrow	CE (%)
1	-1.16943(19)	87.64	-1.17282(5)	95.95
2	-1.17362(5)	97.85	-1.17398(2)	98.78
3	-1.17393(5)	98.59	-1.17412(2)	99.12
4	-1.17412(4)	99.07	-1.17421(1)	99.34
5	-1.17417(5)	99.25	-1.17433(1)	99.64
6	-1.17429(2)	99.54	-1.17436(1)	99.71
7	-1.17433(1)	99.65	-1.17439(1)	99.78
HF limit ^b	-1.133623			
Exact ^c	-1.174476			

^a Energies in hartrees.

^b Ref. [12].

^c Ref. [31].

MANY-ELECTRON SYSTEMS

The ability of the proposed expansion to accurately describe systems with more than two electrons was proved in previous works [22–24]. Here we study the performances of our ansatz more extensively. The lithium atom (Table IV) is an interesting case: This is the first system we studied where it was necessary to use a preexponential factor [see Eq. (3)]: The reason is that the exponential part of the wave function alone is not able to describe the behavior of the outer electron (the 2s electron, in the atomic orbital language). A single term without preexponential factor was highly unstable during the optimization and always gave energies worse than the Hartree–Fock energy. The best results were obtained multiplying the expo-

entials by a single electron–nucleus distance: Having adopted the spin eigenfunction $(\alpha\beta - \beta\alpha)\alpha$, we allowed the third electron to be the “outer” electron, using r_3 as prefactor. Again, the expansion is quickly convergent, albeit less than for the two-electron case. Our best wave function recovers 98.75% of the correlation energy, and this result could be improved by including more and more terms, at the expense, of course, of a more difficult nonlinear optimization problem. To give an idea of the computational effort, for lithium the generation of the entire set of exponential functions took about one week on a modern workstation, roughly 90% of the CPU time spent in the optimization process, and the remaining 10% in the VMC simulation. As previously remarked, this

TABLE IV
Lithium atom, ground state.^a

Terms	Exponential ^b	CE (%)	Jastrow ^b	CE (%)	VQMC ^c	CE (%)
1	-7.46699(12)	75.56	-7.47305(12)	88.95	-7.4731(6)	88.99
2	-7.47248(22)	87.70	-7.47640(4)	96.33		
3	-7.47588(8)	95.18				
4	-7.47641(4)	96.35				
5	-7.47695(4)	97.56				
6	-7.47719(2)	98.07				
7	-7.47729(2)	98.29				
8	-7.47750(2)	98.75				
HF Limit ^d	-7.43274					
Exact ^e	-7.47806					

^a Energies in hartrees.

^b Preexponential factor r_3 .

^c Ref. [32].

^d Ref. [33].

^e Ref. [7].

large number of nonlinear parameters is in part responsible for the compactness of the trial wave function. It is interesting to note that our two-term function gives an energy comparable with a more standard trial wave function form used in VMC: A Slater determinant of Hartree–Fock quality multiplied by a correlation factor (in this case a generalized Jastrow factor [7]). Given the good performances with two-electron systems, we tried to use the Jastrow factor instead of the exponential for the description of the electron–electron part. During the optimization we did not fix the a parameter of the Jastrow factor to satisfy the cusp conditions, as sometimes is common in VMC calculations, since this would have increased the variational energy. A single term with Jastrow factors, with 9 optimizable parameters (all the Jastrow factors are independent), gives a better energy than two exponential terms, with 13 optimizable parameters, and two Jastrow factors are equivalent to 4 exponential terms. However, the optimization of the Jastrow proved to be much more difficult than the simple exponential. Most of the time the addition of another term caused a reduction in the variance of the energy, but an increase in the energy. Even starting from the already optimized parameters of the exponentials proved to be useless, the variance decreased, but the energy increased. Despite considerable effort, we were not able to add the third term of the expansion with a corresponding decrease of the variational energy.

Moving to four-electron systems, we studied the beryllium atom and the lithium hydride molecule. The ground state of beryllium (Table V) is a very interesting case: A single exponential term recovers about 64% of the correlation energy,

almost the same percentage recovered by a Slater determinant times a Schmidt and Moskowitz six-term correlation factor [7], proving the good behavior of our exponential ansatz. Like in the lithium case, we had to use a preexponential factor in order to correctly describe the different electron shells. We chose to put the third and fourth electron in the outer shell, resulting in a prefactor of the form $r_3 r_4$, i.e., we used a one-term function of the form

$$r_3 r_4 e^{k_1 r_1 + k_2 r_2 + k_3 r_3 + k_4 r_4 + k_5 r_{12} + k_6 r_{13} + k_7 r_{14} + k_8 r_{23} + k_9 r_{24} + k_{10} r_{34}} \times [\alpha(1)\beta(2) - \beta(1)\alpha(2)] \times [\alpha(3)\beta(4) - \beta(3)\alpha(4)]. \quad (19)$$

We succeeded in adding three more terms with the same prefactor, recovering 92% of the correlation energy. At this point, we were unable to add more terms of the same form: The selection process always gave a new term with a negligibly small linear coefficient, or a term that, after optimization, worsened the energy. It is also interesting to note that the same optimization problem that we encountered with the Jastrow term in the lithium case was also present here: We have never been able to use more than one Jastrow, which is equivalent to three exponentials.

It is a well-known fact that the near degeneracy of the $2s$ and $2p$ orbitals must be taken into account when expanding the ground state of the beryllium atom in Slater determinants. A similar thing happens with our linear combination of correlated exponentials: The inclusion of a term with a prefactor of the form $x_3 x_4 + y_3 y_4 + z_3 z_4$ recovers 97% of the correlation energy. This term corre-

TABLE V
Beryllium atom, ground state.^a

Terms	Exponential ^b	CE (%)	Jastrow ^b	CE (%)	VQMC ^c	CE (%)
1	−14.63309(57)	63.55	−14.65171(15)	83.25	−14.6332(8)	66.67
2	−14.64424(55)	75.35				
3	−14.65247(24)	84.06				
4	−14.66002(13)	92.02				
5 ^d	−14.66490(13)	97.00				
HF limit ^e	−14.57302					
Exact ^e	−14.66736					

^a Energies in hartrees.

^b Preexponential factor $r_3 r_4$.

^c Ref. [7].

^d Fifth term with preexponential factor $(x_3 x_4 + y_3 y_4 + z_3 z_4)$.

^e Ref. [32].

sponds to the description of two electrons in two $2p$ orbitals.

As a four-electron molecular case, we considered the LiH molecule, using the same singlet spin function we used for the beryllium atom.

Here we have different choices for the preexponential factor. The best results were obtained multiplying the exponentials by the distance r_{3-Li} . In the atomic orbital (AO) language this would mean that an electron has been "assigned" to a $2s$ orbital centered on the lithium atom. Roughly speaking, our trial wave function tries to describe the LiH molecule as composed by the Li and H fragments. The description of the LiH molecule as composed by the Li^+ and H^- ions would correspond to using no prefactors (all electrons occupy $1s$ orbitals), but this choice gives a higher energy (see Table VI).

The use of the Jastrow factor greatly improves the energy of the single term, which is better than the simple exponential term, but again we were

unable to add more terms with this correlation function.

Lastly we considered a five-electron system: the Boron atom in its 2P ground state (Table VII). This case is computationally quite demanding since the antisymmetry requirement generates 48 permutations and furthermore a single term has 15 nonlinear parameters. Nevertheless, we were able to recover 52% of the correlation energy with a single term which includes a preexponential factor $r_3 r_4 x_5$. We wish to point out that, to our knowledge, no calculations using linear expansion of fully correlated Gaussians has been published with more than four electrons.

Using a Jastrow correlation function, the energy improves, as already found for the other systems. During the optimization process of the exponential term, we observed that there are functions with a considerably better energy, but a worse variance of the local energy, than the final wave function obtained from the optimization. We found, for exam-

TABLE VI
LiH ground state, $R = 3.015$ bohr.^a

Terms	Exponential	CE (%)	Jastrow	CE (%)	VQMC ^b	CE (%)
1 ^c	-8.0416(1)	65.65				
1 ^d	-8.0481(4)	73.38	-8.0573(1)	84.79	8.0459(3)	70.77
2 ^d	-8.0516(4)	77.60				
3 ^d	-8.0571(1)	84.15				
HF limit ^e	-7.98735					
Exact ^f	-8.07021(5)					

^a Energies in hartrees.

^b Ref. [34].

^c No preexponential factors.

^d Preexponential factors: r_{3-Li} .

^e Ref. [35].

^f Ref. [36].

TABLE VII
Boron atom, ground state.^a

Terms	Exponential ^b	CE (%)	Jastrow ^b	CE (%)	VQMC ^c	CE (%)
1	-24.5942(5)	52.17	-24.6025(5)	58.81	-24.6113(8)	65.90
HF limit ^d	-24.52906					
Exact ^d	-24.65393					

^a Energies in hartrees.

^b Preexponential factor $r_3 r_4 x_5$.

^c Ref. [7].

^d Ref. [32].

ple, a single exponential term with an energy of $-24.61260(5)$, which recovers 66.90% of the correlation energy, but with $\sigma = 4.38$. The final wave function instead has an energy of $-24.5942(5)$ and $\sigma = 1.07$.

This observation can shed some light on the optimization problems found also for the other systems: It seems that, as the number of electrons increases, the wave function with the best variance of the local energy becomes increasingly different from the function with the best energy, which unfortunately cannot be obtained using the current optimization methods. It is very likely that this is the explanation for the observed difficulties found with Jastrow functions: The addition of a term sometimes is used by the optimization process to decrease the sigma, at the expense of an increase of the energy. This problem apparently becomes more severe as the number of electrons increases: in all many-electron systems we treated, except the boron atom, a single optimized term with Jastrow factor recovered more correlation energy than the Slater determinant plus a generalized Jastrow factor, and this demonstrates the good quality of the ansatz. For the boron case, as previously shown, it is possible to recover more correlation energy, but at the expense of the variance of the local energy.

Conclusions

We have shown that departing from the usual determinantal wave function form can be very fruitful, allowing an accurate and, at the same time, compact description of few-electron atomic and molecular systems. Very few terms are needed to reach an accuracy comparable to more common wave function forms, and the description can be systematically improved by adding more terms using the procedure we described. Comparing the results for atoms, one realizes that the larger the number of electrons, the less correlation energy is recovered by a single exponential term and larger expansions are required to reach the same level of accuracy in the energy.

We investigated a possible improvement of the electron-electron part using a Jastrow factor instead of the simple exponential. We found this choice at the same time promising and disappointing. It surely permits a greater percentage of correlation energy to be recovered per number of non-

linear parameters, and this is a desirable property. However its usefulness is severely hampered by the difficulties found in its optimization. Further investigation is necessary to better understand this behavior; a possible explanation might be that the variance optimization is responsible for this behavior, permitting a higher reduction of the sigma at the expense of the increase of the energy.

ACKNOWLEDGMENTS

Financial support by MURST is gratefully acknowledged. CPU time for this work has been partially granted by the Centro CNR per lo Studio delle Relazioni tra Struttura e Reattività Chimica, Milano.

References

1. Hylleraas, E. A. *Z Physik* 1929, 54, 347–366.
2. James, H. M.; Coolidge, A. S. *J Chem Phys* 1933, 1, 825–835.
3. Jastrow, R. *Phys Rev* 1955, 98, 1479–1484.
4. Sun, Z.; Barnett, R. N.; Lester, Jr., W. A. *J Chem Phys* 1992, 96, 2422–2423.
5. Bueckert, H.; Rothstein, S. M.; Vrbik, J. *Can J Chem* 1992, 70, 366–371.
6. Umrigar, C. J.; Wilson, K. G.; Wilkins, J. W. *Phys Rev Lett* 1988, 60, 1719–1722.
7. Schmidt, K. E.; Moskowitz, J. W. *J Chem Phys* 1990, 93, 4172–4178.
8. Rychlewski, J. *Int J Quant Chem* 1994, 49, 477–494.
9. Boys, S. F. *Proc Roy Soc A* 1960, 258, 402–411.
10. Singer, K. *Proc Roy Soc A* 1960, 258, 412–420.
11. Kozłowski, P. M.; Adamowicz, L. *J Chem Phys* 1991, 95, 6681–6698.
12. Alexander, S. A.; Monkhorst, H. J.; Roeland, R.; Szalewicz, K. *J Chem Phys* 1990, 93, 4230–4235.
13. Cencek, W.; Rychlewski, J. *J Chem Phys* 1995, 102, 2533–2538.
14. Cencek, W.; Rychlewski, J. *J Chem Phys* 1993, 98, 1252–1261.
15. Kozłowski, P. M.; Adamowicz, L. *Int J Quant Chem* 1995, 55, 245–250.
16. Kato, T. *Commun Pure Appl Math* 1957, 10, 151–177.
17. Lester, Jr., W. A.; Krauss, M. *J Chem Phys* 1964, 41, 1407–1413.
18. Ceperley, D.; Chester, G. V.; Kalos, M. H. *Phys Rev B* 1977, 16, 3081–3099.
19. Lester, W. A.; Hammond, B. L. *Ann Rev Phys Chem* 1990, 41, 283–311.
20. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. R.; Teller, A. H.; Teller, E. *J Chem Phys* 1953, 21, 1087–1092.
21. Reynolds, P. J.; Ceperley, D. M.; Alder, B. J.; Lester, Jr., W. A. *J Chem Phys* 1982, 77, 5593–5603.

22. Bressanini, D.; Mella, M.; Morosi, G. *Chem Phys Lett* 1995, 240, 566–570.
23. Bressanini, D.; Mella, M.; Morosi, G. *Chem Phys Lett* 1997, 272, 370–375.
24. Bressanini, D.; Mella, M.; Morosi, G. *Phys Rev A* 1997, 55, 200–205.
25. Poshusta, R. D. *Int J Quant Chem* 1983, 24, 65–77.
26. Kozłowski, P. M.; Adamowicz, L. *Chem Rev* 1993, 93, 2007–2022.
27. Ho, Y. K. *Phys Rev A* 1986, 33, 3584–3587.
28. Frost, A. A. *J Chem Phys* 1942, 10, 240–245.
29. Roothaan, C. C. J.; Weiss, A. W. *Rev Mod Phys* 1960, 32, 194–205.
30. Freund, D. E.; Huxtable, B. D.; Morgan III, J. D. *Phys Rev A* 1984, 29, 980–982.
31. Kolos, W.; Szalewicz, K.; Monkhorst, H. J. *J Chem Phys* 1986, 84, 3278–3283.
32. Davidson, E. R.; Hagstrom, S. A.; Chakravorty, S. J.; Meiser Umar, V.; Froese Fisher, C. *Phys Rev A* 1991, 44, 7071–7083.
33. Jitrik, O.; Bunge, C. F. *Phys Rev A* 1991, 43, 5804–5809.
34. Luchow, A.; Anderson, J. B. *J Chem Phys* 1996, 105, 7573–7578.
35. Handy, N. C.; Harrison, R. J.; Knowles, P. J.; Schaefer III, H. F. *J Phys Chem* 1984, 88, 4852–4854.
36. Chen, B.; Anderson, J. B. *J Chem Phys* 1995, 102, 4491–4494.