
Wave-Function Optimization by Least-Squares Fitting of the Exact Wave Function Sampled by Quantum Monte Carlo

R. BIANCHI, D. BRESSANINI, P. CREMASCHI,
M. MELLA, AND G. MOROSI

Dipartimento di Chimica Fisica ed Elettrochimica and Centro CNR per lo studio delle relazioni tra struttura e reattività chimica. Via Golgi 19, 20133 Milan, Italy

Received June 20, 1994; revised manuscript received September 14, 1994; accepted October 27, 1994

ABSTRACT

The sampling of the exact solution of the Schrödinger equation by quantum Monte Carlo methods allows one to solve the problem of the optimization of linear and nonlinear parameters of a trial wave function by minimization of the distance to the exact wave function in Hilbert space even for those systems whose exact wave function is unknown. The overlap integrals between the basis functions and the exact wave function can be easily estimated within the quantum Monte Carlo formalism. Several observables of the helium atom ground state, computed both within the orbital approximation and by an explicitly correlated basis set, evidence the overall goodness of the wave function optimized according to this criterion. © 1996 John Wiley & Sons, Inc.

Introduction

Analytical wave functions, expressed as a linear combination of basis functions, are usually optimized by the variational method. There are two major drawbacks in this approach: First, the need of integrals including the Hamiltonian operator severely limits the choice of the basis set and forces one to adopt the orbital approximation, excluding very few atomic and molecular systems. Second, the variational method can select the optimal wave function with respect to the energy, but

it is a rather insensitive criterion with respect to a best wave function for other physical properties. In fact, the energy is in error to only the second order, while the approximated wave function is in error to the first order and can give completely erroneous values for properties different from the energy.

Other optimization procedures have been proposed in the past: the minimization of the root-mean-square local energy deviation [1]

$$\delta = \left\{ \int [(\mathbf{H} - E)\psi]^2 d\tau \right\}^{1/2};$$

for wave functions approximated by a single Slater determinant [2], the "best-overlap" criterion, corresponding to the Brueckner [3, 4] orbitals definition; and the "best-density" criterion, satisfied by the natural orbitals. However, both the exact density and the exact wave function are unknown, except for few model systems, so one can only compute approximate natural or Brueckner [5] orbitals. The "best-overlap" criterion, which makes the distance to the exact wave function in Hilbert space minimal, can be extended to trial wave functions of a general type, but, again, one is confronted with the problem of knowing the exact wave function.

The performances of the different optimization criteria were compared computing the helium atom wave function using as a basis set a single 1s Slater orbital. The ζ exponent was optimized by minimization of the energy and of the root-mean-square local energy deviation and by a nonlinear least-squares (NLSQ) fitting of the Koga et al. wave function [6], whose quality allows one to assume it as the exact wave function. The ζ exponents are 1.6875, 1.8044, and 1.6199, respectively. The error (in %), with respect to the exact values computed by Pekeris [7], for several properties computed by the three different wave functions are reported in Table I. The fitting process uniformly weights the whole space and its results are in the best agreement with the exact ones, while the variational

method stresses the importance of the region around the nucleus, which gives the largest contributions to the energy, resulting in a contraction of the Slater function. Even worse are the results computed by the wave function optimized by minimization of the root-mean-square local energy deviation, as a larger ζ value is required in an attempt to improve the electron-nucleus cusp condition: This procedure strongly depends on the regions where two particles are close to each other.

To circumvent the bias introduced by the variational method, Mukherji and Karplus [8] investigated the constrained variational method, a procedure in which the variational energy criterion is supplemented by the requirement that integrals of selected operators agree with the known experimental or theoretical values of the corresponding properties. Using the experimental dipole moment and quadrupole coupling constant of hydrogen fluoride as the additional constraint, they found an energy only 0.0043 Hartree higher than that found in absence of the additional constraint, but this small change in the total energy was accompanied by rather large changes in certain one-electron properties. This result confirms that the variational method is rather insensitive to details of the electronic distribution. However, the need of a priori information on the system under investigation, either by experimental or theoretical data, makes this procedure unsuitable to predict molecular properties from first principles.

The above comparison between different optimization criteria is in favor of the method of minimization of the total square deviation between a trial and the exact wave function. Even if the exact solution of the Schrödinger equation in analytical form is unknown, the quantum Monte Carlo methods [9, 10] can sample it. In this article, we show how this capability can be exploited to extend this optimization method even to systems whose exact analytical wave function is unknown. No restriction of the trial wave function to a single Slater determinant and of the basis set to single-particle functions is introduced.

TABLE I
Error (in %) of helium atom properties computed using a 1s Slater orbital and different wave-function optimization criteria.

Property	Variational	δ	NLSQ fitting
$\langle E \rangle$	1.9	2.4	2.1
$\langle T \rangle$	-1.9	12.1	-9.6
$\langle V \rangle$	1.9	-4.9	5.9
$\langle r_{12}^{-2} \rangle$	29.6	48.2	19.4
$\langle r_{12}^{-1} \rangle$	11.5	19.2	7.0
$\langle r_{12} \rangle$	-8.8	-14.8	-5.0
$\langle r_{12}^2 \rangle$	-16.3	-26.8	-9.1
$\langle r_1^{-1} + r_2^{-1} \rangle$	-0.0	6.9	-4.1
$\langle r_1 + r_2 \rangle$	-4.4	-10.6	-0.4
$\langle r_1^2 + r_2^2 \rangle$	-11.7	-22.8	-4.2
$\langle \delta(r_1) \rangle$	-15.5	3.3	-25.3
$\langle \delta(r_{12}) \rangle$	79.8	119.8	59.0
$\langle T \rangle / \langle V \rangle$	0.0	-6.9	4.0

Method

Let Ψ be the exact wave function, and $\Phi_T(\mathbf{c})$, a normalized trial wave function, where \mathbf{c} is a vector of n parameters:

$$\Delta^2(\mathbf{c}) = \langle \Phi_T(\mathbf{c}) - \Psi | \Phi_T(\mathbf{c}) - \Psi \rangle \quad (1)$$

is the total square deviation between $\Phi_T(\mathbf{c})$ and Ψ that we want to minimize. The minimization of $\Delta^2(\mathbf{c})$ is equivalent to the minimization of

$$\langle \Phi_T(\mathbf{c}) | \Phi_T(\mathbf{c}) \rangle - \langle \Phi_T(\mathbf{c}) | \Psi \rangle - \langle \Psi | \Phi_T(\mathbf{c}) \rangle \quad (2)$$

and, for real and normalized Φ_T , results in the maximization of $\langle \Psi | \Phi_T(\mathbf{c}) \rangle$ since $\langle \Phi_T(\mathbf{c}) | \Phi_T(\mathbf{c}) \rangle$ is constant. Since the QMC algorithm considers Ψ as a probability density, it has the uncommon normalization condition $\int \Psi d\tau = 1$. This is completely unimportant on the algorithm and on the expectation values of the optimized wave function.

The stationary point of $\Delta^2(\mathbf{c})$ is defined by a set of n equations:

$$\left\langle \frac{\partial \Phi_T(\mathbf{c})}{\partial c_i} \middle| \Phi_T(\mathbf{c}) \right\rangle - \left\langle \frac{\partial \Phi_T(\mathbf{c})}{\partial c_i} \middle| \Psi \right\rangle = 0 \quad i = 1, \dots, n. \quad (3)$$

The problem is simplified if we assume that the parameters are linear, i.e.,

$$\Phi_T(\mathbf{c}) = \sum_j c_j \phi_j \quad (4)$$

and the basis functions are real.

In this case, we get a linear system of n equations:

$$\sum_j c_j \langle \phi_i | \phi_j \rangle - \langle \phi_i | \Psi \rangle = 0 \quad i = 1, \dots, n. \quad (5)$$

The solution of this system requires the calculation of the $\langle \phi_i | \phi_j \rangle$ and $\langle \phi_i | \Psi \rangle$ integrals. In general, the $\langle \phi_i | \phi_j \rangle$ overlap integrals are not so difficult to compute as are the integrals including the Hamiltonian operator that the variational method requires: This could enlarge the choice of the basis set to more general functions than the standard product of orbitals, e.g., with the interelectronic coordinates built explicitly into the wave function. As to the $\langle \phi_i | \Psi \rangle$ integrals, their evaluation requires the exact analytical wave function to be known: however, even if this information is missing, these integrals can be estimated by quantum Monte Carlo methods. In these methods, the Schrödinger equation is simulated by an ensemble of random walkers. The N points sampled by the ensemble of walkers are distributed according to the ground-state exact wave function, considered as a distribution function, so that $\langle \phi_i | \Psi \rangle$ integrals

can be easily estimated as

$$\langle \phi_i | \Psi \rangle \cong \frac{1}{N} \sum_{j=1}^N \phi_i(\mathbf{R}_j). \quad (6)$$

In the presence of nonlinear parameters, $\langle \Psi | \Phi_T(\mathbf{c}) \rangle$ can be maximized following Umrigar's method [11], which is by repeated evaluations of the overlap integrals over a fixed set of walkers. QMC is the only method that allows the practical implementation of the least-squares fitting of the exact wave function for real systems.

Results and Discussions

As a test of the method, we optimized a wave function of the ground state of the helium atom using as the basis set the five $1s$ Slater orbitals defined by Clementi and Roetti [12]. The trial function used for the fitting process is

$$\Phi_T = \sum_{i=1}^5 \sum_{j=1}^5 c_{ij} \varphi_i(\mathbf{r}_1) \varphi_j(\mathbf{r}_2). \quad (7)$$

The constraints $c_{ij} = c_{ji}$ were added to impose the correct state symmetry.

A diffusion Monte Carlo [13] simulation was realized using a time step of $0.001 \text{ Hartree}^{-1}$ and 6000 walkers. The simulation was stopped after 1800 blocks, each 2000 steps long, having checked the convergence of all the $\langle \phi_i | \Psi \rangle$ integrals. The standard deviations of these integrals, calculated according to the blocking technique [14], are two to three orders of magnitude less than are the integral values. A series of properties was computed using the wave function computed according to Eq. (5), and the results, reported in Table II, are compared with the values obtained by a full configuration interaction calculation, i.e., using the variational method, and with the exact results computed by Pekeris.

On the whole, the LSQ fitting results are very similar to the variational ones. The average distances of the electrons from the nucleus and their squares are in slightly better agreement with the exact values, suggesting that the fitting method tends to give a better representation of the wave function far from the nucleus, while the variational method describes better the region near the nucleus. The large deviations from the exact values

TABLE II
Helium atom properties calculated using a 5s Slater orbital basis set.

Property	Variational	LSQ fitting	Exact
$\langle E \rangle$	-2.8790	-2.8787	-2.9037
$\langle T \rangle$	2.8806	2.8776	2.9037
$\langle V \rangle$	-5.7595	-5.7564	-5.8074
$\langle r_{12}^{-1} \rangle$	0.9876	0.9898	0.9458
$\langle r_{12} \rangle$	1.3854	1.3826	1.4221
$\langle r_{12}^2 \rangle$	2.4094	2.4011	2.5164
$\langle r_1^{-1} + r_2^{-1} \rangle$	3.3736	3.3731	3.3766
$\langle r_1 + r_2 \rangle$	1.8662	1.8637	1.8589
$\langle r_1^2 + r_2^2 \rangle$	2.4094	2.4011	2.3870
$\langle \delta(r_1) \rangle$	1.8240	1.8621	1.8104
$\langle \delta(r_{12}) \rangle$	0.1563	0.1573	0.1063
$\langle T \rangle / \langle V \rangle$	-0.5001	-0.4999	-0.5000

depend on the basis set that, including only spherical functions, does not allow to describe the angular correlation.

To test the use of basis functions of a more general type than products of one-particle functions, we computed the wave function of the ground state of the helium atom using the six-term explicitly correlated Hylleraas-type functions optimized by Koga [15]. Only the linear parameters

were optimized during a simulation realized by 6000 walkers, at time step of $0.001 \text{ Hartree}^{-1}$ and 1800 blocks, each 2000 steps long; the results are reported in Table III. The variational and the LSQ fitting values are very similar and in reasonable agreement with the exact ones. The $\langle r^n \rangle$ values computed by the two wave functions are, in general, on opposite sides with respect to the exact results. The variational method weights unevenly short- and long-range interactions: It tends to increase the wave function when two particles are near each other and to decrease it at long distances.

On the contrary, the fitting procedure evenly weights the whole space. Then, the least-squares process was carried out varying the nonlinear parameter of the basis set. The NLSQ fitting of the sampled wave function decreases the exponent of the Koga basis set, with a better coverage of the region far from the nucleus, and the agreement between computed and exact properties improves: $\langle r_{12}^{-2} \rangle \langle r_{12} \rangle \langle r_{12}^2 \rangle \langle r_1^2 + r_2^2 \rangle$ (see Table III) are closer than are the variational results to the exact ones, while the error on $\langle r_1 + r_2 \rangle$ is nearly the same. Furthermore, the trends of higher powers of $\langle r^n \rangle$ are consistent with the results in Table II.

Even better values could be obtained by selecting the basis set by the criterion of uniform covering of the whole space, even those regions where

TABLE III
Helium atom properties calculated using the Koga six-term basis set.

Property	Variational	LSQ fitting	NLSQ fitting	Exact
ζ	1.858924	1.858924	1.8389	
$\langle E \rangle$	-2.9035	-2.9034	-2.9033	-2.9037
$\langle T \rangle$	2.9035	2.9045	2.9000	2.9037
$\langle V \rangle$	-5.8069	-5.8079	-5.8033	-5.8074
$\langle r_{12}^{-2} \rangle$	1.4696	1.4616	1.4657	1.4648
$\langle r_{12}^{-1} \rangle$	0.9464	0.9433	0.9446	0.9458
$\langle r_{12} \rangle$	1.4200	1.4265	1.4239	1.4221
$\langle r_{12}^2 \rangle$	2.5043	2.5308	2.5215	2.5164
$\langle r_1^{-1} + r_2^{-1} \rangle$	3.3767	3.3756	3.3739	3.3766
$\langle r_1 + r_2 \rangle$	1.8572	1.8639	1.8607	1.8589
$\langle r_1^2 + r_2^2 \rangle$	2.3759	2.4002	2.3901	2.3870
$\langle \nabla_1 \cdot \nabla_2 \rangle$	-0.1590	-0.1588	-0.1612	-0.1591
$\langle \delta(r_1) \rangle$	1.8124	1.8259	1.8082	1.8104
$\langle \delta(r_{12}) \rangle$	0.1094	0.1088	0.1099	0.1063
$\langle T \rangle / \langle V \rangle$	-0.5000	-0.5001	-0.4997	-0.5000

the variational method is rather insensitive. In fact, the Koga basis set was optimized by minimization of the energy, resulting in a good representation of the space near the nucleus.

In conclusion, the use of QMC sampling allows one to extend the criterion of minimum distance to the exact wave function in Hilbert space even to systems whose exact solution of the Schrödinger equation is unknown, making it of practical use and not only of theoretical interest. This criterion allows one to obtain shape-optimized instead of energy-optimized analytical wave functions and guarantees a wave function better on the average over the whole space.

Within the scope of QMC methods, optimizations of analytical wave functions have been performed by variational quantum Monte Carlo calculations. However, the bias that the minimization of the root-mean-square local energy deviation introduces (see Table I) is such that our results, using only six terms of the Koga's expansion, are better than those obtained by VMC simulations including 50 terms [16].

In the context of quantum Monte Carlo methods, the possibility of obtaining analytical wave functions from the discrete sampling of the configuration space can be exploited to compute approximate values of properties other than the energy, a serious issue in quantum Monte Carlo [17], as it requires the sampling of $\Psi^*\Psi$ instead of Ψ .

Another possible use of the analytical wave function, within the scope of quantum Monte Carlo simulations, is to compute the energy by the mixed estimator algorithm. In any case, the calculation of a property not estimated during a simulation would require one to repeat entirely the simulation, while the analytical wave function, obtained by least-squares fitting of the exact wave function, contains the same amount of information as that of the simulation process, in the limit, of course, of the flexibility of the chosen basis set.

ACKNOWLEDGMENTS

Partial support of this research by the Progetto Finalizzato "Sistemi informatici e calcolo parallelo" of the Italian CNR is gratefully acknowledged.

References

1. H. M. James and A. S. Coolidge, *Phys. Rev.* **51**, 860 (1937).
2. W. Kutzelnigg and V. H. Smith, *J. Chem. Phys.* **41**, 896 (1964); V. H. Smith and W. Kutzelnigg, *Ark. Fys.* **38**, 309 (1968); V. H. Smith, *Nuovo Cim.* **48B**, 443 (1967).
3. K. A. Brueckner, *Phys. Rev.* **96**, 508 (1954).
4. R. K. Nesbet, *Phys. Rev.* **109**, 1632 (1958).
5. (a) N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, *Chem. Phys. Lett.* **167**, 115 (1990); (b) R. Kobayashi, N. C. Handy, R. D. Amos, G. W. Trucks, M. Frisch, and J. A. Pople, *J. Chem. Phys.* **95**, 6723 (1991).
6. T. Koga, Y. Kasai, and A. J. Thakkar, *Int. J. Quantum Chem.* **46**, 689 (1993).
7. C. L. Pekeris, *Phys. Rev.* **115**, 1216 (1959).
8. A. Mukherji and M. Karplus, *J. Chem. Phys.* **38**, 44 (1963).
9. (a) D. Ceperly and B. Alder, *Science* **231**, 555 (1986); (b) W. A. Lester, Jr. and B. L. Hammond, *Annu. Rev. Phys. Chem.* **41**, 283 (1990); (c) B. H. Wells, in *Methods in Computational Chemistry*, S. Wilson, Ed. (Plenum Press, New York, 1987), Vol. 1, p. 311.
10. (a) J. B. Anderson, C. A. Traynor, and B. M. Boghosian, *J. Chem. Phys.* **95**, 7418 (1991); (b) S. Zhang and M. H. Kalos, *Phys. Rev. Lett.* **67**, 3074 (1991); (c) R. Bianchi, D. Bressanini, P. Cremaschi, and G. Morosi, *J. Chem. Phys.* **98**, 7204 (1993).
11. C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1719 (1988).
12. E. Clementi and C. Roetti, *Atomic Data and Nuclear Data Tables* (Academic Press, New York, 1974).
13. J. B. Anderson, *J. Chem. Phys.* **63**, 1499 (1975).
14. H. Flyvbjerg and H. G. Petersen, *J. Chem. Phys.* **91**, 461 (1989).
15. T. Koga, *J. Chem. Phys.* **96**, 1276 (1992).
16. S. A. Alexander, R. D. Coldwell, G. Aissing, and A. J. Thakkar, *Int. J. Quantum Chem. Symp.* **26**, 213 (1992).
17. S. Zhang and M. H. Kalos, *J. Stat. Phys.* **70**, 515 (1993), and references therein.