MOLECULAR PHYSICS, 1992, VOL. 75, NO. 1, 243–253

Nonlinear programming approach to locally constrained variational calculations on the ground state of the helium atom

By FREDERICK W. KING, SUSAN E. KELLY and MARK A. KUEHNE
Department of Chemistry, University of Wisconsin-Eau Claire,
Eau Claire, Wisconsin 54702, USA

(*Received 10 June 1991; accepted 26 July 1991*)

The techniques of nonlinear programming are incorporated into the standard variational method. The set of inequality constraints employed in the nonlinear optimizations is based on the reduced local energy, evaluated at various points of configuration space. These constraints give, indirectly, a means to incorporate the local behaviour of the wavefunction in the variational procedure. A test of the procedure is carried out using a 20-term Hylleraas type wavefunction for the ground state of the helium atom. The impact of the constrained optimization on a variety of expectation values is examined.

1. Introduction

The standard variational technique is the cornerstone of modern quantum chemistry. It is a global procedure and gives no information (except for special cases) on the local accuracy of the wavefunction. Despite this shortcoming, a considerable amount of experience with the procedure has led to the now well known notion that the variational method emphasizes the short to medium range region of configuration space. The near-nuclear and long-range regions of configuration space are often poorly characterized in the variational method. This has obvious consequences for expectation values depending on the latter two regions of configuration space.

The purpose of this paper is to consider a refinement of the standard variational approach, where additional constraints depending *indirectly* on the local behaviour of the wavefunction are incorporated in the calculation. The additional constraint to be imposed utilizes the reduced local energy, defined for an N electron system ($N \geqslant 2$) by [1, 2]

$$E_L(\boldsymbol{r}_1) = \frac{\int \Psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) H \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \mathrm{d}\gamma_1 \mathrm{d}\tau_2 \mathrm{d}\tau_3 \ldots \mathrm{d}\tau_N}{\int \Psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N) \mathrm{d}\gamma_1 \mathrm{d}\tau_2 \mathrm{d}\tau_3 \ldots \mathrm{d}\tau_N}, \quad (1)$$

where Ψ represents an approximate eigenfunction and γ_1 is a spin variable for electron 1. For an exact wavefunction the reduced local energy $E_L(\boldsymbol{r})$ reduces to the exact energy of the system:

$$E_L(\boldsymbol{r}) \rightarrow E_{\text{exact}} \text{ as } \Psi \rightarrow \Psi_{\text{exact}}. \quad (2)$$

A constant $E_L(\boldsymbol{r})$ for all $\boldsymbol{r}$ represents a necessary but not sufficient condition that the exact eigenfunction has been located. Since it is known that an analogous result holds

for the Hartree–Fock formalism, that is [3, 4]

$$E_L^{HF}(\boldsymbol{r}_1) = \frac{\int \Psi^*_{HF}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) H \Psi_{HF}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) d\gamma_1 d\tau_2 d\tau_3 \ldots d\tau_N}{\int \Psi^*_{HF}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) \Psi_{HF}(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N) d\gamma_1 d\tau_2 d\tau_3 \ldots d\tau_N}, \quad (3)$$

and

$$E_L(\boldsymbol{r}) \to E^{HF}_{exact} \text{ as } \Psi_{HF} \to \Psi^{HF}_{exact}, \quad (4)$$

a coordinate independent E_L need not indicate that the exact eigenfunction has been located. In practical calculations, it should be clear which eigenfunction, Ψ_{exact} or Ψ^{HF}_{exact}, one is approaching.

An important pair of questions for the approach taken in this study is why supplement the variational method with a constraint based on $E_L(\boldsymbol{r})$, and how is $E_L(\boldsymbol{r})$ likely to improve (if at all) the eigenfunction in a local sense? Equation (1) is a compromise choice. A superior constraint would be based on the local energy defined by [5–7]

$$E_L(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) = \frac{H\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N)}{\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N)}. \quad (5)$$

Unfortunately, for a system with a large number of electrons, (5) becomes extremely complex because of the multidimensional nature of the function $E_L(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N)$. This would be difficult to implement as a suitable constraint given the current generation of supercomputers. Using $E_L(\boldsymbol{r})$ incorporates the local nature of the eigenfunction, but in a somewhat indirect manner. $E_L(\boldsymbol{r})$ offers the primary advantage that it is computationally manageable. The second question does not have a clear answer. It appears likely that a rigorous theoretical connection between the quantity $|E_L(\boldsymbol{r}) - E_{exact}|$ and the error in the electron density $|\rho(\boldsymbol{r}) - \rho_{exact}(\boldsymbol{r})|$, where $\rho(\boldsymbol{r})$ is based on Ψ and $\rho_{exact}(\boldsymbol{r})$ on Ψ_{exact}, will not be found. What is currently known concerning the aforementioned connection is mostly empirical, based on observation from calculations and some model studies [1, 2, 8–17].

At any particular configuration space point, an accurate $E_L(\boldsymbol{r})$ *does not* imply an accurate $\rho(\boldsymbol{r})$. $E_L(\boldsymbol{r})$ may be rather inaccurate at some point of configuration space, and the electron density may be rather accurate at the point. This observation is either implicit or explicit in a number of papers [1, 8, 9, 11, 15, 17]. The near-nuclear region is a good example where this problem arises. Intuitively, it would be expected that if $E_L(\boldsymbol{r})$ is inaccurate for a sufficiently large collection of points in some region of configuration space, then $\rho(\boldsymbol{r})$ would likewise be inaccurate. Also, it might be conjectured that the converse is true. Since no relationship between $E_L(\boldsymbol{r})$ and $\rho(\boldsymbol{r})$ is known, it is impossible to say how closely the inaccuracies of $E_L(\boldsymbol{r})$ and $\rho(\boldsymbol{r})$ will correlate over a given region of configuration space. The empirical evidence available [1, 2, 8–13, 17] does appear to support the above working hypothesis. There is a clear need for detailed investigation on realistic model systems that can be solved analytically.

A principal objective of the present investigation is to explore how the nonlinear programming modification of the standard variation method affects a variety of expectation values, emphasizing different regions of configuration space. The ground state of the helium atom was selected for the present study for two reasons. The small number of electrons involved allowed for a considerable amount of exploration with the nonlinear programming phase of the calculations described in section 2. For the

helium atom, there is an extensive literature available on the accurate evaluation of a range of properties [19–21]. These will provide an important comparison for the results of the present study.

2. Theory

The problem investigated in this study is the following optimization task. Determine the minimum of the quantity

$$E = \frac{\int \Psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N) H \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N) \mathrm{d}\tau_1 \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_N}{\int \Psi^*(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N) \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots \boldsymbol{r}_N) \mathrm{d}\tau_1 \mathrm{d}\tau_2 \ldots \mathrm{d}\tau_N}, \tag{6}$$

with respect to parameters which appear in the expansion of the trial wavefunction Ψ, subject to the following set of constraints,

$$\begin{aligned} E - E_{\mathrm{L}}(r_1) &\leqslant a_1 \\ E - E_{\mathrm{L}}(r_2) &\leqslant a_2 \\ &\vdots \\ E - E_{\mathrm{L}}(r_i) &\geqslant a_i \\ E - E_{\mathrm{L}}(r_j) &\geqslant a_j. \end{aligned} \tag{7}$$

The optimization task presented by (6) and (7) is a nonlinear programming problem. The angle integrated form (over $\mathrm{d}\Omega_1$) of (1) has been employed for the constraints in (7).

The initial choice of E employed in (7) is the value obtained from the minimization of the objective function, (6), without the imposition of the constraints given in (7). More complicated constraints could be imposed, such as those depending on the pth reduced local energy [22], where integration is carried out over p electron coordinates. Such constraints would add greatly to the complexity of the optimization problem, thereby increasing computer time requirements enormously.

Nonlinear programming problems are much more difficult to solve than the corresponding linear programming problems [23, 24]. The difficulty associated with following a boundary of a nonlinearly constrained region creates major complications. Probably for these reasons, nonlinear programming techniques have been utilized in a very limited way in quantum chemistry.

In a previous work [12], Rosen's projected gradient method [25, 26] was utilized to do nonlinear programming calculations. Rosen's method, however, has limitations when the inequality constraints are nonlinear. Because of the general importance of nonlinear programming techniques, and the diverse fields to which the technique applies [27], the development of refined algorithms continues to be researched [28, 29]. This growth is likely to continue well into the future. The method used to solve the optimization problem in this study is the generalized reduced gradient approach. This algorithm is fully described elsewhere [29].

3. Computational details

The wavefunction employed in the present calculations was the Hylleraas function

$$\Psi = e^{-ks/2} \sum_{i=1}^{N} C_i s^{l_i} t^{m_i} u^{n_i}, \tag{8}$$

where $s = r_1 + r_2$, $t = r_1 - r_2$, $u = r_{12}$, k is a constant, C_i are the variationally determined expansion coefficients, and N is the number of terms in the basis set. The results reported in the next section were determined using $k = 3{\cdot}50$. Refinements could certainly be achieved using slightly different values of this constant. For the set $\{l, m, n\}$, the selection employed by Hart and Hertzberg [30] was utilized. This choice has 20 terms, and leads to a fairly accurate ground state energy and reasonably accurate values for a number of expectation values, based on comparison with the results of very elaborate calculations on the He atom [19–21]. The set of terms $\{l, m, n\}$ employed is (0, 0, 0), (0, 0, 1), (0, 2, 0), (1, 0, 0), (2, 0, 0), (0, 0, 2), (1, 0, 1), (0, 2, 1), (0, 0, 3), (0, 2, 2), (1, 2, 0), (3, 0, 0), (0, 2, 4), (0, 0, 4), (0, 0, 5), (0, 2, 3), (2, 2, 0), (4, 0, 0), (1, 2, 1) and (0, 4, 0).

In some respects, a smaller basis set would have been much more manageable in the nonlinear optimization phase of the project. A principal objective was to examine to what extent various expectation values downgrade or improve in the course of the constrained optimizations when the starting (unconstrained) wavefunction is already *fairly* accurate. To test the aforementioned, a basis set of reasonable size is required.

The nonlinear programming phase of the calculations was carried out in three parts, each focusing on a different region of configuration space. For expectation values which depend on the medium range region of configuration space, the unconstrained variational calculation is expected to produce reasonably accurate values. Properties sensitive to the near-nuclear region or the long-range region of configuration are generally expected to be less accurate. Constraints were selected in either the near-nuclear, medium-range or long-range regions for the constrained variational calculations.

The starting point for each constrained optimization problem was the wavefunction obtained from the unconstrained optimization. The nonlinear programming phase of the calculations was carried out in a stepwise fashion. The inequality constraints were tightened until a solution of the nonlinear programming problem could not be obtained. Some trial explorations from other starting points were also carried out.

The placement of constraints is a matter of trial and error. It is driven principally by two factors: the region of configuration space where attempts are being made to refine $E_L(r)$, and the possibility of actually finding a solution of the nonlinear programming problem. Constraints with very tight limits on the allowed error in $E_L(r)$ very quickly diminish the possibilities of finding a solution to the optimization problem.

4. Results

Table 1 presents a summary of the wavefunctions obtained from the unconstrained variational calculation, and from the nonlinear programming calculations. The number of significant digits of precision included for the coefficients in table 1 is to allow the tabulated expectation values to be reproduced, and does not signify that the eigenvector evaluations have converged to this level of precision. For the con-

Table 1. Coefficients for the wavefunctions and constraints employed in the nonlinear optimization.

Unconstrained	Short-range constrained	Intermediate-range constrained	Long-range constrained
1.	1.	1.	1.
0·401612556	0·389203095	0·402566124	0·391281852
0·185974384	0·205685720	0·186121137	0·192942284
$-0·177329286$	$-0·172987760$	$-0·177697271$	$-0·171907794$
$0·480613099 \times 10^{-1}$	$0·484402360 \times 10^{-1}$	$0·481029655 \times 10^{-1}$	$0·472892480 \times 10^{-1}$
$-0·126240223$	$-0·116495341$	$-0·127004477$	$-0·126353979$
$0·801465230 \times 10^{-2}$	$0·803900131 \times 10^{-2}$	$0·801358213 \times 10^{-2}$	$0·801288073 \times 10^{-2}$
$-0·991849913 \times 10^{-1}$	$-0·112616859$	$-0·992208969 \times 10^{-1}$	$-0·981088915 \times 10^{-1}$
$0·397280426 \times 10^{-1}$	$0·392200812 \times 10^{-1}$	$0·398230223 \times 10^{-1}$	$0·412150630 \times 10^{-1}$
$0·294746063 \times 10^{-1}$	$0·270897615 \times 10^{-1}$	$0·296830660 \times 10^{-1}$	$0·211370391 \times 10^{-1}$
$0·119490872 \times 10^{-1}$	$0·194699142 \times 10^{-1}$	$0·116213806 \times 10^{-1}$	$0·151328858 \times 10^{-1}$
$-0·885378943 \times 10^{-2}$	$-0·881699544 \times 10^{-2}$	$-0·885080420 \times 10^{-2}$	$-0·883763155 \times 10^{-2}$
$0·290328659 \times 10^{-3}$	$0·295099929 \times 10^{-3}$	$0·288894779 \times 10^{-3}$	$0·115224703 \times 10^{-3}$
$-0·706067938 \times 10^{-2}$	$-0·707999836 \times 10^{-2}$	$-0·705334022 \times 10^{-2}$	$-0·707572363 \times 10^{-2}$
$0·482291177 \times 10^{-3}$	$0·482000038 \times 10^{-3}$	$0·482548474 \times 10^{-3}$	$0·472567117 \times 10^{-3}$
$-0·497295936 \times 10^{-2}$	$-0·476600389 \times 10^{-2}$	$-0·499597160 \times 10^{-2}$	$-0·277838461 \times 10^{-2}$
$-0·269043869 \times 10^{-2}$	$-0·265799950 \times 10^{-2}$	$-0·268972695 \times 10^{-2}$	$-0·247896305 \times 10^{-2}$
$0·696632143 \times 10^{-3}$	$0·697600097 \times 10^{-3}$	$0·696745477 \times 10^{-3}$	$0·700200273 \times 10^{-3}$
$0·498493467 \times 10^{-2}$	$0·504799895 \times 10^{-2}$	$0·498103981 \times 10^{-2}$	$0·507587261 \times 10^{-2}$
$0·418332999 \times 10^{-3}$	$0·417199947 \times 10^{-3}$	$0·418304504 \times 10^{-3}$	$0·419279795 \times 10^{-3}$
	$E_L(r = 0·04) < -2·903224$	$E_L(r = 0·06) < -2·902724$	$E_L(r) < -2·900724$
	$E_L(r = 0·12) < -2·903224$	$E_L(r = 0·60) < -2·902724$	$E_L(r) > -2·906724$
	$E_L(r = 0·20) < -2·903224$	$E_L(r = 1·2) < -2·902924$	for $r = 1·0$, 2·0, 3·0,
	$E_L(r = 0·04) > -2·904224$	$E_L(r = 0·06) > -2·904224$	4·0, 5·0, 6·0, 7·0
	$E_L(r = 0·12) > -2·904224$	$E_L(r = 0·60) > -2·904724$	
	$E_L(r = 0·20) > -2·904224$	$E_L(r = 1·2) > -2·904524$	

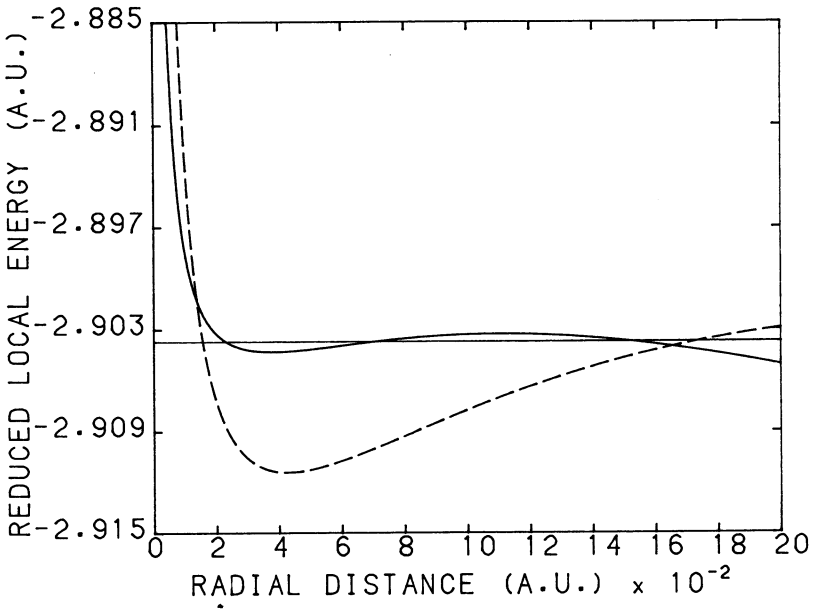

Figure 1. Reduced local energy versus the radial coordinate for the near-nuclear region. The exact reduced local energy should follow the horizontal line at −2·903724. The dashed curve represents $E_L(r)$ derived from the unconstrained optimized wavefunction, and the solid curve represents $E_L(r)$ derived from the short range constrained optimized wavefunction.

strained optimizations, the inequality constraints employed are listed at the bottom of table 1. The resulting $E_L(r)$ for each wavefunction is displayed in figures 1–3. The horizontal line at −2·903724 hartree on each plot represents the behaviour for the exact reduced local energy.

In order to assess the impact of the nonlinear constrained optimizations, a number of different expectation values have been evaluated. The notational abbreviations

$$\langle O_i \rangle = \frac{\langle \Psi | \Sigma_{i=1}^{2} O_i | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{9}$$

$$\langle O_{12} \rangle = \frac{\langle \Psi | O_{12} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{10}$$

have been employed. The moments $\langle r_i^n \rangle$ for $n = -2$ to 6, $\langle r_{12}^n \rangle$ for $n = -2$ to 3, $\langle \nabla_1 \cdot \nabla_2 \rangle$, $\langle \delta(r_i) \rangle$, $\langle \delta(r_{12}) \rangle$ and the energy have been evaluated. δ denotes the Dirac delta function. $\langle \delta(r_i) \rangle$ gives the electron density at the nucleus. The expectation value $\langle \nabla_1 \cdot \nabla_2 \rangle$ is required for the evaluation of the specific mass shift and the transition isotope shift, and is particularly sensitive to electron correlation effects. This set of expectation values were selected because different regions of configuration space are sampled. In addition, results from more elaborate wavefunctions are available for comparison. The results for the expectation values are collected in table 2. The 'exact' values reported in table 2 that are taken from the work of Pekeris [19] and Drake [31] are accurate to the number of digits reported. The Thakkar–Smith [21] entries have an uncertainty that probably resides in the last quoted digit, based on observations

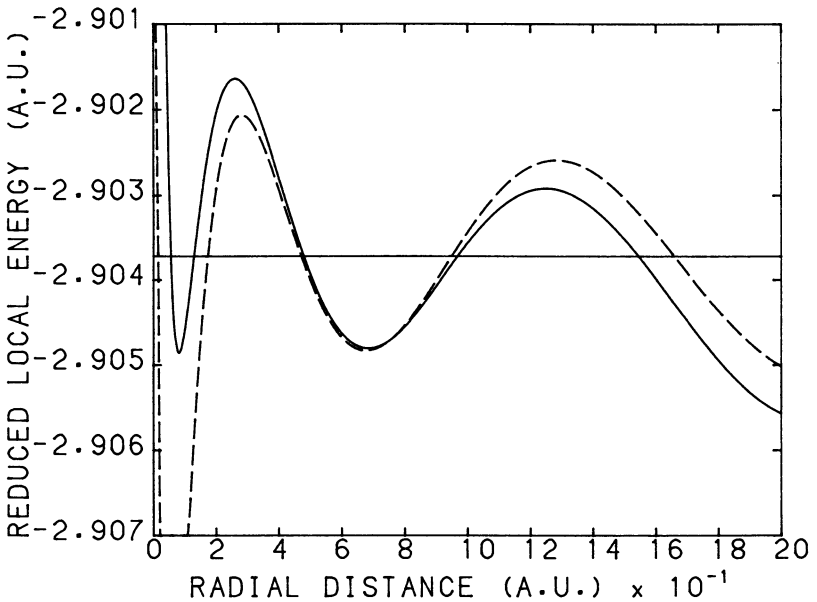

Figure 2. Reduced local energy versus the radial coordinate for the intermediate region of configuration space. The dashed curve represents $E_L(r)$ derived from the unconstrained optimized wavefunction, and the solid curve represents $E_L(r)$ derived from the intermediate-range constrained optimized wavefunction.

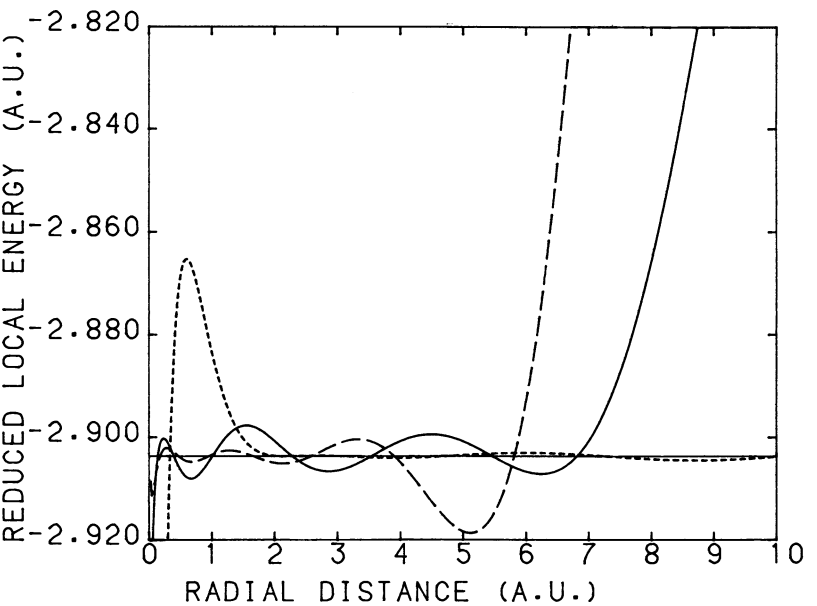

Figure 3. Reduced local energy versus the radial coordinate for the long-range region of configuration space. The longer-dashed curve represents $E_L(r)$ derived from the unconstrained wavefunction, and the solid curve represents $E_L(r)$ derived from the long range constrained optimized wavefunction. The shorter-dashed curve represents $E_L(r)$ determined from an alternative constrained wavefunction.

Table 2. Expectation values evaluated from the constrained and unconstrained variational wave functions.

Expectation value	Unconstrained	Constrained short-range	Constrained intermediate-range	Constrained long-range	Exact[a]
$\langle r_i^{-2} \rangle$	12·0346	12·0514	12·0346	12·0442	12·0348
$\langle r_i^{-1} \rangle$	3·37663	3·37568	3·37665	3·37725	3·37663
$\langle r_i \rangle$	1·85894	1·85956	1·85892	1·85893	1·85894
$\langle r_i^2 \rangle$	2·38689	2·38238	2·38690	2·38636	2·38697
$\langle r_i^3 \rangle$	3·93549	3·90006	3·93568	3·93209	3·93589
$\langle r_i^4 \rangle$	7·94529	7·76736	7·94608	7·92995	7·94707
$\langle r_i^5 \rangle$	19·0014	18·1956	19·0031	18·9376	19·0075
$\langle r_i^6 \rangle$	52·5565	48·9585	52·5522	52·2766	
$\langle r_{12}^{-2} \rangle$	1·46569	1·46108	1·46569	1·47191	1·46477
$\langle r_{12}^{-1} \rangle$	0·945841	0·944370	0·945877	0·947207	0·945818
$\langle r_{12} \rangle$	1·42205	1·42220	1·42202	1·42110	1·42207
$\langle r_{12}^2 \rangle$	2·51634	2·51104	2·51636	2·51477	2·51644
$\langle r_{12}^3 \rangle$	5·30770	5·26937	5·30825	5·30860	5·30800
$\langle \nabla_1 \cdot \nabla_2 \rangle$	−0·159202	−0·156965	−0·159157	−0·155438	−0·159069
$\langle \delta(r_i) \rangle$	3·62057	3·63950	3·62058	3·62635	3·62086
$\langle \delta(r_{12}) \rangle$	0·107368	0·107345	0·107318	0·108702	0·106345
η	1·000021	1·004538	0·999817	1·001077	
Energy	−2·903711	−2·903498	−2·903711	−2·903646	−2·903724

[a] The exact values are taken from reference [19], except $\langle r_i^3 \rangle$, $\langle r_i^4 \rangle$ and $\langle r_{12}^3 \rangle$ are from reference [21], $\langle r_i^5 \rangle$ is from reference [20], and $\langle \delta(r_i) \rangle$ is from reference [31].

of the convergence pattern of their calculations. For the Chong–Weinhold [20] entry, the number of digits of precision is uncertain.

Also reported in table 2 is the scale factor η, defined by

$$\eta = \frac{-\frac{1}{2}\langle V \rangle}{\langle T \rangle}, \tag{11}$$

where $\langle V \rangle$ and $\langle T \rangle$ are the potential energy and kinetic energy, respectively. All expectation values reported in table 2 have been appropriately scaled using the values of η presented in table 2.

5. Discussion

Since the function $E_L(r)$ is playing a central role in the nonlinear programming problem, it is appropriate to examine how it changes in the constrained optimization. The reference point in each case is the exact E_L, which is the constant function shown in each figure. Also shown in each figure is $E_L(r)$ resulting from the unconstrained optimization.

At short-range, figure 1 shows the constrained optimization has produced an $E_L(r)$ which approaches the exact E_L more closely on the average, than does the $E_L(r)$ resulting from the unconstrained wavefunction. Both the constrained and unconstrained wavefunctions give a poor $E_L(r)$ at distances very close to the nucleus. Neither wavefunction satisfies the cusp condition. At intermediate distances, figure 2 indicates that the unconstrained and constrained wavefunctions lead to $E_L(r)$ functions that are fairly close to one another. In some parts of the intermediate region the

unconstrained wavefunction gives a better $E_L(r)$, and in other parts, it is the constrained wavefunction that does better. For the long-range region, the constrained wavefunction leads to a better $E_L(r)$ in roughly the range $r \sim 3$–7 a.u., whereas the unconstrained wavefunction gives a better $E_L(r)$ in roughly the range $r \sim 0$–3 a.u.

During the course of trial explorations, a constrained wavefunction was discovered whose resulting $E_L(r)$ function is shown in figure 3 as the curve with short dashes. This $E_L(r)$ is clearly superior to the results from the other functions for $r \sim 2$–10 a.u., but rather poor in the region $r \sim 0$–2 a.u. Attempts to improve this wavefunction so that $E_L(r)$ was improved in the region $r \sim 0$–2 a.u. were not successful.

Of particular interest in this study is what changes take place in the expectation values when the nonlinear optimization is carried out. This should provide some evidence as to whether or not $E_L(r)$ will be useful as a criterion which can be used to improve the wavefunction. One key factor needs to be kept in mind when comparing the expectation values evaluated from the different wavefunctions employed in this work. The unconstrained wavefunction employed as a reference produces results of fairly high quality, so if improvements are obtained, they will be fairly small. Two linked questions arise at this point. Is it important to obtain refinements at this high level of precision? If it is, why not simply go to basis sets of larger size? The answer to the first of these two questions is driven by the very high precision experimental data that is now available for a number of systems, particularly few-electron systems [32]. Quantities like the hyperfine coupling constant have recently been measured in ion-trap experiments to approximately eleven significant digits [33]. Computational attempts to reproduce these results reveal significant features about the theoretical formulations. Transition isotope shifts are particularly sensitive to the quality of the wavefunctions [33], because they depend on the difference of values which are very close together. Recent work on the helium atom [31, 35–40] also shows the importance of high precision calculations. Accepting the case that improving the wavefunction to high precision serves a useful purpose, expansion of the size of the basis set is certainly a means to achieve a higher quality wavefunction. There are some well known drawbacks to this approach, even for simple systems. CPU costs increase significantly for large basis sets, and numerical problems may also arise. In addition, it can be difficult to tailor a wavefunction for a particular expectation value, unless specialized basis functions are employed. It therefore seems desirable to explore other procedures that may be able to refine the wavefunction.

It is useful to remember that the integrands for the expectation values may depend on a wide range of points in configuration space. For example, $\langle r_i^2 \rangle$ depends most sensitively on the region $r \sim 0{\cdot}2$–$4{\cdot}0$ a.u.; $\langle r_i^6 \rangle$ depends on the region $r \sim 1{\cdot}0$–$6{\cdot}0$ a.u. Examples for other moments can be found elsewhere [12, 18]. This means that for the accurate evaluation of higher moments, both the intermediate and long range regions of the wavefunctions must be accurate. Since the moments are global characterizations of the wavefunction, it is possible that inaccuracies in one part of configuration space may offset errors from another region, though it is expected in the present study that this is not a significant consideration. Similar comments apply to the evaluation of expectation values such as $\langle r_i^{-2} \rangle$. Although the near-nuclear region of configuration space is emphasized, the near-intermediate region is also sampled.

The best results in this investigation were obtained with the nonlinear optimization in the intermediate range. Of the 16 expectation values reported, nine are improved, two are unchanged, one case cannot be decided for lack of an 'exact'

comparison value, and four values are worse. Of the latter four values, the difference is 2–3 in the sixth digit of precision. Some of the improvements are small but, in several cases, the unconstrained wavefunction leads to expectation values in very close agreement with the exact values.

For the wavefunction optimized for the near-nuclear region, all but one of the expectation values are poorer than those evaluated from the unconstrained wavefunction. The loss of accuracy for the expectation values emphasizing the intermediate and long range regions of configuration space is not unexpected, as the near-nuclear constrained optimization downgrades the wavefunction in these two regions. Expectation values such as $\langle r_i^{-2} \rangle$ and $\langle r_i^{-1} \rangle$, which are sensitive to the region near the nucleus, are close to the exact values, but just a little worse than the results arising from the unconstrained optimization. It is probably the case that the near-intermediate region of configuration space makes a sufficient contribution to the expectation values of $\langle r_i^{-2} \rangle$ and $\langle r_i^{-1} \rangle$, and this intermediate region is described better by the unconstrained wavefunction versus the short-range constrained wavefunction. Hence there is a slight loss of accuracy for the expectation values.

For the long-range optimized wavefunction, the expectation values are a little worse. Many of the expectation values, even the high moments, depend on the intermediate range region to some extent. The constrained optimization downgrades this intermediate region during the optimization procedure. The expectation values obtained from the wavefunction whose $E_L(r)$ is rather excellent from $r \sim 2$–10 a.u. in figure 3, were somewhat poorer than those resulting from the long-range constrained wave function. It is conjectured that these poorer quality expectation values arise from the inferior description of the intermediate range region. Note the poor $E_L(r)$ in this region displayed in figure 3. Since this particular wavefunction is inferior to the long-range constrained optimized wavefunction, the details for this wavefunction are omitted from tables 1 and 2.

Given the above observations, it appears desirable to augment any constrained optimization in the near-nuclear or long-range regions, with sufficient inequality constraints centred in the intermediate region, so as to maintain, or simultaneously improve the accuracy in the latter region. The difficulty with this strategy is that if the binding inequality constraints are set rather tightly, it is not possible to find a solution to the nonlinear programming problem. This situation could be circumvented by carefully adding additional basis functions. This was not undertaken in the present investigation, as it would be rather difficult to assign whether improvements in the expectation values, relative to those obtained from a reference unconstrained optimized wavefunction, were the result of improvement in the local behaviour of the constrained wavefunction, the result of the additional basis functions, or some combination of these two factors.

The CPU costs required to implement the nonlinear programming phase of the calculations are a significant fraction of the total cost of the calculation. The costs could have been greatly diminished if less computational exploration had been carried out. The costs could be further lowered if only a restricted region of configuration space had been sampled.

The results of this study show that, for the system studied herein, the nonlinear programming modification of the standard variational method can be used to improve expectation values, if attenuation is focused on the intermediate range region of configuration space. The reduced local energy does function as a useful *indirect* reflection of the local errors in the wavefunction.

It would be extremely desirable if a theoretical analysis could be carried out to

unravel, at least in an approximate way, the relationship of $|E_L(r) - E_{exact}|$ to $\Delta\rho(r)$, the error in the electron density. Such analysis might provide some insights into the results obtained in the present study. Calculations on model systems might offer the best hope to achieve this goal.

Acknowledgment is made to the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for support of this research. The Camille and Henry Dreyfus Foundation is also thanked for their support of this project. The authors acknowledge the assistance of Ken Dykema in some aspects of the nonlinear programming phase of the investigation.

References

[1] THOMAS, G. F., JAVOR, F., and ROTHSTEIN, S. M., 1976, *J. chem. Phys.*, **64,** 1574.
[2] JAVOR, F., THOMAS, G. F., and ROTHSTEIN, S. M., 1977, *Int. J. Quantum Chem.*, **11,** 59.
[3] COHEN, L., and FRISHBERG, C., 1976, *Phys. Rev.* A, **13,** 927.
[4] NAKATSUJI, H., 1976, *Phys. Rev.* A, **14,** 41.
[5] FROST, A. A., 1942, *J. chem. Phys.*, **10,** 240.
[6] FROST, A. A., KELLOGG, R. E., and CURTIS, E. C., 1960, *Rev. mod. Phys.*, **32,** 313.
[7] BARTLETT, J. H., 1955, *Phys. Rev.*, **98,** 1067.
[8] KING, F. W., LEGORE, M. A., and KELLY, M. K., 1981, *J. chem. Phys.*, **75,** 809.
[9] KING, F. W., KELLY, M. K., and LEGORE, M. A., 1982, *J. chem. Phys.*, **76,** 574.
[10] GRELLAND, H. H., and ALMLÖF, J., 1982, *Int. J. Quantum Chem.*, **21,** 885.
[11] KING, F. W., NEMEC, L. G., and KELLY, M. K., 1983, *Molec. Phys.*, **50,** 1285.
[12] KING, F. W., and DALKE, B. D., 1983, *J. chem. Phys.*, **78,** 3143.
[13] KING, F. W., and POITZSCH, M. E., 1984, *Molec. Phys.*, **51,** 835.
[14] KING, F. W., KELLY, M. K., LEGORE, M. A., and POITZSCH, M. E., 1984, *Comput. Phys. Commun.*, **32,** 215.
[15] CASIDA, M., 1985, *Int. J. Quantum Chem.*, **27,** 451.
[16] THOMAS, G. F., 1986, *Int. J. Quantum Chem.*, **29,** 867.
[17] DEAL, W. J., 1989, *Int. J. Quantum Chem.*, **35,** 513.
[18] KING, F. W., DYKEMA, K. J., and DALKE, B. D., unpublished.
[19] PEKERIS, C. L., 1959, *Phys. Rev.*, **115,** 1216.
[20] CHONG, D. P., and WEINHOLD, F., 1973, *Can. J. Chem.*, **51,** 260.
[21] THAKKAR, A. J., and SMITH, JR., V. H., 1977, *Phys. Rev.* A, **15,** 1.
[22] GRELLAND, H. H., 1981, *Int. J. Quantum. Chem.*, **19,** 847.
[23] HIMMELBLAU, D. M., 1972, *Applied Nonlinear Programming* (McGraw-Hill).
[24] MCCORMICK, G. P., 1983, *Nonlinear Programming* (Wiley).
[25] ROSEN, J. B., 1960, *J. Soc. indust. appl. Math.*, **8,** 181.
[26] ROSEN, J. B., 1961, *J. Soc. indust. appl. Math.*, **9,** 514.
[27] LASDON, L. S., and WAREN, A. D., 1980, *Operat. Res.*, **28,** 1029.
[28] LASDON, L. S., and WAREN, A. D., 1978, *Design and Implementation of Optimization Software*, edited by H. J. Greenberg (Sijthoff and Noordhoff, Alphen ann der Rijn).
[29] LASDON, L. S., WAREN, A. D., JAIN, A., and RATNER, M., 1978, *ACM Trans. math. Software*, **4,** 34.
[30] HART, J. F., and HERZBERG, G., 1957, *Phys. Rev.*, **106,** 79.
[31] DRAKE, G. W. F., 1988, *Nucl. Inst. Meth. Phys. Res.* B., **31,** 7.
[32] See, for example, WINELAND, D. J., ITANO, W. M., BERGQUIST, J. C., and BOLLINGER, J. J., editors, 1988, *Trapped Ions and Laser Cooling II*, NIST Technical Note 1324.
[33] WINELAND, D. J., BOLLINGER, J. J., and ITANO, W. M., 1983, *Phys. Rev. Lett.*, **50,** 628.
[34] KING, F. W., 1986, *Phys. Rev.* A, **34,** 4543.
[35] DRAKE, G. W. F., and MAKOWSKI, A. J., 1988, *J. opt. Soc. Am.* B, **5,** 2207.
[36] DRAKE, G. W. F., 1987, *Phys. Rev. Lett.*, **59,** 1549.
[37] ZHAO, P., LAWALL, J. R., and PIPKIN, F. M., 1991, *Phys. Rev. Lett.*, **66,** 592.
[38] LICHTEN, W., SHINER, D., and ZHOU, Z.-X., 1991, *Phys. Rev.* A, **43,** 1663.
[39] HESSELS, E. A., DECK, F. J., ARCUNI, P. W., and LUNDEEN, S. R., 1990, *Phys. Rev. Lett.*, **65,** 2765.
[40] DRAKE, G. W. F., 1990, *Phys. Rev. Lett.*, **65,** 2769.