Energies of the first row atoms from quantum Monte Carlo

M.D. Brown, J.R. Trail, P. López Ríos, and R.J. Needs

Theory of Condensed Matter Group, Cavendish Laboratory,
J.J. Thomson Avenue, Cambridge, CB3 0HE, UK

(Dated: June, 2007)

All-electron variational and diffusion quantum Monte Carlo calculations of the ground state energies of the first row atoms (Li to Ne) are reported. We use trial wavefunctions of four types: single determinant Slater-Jastrow wavefunctions; multi-determinant Slater-Jastrow wavefunctions; single determinant Slater-Jastrow wavefunctions with backflow transformations; multi-determinant Slater-Jastrow wavefunctions with backflow transformations. At the diffusion quantum Monte Carlo level and using our best trial wavefunctions we recover 99% or more of the correlation energy for Li, Be, B, C, N, and Ne, 97% for O, and 98% for F.

PACS numbers:

I. INTRODUCTION

Quantum Monte Carlo (QMC) methods can yield highly accurate energies for atoms, molecules and solids,\textsuperscript{1,2} offering an alternative to quantum chemistry approaches such as Configuration Interaction and Coupled Cluster. Although the cost of QMC calculations is large, the scaling with system size is much better than for accurate quantum chemistry approaches, and QMC methods have been applied to systems with of order $10^3$ electrons.\textsuperscript{3}\textsuperscript{,}4,5,6

The two most commonly used continuum QMC methods are variational quantum Monte Carlo (VMC) and diffusion quantum Monte Carlo (DMC). Stochastic integration is used in VMC to evaluate expectation values with a trial wavefunction, to a precision which can be systematically improved by increasing the number of configurations sampled. DMC is a stochastic projector method in which a wavefunction is evolved in imaginary time towards the ground state. Exact projector methods suffer contamination into the trial wavefunction, so that $\Psi$ will be a multideterminant representation of a CSF, and backflow transformations recover a substantial additional fraction of the correlation energy at both the VMC and DMC levels. Our best wavefunctions recover 99% or more of the correlation energy at the DMC level for Li, Be, B, C, N, and Ne, and more than 97% for O and 98% for F. This represents a substantial advance upon previous DMC calculations.

II. WAVEFUNCTION FORM AND OPTIMIZATION

Our multi-CSF SJ wavefunctions with backflow transformations can be written in the form

$$\Psi(\{r_i\}, p) = e^{J(\{r_i\}, b)} \sum_{n=1}^{N_{CSF}} \alpha_n \Phi_n(\{x_i\}),$$ \hspace{1cm} (1)

where each $x_i = x_i(\{r_i\}, c)$ results from the backflow transformation of the electronic coordinates $\{r_i\}$, each $\Phi_n$ is a multideterminant representation of a CSF, and $p = (a, b, c)$ are the variable wavefunction parameters: the CSF coefficients $a$, the Jastrow parameters $b$, and the backflow parameters $c$. It is worth noting that the Jastrow factor normally introduces spin-contamination\textsuperscript{14} and the backflow transformation angular momentum contamination into the trial wavefunction, so that $\Psi$ will not normally be a spin/spatial angular momentum eigenstate. However, the VMC and DMC energies still provide upper bounds on the true ground state energies.

The general form of our Jastrow factor $J(\{r_i\}, b)$ is described in an earlier paper.\textsuperscript{15} In this study we used Jastrow factors containing electron-electron, electron-nucleus, and electron-electron-nucleus terms. Our backflow transformations also consisted of electron-electron, electron-nucleus, and electron-electron-nucleus terms, as...
described in Ref. [2]. For both the Jastrow and backflow functions, the expansion orders and spin dependencies were chosen by increasing the variational freedom until no further lowering of the VMC energy was observed. The maximum number of variable parameters in the trial wavefunctions varied from 146 for Li to 252 for B, N, C, O, and F. The number of variable parameters is therefore only weakly dependent on the atomic number.

We used the multiconfigurational Hartree-Fock (MCHF) atomic-structure package ATSP2K[13] to construct the $\Phi_\nu$ functions appearing in Eq. (1). This provides the lowest energy solution of the many-electron problem where the solution is limited to an ‘active space’ (AS) specified by an allowed set of subshell occupation numbers, or configurations. The wavefunction may formally be written as a sum of CSFs,

$$|LS\rangle_{AS} = \sum_\nu a_\nu |\nu LS\rangle,$$  \hspace{1cm} (2)

where $\nu$ specifies a configuration (for example, $1s^22s^22p^3$) and seniority number[12]. Each CSF, $|\nu LS\rangle$, is defined by a radial orbital for each subshell, and by the total spatial/spin angular momentum quantum numbers, usually denoted $2S+1L$.

An additional freedom exists because the underlying symmetry of the Hamiltonian results in degenerate solutions corresponding to different projection angular momentum eigenvalues, $L_z$ and $S_z$. These additional quantum numbers are not required to specify the total energy, the coefficients $a_\nu$, and the mean field equations for the radial orbitals, and they are not provided by the ATSP2K code. As a result it is not straightforward to obtain an explicit form for the CSFs which is suitable for use in QMC calculations. For the calculations presented here an appropriate representation of each CSF, $\Phi_\nu$, was constructed as follows.

First we note that two separate angular momentum eigenstates containing $n_1$ and $n_2$ electrons may be coupled together to generate angular momentum eigenstates containing $n_1 + n_2$ electrons[17]. Starting with single electrons and performing such a coupling iteratively for every allowed combination provides all of the total spin/angular momentum eigenstates possible for a given configuration as sums of determinants of spin-orbitals. Due care was taken to avoid redundant computation, to respect antisymmetry, and to employ the same coupling order/sign conventions as used in ATSP2K, allowing the efficient construction of each CSF as a sum of determinants of one-electron spin-orbitals.

A different wavefunction results from each choice of $L_z$ and $S_z$. The spin independence of the Hamiltonian allows a direct conversion of the sum of determinants of complex spin-orbitals into a sum of products of two determinants of real orbitals, one for spin up and one for spin down electrons. This procedure provides a representation of each CSF in the form

$$\Phi_{\nu LS} = \sum_j a_j D^j_\nu (\nu LS) D^{\dagger}_j (\nu LS),$$  \hspace{1cm} (3)

where the determinants are of real single-particle orbitals, and such that $\sum_\nu a_\nu \Phi_{\nu LS}$ has the same energy expectation value as the original MCHF wavefunction. Note that although $\Phi_{\nu LS}$ has total spatial angular momentum $L$ and projected spin angular momentum $S_z$, it is not necessarily an eigenstate of $L_z$ or $S^2$.

It should be clear that the $\Phi_\nu$ are not uniquely specified by $(\nu,2S+1L)$ due to the spherical symmetry of the underlying Hamiltonian. In what follows we make the specific choice $|L_z| = L$, as this provides the smallest multideterminant expansions, and $|S_z| = S$, as this provides that same number of spin up/down electrons as predicted by Hund’s first rule.

To define the multideterminant part of Eq. (1) we take $2S+1L$ from the lowest energy HF state, and define an AS using the following rules for allowed excitations from the HF ground state configuration: only single and double (SD) electronic excitations were allowed, excitations of the 1s core electrons were not allowed, and excitations were allowed into orbitals with quantum numbers $n \leq 7$ and $l \leq 5$. A MCHF calculation was then performed for the chosen AS, and the resulting CSFs ordered by weight, $a_\nu^2$. All but the $N_{CSF}$ largest-weight configurations were then discarded, and a MCHF calculation was performed using this reduced AS($N_{CSF}$) active space. The percentages of the correlation energies recovered at the MCHF level are shown in Fig. 1 for $N_{CSF} \leq 100$.

Li is an exception: the above rules define an active space that is equivalent to HF because of the absence of 1s excitations. In order to generate a nontrivial AS for Li we allowed excitations of all electrons into orbitals with quantum numbers $n \leq 5$ and $l \leq 5$.
We chose $N_{CSF} = 20$ for the multideterminant expansions used in the QMC calculations (corresponding to a number of determinants ranging from 83 for Li to 499 for Ne). Although the application of this selection criterion to the first row atoms is necessarily somewhat arbitrary, it is not without justification. Figure I shows that for around 20 CSFs the fraction of correlation energy recovered within MCHF is roughly equivalent from Be to Ne. Including further orbitals in the initial ‘large’ AS, including triple excitations, or allowing core excitations, did not significantly change the MCHF energies for $N_{CSF} \leq 20$, with the exceptions of Li and Be for which core excitations improved the MCHF energies significantly for small $N_{CSF}$, but had a negligible impact on the DMC results.

The variable parameters $p$ were optimized by minimizing the VMC energy, using the scheme developed very recently by Umrigar et al. This scheme is a generalization of the method of Nightingale and Melik-Alaverdian to include non-linear parameters in the trial wavefunction. In the past, variance minimization techniques were normally used because they are robust for optimizing Jastrow factors, but in fact they have been found to behave poorly when optimizing linear coefficients. The method of Nightingale and Melik-Alaverdian is very robust for optimizing linear coefficients, and Umrigar et al. have developed an energy minimization scheme which is robust for both linear and non-linear parameters. All of the wavefunction parameters were optimized simultaneously in our calculations. In some cases we obtained significantly lower VMC and DMC energies with energy minimization than with variance minimization when optimizing the CSF coefficients $a$, and the performance for the Jastrow and backflow parameters ($b$ and $c$) was also slightly improved.

### III. RESULTS

We used the CASINO code for all of our VMC and DMC calculations. The wavefunction optimizations were performed using $10^6$ statistically independent electronic configurations. The target population of walkers in the DMC calculations was about 2000 in each case, which

![FIG. 2: (Color online) The variation of the DMC energy (in Hartrees) with timestep for the C atom, using a MDBF wavefunction. The statistical error bars on the data are shown.](image-url)
are similar to those reported by other authors.

wavefunctions in DMC calculations using single determinant SJ trial of 40%, while the DMC one covers only 12%. Our best gives the same information at the DMC level.
pansion and backflow transformations were added. Fig. 4
mized wavefunctions within VMC as the multi-CSF ex-
thing the most accurate correlation energy percentages (0.0015 a.u.).
step (0.096 a.u.) to 99.8(1)% for the smallest time step
acceptance ratios varied from 93(1)% for the largest time
our calculations for the first row atoms. For C, the DMC
shown in Fig. 2, and the shape of this curve is typical of
Table I gives the VMC and DMC energies, including
the most accurate correlation energy percentages
achieved for each atom. The “exact” non-relativistic energies,
assuming a point nucleus of infinite mass, were taken from Refs. 22,23. Fig. 3 shows the percentage of
each atom’s correlation energy recovered by the optimized wavefunctions within VMC as the multi-CSF expansion
and backflow transformations were added. Fig. 3 gives the same information at the DMC level.

The general shapes of the curves in Figs. 3 and 4 are similar, but note that the VMC graph covers a range of
40%, while the DMC one covers only 12%. Our best VMC results surpass the accuracy previously achieved
in DMC calculations using single determinant SJ trial wavefunctions.22 27. Our single determinant DMC results
are similar to those reported by other authors.22 27. For Li and Be our best DMC energies are within error bars of
some previous DMC calculations, but for the other atoms our results are clearly superior. Flad et al.28 performed
multideterminant DMC calculations for the atoms B to F, but we have used larger expansions and our results
are already superior at the DMC-MD level. Casula and Sorella performed DMC calculations for the atoms Li-
Ne with correlated geminal pairing wavefunctions, but our energies are substantially better than theirs. VMC
and DMC calculations using backflow wavefunctions have been reported for Li and Ne.29 Lüchow and Fink9 per-
fomed a DMC calculation for N using a pair natural orbital configuration interaction trial wavefunction, re-
triving 98.2% of the correlation energy, which is very similar to our DMC-MD result, but our DMC-MDBF
energy is lower. Our results for Li and Be are surpassed by Hylleraas variational calculations.22 23. We should also
note the existence of high quality non-variational results such as those from explicitly correlated r12 Configuration
Interaction and Coupled Cluster methods.31,32

Each of the VMC and DMC energies for Li corresponds to recovering well over 90% of the correlation energy. The
single-determinant HF nodal surface of Li is extremely accurate, but incorporating backflow improves the DMC
energy noticeably. Backflow gives a significant improvement for Be at the VMC level, but a very modest
improvement within DMC. The ground state wavefunction for Be contains a substantial admixture of the 1s2p2
excited state configuration which significantly alters the nodal surface. The single determinant 1s2p2 HF nodal
surface divides the configuration space into four nodal pockets, while the exact wavefunction has only two nodal
pockets.33,34. It does not appear possible to achieve an accurate DMC energy for Be with a trial wavefunction
containing four nodal pockets. The continuous backflow transformations we use are incapable of changing the
number of nodal pockets and consequently the single determinant DMC energy for Be is very poor. When an appropriate
amount of the 1s2p2 configuration is included in the trial wavefunction, the DMC energy improves dra-

![Graph 3: Correlation energy percentage vs. atom](image3)

**FIG. 3:** (Color online) The percentage of the correlation energy recovered for each atom within VMC, using a single determinant SJ wavefunction (VMC) with the addition of backflow (VMC-BF), multiple determinants (VMC-MD), and both together (VMC-MDBF). The statistical error bars (not shown) are smaller than the symbols.

![Graph 4: Correlation energy percentage vs. atom](image4)

**FIG. 4:** (Color online) The percentage of the correlation energy recovered for each atom within DMC, using a single determinant SJ wavefunction (DMC) with the addition of backflow (DMC-BF), multiple determinants (DMC-MD), and both together (DMC-MDBF). The statistical error bars (not shown) are smaller than the symbols.
matically, and backflow contributes a small additional improvement.

The percentages of the correlation energy recovered for N at the VMC level from the single determinant backflow and 20 CSF wavefunctions are almost the same, while for higher atomic numbers backflow gives more correlation energy than the CSFs. Fig. 1 shows that the percentages of the correlation energy recovered within the MCHF calculations for O and N increase relatively slowly with the number of CSFs, which may be connected with the lower percentages of correlation energy recovered in the MD and MDBF calculations for these atoms. Note that, apart from Li and Be where the DMC-MD energies with 20 CSFs are already excellent, backflow adds substantially to the correlation energy recovered. It seems that backflow and multi-CSFs improve the single determinant wavefunctions in rather different ways, and therefore it can be very advantageous to combine them.

One might expect that the VMC-MDBF energies would always be lower than the VMC-BF and VMC-MD energies, as the wavefunction parameters are obtained by minimizing the VMC energy. Table 1 shows, however, that this does not hold for Ne. The most likely reason for this is that the HF and MCHF orbitals are different, so that the single determinant wavefunction does not correspond to a particular choice of the MD wavefunction parameter values. This also indicates that further reductions in the VMC energy could be obtained by optimizing the orbitals along with the Jastrow and backflow parameters and the CSF coefficients.

IV. DISCUSSION

As can be seen from Fig. 4, we recover 99% or more of the correlation energy for all the atoms except O and F. The different freedoms introduced by the multi-CSF expansion and backflow transformations are both vital in achieving the best energies, indicating that they capture different aspects of the wavefunction correlation. Backflow appears to describe parts of the dynamic correlation beyond the SJ wavefunction while the low-energy CSFs beyond the HF ground state configuration describe static correlation energy. Our results contribute to the growing evidence that adding backflow transformations to SJ wavefunctions can significantly improve the nodal surface for systems from atoms to solids.

It should be noted that, while it might be difficult to substantially improve the Jastrow factors and backflow transformations, we could easily include many more CSFs in the trial wavefunctions without a prohibitive increase in computational expense. The orbitals in the CSFs could also be optimized within VMC. The energies reported here do not therefore represent the current practical limits of the VMC and DMC methods for all-electron atoms.

For general systems, the number of CSFs required to obtain a significant improvement in the energy increases rapidly with the number of electrons \( N \). The number of required parameters in the backflow functions only increases (roughly linearly) with the number of inequivalent atoms. The cost of moving an electron in a VMC/DMC calculation scales approximately as \( N^2 \). In a VMC-BF/DMC-BF calculation the cost scales as \( N^3 \) because all of the orbitals must be evaluated at the positions \( \{ x_i \} \) when an electron is moved. The cost of a VMC-MD/DMC-MD calculation scales as \( N^2 M \), where \( M \) is the number of determinants, and the cost of a VMC-MDBF/DMC-MDBF calculation scales as \( N^3 M \). The number of moves required to obtain a fixed statistical error bar in the energy is smaller for a more accurate trial wavefunction, so that the pre-factors in the cost for the more sophisticated wavefunction forms are smaller. As an example, for the C atom the relative computational costs to achieve the same error bar are 1.0/3.6/6.2/17.8 for the VMC SD/SDBF/MD/MDBF calculations, while in DMC the corresponding cost ratios are 9.0/13.7/15.1/31.9. Overall, we expect that we can maintain the level of accuracy achieved here for small molecules.

We obtained superior results with the energy minimization scheme of Umrigar et al.\textsuperscript{19} than with variance minimization procedures.\textsuperscript{19} The energy minimization method is particularly advantageous when multiple determinants are used because it is well-suited to optimizing linear parameters, whereas variance minimization encounters some difficulties in optimizing CSF coefficients. The combination of energy minimization and the MD and BF wavefunction forms has led to substantial reductions in the VMC energies, which are very important because the variance of the DMC energy has been found to be proportional to the error in the VMC energy.\textsuperscript{35} Although reductions in the DMC energies from the improved trial wavefunctions are smaller than in VMC, they represent substantial fractions of the remaining correlation energy. Our best DMC energies for Li and Be are within error bars of previously published DMC energies, but our results for the higher atomic number atoms are markedly superior.

Acknowledgments

Financial support was provided by the Engineering and Physical Sciences Research Council (EPSRC) of the UK. Computing resources were provided by the University of Cambridge High Performance Computing Service (HPCS).

---

1 W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, Rev. Mod. Phys. 73, 33 (2001).

2 C.J. Umrigar and M.P. Nightingale, eds., Quantum Monte
Carlo Methods in Physics and Chemistry, NATO ASI Ser. C 525 (Kluwer, Dordrecht, 1999).
3. A. Szabo and N.S. Ostlund, Modern Quantum Chemistry (Dover, 1996).
4. R. Mazzono, A. Ma, M.D. Towler and R.J. Needs, Phys. Rev. Lett. 98, 025701 (2007).
5. J.B. Anderson, J. Chem. Phys. 63, 1499 (1975).
6. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. 45, 566 (1980).
7. P. López Ríos, A. Ma, N.D. Drummond, M.D. Towler and R.J. Needs, Phys. Rev. E 74, 066701 (2006).
8. R. Maezono, A. Ma, M.D. Towler and R.J. Needs, J. Chem. Phys. 124, 224104 (2006).
9. J.B. Anderson, J. Chem. Phys. 63, 1499 (1975).
10. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. 45, 566 (1980).
11. P. López Ríos, A. Ma, N.D. Drummond, M.D. Towler and R.J. Needs, Phys. Rev. E 74, 066701 (2006).
12. N.D. Drummond, P. López Ríos, A. Ma, J.R. Trail, G.G. Spink, M.D. Towler and R.J. Needs, J. Chem. Phys. 124, 224104 (2006).
13. J.B. Anderson, J. Chem. Phys. 63, 1499 (1975).
14. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. 45, 566 (1980).
15. R.P. Feynman, Phys. Rev. 94, 262 (1954).
16. R.P. Feynman and M. Cohen, Phys. Rev. 102, 1189 (1956).
17. C. Froese Fischer, T. Brage and P. Jönsson, Computational Atomic Structure: An MCHF approach (IOP Publishing Ltd, 1997).
18. G. Gaigalas, Z. Rudzikas and C. Froese Fischer, At. Data Nucl. Tables 70, 1 (1998).
19. C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella and R.G. Hennig, Phys. Rev. Lett. 98, 110201 (2007).
20. J. Toulouse and C.J. Umrigar, J. Chem. Phys. 126, 084102 (2007).
21. M.P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. 87, 043401 (2001).
22. M. Puchalski and K. Pachucki, Phys. Rev. A 73, 022503 (2006).
23. S.J. Chakravorty, S.R. Gwaltney, E.R. Davidson, F.A. Parpia and C.F. Fischer, Phys. Rev. A 47, 3649 (1993).
24. R.J. Needs, M.D. Towler, N.D. Drummond and P. López Ríos, CASINO user’s guide, version 2.0.0 (2006).
25. A. Lüchow and J.B. Anderson, J. Chem. Phys. 105, 7573 (1996).
26. P. Langfelder, S.M. Rothstein and J. Vrbik, J. Chem. Phys. 107, 8525 (1997).
27. K. Hongo, Y. Kawazoe and H. Yasuhara, Materials Transactions 47, 2612 (2006).
28. H. Flad, M. Caffarel and A. Savin, in Recent Advances in Quantum Monte Carlo Methods, edited by W.A. Lester, Jr. (World Scientific, Singapore, 1997).
29. A. Lüchow and R.F. Fink, J. Chem. Phys. 113, 8457 (2000).
30. J. Komasa, J. Rychlewski and K. Jankowski, Phys. Rev. A 65, 042507 (2002).
31. J. Noga, P. Valiron and W. Klopper, J. Chem. Phys. 115, 2022 (2001).
32. W.A. Glauser, W.R. Brown, W.A. Lester, Jr., D. Bressanini, B.L. Hammond and M.L. Koszykowski, J. Chem. Phys. 97, 9200 (1992).
33. D. Bressanini, D.M. Ceperley and P.J. Reynolds, in Recent Advances in Quantum Monte Carlo Methods, 2nd ed. edited by W.A. Lester, Jr., S.M. Rothstein and S. Tanaka (World Scientific, Singapore, 2002).
34. C.J. Umrigar, K.G. Wilson and J.W. Wilkins, Phys. Rev. Lett. 60, 1719 (1988).
35. P.R.C. Kent, R.J. Needs and G. Rajagopal, Phys. Rev. B 59, 12344 (1999).
36. R.J. Needs, M.D. Towler and R.J. Needs, Phys. Rev. B 72, 085124 (2005).
37. D.M. Ceperley, J. Stat. Phys. 43, 815 (1986).
38. M. Puchalski and K. Pachucki, Phys. Rev. A 73, 022503 (2006).
39. R.P. Feynman and M. Cohen, Phys. Rev. 102, 1189 (1956).
40. C. Froese Fischer, T. Brage and P. Jönsson, Computational Atomic Structure: An MCHF approach (IOP Publishing Ltd, 1997).
41. G. Gaigalas, Z. Rudzikas and C. Froese Fischer, At. Data Nucl. Tables 70, 1 (1998).
42. C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella and R.G. Hennig, Phys. Rev. Lett. 98, 110201 (2007).
43. J. Toulouse and C.J. Umrigar, J. Chem. Phys. 126, 084102 (2007).
44. M.P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. 87, 043401 (2001).
45. M.D. Ceperley, J. Stat. Phys. 43, 815 (1986).
46. A. Ma, N.D. Drummond, M.D. Towler and R.J. Needs, Phys. Rev. E 71, 066704 (2005).