Nodal Pulay terms for accurate diffusion quantum Monte Carlo forces

A. Badinski,[1] P. D. Haynes,[1,2] and R. J. Needs[1]
[1]*Theory of Condensed Matter Group, Cavendish Laboratory, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*
[2]*Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom*
(Received 2 August 2007; published 13 February 2008)

Exact expressions are derived for forces within the mixed and pure fixed-node diffusion quantum Monte Carlo (DMC) methods. These expressions include the "nodal terms" which arise from the discontinuity in the gradient of the DMC wave function at the nodal surface. We devise a practical scheme for estimating the nodal terms, and demonstrate that their inclusion leads to very accurate forces.

DOI: 10.1103/PhysRevB.77.085111 PACS number(s): 02.70.Ss, 31.15.V−, 71.10.−w, 71.15.−m

I. INTRODUCTION

The diffusion quantum Monte Carlo (DMC) method is the most accurate approach available for calculating the total energies of solids and large molecules.[1] Energy gradients are also of great significance in quantum mechanical calculations, because they give the forces on atoms which may be used to relax structures and perform molecular dynamics simulations. Progress in such calculations using DMC methods has, however, been held up by difficulties encountered in evaluating forces accurately and efficiently.

The DMC method is based on imaginary time evolution, which projects out the lowest energy many-body wave function.[1] The fermionic symmetry is maintained by fixing the nodal surface (the surface on which the wave function is zero and across which it changes sign) to be that of an antisymmetric trial wave function, Ψ_T. The nodal surface divides the wave function into nodal pockets, and the DMC algorithm gives the lowest energy solution Φ within each pocket.

The Hellmann-Feynman theorem (HFT) implies that the force is given by the expectation value of the gradient of the Hamiltonian with respect to the relevant parameter, λ, when the wave function is an exact eigenstate.[2,3] Standard fixed-node DMC samples the "mixed" probability distribution $\Psi_T\Phi$. It is straightforward to evaluate the HFT expression for the energy gradient within DMC, but it does not give the exact gradient of the DMC energy unless Ψ_T is exact. If Ψ_T is not exact, the correct energy gradient is obtained only when the Pulay correction terms[4] are included, which contain the gradient of the wave function with respect to λ. Mixed DMC calculations of forces including approximate Pulay terms have been reported in Refs. 5 and 6.

One approach to reducing the size of the Pulay terms is to evaluate the "pure" estimate of the HFT operator[7] by sampling the probability distribution $\Phi\Phi$. This approach still does not produce the exact gradient of the DMC energy unless the trial nodal surface is exact, because it neglects a Pulay term which can be written as an integral over the nodal surface. The existence of this nodal term in the pure estimate was pointed out in Ref. 8, and an explicit expression for it was given in Ref. 9. However, a practical scheme for evaluating this nodal term was not developed and it has been neglected in force calculations.[7,10] Here, we show that the gradients of both the mixed and pure estimates of the energy contain nodal Pulay terms, and we describe and test a practical scheme for estimating them.

This paper is organized as follows: In Sec. II, we introduce the DMC energy when integrating over a single nodal pocket. In Sec. III, we derive exact expressions for the first derivative of the DMC energy, and in Sec. IV, we give practical expressions for estimating them. In Sec. V, we present and discuss the results obtained for a test system, and we draw our conclusions in Sec. VI.

II. DIFFUSION MONTE CARLO ENERGY AND HAMILTONIAN

The fixed-node approximation is equivalent to placing an infinite potential barrier everywhere on the nodal surface of the trial wave function Ψ_T. The infinite potential barrier has no effect on the energy if the trial nodal surface is exact, but if it is inexact, the energy is always raised. It follows that the DMC energy is always greater than or equal to the exact ground state energy. In addition, it turns out that all the nodal pockets of the exact ground state are equivalent[11] and, if Ψ_T has this tiling property, we need sample only one of its pockets.

Consider a single nodal pocket Ω on whose boundary Γ the trial function Ψ_T vanishes. Assume that the tiling property holds and that $\Psi_T>0$ inside Ω. The DMC wave function Φ for the pocket can be written as

$$\Phi = \Theta(\Psi_T)\Phi_E, \tag{1}$$

where Θ is the Heaviside function, and the envelope function Φ_E also vanishes on Γ. Φ_E is positive within Ω, is differentiable (at least up to second order) throughout Ω and on Γ, and satisfies the equation

$$\Theta(\Psi_T)\hat{H}\Phi_E = \Theta(\Psi_T)E_D\Phi_E \tag{2}$$

throughout space. Thus, $\Phi=\Phi_E$ throughout Ω and on Γ, $\Phi=0$ outside Ω, and Φ has a kink at the nodal surface, as illustrated in Fig. 1.

To obtain an expression for $\hat{H}\Phi$, we use the Laplacian of Φ,

$$\nabla^2\Phi = \Theta(\Psi_T)\nabla^2\Phi_E + 2\delta(\Psi_T)\,\nabla\,\Phi_E\cdot\nabla\Psi_T + \delta(\Psi_T)\Phi_E\nabla^2\Psi_T + \delta'(\Psi_T)\Phi_E|\nabla\Psi_T|^2, \tag{3}$$

where δ is the Dirac delta function and δ' is its first derivative. Since Ψ_T is positive within Ω, on Γ, $\nabla\Psi_T$ is parallel to

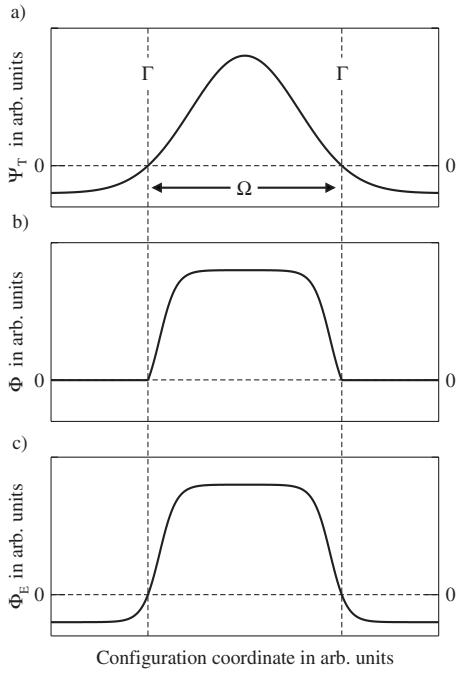

FIG. 1. Schematic illustration of wave functions Ψ_T, Φ, and Φ_E (arb. units) plotted over a one-dimensional electron configuration coordinate (arb. units): (a) the trial wave function Ψ_T whose zeros define the nodal surfaces Γ and, therefore, the nodal pocket Ω; (b) the corresponding DMC wave function Φ for the nodal pocket Ω where Φ is nonzero inside Ω, zero outside Ω, and its first derivative is discontinuous at Γ; and (c) a wave function Φ_E constructed to equal Φ inside Ω, but which is differentiable (at least up to second order) on Γ.

the normal vector to the nodal surface pointing *into* Ω. As the gradients $\nabla\Phi_E$ and $\nabla\Psi_T$ must be parallel on the nodal surface, it follows that $\nabla\Phi_E \cdot \nabla\Psi_T$ can be replaced by $|\nabla\Phi_E||\nabla\Psi_T|$. The limiting behavior of Φ_E on approaching Γ from inside the pocket can be found using these results and a multidimensional generalization of l'Hôpital's rule, giving

$$\frac{\nabla\Phi_E}{\Phi_E} = \frac{\nabla\Psi_T}{\Psi_T} \quad \text{on } \Gamma. \tag{4}$$

Using Eqs. (3) and (4) and the standard identity

$$x^n \frac{d^n \delta(x)}{dx^n} = (-1)^n n! \delta(x), \tag{5}$$

which is valid inside integrals, we find

$$\hat{H}\Phi = \Theta(\Psi_T)\hat{H}\Phi_E - \frac{1}{2}\delta(\Psi_T)\left[\frac{|\nabla\Psi_T|^2}{\Psi_T} + \nabla^2\Psi_T\right]\Phi_E. \tag{6}$$

The term $-\frac{1}{2}\delta(\Psi_T)(\nabla^2\Psi_T)\Phi_E$ does not contribute to the expressions that follow because $(\nabla^2\Psi_T)\Phi_E$ is zero on the nodal surface where the δ function is evaluated.

Using Eqs. (1), (2), and (6), one can define an effective Hamiltonian including the δ function term from Eq. (6), in which the δ function term could be interpreted as an infinite "nodal potential" causing the wave function Φ to vanish on the nodal surface. We have, however, omitted such a definition since it is not relevant here.

The DMC energy for the nodal pocket is written as

$$E_D = \frac{\int \Psi \hat{H} \Phi dV}{\int \Psi \Phi dV}, \tag{7}$$

where both Ψ and Φ vanish on Γ. This expression includes both the mixed ($\Psi = \Psi_T$) and pure ($\Psi = \Phi_E$) estimates, which give the same DMC energy. The δ function term in Eq. (6) is nonzero only on Γ, where Ψ vanishes, and, therefore, gives no contribution to Eq. (7). As the integrand in Eq. (7) is zero outside the nodal pocket, we extend the region of integration in Eq. (7) throughout space. As a result of this and the specific construction of Φ, one can show that $\hat{H}$ is Hermitian in the following equations. This would generally not be the case if the region of integration were limited to Ω (see Ref. 9).

III. DERIVATIVE OF THE DIFFUSION MONTE CARLO ENERGY

We consider now a general parameter λ, e.g., a nuclear coordinate or an electric field, which is used to vary the Hamiltonian and upon which both the nodal surface (i.e., Ψ_T) and wave function Φ depend. The derivative of the DMC energy with respect to λ is

$$\frac{dE_D}{d\lambda} = \frac{\int \Psi \frac{d\hat{H}}{d\lambda}\Phi dV}{\int \Psi\Phi dV} + \frac{\int \frac{d\Psi}{d\lambda}(\hat{H} - E_D)\Phi dV}{\int \Psi\Phi dV} + \frac{\int \Psi(\hat{H} - E_D)\frac{d\Phi}{d\lambda} dV}{\int \Psi\Phi dV}. \tag{8}$$

We now show that the second term on the right hand side of Eq. (8) can be written as an integral over the nodal surface. Using Eqs. (2) and (6), which are valid throughout space, and the standard identity,

$$\delta(\Psi_T(\mathbf{r})) = \int_\Gamma \frac{\delta(\mathbf{r} - \mathbf{r}')}{|\nabla\Psi_T(\mathbf{r}')|} dS', \tag{9}$$

we obtain

$$-\frac{1}{2}\frac{\int_\Gamma \frac{d\Psi(\mathbf{r}')}{d\lambda}|\nabla\Psi_T(\mathbf{r}')|\frac{\Phi_E(\mathbf{r}')}{\Psi_T(\mathbf{r}')} dS'}{\int \Psi\Phi dV}. \tag{10}$$

With Eq. (4), we then find

$$\frac{\int \frac{d\Psi}{d\lambda}(\hat{H}-E_D)\Phi dV}{\int \Psi\Phi dV} = -\frac{1}{2}\frac{\int_\Gamma |\nabla\Phi_E|\frac{d\Psi}{d\lambda}dS}{\int \Psi\Phi dV}. \quad (11)$$

We now use Eqs. (8) and (11), and the fact that $\hat{H}$ is Hermitian to obtain expressions for the energy gradient within mixed and pure DMC. For the mixed estimator, we set $\Psi = \Psi_T$ and obtain

$$\frac{dE_D}{d\lambda} = \frac{\int \Psi_T \frac{d\hat{H}}{d\lambda}\Phi dV}{\int \Psi_T\Phi dV} - \frac{1}{2}\frac{\int_\Gamma |\nabla\Phi_E|\frac{d\Psi_T}{d\lambda}dS}{\int \Psi_T\Phi dV} + \frac{\int \frac{d\Phi}{d\lambda}(\hat{H}-E_D)\Psi_T dV}{\int \Psi_T\Phi dV}. \quad (12)$$

For the pure estimator, we set $\Psi=\Phi_E$ in Eqs. (8) and (11), and obtain

$$\frac{dE_D}{d\lambda} = \frac{\int \Phi_E \frac{d\hat{H}}{d\lambda}\Phi dV}{\int \Phi_E\Phi dV} - \frac{1}{2}\frac{\int_\Gamma \frac{|\nabla\Phi_E|^2}{|\nabla\Phi_E|}\frac{d\Phi_E}{d\lambda}dS}{\int \Phi_E\Phi dV}. \quad (13)$$

If the trial nodal surface is independent of λ, then $d\Phi_E/d\lambda = d\Psi_T/d\lambda = 0$ on the nodal surface and the nodal terms in Eqs. (12) and (13) are zero. Equation (13) is equivalent to Eq. (8) of Ref. 9.

The nodal term in Eq. (13) has a simple physical interpretation. We can identify $\frac{d\Phi_E}{d\lambda}/|\nabla\Phi_E|$ evaluated at a point on the nodal surface as the rate at which the nodal surface of Φ_E moves as λ is varied, and $\frac{1}{2}|\nabla\Phi_E|^2$ evaluated on the nodal surface as the kinetic energy density. Therefore, the nodal surface integral is the rate of change of the kinetic energy arising from the kink in Φ at the nodal surface.

IV. ESTIMATION OF THE NODAL TERMS

It is not clear how to evaluate the nodal terms by a Monte Carlo integration over the nodal surface, but we can make progress by expressing them as volume integrals. For mixed DMC, we start from Eq. (11), and use the Hermiticity to act as $\hat{H}$ to the left in the second term on the right hand side of Eq. (8), obtaining

$$\frac{\int \Psi_T\Phi\frac{1}{\Psi_T}(\hat{H}-E_D)\frac{d\Psi_T}{d\lambda}dV}{\int \Psi_T\Phi dV}, \quad (14)$$

which can be evaluated straightforwardly as an average over the mixed distribution, $\Psi_T\Phi$.

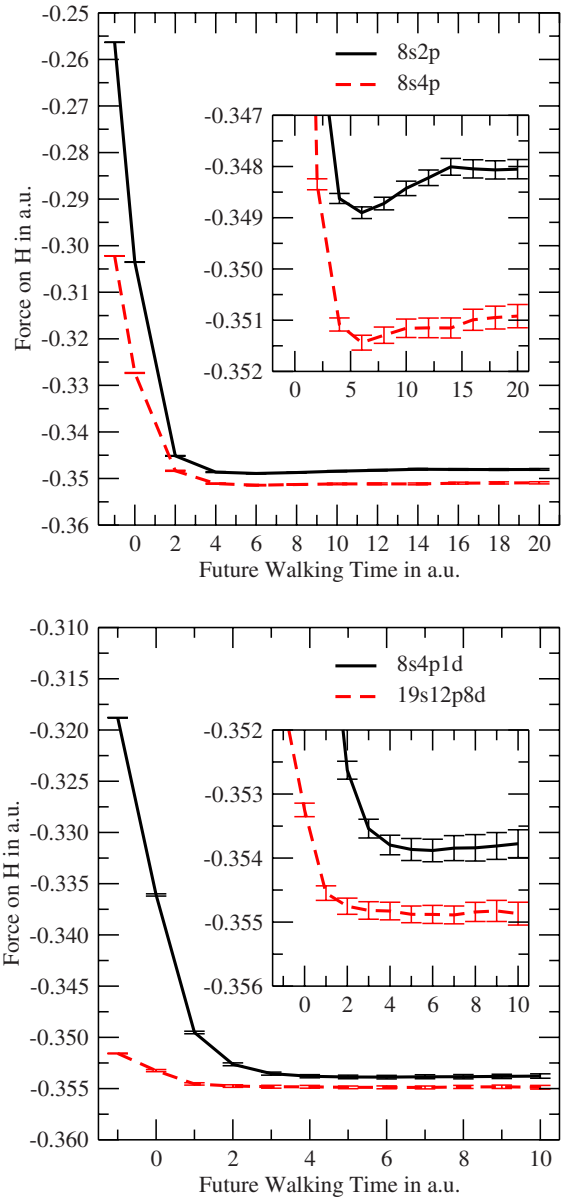

FIG. 2. (Color online) The graphs show the future walking HFT forces (a.u.) on the H atom of the GeH molecule plotted against the future walking projection time. The bond length is 2.929 a.u. The mixed DMC forces correspond to the zero future walking projection time, and the VMC forces are plotted at −1 a.u. to guide the reader. Results for different basis sets are presented. The upper graph contains results obtained with the very poor $8s2p$ basis (solid line) and with the poor $8s4p$ basis (dashed line). The lower graph contains results obtained with the intermediate $8s4p1d$ basis (solid line) and with the very good $19s12p8d$ basis (dashed line). Note that the range of the projection time is doubled for the poor and very poor basis sets.

There is, however, an additional problem in mixed DMC, because the third term on the right hand side of Eq. (12), which is a volume term, cannot be evaluated straightforwardly because the quantity being averaged involves the derivative of the DMC wave function $d\Phi/d\lambda$. We can, how-

ever, use the approximation introduced by Reynolds *et al.*,[5] and used subsequently in Ref. 6, that

$$\frac{1}{\Phi}\frac{d\Phi}{d\lambda} \simeq \frac{1}{\Psi_T}\frac{d\Psi_T}{d\lambda}, \tag{15}$$

which introduces an error of first order in $(\Psi_T - \Phi)$.

To evaluate the pure DMC nodal term, we first note that the nodal surfaces of Φ_E and Ψ_T must remain coincident for all values of λ, and hence,

$$\frac{1}{|\nabla\Phi_E|}\frac{d\Phi_E}{d\lambda} = \frac{1}{|\nabla\Psi_T|}\frac{d\Psi_T}{d\lambda} \quad \text{on } \Gamma. \tag{16}$$

Using identities (4) and (16), one can rewrite the mixed and pure DMC nodal terms of Eq. (11) as

$$-\frac{1}{2}\frac{\int_\Gamma \Psi\Phi\frac{|\nabla\Psi_T|}{\Psi_T}\frac{1}{\Psi_T}\frac{d\Psi_T}{d\lambda}dS}{\int \Psi\Phi dV}, \tag{17}$$

with $\Psi=\Psi_T$ for mixed and $\Psi=\Phi_E$ for pure DMC. This expression shows that the mixed and pure DMC nodal terms can be written as averages of a quantity involving only Ψ_T.

The pure nodal term of Eq. (17) with $\Psi=\Phi_E$ cannot be evaluated straightforwardly, but it may be evaluated approximately using extrapolation[12] or another suitable method. The standard extrapolation formula relates the pure (P), mixed (M), and variational (V) estimates by $P \simeq 2M - V$.

We must, therefore, evaluate the expression for the variational nodal term,

$$\frac{\int_\Gamma \Psi_T\Psi_T\frac{|\nabla\Psi_T|}{\Psi_T}\frac{1}{\Psi_T}\frac{d\Psi_T}{d\lambda}dS}{\int \Psi_T\Psi_T dV}. \tag{18}$$

Since both $d\Psi_T/d\lambda$ and $|\nabla\Psi_T|$ are continuous on the nodal surface, the surface integrals for two bordering pockets must cancel. Therefore, the integral over the nodal surface of all nodal pockets must be zero. However, since we assumed that the tiling property holds, so that each nodal pocket is equivalent, the surface integral in Eq. (18) is also zero for each nodal pocket. We explicitly tested this result, as reported in Appendix C. As the variational nodal term is zero, the pure DMC nodal term can be approximated as twice the mixed DMC nodal term. The extrapolation approximation introduces an error of second order in $(\Psi_T - \Phi)$.

Our final expression for the pure DMC energy derivative is then

$$\frac{dE_D}{d\lambda} \simeq \frac{\int \Phi\frac{d\hat{H}}{d\lambda}\Phi dV}{\int \Phi\Phi dV} + 2\frac{\int \Psi_T\Phi\frac{1}{\Psi_T}(\hat{H}-E_D)\frac{d\Psi_T}{d\lambda}dV}{\int \Psi_T\Phi dV}. \tag{19}$$

Note that the first term in Eq. (19) may be written as an average over the pure distribution, $\Phi\Phi$, while the second is an average over the mixed distribution, $\Psi_T\Phi$. Practical methods for sampling the pure distribution, such as future walking,[17] may also provide the mixed distribution, in which case the occurrence of both distributions in Eq. (19) is not a significant disadvantage. Equation (19) contains an error of order $(\Psi_T - \Phi)^2$ from the use of extrapolation. In practice, it may be necessary to use approximations to $d\Psi_T/d\lambda$, and any error in this quantity may result in a first order error in the nodal term. The nodal term is, however, expected to be small, in which case the resulting first order error should be small.

V. TESTS ON GERMANIUM HYDRIDE

We have evaluated the forces on the atoms of a germanium hydride dimer (GeH). As the nodal terms arise from the kinetic energy, it is sufficient to consider a model without electron-electron interactions. Neglecting this interaction substantially reduces the equilibrium bond length, but the kinetic energy is still realistic. The advantage of studying a noninteracting system is that we can control the quality of the trial wave function more systematically. We have chosen basis sets to cover the range of errors in the trial wave function that might be encountered in practical calculations for the fully interacting system. We use single-determinant trial wave functions obtained with four different Gaussian basis sets. These vary in quality from very good (19*s*, 12*p*, 8*d* functions), intermediate (8*s*4*p*1*d*), poor (8*s*4*p*), to very poor (8*s*2*p*).

We used local pseudopotentials to represent the Ge^{4+} and H^+ ions. The Ge pseudopotential was taken as 1/4 of the *s* component and 3/4 of the *p* component of the Hartree-Fock pseudopotential from Ref. 13, while we used the *s* component of the H pseudopotential. The noninteracting orbitals were calculated using the GAMESS-US (Ref. 15) code. The variational, mixed, and pure DMC forces are defined by specifying the trial wave function Ψ_T and its derivative $d\Psi_T/d\lambda$. In our calculations, we have specified $\Psi_T(\lambda)$ by precisely minimizing the variational energy within the given basis. Because the molecular orbital coefficients are calculated by minimizing the energy, their variation with λ does not contribute to the first derivative of the variational energy, and we have neglected their influence on the mixed and pure DMC forces. We have, therefore, evaluated $d\Psi_T/d\lambda$ from the derivatives of the Gaussian basis functions with respect to the bond length. This approach to estimating $d\Psi_T/d\lambda$ is inexpensive and can be applied to large systems. It can also be readily extended to interacting systems by including the derivatives of the ion-centered terms in the Jastrow factor.

We performed fixed-node DMC calculations using the CA-

SINO code.[14] For a bond length of 2.929 a.u., which is close to the experimental value of 3.003 a.u.,[16] the errors in the DMC energies are approximately zero for the most accurate trial wave function, 0.002 35(2) a.u. for the intermediate trial wave function, 0.005 66(2) a.u. for the poor trial wave function, and 0.009 25(1) a.u. for the very poor trial wave function. The variational Monte Carlo (VMC) errors are roughly five times larger than the corresponding DMC errors. Pure estimates were evaluated using extrapolation[12] and future walking.[17] Tests of the convergence of the future walking results are described in Appendix A.

It is useful to think about the error in a force in terms of the associated error in the bond length. To convert errors in forces to errors in bond lengths, we use the experimental vibrational frequency of GeH of 1908 cm^{-1} to obtain the force constant, finding that an error in the magnitude of the force of $\Delta F=0.001$ a.u. is equivalent to an error in the bond length of $\Delta a=0.0073$ a.u. It is worth noting that the equilibrium bond lengths obtained in less costly methods, such as density functional theory (DFT), are often rather accurate. However, the error in DFT calculations depends on the functional used and the system studied. A reasonable value for the error in the equilibrium bond length which might be obtained in a high quality DFT calculation for an *sp*-bonded system is about 0.01 a.u.[18] A useful DMC calculation of equilibrium bond lengths should, therefore, surpass this level of accuracy. DFT results often give poor descriptions of bond breaking, and for such a process, it is likely that substantially greater accuracy can be achieved with DMC calculations. In any case, it is important to have DMC forces which represent accurate derivatives of the DMC energy.

Table I gives results for the various terms in the forces on the H and Ge atoms. The accurate energy gradients $-dE_D/d\lambda$ were obtained from a set of DMC energy calculations at different bond lengths, as described in Appendix B. The mixed estimates of the HFT operator (A) give very poor forces for all but the most accurate trial wave function, even the intermediate trial wave function gives a bond-length error of $\Delta a=0.14$ a.u. from the force on the H atom and $\Delta a=0.09$ a.u. from the force on the Ge atom. Note that the sum of forces on the H and Ge atoms is not zero unless the HFT and Pulay forces are evaluated exactly. The nodal term (B) is larger for the poorer trial wave functions, but the Pulay terms are dominated by the volume term (C). The mixed estimate of the total force ($A+B+C$) is much more accurate than the mixed HFT term alone, but for the Ge atom with the poorest trial wave function, Δa is still 0.034 a.u.

The pure estimates of the HFT force from extrapolation and future walking (D_e and D_f) differ significantly for the poorer trial wave functions, indicating that the higher order terms absent in the extrapolation are not negligible. Overall, the pure estimates of the HFT force (without Pulay corrections) are of similar quality to the total mixed estimates of force (including Pulay terms). The nodal terms in the pure estimates ($2B$) are quite significant, in the worst case (poor trial wave function and the Ge atom) amounting to $\Delta a=0.061$ a.u. Including the nodal term substantially improves the pure estimates of the total forces (D_e+2B and D_f+2B) on both the H and Ge atoms. In the worst cases, the errors amount to $\Delta a=-0.023$ a.u. for the extrapolated HFT estimate and only $\Delta a=0.007$ a.u. for the future walking estimate. The future walking estimates of the pure DMC forces obtained using our formula in Eq. (19) are, therefore, useful by the criterion suggested above (error in bond length of less than 0.01 a.u.), for each of the four wave functions used. The estimates for forces obtained using the previously available expressions are, however, only satisfactory for trial wave functions of higher quality than those which are normally available in quantum Monte Carlo calculations.

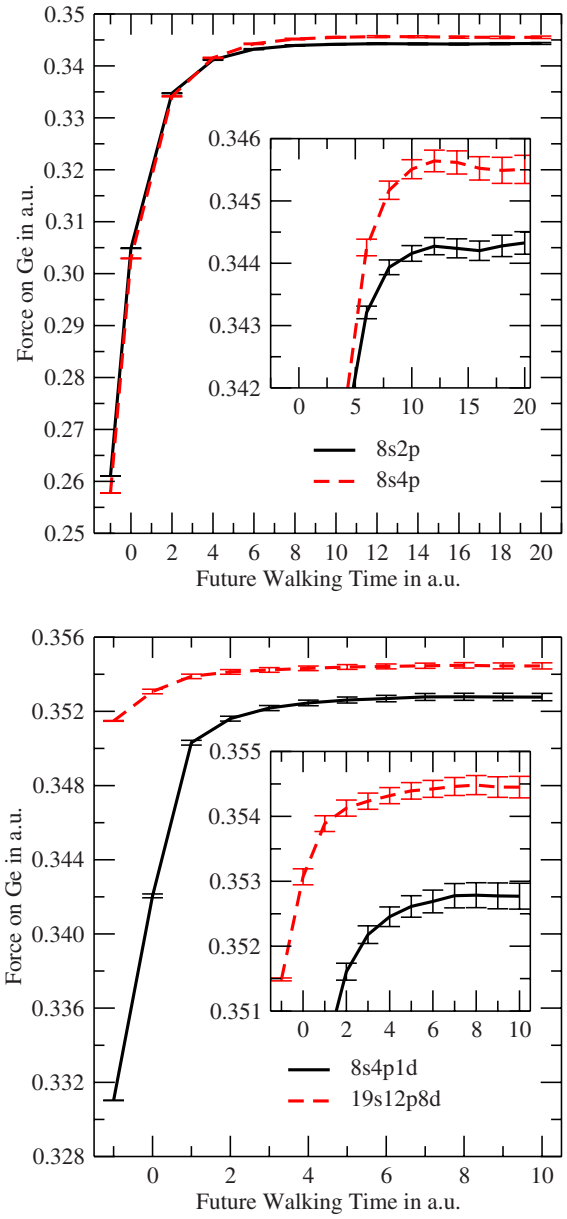

FIG. 3. (Color online) Future walking HFT forces (a.u.) on the Ge atom. The other information is the same as described in the caption of Fig. 2.

VI. CONCLUSIONS

In conclusion, we have obtained exact expressions for the forces within mixed and pure DMC. These expressions con-

tain terms which can be written as integrals over the nodal surface. The mixed DMC nodal term can be evaluated straightforwardly, and the pure DMC nodal term is approximately twice the mixed one. Using these results, we have obtained an expression for the total force within the pure DMC method [Eq. (19)] containing an approximate nodal term that can be evaluated straightforwardly. The nodal term has not been included in previous pure estimate force calculations. The extrapolation we have used in evaluating the pure nodal term leads to an error of second order in the error in the trial wave function. Tests for the GeH dimer have shown that the pure nodal term can be significant and that including it can significantly improve the pure estimate of the force. We also calculated the total forces with the mixed DMC method relying on the first order approximation of Reynolds *et al.*[5] of Eq. (15). Our tests indicate that this approximation for the mixed estimation of the force is more severe than the second order approximation used for the pure estimation. As mentioned above, we have used very high quality derivatives of the trial wave function to evaluate the Pulay terms, but if the quality were lower, it would be even more advantageous to use the pure estimate, as it depends much less sensitively on the estimate of $d\Psi_T/d\lambda$. Pure DMC estimates are more costly to evaluate than mixed ones, but our results indicate that, when the nodal term is included, they can give very accurate forces.

ACKNOWLEDGMENTS

We are grateful to John Trail, Daniel Corbett, and Zoltan Radnai for helpful discussions. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) of the United Kingdom and the Royal Society. Computing resources were provided by the University of Cambridge High Performance Computing Service (HPCS).

APPENDIX A: FUTURE WALKING DATA

In this appendix, we demonstrate that our future walking HFT estimates are well converged with respect to the future walking projection time. In Fig. 2, future walking HFT forces on the H atom of the GeH molecule are plotted against the future walking projection time. The HFT forces at zero projection time correspond to the mixed DMC estimates, and the VMC HFT forces are plotted at −1 a.u. to guide the reader. Although the future walking estimates are only formally exact for an infinite future walking projection time, we find that a relatively small future walking projection time is sufficient to obtained well-converged results. When using the very poor or poor basis set, the future walking estimates reach a plateau for projection times larger than 15 a.u. (upper graph in Fig. 2). For the intermediate or very good basis set, a plateau occurs for projection times larger than 5 a.u. (lower graph in Fig. 2). When looking at the corresponding two graphs for the future walking HFT forces on the Ge atom in the GeH molecule in Fig. 3, we see a plateau that starts at around 12 a.u. for the two poor basis sets, and a plateau that starts at 7 a.u. for the intermediate and the very good basis sets. These findings suggest that poorer basis sets require longer future walking projection times. To account for the different convergence behaviors of the future walking HFT estimates given in Table I, we choose a projection time of 10 a.u. for the intermediate and very good basis sets, and 20 a.u. for the two poor basis sets.

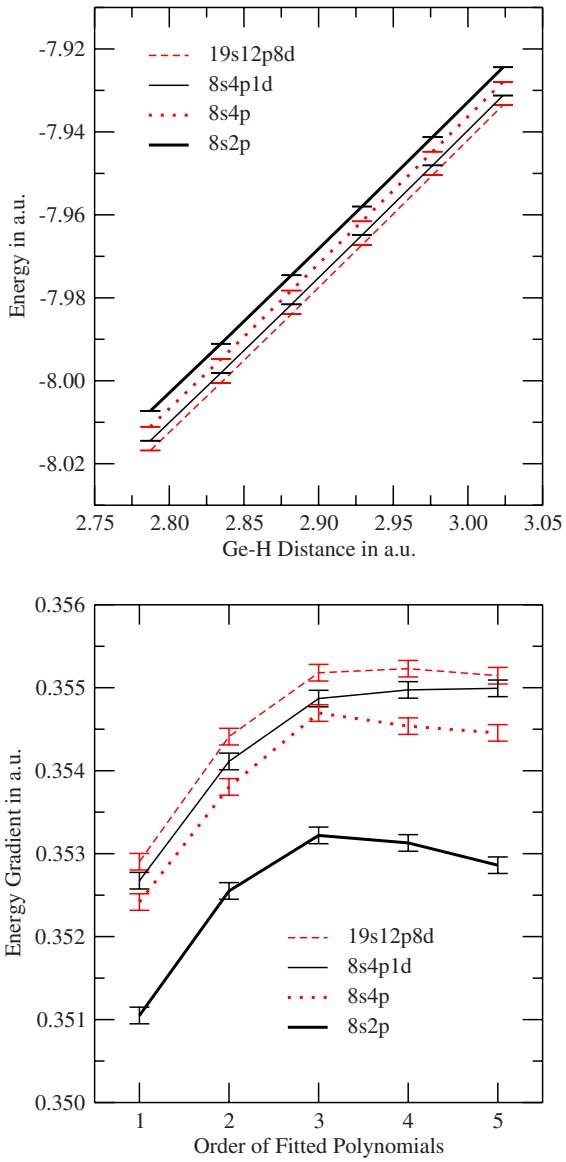

FIG. 4. (Color online) The upper graph shows the DMC total energy (a.u.) evaluated at six different geometries for the four different basis sets. In this graph, a cubic polynomial form is fitted to the six energy data points for each basis set. The lower graph shows the dependency of the gradient of the energy on the underlying fitted polynomial form. The gradient of the energy is obtained from the derivative of the polynomial evaluated at 2.929 a.u. The graph shows the gradient of the energy in a.u. against the order of the fitted polynomials.

APPENDIX B: ESTIMATING THE EXACT ENERGY GRADIENT

To obtain the energy gradients $-dE_D/d\lambda$ in Table I of the main text, we calculated the DMC total energy at bond

TABLE I. Terms in the forces (a.u.) on the H and Ge atoms of the GeH dimer evaluated at a bond length of 2.929 a.u. The mixed estimate of the HFT term [first term of the right hand side of Eq. (12)] is denoted by A, while B denotes the nodal term evaluated using Eq. (14), and C is the third term of Eq. (12) evaluated with the approximation of Reynolds *et al.* (Ref. 5) [Eq. (15)]. $A+B+C$ is the mixed estimate of the total force. The pure estimate of the HFT term [first term of the right hand side of Eq. (19)] is denoted by D_e when estimated with the extrapolation method, and denoted by D_f when calculated with the future walking method. $2B$ is the estimated pure nodal term [second term of the right hand side of Eq. (19)]. D_e+2B and D_f+2B are the pure estimates of the total force. Accurate DMC forces for each trial wave function are given by $-dE_D/d\lambda$.

	Ψ_T	A	B	C	$A+B+C$	D_e	D_f	D_e+2B	D_f+2B	$-dE_D/d\lambda$
H	$19s12p8d$	−0.3532(1)	0.0001(1)	−0.0015(0)	−0.3547(2)	−0.3549(2)	−0.3549(1)	−0.3548(3)	−0.3547(3)	−0.3552(1)
	$8s4p1d$	−0.3358(1)	−0.0009(1)	−0.0180(0)	−0.3546(2)	−0.3534(1)	−0.3535(2)	−0.3551(3)	−0.3552(3)	−0.3549(1)
	$8s4p$	−0.3271(1)	−0.0020(1)	−0.0310(0)	−0.3601(1)	−0.3520(1)	−0.3509(2)	−0.3560(3)	−0.3549(3)	−0.3546(1)
	$8s2p$	−0.3035(1)	−0.0025(1)	−0.0491(0)	−0.3550(1)	−0.3506(1)	−0.3481(2)	−0.3555(2)	−0.3530(3)	−0.3532(1)
Ge	$19s12p8d$	0.3531(1)	0.0001(1)	0.0021(0)	0.3552(2)	0.3547(1)	0.3545(1)	0.3548(3)	0.3547(3)	0.3552(1)
	$8s4p1d$	0.3421(1)	0.0010(1)	0.0089(0)	0.3520(1)	0.3531(1)	0.3527(2)	0.3552(2)	0.3548(3)	0.3549(1)
	$8s4p$	0.3027(1)	0.0042(1)	0.0517(0)	0.3586(1)	0.3476(1)	0.3455(2)	0.3560(2)	0.3539(3)	0.3546(1)
	$8s2p$	0.3048(1)	0.0039(1)	0.0398(0)	0.3485(1)	0.3486(1)	0.3443(2)	0.3564(2)	0.3522(2)	0.3532(1)

lengths of 2.787, 2.835, 2.882, 2.929, 2.976, and 3.024 a.u. The total energies are plotted in the upper graph of Fig. 4. We fitted these energies to polynomial forms and evaluated the energy gradients as the derivatives of the polynomials at 2.929 a.u. To investigate the sensitivity of the energy gradients to the quality of the underlying fitted polynomial forms, we tested polynomials from first to fifth order. The lower graph of Fig. 4 shows the energy gradients for the four basis sets against the order of the fitted polynomials. We see that, for the very good and intermediate basis sets, the differences between the gradients obtained from polynomials of third and higher order are very small. For the two poor basis sets, the gradients change slightly between the third and fifth orders by a little more than one statistical error bar. We noted that while the third and fourth order fits are almost identical, the fifth order fit showed some small oscillations, indicating that overfitting had occurred. We, therefore, decided to obtain our best estimates by averaging the energy gradients obtained from the third and fourth order polynomials.

APPENDIX C: NODAL TERM IN VARIATIONAL MONTE CARLO

We verified the statement that the surface integral in Eq. (18) is zero for each nodal pocket by evaluating an equivalent expression to Eq. (18) within VMC. To do so, we transform the numerator of Eq. (18) as

$$\begin{aligned}\int_\Gamma |\nabla\Psi_T|\frac{d\Psi_T}{d\lambda}dS &= -\int_\Gamma \frac{d\Psi_T}{d\lambda}\nabla\Psi_T\cdot d\mathbf{S}\\ &= -\int_\Omega \nabla\cdot\left[\frac{d\Psi_T}{d\lambda}\nabla\Psi_T\right]dV\\ &= -\int_\Omega \Psi_T\Psi_T\left[\frac{1}{\Psi_T}\frac{d\Psi_T}{d\lambda}\frac{\nabla^2\Psi_T}{\Psi_T}\right.\\ &\quad\left.+\frac{\nabla\Psi_T}{\Psi_T}\cdot\frac{1}{\Psi_T}\nabla\frac{d\Psi_T}{d\lambda}\right]dV, \qquad \text{(C1)}\end{aligned}$$

where all quantities can straightforwardly be calculated in VMC. We evaluated $d\Psi_T/d\lambda$ via finite differences by using a very small displacement of the nuclei. For three different single particle Gaussian basis sets $8s4p$, $8s4p1d$, and $19s12p8d$, we calculated the VMC nodal term on the H and Ge atoms of the GeH molecule at a bond length of 2.929 a.u. For all three basis sets, we find that the VMC nodal terms are zero within or close to one standard error of 0.0005 a.u. when sampling all nodal pockets. When the sampling is limited to a single nodal pocket, the VMC nodal term remains zero within or close to one standard error of 0.0006 a.u.

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[2] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), p. 285.
[3] R. P. Feynman, Phys. Rev. **56**, 340 (1939).
[4] P. Pulay, Mol. Phys. **17**, 197 (1969).
[5] P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester, Jr., Int. J. Quantum Chem. **29**, 589 (1986).
[6] R. Assaraf and M. Caffarel, J. Chem. Phys. **113**, 4028 (2000); R. Assaraf and M. Caffarel, *ibid.* **119**, 10536 (2003); M. Casalegno, M. Mella, and A. M. Rappe, *ibid.* **118**, 7193 (2003); M. W. Lee, M. Mella, and A. Rappe, *ibid.* **122**, 244103 (2005).
[7] S. Chiesa, D. M. Ceperley, and S. Zhang, Phys. Rev. Lett. **94**, 036404 (2005).
[8] K. C. Huang, R. J. Needs, and G. Rajagopal, J. Chem. Phys. **112**, 4419 (2000).
[9] F. Schautz and H.-J. Flad, J. Chem. Phys. **112**, 4421 (2000).
[10] A. Badinski and R. J. Needs, Phys. Rev. E **76**, 036707 (2007).

[11] D. M. Ceperley, J. Stat. Phys. **63**, 1237 (1991); see also W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, Phys. Rev. B **60**, 4558 (1999).
[12] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer, Berlin, 1979), p. 145.
[13] J. R. Trail and R. J. Needs, J. Chem. Phys. **122**, 174109 (2005); **122**, 014112 (2005); see also www.tcm.phy.cam.ac.uk/~mdt26/casino2_pseudopotentials.html
[14] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, *CASINO2.1 User's Manual* (University of Cambridge, Cambridge, 2007).
[15] M. W. Schmidt, K. K. Baldridge, and J. A. Boatz, J. Comput. Chem. **14**, 1347 (1993).
[16] L. Klynning and B. Lindgren, Ark. Fys. **32**, 575 (1966).
[17] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Comput. Phys. **96**, 258 (1991).
[18] A. Neugebauer and G. Hafelinger, THEOCHEM **578**, 229 (2002).