



Accurate Forces in Diffusion Monte Carlo - including Nodal Pulay Terms

Alexander Badinski¹, Peter Haynes^{1,2}, Richard Needs²

¹Cavendish Laboratory, University of Cambridge, UK

²Physics Department, Imperial College London, UK



“actual” DMC Wave Function

Fixed-node approximation to solve the fermion sign problem
 \Leftrightarrow infinite potential barrier where $\Psi_T = 0$

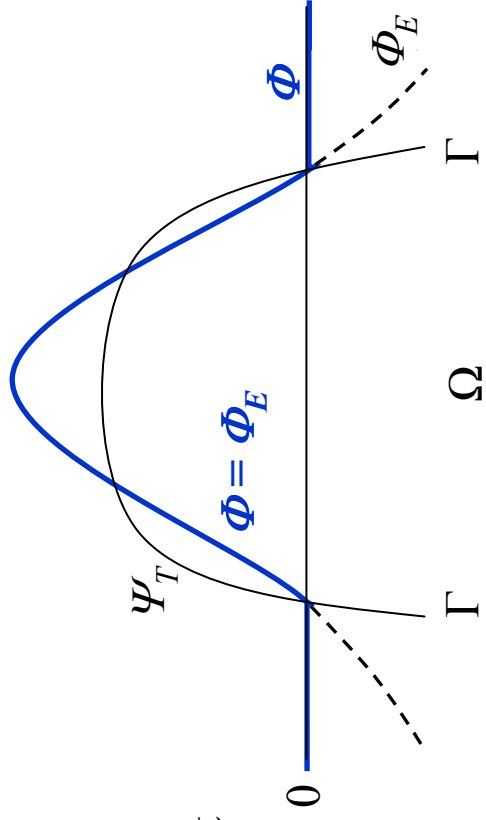
Tiling Theorem¹ \Rightarrow nodal pockets of the exact ground state are equivalent
 need to sample only one pocket

Ψ_T is trial wave function
 Ω is a single nodal pocket
 Γ is boundary of Ω where $\Psi_T = 0$

DMC wave function for one nodal pocket
 valid throughout space

$$\Phi = \Theta(\Psi_T) \Phi_E$$

where Φ_E is differentiable and
 $\Theta(\Psi_T) = 1$ for $\Psi_T > 0$ and
 $\Theta(\Psi_T) = 0$ for $\Psi_T < 0$ and



Note: $\nabla \Phi_E$ and $\nabla \Psi_T$ are parallel $\Rightarrow \nabla \Phi_E \cdot \nabla \Psi_T = |\nabla \Phi_E| |\nabla \Psi_T|$ (1)

¹ D.M. Ceperley, J. Stat. Phys. Vol. 63 1237 (1991)

δ - Function Term

Applying the Laplacian ∇^2 onto Φ

$$\nabla^2 \Phi = \Theta(\Psi_T) \nabla^2 \Phi_E + 2\delta(\Psi_T) \nabla \Phi_E \cdot \nabla \Psi_T + \delta(\Psi_T) \Phi_E \nabla^2 \Psi_T + \delta'(\Psi_T) \Phi_E |\nabla \Psi_T|^2$$

Using $\nabla \Phi_E \cdot \nabla \Psi_T = |\nabla \Phi_E| |\nabla \Psi_T|$ from Eq. (1),

the standard identity
$$x \frac{d\delta(x)}{dx} = -\delta(x) \tag{2}$$

and L'Hôpital in multi-dim.
$$\frac{\Phi_E}{\Psi_T} = \frac{|\nabla \Phi_E|}{|\nabla \Psi_T|} \text{ on } \Gamma \tag{3}$$

gives
$$\nabla^2 \Phi = \Theta(\Psi_T) \nabla^2 \Phi_E + \delta(\Psi_T) \left(\frac{|\nabla \Psi_T|^2}{\Psi_T} + \nabla^2 \Psi_T \right) \Phi_E$$

Transfer results to Hamiltonian \hat{H}

$$\hat{H}\Phi = \Theta(\Psi_T) \hat{H}\Phi_E - \frac{1}{2} \delta(\Psi_T) \left[\frac{|\nabla \Psi_T|^2}{\Psi_T} + \nabla^2 \Psi_T \right] \Phi_E \tag{4}$$

Note: one could define an effective Hamiltonian that includes the δ fct. term \Rightarrow interpret δ fct. term as infinite “nodal potential“ as required to force Φ to be zero on Γ ... but not relevant here



DMC Energy Gradient

$$\text{DMC energy: } E_D = \frac{\int_{\Omega} \Psi \hat{H} \Phi \, dV}{\int_{\Omega} \Psi \Phi \, dV}$$

for **mixed DMC** ($\Psi = \Psi_T$) and **pure DMC** ($\Psi = \Phi_E$)

- the δ function term does not contribute to E_D
- no difference whether region of integration in expressions for E_D is over Ω over all space \Rightarrow choose all space!

Take the derivative of E_D wrt parameter λ (nucleus coordinate, electric field, etc.)

$$\begin{aligned} \frac{\partial E_D}{\partial \lambda} &= \underbrace{\frac{\int \Psi \frac{\partial \hat{H}}{\partial \lambda} \Phi \, dV}{\int \Psi \Phi \, dV}} + \underbrace{\frac{\int \Psi (\hat{H} - E_D) \frac{\partial \Phi}{\partial \lambda} \, dV}{\int \Psi \Phi \, dV}} + \underbrace{\frac{\int \frac{\partial \Psi}{\partial \lambda} (\hat{H} - E_D) \Phi \, dV}{\int \Psi \Phi \, dV}} \\ &= \mathbf{T_{HFT} \text{ (Hellmann- Feynman Term)}} = \mathbf{T_V \text{ (Volume Term)}} = \mathbf{T_N \text{ (Noda Term)}} \end{aligned}$$

Note that the integration limit was independent of λ

T_{HFT} : is **mixed estimator** ($\Psi = \Psi_T$) or **pure estimator** ($\Psi = \Phi_E$)



Rate of Change of Γ wrt λ

We can write the total differential of Ψ_T as

$$\mathbf{d}\Psi_T(\lambda, \mathbf{r}) = \frac{\partial \Psi_T}{\partial \lambda} \mathbf{d}\lambda + \nabla \Psi_T \cdot \mathbf{d}\mathbf{r}$$

We set $\mathbf{d}\Psi_T = 0$ and divide by $|\nabla \Psi_T| \mathbf{d}\lambda$

$$0 = \frac{1}{|\nabla \Psi_T|} \frac{\partial \Psi_T}{\partial \lambda} + \frac{\nabla \Psi_T \cdot \mathbf{d}\mathbf{r}}{|\nabla \Psi_T| \mathbf{d}\lambda}$$

Choose $\mathbf{d}\mathbf{r}_{\parallel}$ parallel and $\mathbf{d}\mathbf{r}_{\perp}$ orthogonal to normal $\frac{-\nabla \Psi_T}{|\nabla \Psi_T|}$ of Γ :

$$\text{Hence, } \frac{\partial \Psi_T}{\partial \lambda} \frac{1}{|\nabla \Psi_T|} = \frac{\mathbf{d}\mathbf{r}_{\parallel}}{\mathbf{d}\lambda}$$

lhs can be interpreted as the rate of change of the NS of Ψ_T radial to Γ as λ moves

This rate of change must be the same for Ψ_T and for Φ_E i.e.,

$$\frac{\partial \Psi_T}{\partial \lambda} \frac{1}{|\nabla \Psi_T|} = \frac{\partial \Phi_E}{\partial \lambda} \frac{1}{|\nabla \Phi_E|} \quad \text{on } \Gamma \tag{5}$$

Volume Term T_V

For our specific form of Φ , one can show that \hat{H} is Hermitian in all our expressions, i.e.

$$T_N = \frac{\int \Psi(\hat{H} - E_D) \frac{\partial \Phi}{\partial \lambda} dV}{\int \Psi \Phi dV} = \frac{\int \frac{\partial \Phi}{\partial \lambda} (\hat{H} - E_D) \Psi dV}{\int \Psi \Phi dV}$$

This is not obvious as \hat{H} is not Hermitian when integrals are over Ω

For **mixed DMC** ($\Psi = \Psi_T$)

$$\text{use Reynolds' approximation}^1 \quad \frac{1}{\Phi} \frac{\partial \Phi}{\partial \lambda} \approx \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda}$$

$$T_N(\text{mix}) = \frac{\int \Psi_T \Phi \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda} (\hat{H} - E_D) \frac{\Psi_T}{\Psi_T} dV}{\int \Psi_T \Phi dV}$$

For **pure DMC** ($\Psi = \Phi_E$)

$$T_N(\text{pure}) = 0$$

¹P. Reynolds, et al. Internat. J. Quant. Chem. 29 589 (1986)



Nodal Term T_N

Inserting Eq. (4) and using $\hat{H}\Phi_E = E_D\Phi_E$ gives

$$\mathbf{T}_N = \frac{\int \frac{\partial \Psi}{\partial \lambda} (\hat{H} - E_D) \Phi \mathbf{dV}}{\int \Psi \Phi \mathbf{dV}} = -\frac{1}{2} \frac{\int \frac{\partial \Psi}{\partial \lambda} \delta(\Psi_T) \left[\frac{|\nabla \Psi_T|^2}{\Psi_T} + \nabla^2 \Psi_T \right] \Phi_E \mathbf{dV}}{\int \Psi \Phi \mathbf{dV}}$$

Using standard identity $\delta(\Psi_T) = \int_{\Gamma} \frac{\delta(\mathbf{r}-\mathbf{r}')}{|\nabla \Psi_T|} \mathbf{dS}$ and L'Hôpital $\Psi_T^{-1} \Phi_E = |\nabla \Psi_T|^{-1} |\nabla \Phi_E|$

$$\mathbf{T}_N \text{ (mix)} = -\frac{1}{2} \frac{\int_{\Gamma} \frac{\partial \Psi_T}{\partial \lambda} |\nabla \Phi_E| \mathbf{dS}}{\int \Psi_T \Phi \mathbf{dV}}$$

$$\mathbf{T}_N \text{ (pure)} = -\frac{1}{2} \frac{\int_{\Gamma} |\nabla \Phi_E|^2 \frac{1}{|\nabla \Phi_E|} \frac{\partial \Phi_E}{\partial \lambda} \mathbf{dS}}{\int \Phi_E \Phi \mathbf{dV}}$$

Interpretation of T_N (pure)

- $|\nabla \Phi_E|^2$ is kinetic energy density and $\frac{1}{|\nabla \Phi_E|} \frac{\partial \Phi_E}{\partial \lambda}$ is rate of change of Γ wrt λ
- T_N is rate of change of kinetic energy arising from kink in Φ at Γ
(lin. response of kinetic energy wrt variations of NS)
- T_N is a single particle expression!



Calculate Nodal Term (I)

It is not straightforward to evaluate mixed & pure DMC Nodal Terms!

- To find an expression for the **mixed DMC** nodal term, use that \hat{H} is Hermitian (proof omitted), i.e.

$$T_N = \frac{\int \frac{\partial \Psi_T}{\partial \lambda} (\hat{H} - E_D) \Phi dV}{\int \Psi_T \Phi dV} = \frac{\int \Psi_T \Phi \frac{1}{\Psi_T} (\hat{H} - E_D) \frac{\partial \Psi_T}{\partial \lambda} dV}{\int \Psi_T \Phi dV}$$

- To calculate the **pure DMC** nodal term, we note that the pure and mixed DMC Nodal Terms T_N can be rewritten using Eq.(3) and (5),

$$T_N = -\frac{1}{2} \frac{\int_{\Gamma} \Psi \Phi \frac{|\nabla \Psi_T|}{\Psi_T} \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda} dS}{\int \Psi \Phi dV}$$

for mixed DMC ($\Psi = \Psi_T$) and pure DMC ($\Psi = \Phi_E$)

$$\text{Averaged quantity } Q = \frac{|\nabla \Psi_T|}{\Psi_T} \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda}$$

- is same for mixed and pure DMC
- only contains Ψ_T



Calculate Nodal Term (II)

Use extrapolation formula which is exact to 2nd order in error of Ψ_T

$$\langle Q \rangle_{pure} = 2\langle Q \rangle_{mixed} - \langle Q \rangle_{VMC} \quad \text{with} \quad Q = \frac{|\nabla \Psi_T|}{\Psi_T} \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda}$$

(which is also valid when integrals are over surface)

The average of Q in VMC is zero

$$T_N = \frac{1 \int_{\Gamma} \Psi_T \Psi_T \frac{|\nabla \Psi_T|}{\Psi_T} \frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda} \mathbf{dS}}{2 \int \Psi_T \Psi_T \mathbf{dV}} = \frac{1 \int_{\Gamma} |\nabla \Psi_T| \frac{\partial \Psi_T}{\partial \lambda} \mathbf{dS}}{2 \int \Psi_T \Psi_T \mathbf{dV}}$$

$$= \frac{1 \int_{\Gamma} \frac{\partial \Psi_T}{\partial \lambda} \nabla \Psi_T \cdot \mathbf{dS}}{2 \int \Psi_T \Psi_T \mathbf{dV}} = 0$$

Since $\frac{\partial \Psi_T}{\partial \lambda}$ is continuous on Γ , the surface integrals of two bordering pockets cancel.
 \Rightarrow integral over Γ of all pockets simultaneously is zero
 \Rightarrow surface integral over each pocket is zero (tiling theorem)
 (We verified this numerically, too!)

Hence we can calculate the pure DMC nodal term as

$$T_N \text{ (pure DMC)} \approx 2 T_N \text{ (mixed DMC)}$$

Summary so far

$$\text{Mixed DMC } \frac{\partial E_D}{\partial \lambda} = \frac{\int \Psi_T \frac{\partial \hat{H}}{\partial \lambda} \Phi \, dV}{\int \Psi_T \Phi \, dV} + \frac{\int \Psi_T \Phi \left[\frac{1}{\Psi_T} \frac{\partial \Psi_T}{\partial \lambda} \frac{1}{\Psi_T} (\hat{H} - E_D) \Psi_T \right] \, dV}{\int \Psi_T \Phi \, dV} + \frac{\int \Psi_T \Phi \frac{1}{\Psi_T} (\hat{H} - E_D) \frac{\partial \Psi_T}{\partial \lambda} \, dV}{\int \Psi_T \Phi \, dV} + O(\Delta \Psi_T)$$

$$\text{Pure DMC } \frac{\partial E_D}{\partial \lambda} = \frac{\int \Phi_E \frac{\partial \hat{H}}{\partial \lambda} \Phi \, dV}{\int \Phi_E \Phi \, dV} + 2 \frac{\int \Psi_T \Phi \frac{1}{\Psi_T} (\hat{H} - E_D) \frac{\partial \Psi_T}{\partial \lambda} \, dV}{\int \Psi_T \Phi \, dV} + O(\Delta \Psi_T^2)$$

- if NS is independent of $\lambda \Leftrightarrow \frac{\partial \Psi_T}{\partial \lambda} = 0$ on $\Gamma \Rightarrow \mathbf{T}_N$ is zero
- if Ψ_T exact $\Rightarrow \mathbf{T}_N$ is zero
(not obvious but can be shown using the tiling theorem)

Computational Details



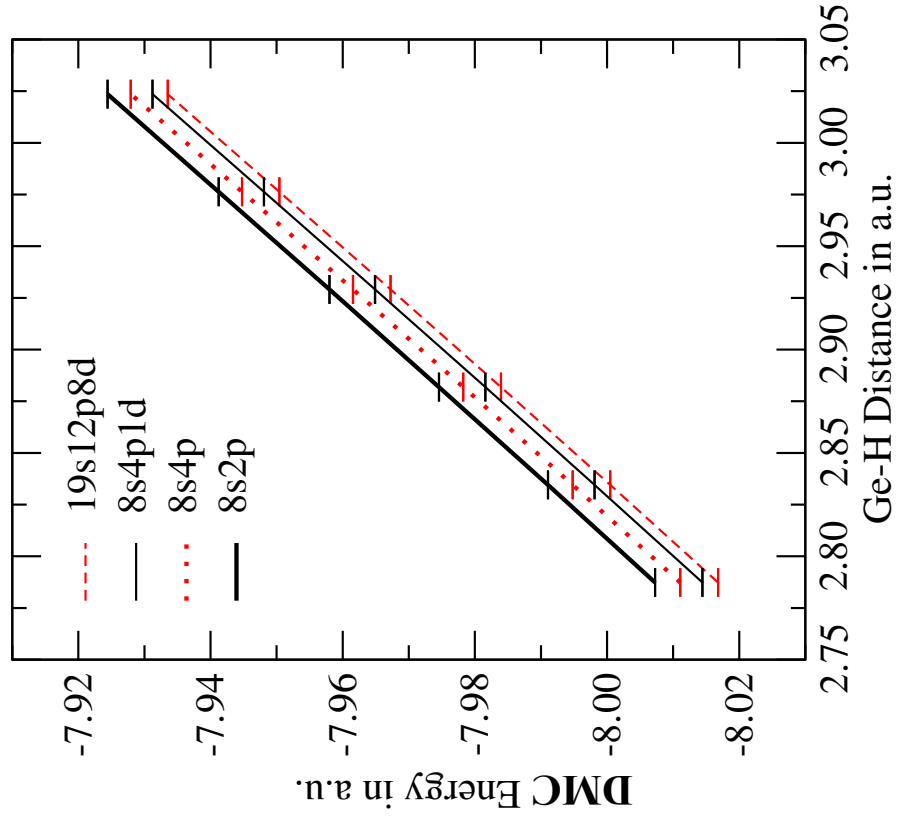
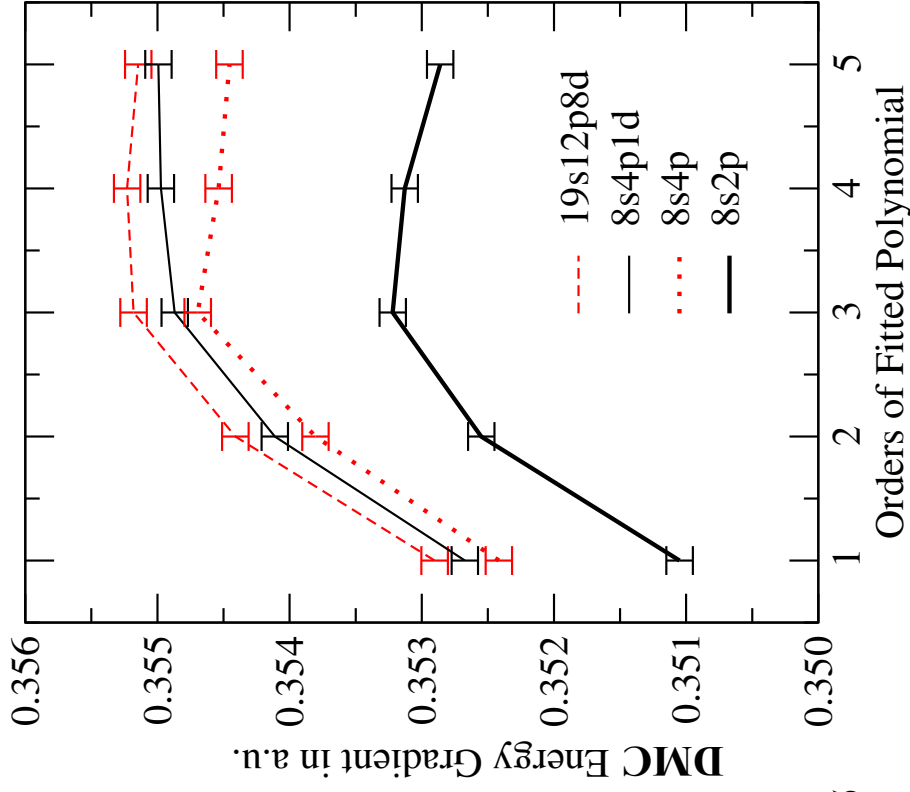
- Study **Germanium hydride dimer** (GeH)
- **no electron-electron interaction** (nodal terms arise from kinetic energy!)
- use **Slater det.** with non-interacting MOs calculated using GAMESS-US
- 4 basis sets: very good (*19s12p8d*), good (*8s4p1d*), poor (*8s4p*), very poor (*8s2p*)
cover range of errors in Ψ_T for practical calculations
- local pseudopotentials (constructed to emulate non-local pseudopotentials...)
- first derivative of Ψ_T from finite difference
- use extrapolation and **Future walking**¹ to calculate T_{HFT}

$$\langle A \rangle_{\phi^2} = \frac{\sum_k A(\mathbf{r}_k) q(\mathbf{r}_k)}{\sum_k q(\mathbf{r}_k)} \quad \text{with} \quad q(z_0) = \lim_{t \rightarrow \infty} \int f(x, z_0, t) dx$$

- for reference, calculate energy gradient from potential energy curve...

For convenience: convert changes in forces ΔF to changes in bond length Δa :

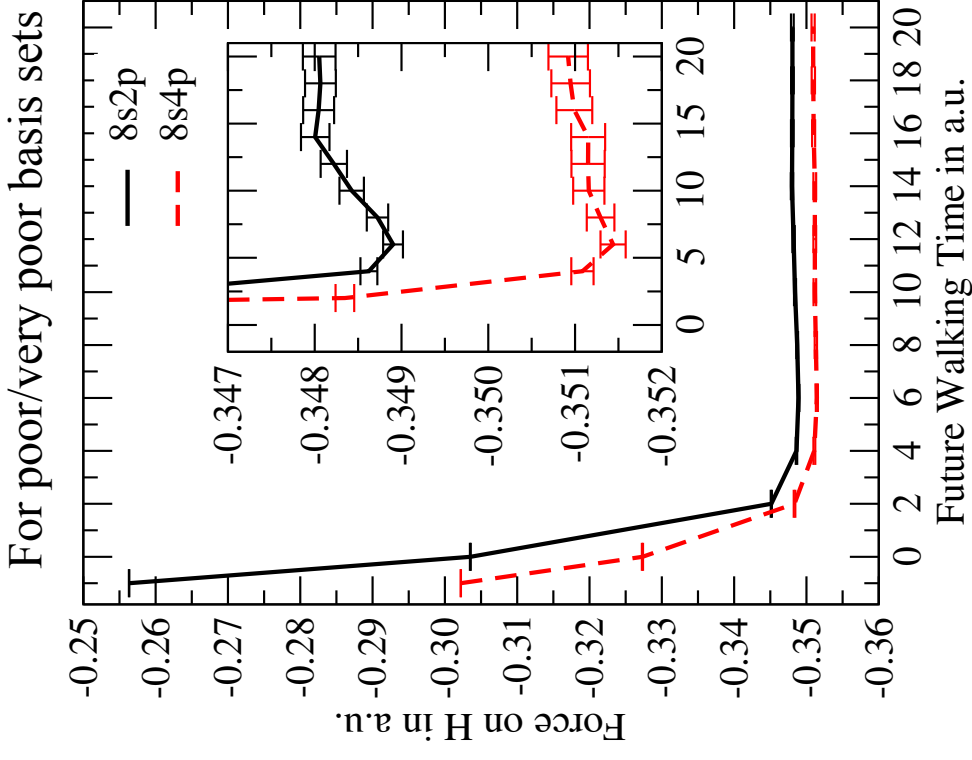
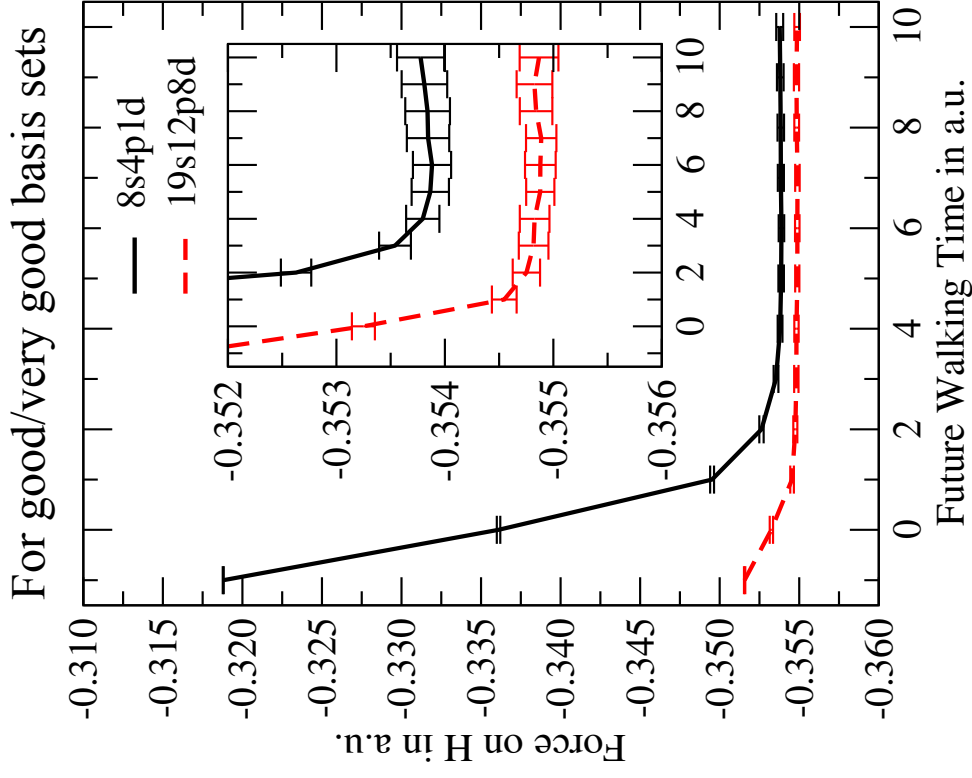
$$\Delta F = 0.001 \text{ a.u.} \approx \Delta a = 0.01 \text{ a.u.}$$



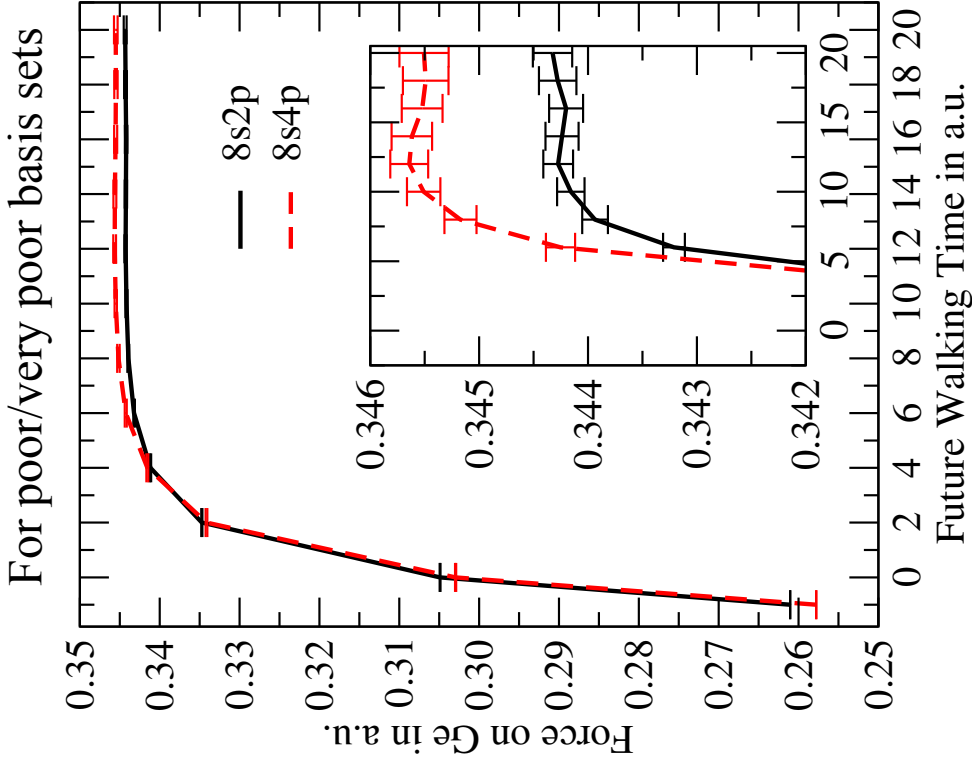
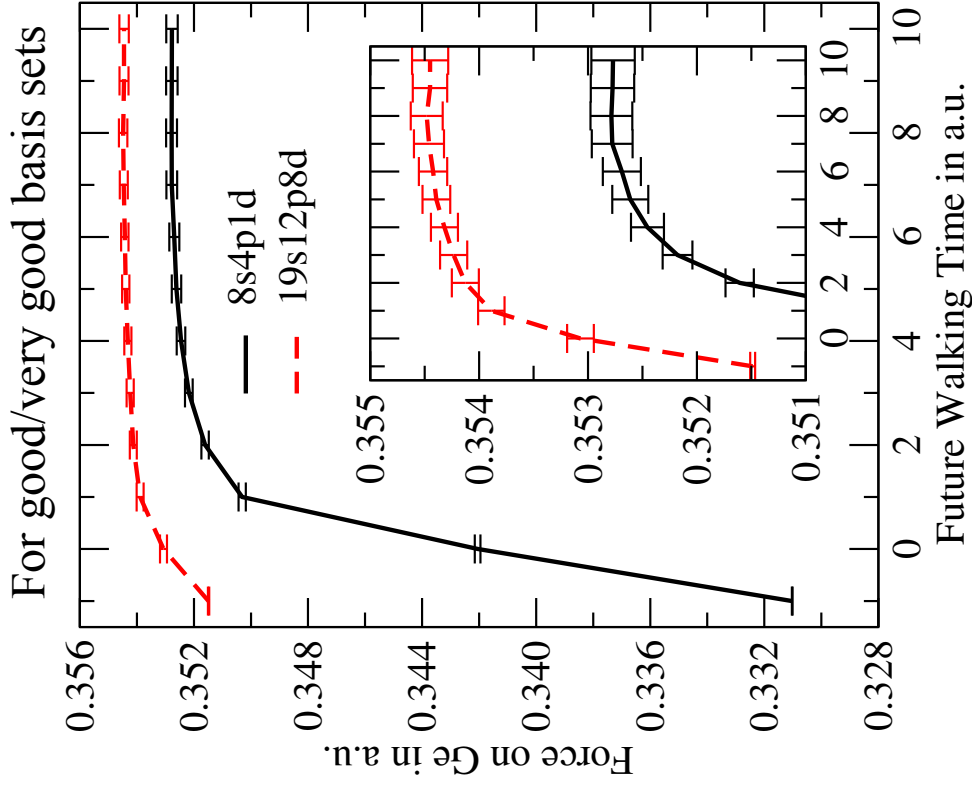
evaluated at 2.929 a.u.

(3.003 a.u. is expt. Bond length)

Future Walking Estimate (on H)



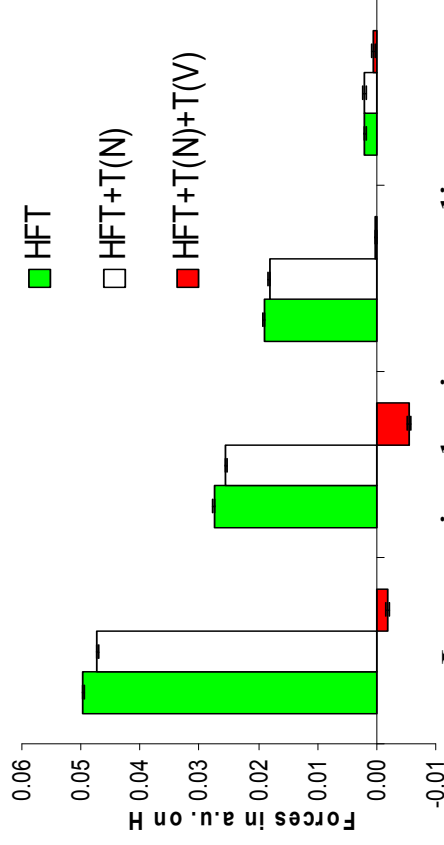
Future Walking Estimate (on Ge)



Mixed DMC Force

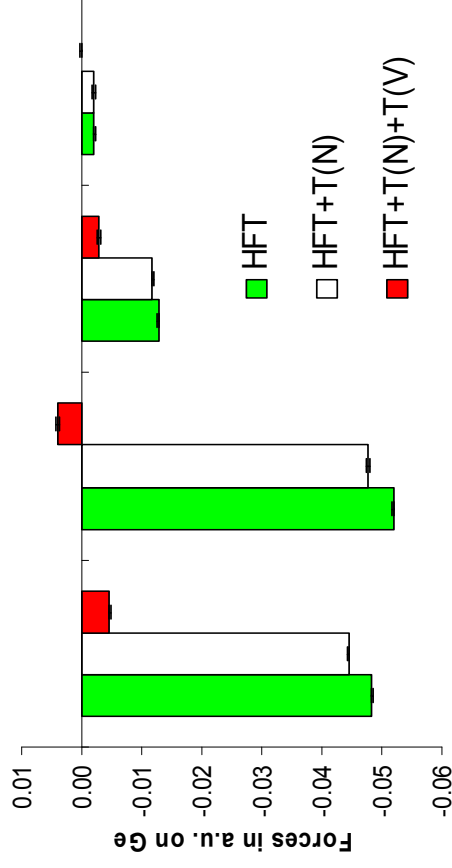


Forces on H atom - energy gradient



Increasing basis set quality

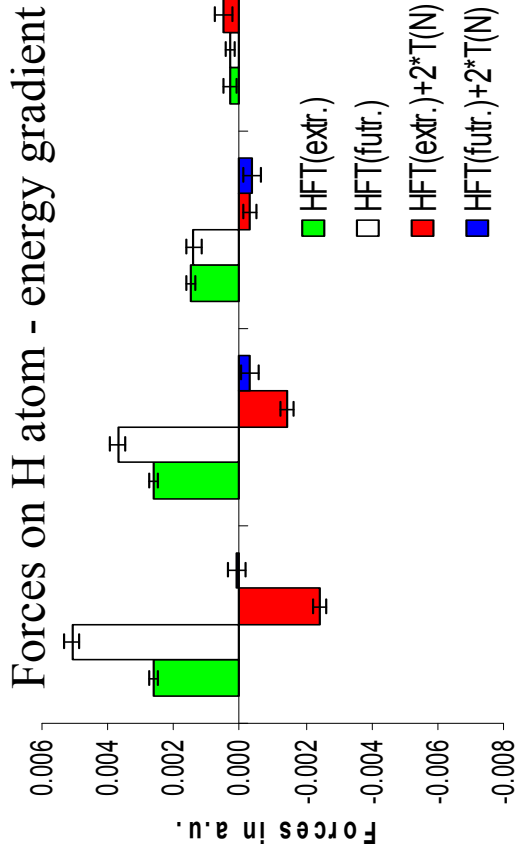
Forces on Ge atom – energy gradient



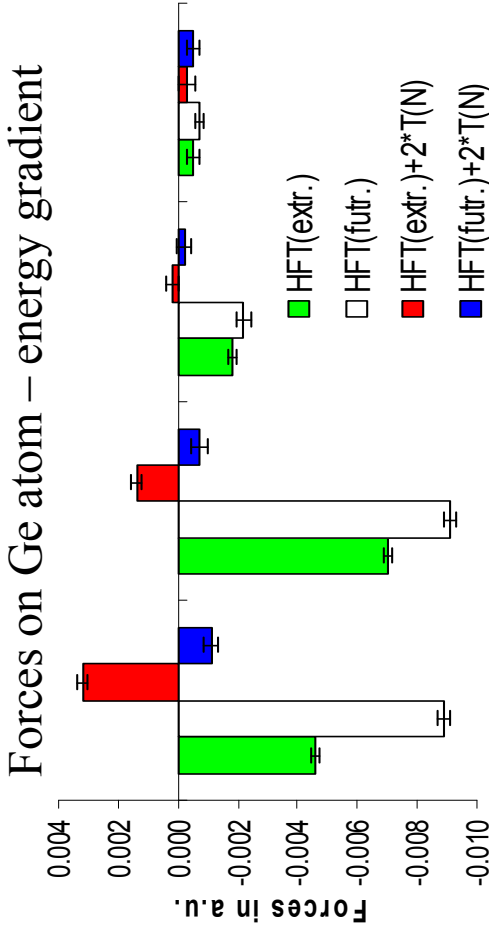
Increasing basis set quality

- significant basis set dependence
- adding volume term T(V) significantly improves mixed HFT force
- adding nodal term T(N) slightly improves mixed HFT force
- Remaining errors: 1) Reynolds' approx. (1st order in error of Ψ_T)
2) extrapolation approx. in nodal term (2nd order in error of Ψ_T)

Pure DMC Force



Increasing basis set quality



Increasing basis set quality

- significant basis set dependence
- Future walking estimate of HFT seem not better than extrapolation, but ...
- adding T(N) to HFT(extr.) “overshoots” correction for poor basis sets
- adding T(N) to HFT(futr.) gives exact total forces within (close to) one standard error
- pure HFT forces (without nodal term correction) as good as total mixed forces \Rightarrow total pure DMC force (with nodal terms) more accurate than total mixed DMC force



Conclusions

- We derived **exact expressions for forces** within mixed and pure DMC including Pulay surface terms previously neglected in pure DMC calculations
- In mixed DMC: **Pulay surface terms** can be calculated straightforward
In pure DMC: Pulay surface terms are approximated as twice the mixed terms
- Tests for the GeH indicate that mixed and pure DMC **nodal terms are significant** and including them significantly improves forces!
- Pure DMC forces seem more accurate than mixed DMC forces (in particular when the quality Ψ_T is less good)
- Pure DMC forces are slightly more costly than mixed once, but seems worth!
- **Outlook:** apply to larger systems, use with non-local pseudopotentials

Acknowledgments



Many thanks to John Trail and Zoltan Radnai

Financial Support: **Engineering and Physical Sciences Research Council** of the United Kingdom