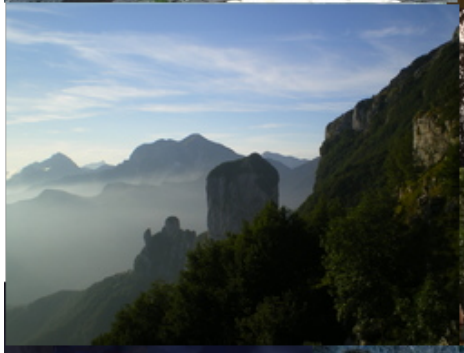
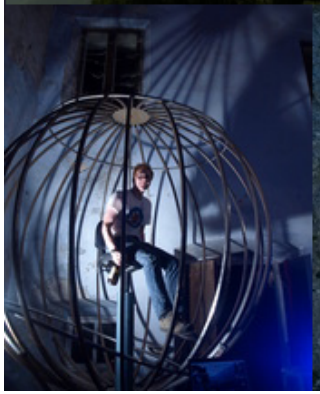







The Towler Institute

2016 International Summer School

Quantum Monte Carlo and the CASINO program X

Vallico Sotto, Tuscany, Italy 23rd - 30th July 2016

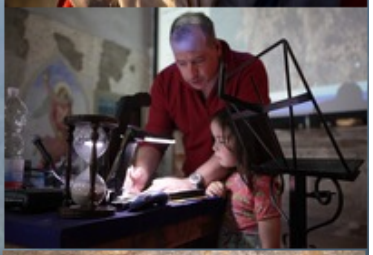
vallico.net/tti/tti.html

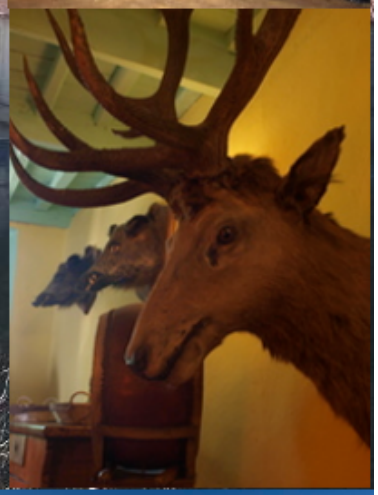
email : mdt26 at cam.ac.uk





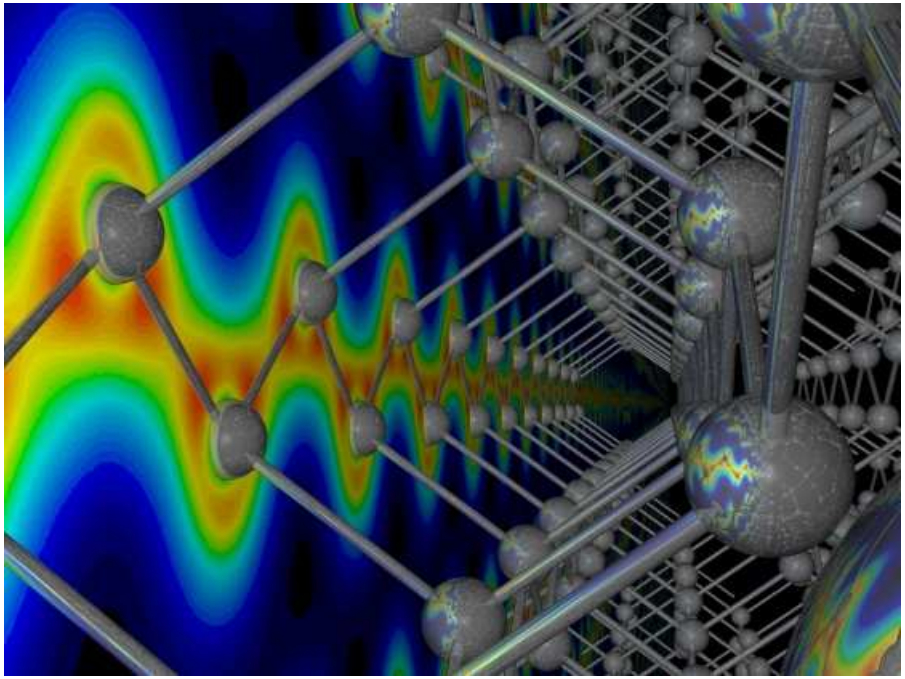






QUANTUM MONTE CARLO

Forces and dynamics. Expectation values other than the energy.



Mike Towler

TCM Group, Cavendish Laboratory, University of Cambridge

QMC web page: `vallico.net/casinoqmc`

Email: `mdt26@cam.ac.uk`

Derivatives of the energy

$$E(F_i) = E(0) + F_i \underbrace{\left[\frac{\partial E}{\partial F_i}\right]_{F_i=0}}_{\text{dipole moment}} + \frac{1}{2}\sum_{j=1}^{3} F_i F_j \underbrace{\left[\frac{\partial^2 E}{\partial F_i F_j}\right]_{F_i=0,F_j=0}}_{\text{dipole polarizability tensor}} + \cdots$$

- Many quantities of physical interest can be formulated as an energy derivative and thus an ability to calculate derivatives accurately in QMC considerably enhances the scope of the method.
- Then can calculate e.g. dipole moment without need to sample pure distribution (by applying a small electric field in each direction and observing the change in the energy).
- Perturbation could be electric field (giving dipole moment, dipole polarizability tensor etc..) or displacement of nuclear position (giving forces, etc.) or combination (e.g. intensity of peaks in IR spectrum depends on change in dipole moment corresponding to a change in geometry).
- Energy derivatives can be computed numerically (by finite differencing) or analytically (by differentiating the appropriate energy expressions).

Forces

- Average components of force on atoms given by first derivatives of the energy with respect to nuclear displacements

$$\bar{F}_{R_i} = -\frac{\partial E_0(\mathbf{R})}{\partial R_i}$$

  where $\mathbf{R}$ represents the $3N_n$ nuclear coords. R_i is displacement of particular nucleus in direction of desired force component.

- Define local force as

$$F_{R_i}(\mathbf{x}, \mathbf{R}) = -\frac{\partial V(\mathbf{x}, \mathbf{R})}{\partial R_i}$$

  $\mathbf{x}$ represents the $3N_e$ electronic coordinates and V the potential energy operator.

- If the Hellmann-Feynman theorem applies, then *average* forces can be rewritten as the statistical average of the local force over the exact distribution $\Psi_0^2(\mathbf{x})$.

$$\bar{F}_{R_i} = \langle F_{R_i}(\mathbf{x}, \mathbf{R}) \rangle_{\Psi_0^2(\mathbf{x})}$$

  Usually it doesn't (since things other than the potential energy depend on R).

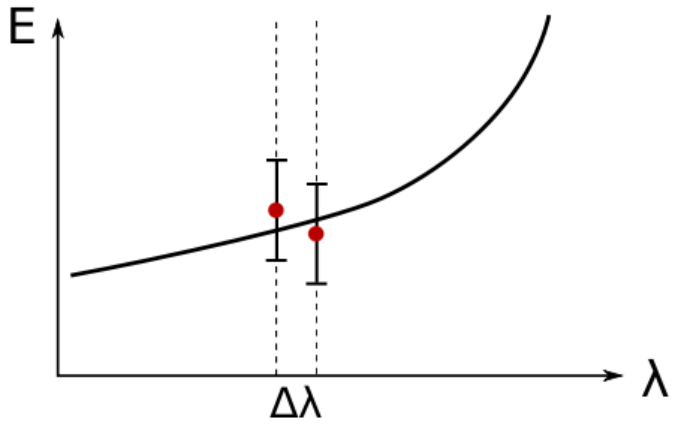
- Forces $\longrightarrow$ equilibrium geometries, potential energy surfaces, vibrational frequencies, dynamics. Useful! But forces are difficult to calculate in QMC..

Finite differences

Approximate derivative by standard two-point finite difference formula:

$$\frac{dE}{dq_i} \approx \frac{E(\delta q_i) - E(0)}{\delta q_i}$$

- Traditional *ab initio* quantum chemistry methods yield approximate values of properties that can be determined to arbitrary precision (primarily governed by the degree to which the wave functions/density can be converged).
- Important to distinguish between accuracy and precision. Accuracy is determined by the approximation (Hartree-Fock, CI etc..). Monte Carlo methods are very accurate but their precision is limited by statistical uncertainty.
- This limited precision can lead to problems when taking finite differences. *Typical statistical uncertainty in the derivative is 20 to 100 (or greater) times that of the uncertainty in the energies.*



Finite differences and correlated sampling

- One solution to this problem is to use *correlated sampling* : construct Monte Carlo algorithm so that the statistical fluctuations in $E(0)$ and $E(\delta q_i)$ are correlated.
- In the limiting case where the two local energies $E_L(\mathbf{x};0)$ and $E_L(\mathbf{x};\delta q_i)$ are related pointwise by a constant offset, correlated sampling of the *difference* would have zero statistical uncertainty.
- In reality the variance of the difference increases as the perturbation increases, because the correlation between the perturbed and unperturbed systems decreases.
- Use VMC walk to generate ensemble of configurations $\mathbf{X_k^{(i)}}$ for the unperturbed system. Then get energies of perturbed system using the same configurations and reweighting the energies (in similar way to weighted varmin):

$$E(0) \approx \frac{1}{MN_c}\sum_{i=1}^{M}\sum_{k=1}^{N_c} E_L(\mathbf{X}_k^{(i)};0)$$

$$E(\delta q_i) \approx \frac{\sum_{i=1}^{M}\sum_{k=1}^{N_c} E_L(\mathbf{X}_k^{(i)};\delta q_i) W(\mathbf{X}_k^{(i)};\delta q_i)}{\sum_{i=1}^{M}\sum_{k=1}^{N_c} W(\mathbf{X}_k^{(i)};\delta q_i)}$$

$$W(\mathbf{X}_k^{(i)};\delta q_i) = \frac{\Psi^2(\mathbf{X}_k^{(i)};\delta q_i)}{\Psi^2(\mathbf{X}_k^{(i)};0)} \quad \text{(Weights)}$$

Finite differences in practice

- Accuracy and precision of the correlated sampling degrades rapidly upon increase of the difference between the two geometries.
- This difficulty is reduced if the trial wave function for all the geometries different from the reference one is reoptimized, and a "warp" coordinate transformation is used. See Filippi and Umrigar *Phys. Rev. B* **61**, R16291 (2000).
- However, computational cost needed to compute a D-dimensional gradient is, at least, $D+1$ times the one for a single energy estimate while the cost for the Hessian matrix scales as $D(D+1)/2$.

Serious problem for more than two atoms!

Hellmann-Feynman theorem (HFT) in VMC

Total energy $E_{\rm VMC} = \frac{\langle\Psi_T|\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle}$ Total force $F_{\rm VMC} = -\nabla_{\mathbf{R}} E_{\rm VMC}$

If the wave function Ψ_T is the exact one Ψ_0 then the force is the expectation value of the gradient of $\hat{H}$, i.e. $-F^{\rm HFT}_{\rm VMC} = \nabla_{\mathbf{R}} E_{\rm VMC} = \frac{\langle\Psi_0|\nabla\hat{H}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle}$

Proof

$$\begin{aligned}
\nabla_{\mathbf{R}} E_{\rm VMC} &= \frac{\langle\Psi_T|\nabla_{\mathbf{R}}\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} + \frac{\langle\Psi'_T|\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} + \frac{\langle\Psi_T|\hat{H}|\Psi'_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \\
&\quad -\frac{\langle\Psi_T|\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle^2}(\langle\Psi'_T|\Psi_T\rangle + \langle\Psi_T|\Psi'_T\rangle) + \sum_i \frac{\partial E}{\partial c_i}\frac{\partial c_i}{\partial \mathbf{R}} \\
&= \underbrace{\frac{\langle\Psi_T|\nabla_{\mathbf{R}}\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle}}_{F_R^{\rm HFT}} + \underbrace{2\frac{\langle\Psi'_T|\hat{H}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} - 2E\frac{\langle\Psi'_T|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle}}_{F_R^{\Psi}} + \underbrace{\sum_i \frac{\partial E}{\partial c_i}\frac{\partial c_i}{\partial \mathbf{R}}}_{F_R^{c}}
\end{aligned}$$

F_R^{Ψ} and F_R^c go to zero as $\Psi_T \longrightarrow \Psi_0$, hence the theorem. The use of the HFT is thus an approximation in VMC where we only have an inexact trial wave function.

VMC force is sum of three terms

$$\nabla_{\mathbf{R}} E_{\mathrm{VMC}} = \underbrace{\frac{\langle \Psi_T | \nabla_{\mathbf{R}} \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\mathrm{HFT}}} + \underbrace{2\frac{\langle \Psi_T' | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} - 2E_{\mathrm{VMC}} \frac{\langle \Psi_T' | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\Psi}} + \underbrace{\sum_i \frac{\partial E_{\mathrm{VMC}}}{\partial c_i} \frac{\partial c_i}{\partial \mathbf{R}}}_{F_R^{c}}$$

- F_R^{HFT} is the Hellmann-Feynman term.
- F_R^{Ψ} (the *Pulay* term) involves partial derivatives of Ψ_T with respect to the nuclear positions acting only on the part of the trial wave function that explicitly depends on $\mathbf{R}$ e.g. atomic basis function centres or electron-nuclear terms in the Jastrow factor.

  Therefore best to choose basis functions that do not depend directly on $\mathbf{R}$*, although sometimes this is unavoidable.*

- F_R^{c} accounts for the action of $\nabla_{\mathbf{R}}$ on the parameters that only indirectly couple with the nuclear positions (e.g. orbital coefficients).

  Expect that $\frac{\partial E}{\partial c_i}$ *derivatives in* F_R^{c} *term will be near-zero if wave function is optimized through energy minimization. Up to very recently has unfortunately been more usual to minimize the variance, though optimal parameters in each case usually pretty similar. The parameters also contain statistical noise.*

Expression for the force in VMC

$$\nabla_{\mathbf{R}} E_{\mathrm{VMC}} = \underbrace{\frac{\langle \Psi_T | \nabla_{\mathbf{R}} \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\mathrm{HFT}}} + \underbrace{2\frac{\langle \Psi_T' | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} - 2E_{\mathrm{VMC}} \frac{\langle \Psi_T' | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\Psi}} + \underbrace{\sum_i \frac{\partial E_{\mathrm{VMC}}}{\partial c_i} \frac{\partial c_i}{\partial \mathbf{R}}}_{F_R^{c}}$$

Expression for the force in VMC

$$\nabla_{\mathbf{R}} E_{\mathrm{VMC}} = \underbrace{\frac{\langle \Psi_T | \nabla_{\mathbf{R}} \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\mathrm{HFT}}} + \underbrace{2\frac{\langle \Psi_T' | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} - 2E_{\mathrm{VMC}} \frac{\langle \Psi_T' | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}}_{F_R^{\Psi}} + \underbrace{\sum_i \frac{\partial E_{\mathrm{VMC}}}{\partial c_i} \frac{\partial c_i}{\partial \mathbf{R}}}_{F_R^{c}}$$

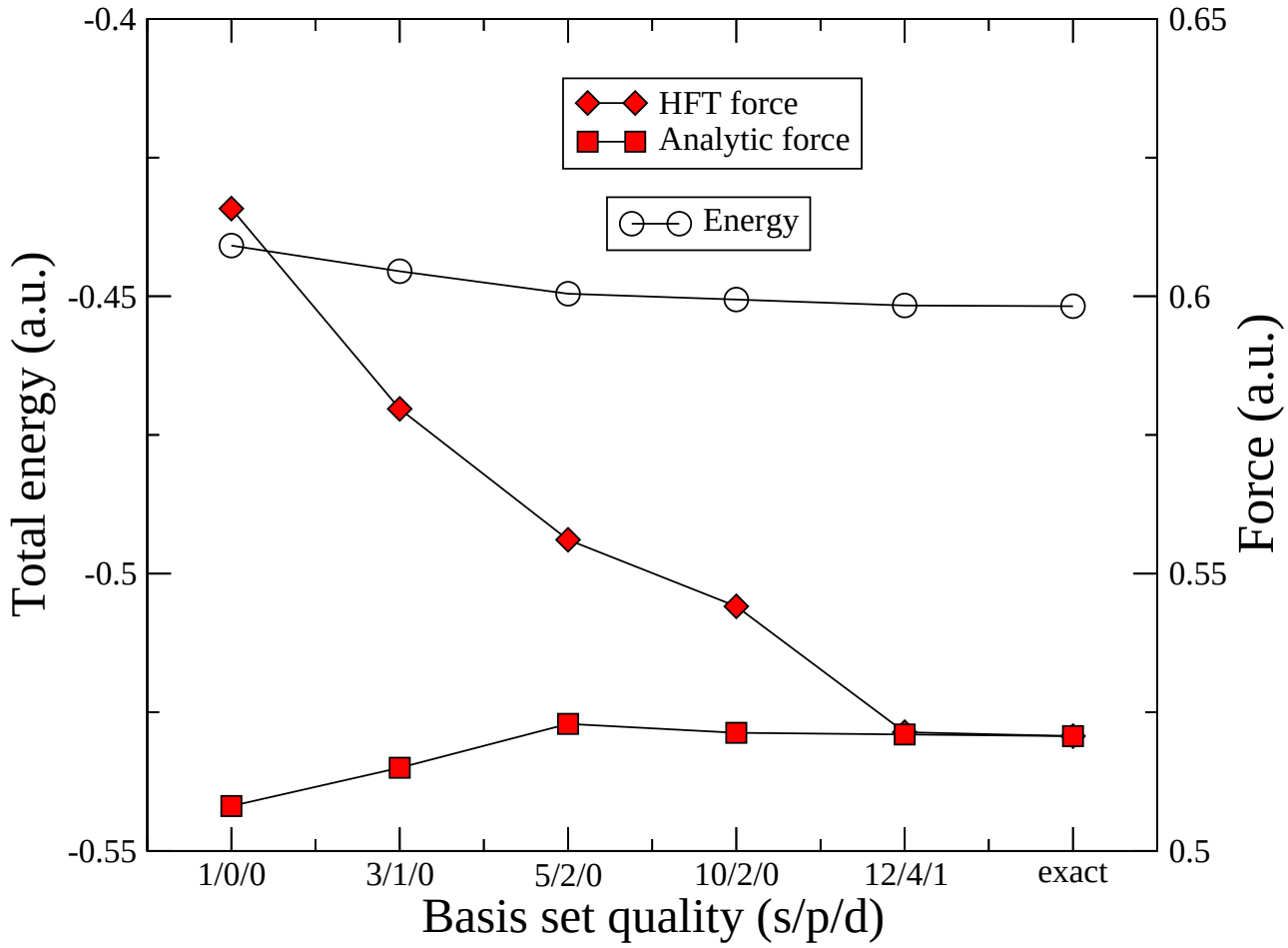
$$F_R^{\mathrm{HFT}} = \frac{\langle \Psi_T | \sum_{\alpha,i} Z_\alpha \frac{\mathbf{r}_i - \mathbf{R}_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|^3} + \sum_{\alpha,\beta} Z_\alpha Z_\beta \frac{\mathbf{R}_\alpha - \mathbf{R}_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|^3} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}$$

$$F_R^{\Psi} = \frac{\langle \Psi_T | 2(E_L - E_{\mathrm{VMC}}) \frac{\Psi_T'}{\Psi_T} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle}$$

$$F_R^{c} = 0 \quad \text{(assumption)}$$

Hellmann-Feynman theorem in Hartree-Fock/DFT

Hartree-Fock HFT and analytic derivatives for H_2^+ at R=1a.u.



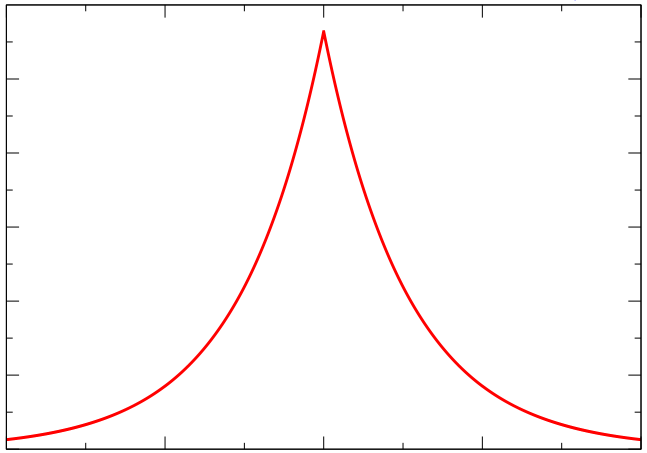
HFT is valid in Hartree-Fock and DFT with a plane-wave basis set (which doesn't depend on nuclear positions). For basis sets with explicit dependence on nuclear positions (such as Gaussians) then the HFT is only valid in the limit of a complete basis set. Need to compute additional Pulay terms to get correct forces.

Infinite variance problem

Define the *local HFT force* as $F^{\mathrm{HFT}}_{L,i} = (-\nabla \hat{H})_i$ for $i = 1 \ldots 3N$, where

$$\nabla \hat{H} = \sum_{\alpha,i} Z_\alpha \frac{\mathbf{r}_i - \mathbf{R}_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|^3} + \sum_{\alpha,\beta} Z_\alpha Z_\beta \frac{\mathbf{R}_\alpha - \mathbf{R}_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|^3}$$

Then the *variance* of the local force is given by $\sigma(F^{\mathrm{HFT}}_{L,i})^2 = \langle \left(F^{\mathrm{HFT}}_{L,i}\right)^2 \rangle - \langle F^{\mathrm{HFT}}_{L,i} \rangle^2$.



Consider H atom with wave function $\Psi = e^{-ar}$. Average force clearly zero for atom, but consider variance of Hellmann-Feynman derivative in x direction, i.e. $\sigma^2 = \langle \frac{x^2}{r^6} \rangle - \langle \frac{x}{r^3} \rangle^2$. After conversion to spherical polars with $x = r \sin\theta \cos\theta$

$$\sigma^2 = \int \sin^3(\theta) \cos^2(\phi)\, d\theta d\phi \int \frac{e^{-2ar}}{r^2}\, dr - \cdots$$

Thus the variance of the Hellmann-Feynman force is infinite!

Renormalization

Assaraf and Caffarel, *J. Chem. Phys.* **119**, 10536 (2003).

$$\nabla E_{\mathrm{VMC}} = \frac{1}{\langle \Psi_T | \Psi_T \rangle} \langle \Psi_T | \hat{H}' + \frac{(\hat{H} - E_L)\Psi'_T}{\Psi_T} + 2\frac{\Psi'_T}{\Psi_T}(E_L - E_{\mathrm{VMC}}) | \Psi_T \rangle$$

Basic idea is to construct a "renormalized" or improved observable that has the same average as the original one but a lower variance.

- In this way, can remove the pathological part responsible for the infinite variance of the Hellmann-Feynman estimator.
- Better in large systems since it allows simultaneous computation of all the gradient components in a single run, without the burden of optimizing $D + 1$ wave functions.

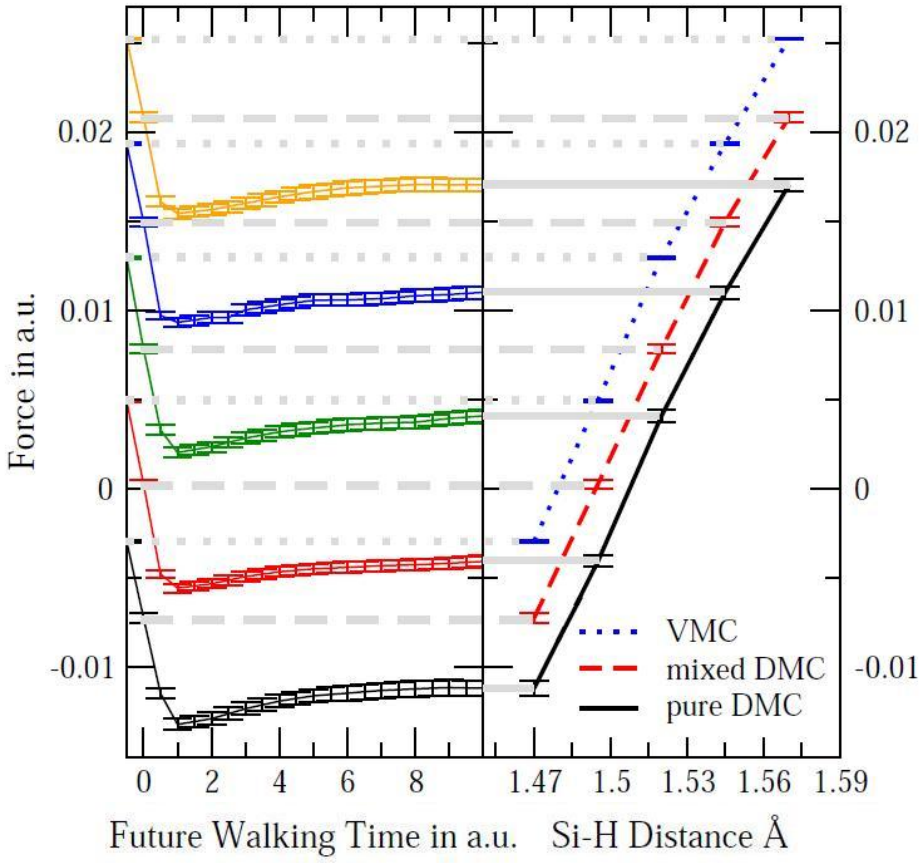
Other Possibilities

- Filter out s-component of density (Chiesa *et al.*, *PRL*, **94**, 036404 (2005).)
- Use pseudopotentials that remove singularity at origin - no more infinite variance problem. Difficult to differentiate non-local pseudopotential operator (but this has now been done in CASINO, see Badinski and Needs, "*Accurate forces in QMC calculations with non-local pseudopotentials*", PRE **76**, 036707 (2007)).

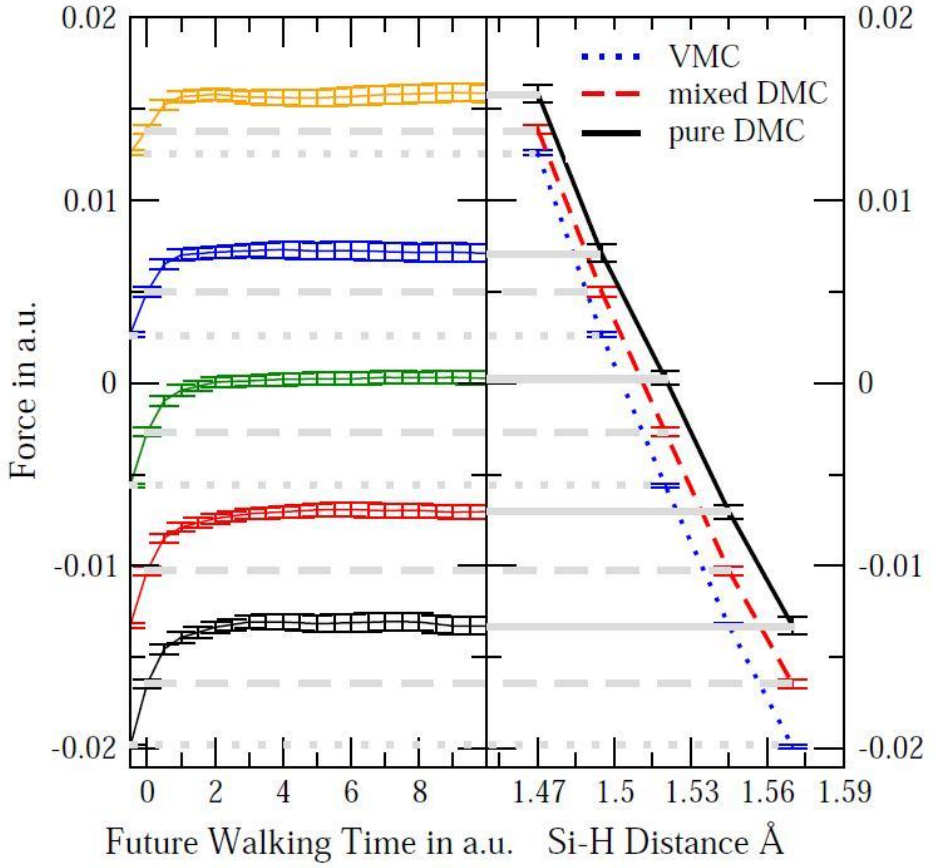
Forces in DMC

- Taking the gradient of the fixed-node DMC energy leads to additional Pulay-like terms, and the estimate of the force is unbiased when these terms are included. The additional terms include gradients of the DMC wave function Φ which cannot be straightforwardly evaluated.
- Can replace these terms by similar ones involving the gradient of the trial function Ψ_T. Though approximate, this scheme should give good results if Ψ_T is accurate enough.
- Can generate the "pure" distribution $\Phi\Phi$ using e.g. approximate extrapolated estimation method, future walking, or reptation MC. Claimed in literature that HFT force calculated with the pure distribution is equal to the exact negative DMC energy gradient, but it is clear that this cannot be correct. In fact there is an additional "nodal term" which is only zero when the nodal surface of Ψ_T is exact.

Forces in CASINO from version 2.1.1



Si atom in SiH



H atom in SiH

- Badinski/Needs derive gradients of non-local pseudopotential operator for FPLA (standard) and T-move localization schmes.
- They calculate equilibrium bond lengths and vibrational frequencies in good agreement with experiment with an accuracy comparable to or greater than DFT-PBE, MP2, CC for the five molecules H_2, LiH, SiH, SiH_4 and GeH.

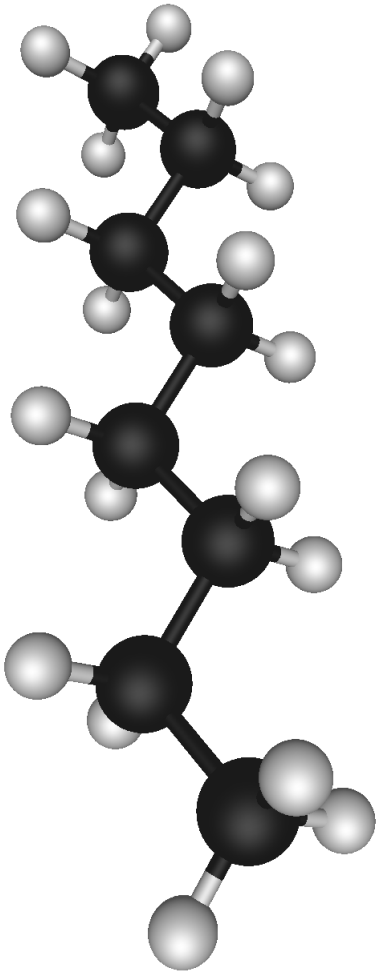
input keyword: forces

Moving atoms

Quantum Monte Carlo calculations normally done with *static nuclei*.

Hardly surprising since DMC is c. 1000 times more expensive than DFT, and in all fairness, even DFT calculations are too expensive to do *really* interesting things with fully *ab initio* molecular dynamics.

Nevertheless, real nuclei do move. Is there any scope whatever for this in DMC?



Implemented in CASINO:

- Grossman-Mitas molecular dynamics.
  What is this and is it of any use?

Selected other things in the literature

- Ceperley - DMC with quantum nuclei
- Coupled electron-ion Monte Carlo (CEIMC)
- Attaccalite and Sorella method

Has Mike had any ideas?

- Maybe. Maybe not.

Can QMC help with simpler models?

- Parametrized force fields? GAP potentials?, LOTF?

Time dependence: the time-dependent Schrödinger equation

Solve $i\hbar\frac{\partial}{\partial t}\Psi(x,t) = H\Psi(x,t)$ by separation of variables to give the following particular solutions (which have the counterintuitive property of predicting time-independent observables):

$$\Psi(x,t) = \phi_E(x)e^{-\frac{i}{\hbar}Et} \qquad |\Psi(x,t)|^2 = |\phi_E(x)|^2$$

Where has the time gone? It is restored to us by a general solution to the TDSE - an arbitrary superposition of the particular solutions:

$$\begin{aligned}\Psi(x,t) &= \sum_{n=1}^{\infty} a_n\phi_n(x)e^{-\frac{i}{\hbar}E_nt} && \text{(discrete spectrum)}\\ &= \int_0^{\infty} a(E)\phi_E(x)e^{-\frac{i}{\hbar}Et}\,\mathrm{d}E && \text{(continuous spectrum)}\end{aligned}$$

Quite generally, a *wave packet* - a superposition of states having different energies - is required in order to have a time-dependence in the probability density and in other observable quantities, such as the average position or momentum of a particle. Simplest example: a linear combination of just two particular solutions $\Psi(x,t) = a\phi_E(x)e^{-\frac{i}{\hbar}Et} + b\phi_{E'}(x)e^{-\frac{i}{\hbar}E't}$. The probability density is given by:

$$|\Psi(x,t)|^2 = |a|^2|\phi_E(x)|^2 + |b|^2|\Psi_{E'}(x)|^2 + 2\mathrm{Re}\left\{a^*b\phi_E^*(x)\phi_{E'}(x)\mathrm{e}^{-i\frac{(E'-E)t}{\hbar}}\right\}$$

All the time-dependence is contained in the interference term.

Time dependence: the approximations

TDSE generally expensive to solve directly (exponential scaling) - simplification required. Expand Ψ in complete set of solutions of *time-independent* SE for all possible nuclear configurations R:

$$\Psi(\mathbf{r}, \mathbf{R}; t) = \sum_{l}^{\infty} \psi_l(\mathbf{r}; \mathbf{R}) \chi_l(\mathbf{R}; t)$$

where nuclear wave functions χ are essentially 'time-dependent expansion coefficients'. Insert into *time-dependent* SE and manipulate - end up with set of coupled differential equations:

$$\left[-\sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + E_k(\mathbf{R}_I) \right] \chi_k + \sum_l C_{kl} \chi_l = i\hbar \frac{\partial}{\partial t} \chi_k$$

where the exact nonadiabatic coupling operator C_{kl} is:

$$C_{kl} = \int \psi_k^* \left[-\sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 \right] \psi_l \, \mathbf{dr} + \frac{1}{M_I} \sum_I \left\{ \int \psi_k^* [-i\hbar \nabla_I] \psi_l \, \mathbf{dr} \right\} [-i\hbar \nabla_I]$$

Diagonal contribution C_{kk} depends only on single wave function ψ_k and thus represents correction to eigenvalue of electronic SE in this kth state. **Adiabatic approximation** considers only these diagonal terms, and for real wfns the green momentum term is zero, so we get complete decoupling:

$$\left[-\sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + E_k(\mathbf{R}_I) + C_{kk}(\mathbf{R}_I) \right] \chi_k = i\hbar \frac{\partial}{\partial t} \chi_k$$

Then can write $\Psi(\mathbf{r}, \mathbf{R}; t) \approx \psi_k(\mathbf{r}; \mathbf{R}) \chi_k(\mathbf{R}; t)$ as direct product of electronic and nuclear wave functions. Neglecting diagonal coupling term gives the familiar **Born-Oppenheimer approximation**.

Time dependence: semi-classical molecular dynamics

To make the method *semi-classical*, we need to get rid of the nuclear wave function χ and replace it with classical point particles. **Bohm method**: write χ in polar form $A(\mathbf{R})\exp(iS(\mathbf{R})/\hbar)$, insert in the equation defining the BO approximation on previous slide, and separate real and imaginary parts. One part is a continuity equation which keeps the nuclear probability density normalized, the other is an equation which reduces to the Hamilton-Jacobi equation of classical mechanics if you delete the term involving $\hbar$ (er.. '*taking the classical limit*'). Transform this to Newtonian form and you have:

$$M_I\ddot{\mathbf{R}}_I(t) = -\nabla V_k^{BO}(\mathbf{R}_I(t))$$

Thus nuclei move according to *classical* mechanics in an effective potential V_k^{BO} which is given by the Born-Oppenheimer potential energy surface E_k obtained by solving simultaneously the *time-independent* electronic Schrödinger equation for the kth state at the nuclear configuration R.

Alternative methods

AIMD	Nuclei	Electronic structure
Born-Oppenheimer	$M_I\ddot{\mathbf{R}}_I(t) = -\nabla_I \min_{\{\phi_i\}}\{\langle\Psi_0\|H_e\|\Psi_0\rangle\}$	$0 = -H_e\phi_i + \sum_j \Lambda_{ij}\phi_j$
Car-Parrinello	$M_I\ddot{\mathbf{R}}_I(t) = -\nabla_I\langle\Psi_0\|H_e\|\Psi_0\rangle$	$\mu\ddot{\phi}_i(t) = -H_e\phi_i + \sum_j \Lambda_{ij}\phi_j$
Ehrenfest	$M_I\ddot{\mathbf{R}}_I(t) = -\nabla_I\langle\Psi_0\|H_e\|\Psi_0\rangle$	$i\hbar\dot{\Psi}_0(t) = H_e\Psi_0$

Need to calculate forces!

CP and E propagate wave function dynamically and thus don't require explicit minimization of the total energy. Arguments about whether PC or BO is better are very boring.

Ab initio molecular dynamics

- AIMD based on DFT is an exceptionally powerful simulation tool for dynamical studies of several hundred atoms for times of order 100 ps (e.g. for water properties, biological systems, high-pressure systems, phase diagrams).
- Using AIMD many properties can be predicted for wide range of ρ, P and T. But some properties need accuracy beyond DFT (van der Waals' bonded systems, transition-metal compounds, calculation of excitation energies and energy gaps, accurate energy barriers in chemical reactions, bond dissociation processes etc. etc.).

Ab initio molecular dynamics

- AIMD based on DFT is an exceptionally powerful simulation tool for dynamical studies of several hundred atoms for times of order 100 ps (e.g. for water properties, biological systems, high-pressure systems, phase diagrams).
- Using AIMD many properties can be predicted for wide range of ρ, P and T. But some properties need accuracy beyond DFT (van der Waals' bonded systems, transition-metal compounds, calculation of excitation energies and energy gaps, accurate energy barriers in chemical reactions, bond dissociation processes etc. etc.).

How do we improve the accuracy of these properties over DFT?

Ab initio molecular dynamics

- AIMD based on DFT is an exceptionally powerful simulation tool for dynamical studies of several hundred atoms for times of order 100 ps (e.g. for water properties, biological systems, high-pressure systems, phase diagrams).
- Using AIMD many properties can be predicted for wide range of ρ, P and T. But some properties need accuracy beyond DFT (van der Waals' bonded systems, transition-metal compounds, calculation of excitation energies and energy gaps, accurate energy barriers in chemical reactions, bond dissociation processes etc. etc.).

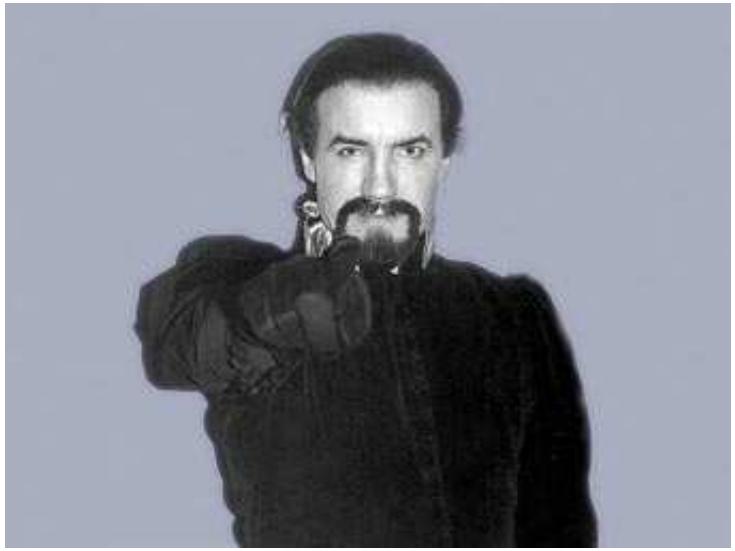
How do we improve the accuracy of these properties over DFT?

- Need to use a more accurate technique : options include QMC, GW, BSE, post-HF etc.. Good for static calculations but presumably very expensive for dynamics.

Let's try to couple DMC with AIMD somehow..

DMC-MD work first published by these people

Jeff Grossman (Lawrence Livermore)

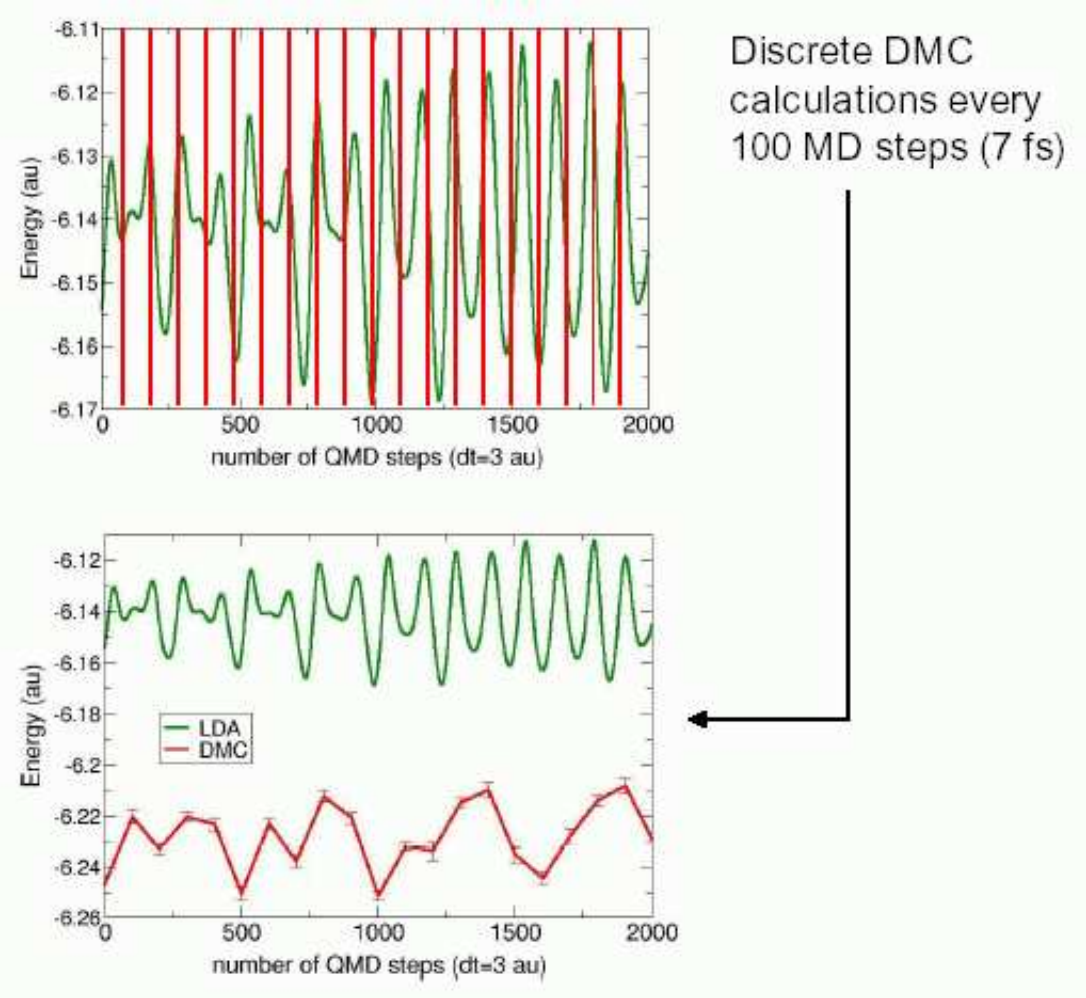


Lubos Mitas (North Carolina State University)

"*Efficient quantum Monte Carlo energies for molecular dynamics simulations*"
Phys. Rev. Lett., **94**, 056403 (2005).

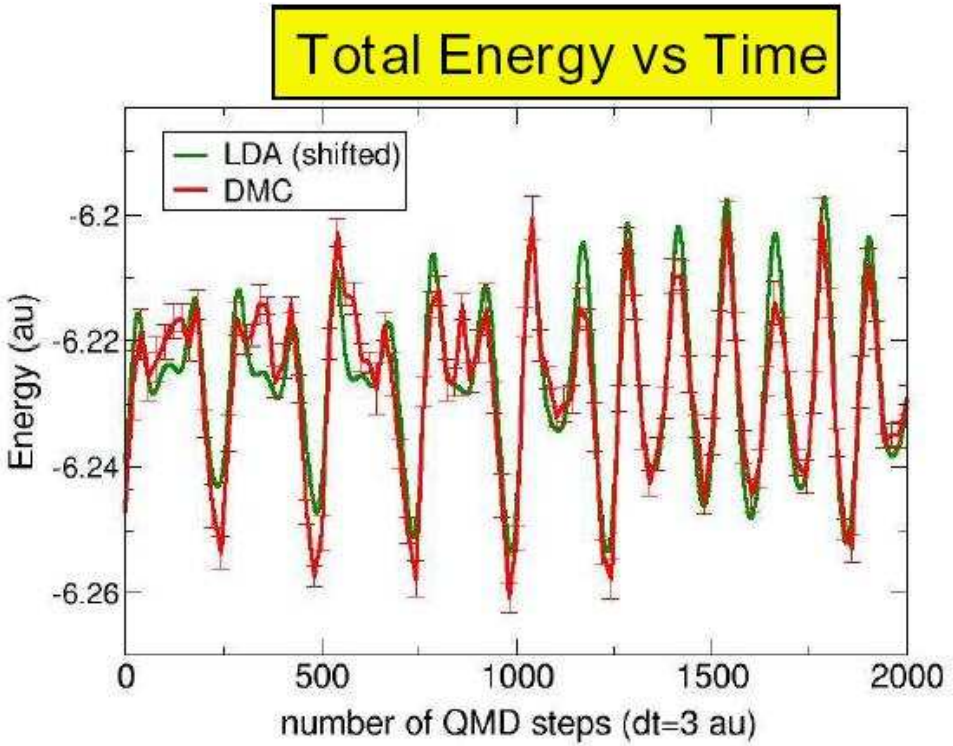
Snapshot approach

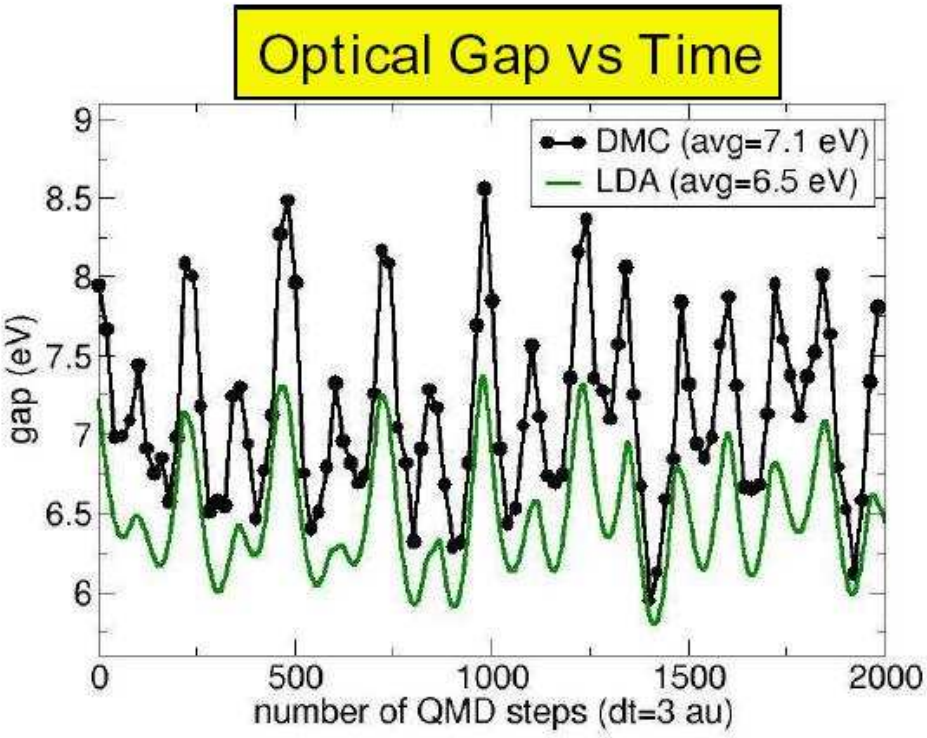
- Take snapshots of AIMD trajectory every so often and do full DMC calculation for each point e.g. silane molecule for $T \approx 1500$K.



- Very expensive! Discrete sampling may miss important events.

QMC energies in an AIMD simulation





Is there a more efficient way to use QMC in AIMD?

In DMC, walkers evolve in $3N_e$ dimensions, and the wave function is sampled according to the imaginary time-dependent Schrödinger equation.

$$\phi(r,\tau) = \exp(-\tau\hat{H})\phi(r,\tau=0)$$

But can the stochastic electronic propagation in DMC be coupled to MD?

Is there a more efficient way to use QMC in AIMD?

In DMC, walkers evolve in $3N_e$ dimensions, and the wave function is sampled according to the imaginary time-dependent Schrödinger equation.

$$\phi(r,\tau) = \exp(-\tau\hat{H})\phi(r,\tau=0)$$

But can the stochastic electronic propagation in DMC be coupled to MD?

Analogy with DFT-MD for coupling electrons and ions:

Born-Oppenheimer dynamics

$$\Phi = \min\left[E_{KS}\right]$$

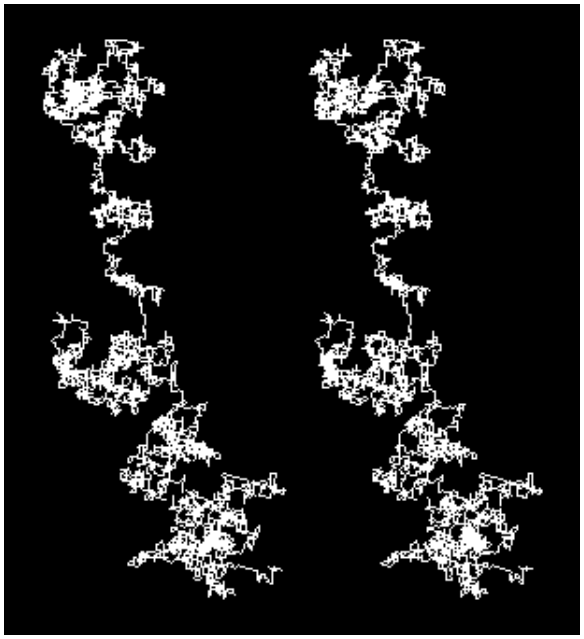
$$M_I\ddot{R}_I = -\nabla_I\Phi$$

Car-Parrinello dynamics

$$\mu\ddot{\Psi}_i = -\frac{\partial E_{KS}\left[\Psi,\mathbf{R}\right]}{\partial\Psi_i} - \sum\Lambda_{ij}\Psi_j$$

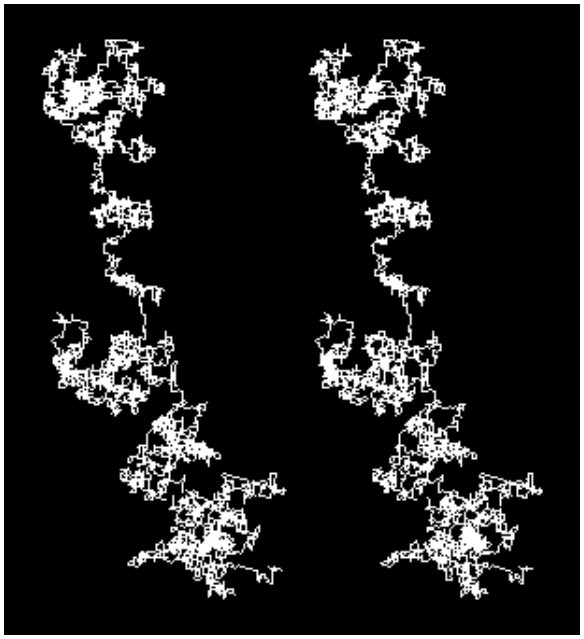
$$M_I\ddot{R}_I = -\frac{\partial E_{KS}\left[\Psi,\mathbf{R}\right]}{\partial R_I}$$

Basic idea



How far does an electron move
in a typical DMC time step?

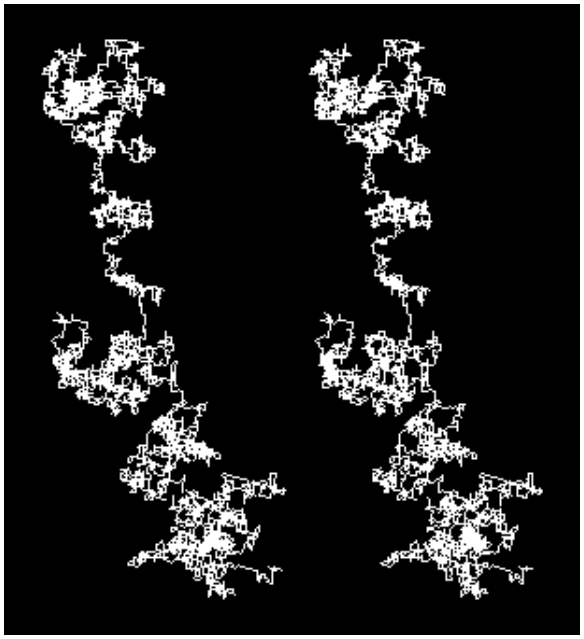
Basic idea



How far does an electron move
in a typical DMC time step?

$0.01 \ldots 0.1$ au

Basic idea

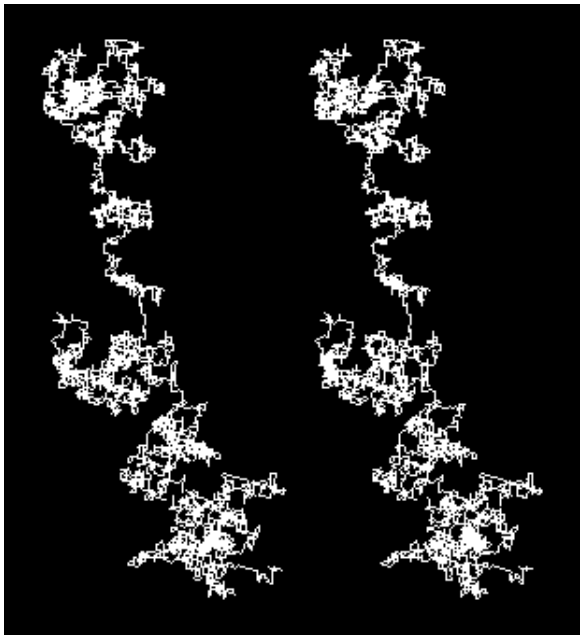


How far does an electron move
in a typical DMC time step?

$0.01 \ldots 0.1$ au

How far does an ion move
in a typical MD time step?

Basic idea



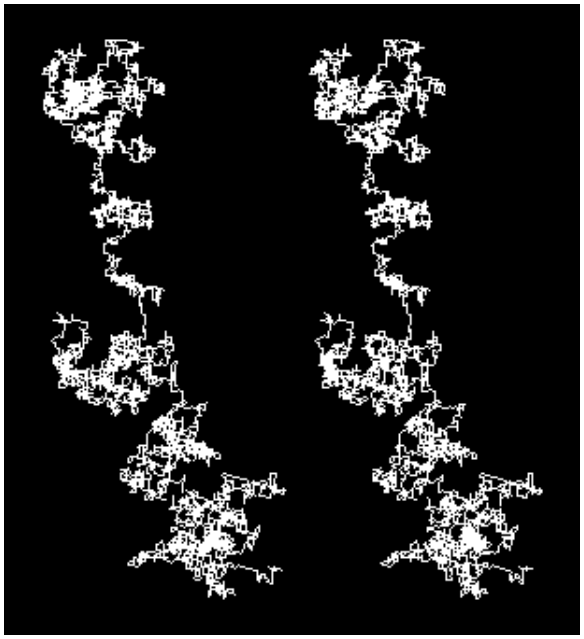
How far does an electron move
in a typical DMC time step?

$0.01 \ldots 0.1$ au

How far does an ion move
in a typical MD time step?

$0.0001 \ldots 0.001$ au

Basic idea



How far does an electron move
in a typical DMC time step?

$0.01 \ldots 0.1$ au

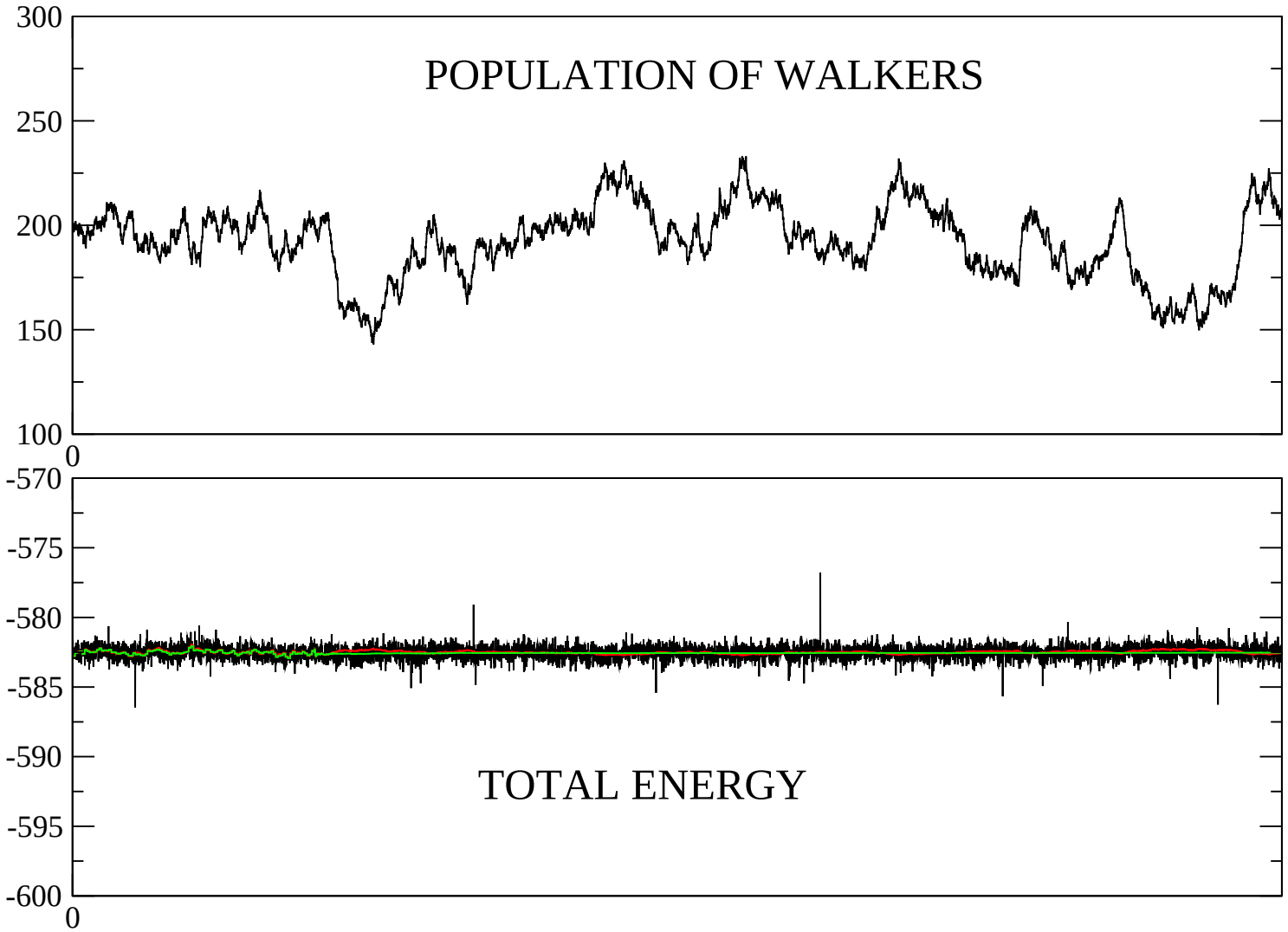
How far does an ion move
in a typical MD time step?

$0.0001 \ldots 0.001$ au

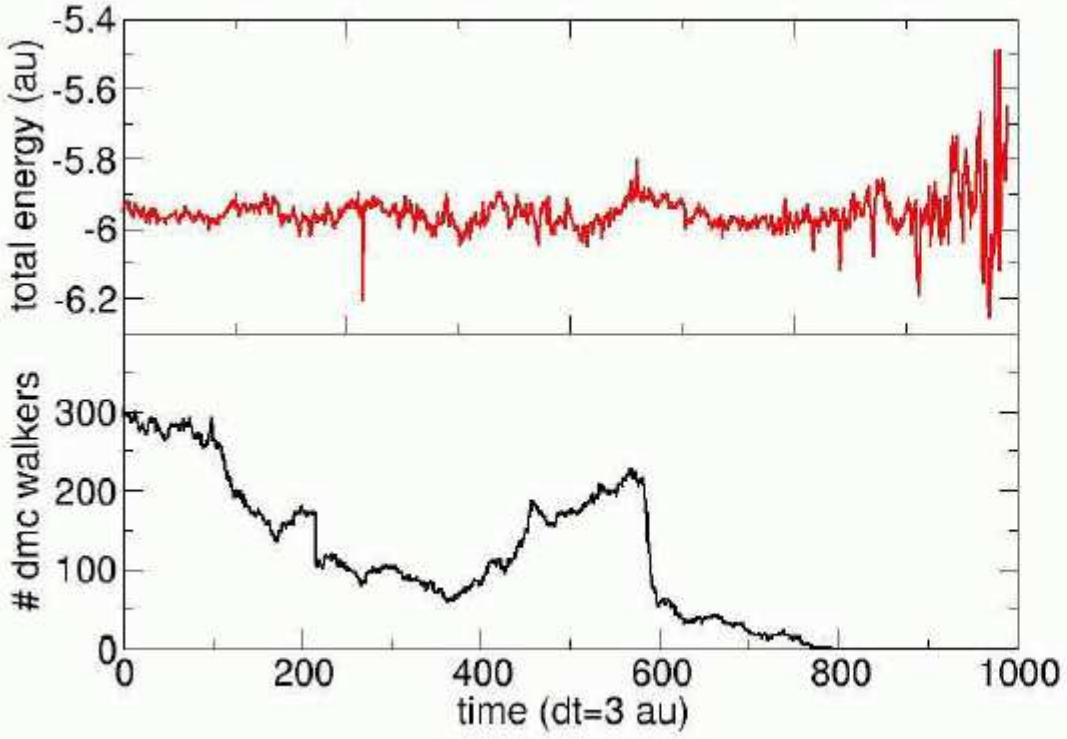
So, instead of discrete sampling of each point with new DMC run, calculate DMC energies 'on the fly' during the dynamic simulation. Continuously update the DMC walkers so that they correctly represent the evolving wave function.

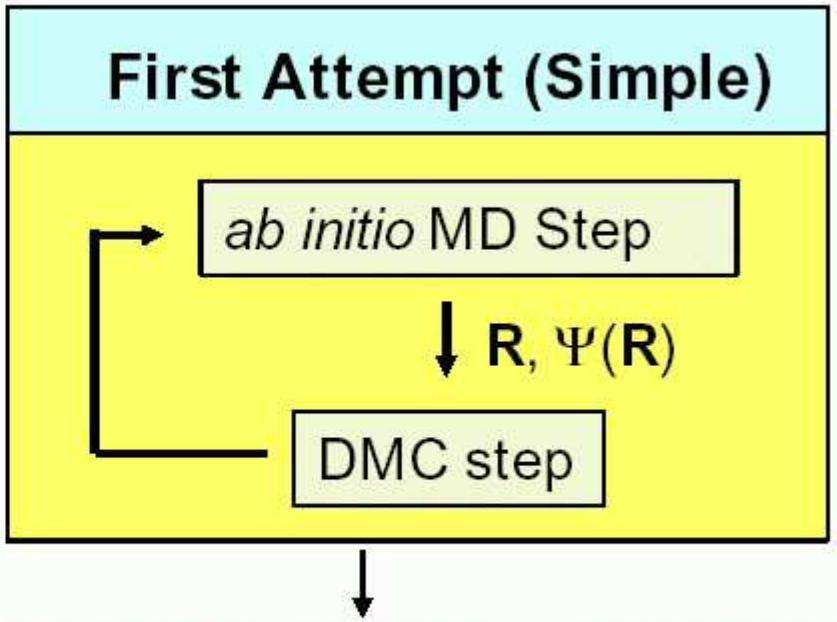
- Evolution of both configuration spaces is coupled : as the ionic dynamical trajectories evolve, so does the population of DMC electrons.
- The slow evolution in AIMD is highly advantageous in DMC!

What a DMC population *should* look like



Try coupled DMC-MD as described above





Unstable walker population

Oh dear!

Considerations

- Wave function changes as ions move : bias in walker distribution?

wrong wave function at the beginning of each step

Considerations

- Wave function changes as ions move : bias in walker distribution?

  wrong wave function at the beginning of each step

- DMC dynamics not right : problem with rejection step?

  detailed balance may no longer hold

Considerations

- Wave function changes as ions move : bias in walker distribution?

  wrong wave function at the beginning of each step

- DMC dynamics not right : problem with rejection step?

  detailed balance may no longer hold

- Dynamics of orbital occupations : orbital rotation, state crossings?

  swapping during dynamics

Considerations

- Wave function changes as ions move : bias in walker distribution?

  wrong wave function at the beginning of each step

- DMC dynamics not right : problem with rejection step?

  detailed balance may no longer hold

- Dynamics of orbital occupations : orbital rotation, state crossings?

  swapping during dynamics

- Wave function changes as ions move : node crossings?

  electrons can cross nodes and nodes can cross electrons

Fix the algorithm

DMC wave function bias (walker distribution)

At MD time step $t_{MD} = 0$ the DMC walker distribution is $f(\mathbf{R}, \tau; \hat{H}(0))$. At next MD step $t_{MD} = 1$ should be sampling $f(\mathbf{R}, \tau + \Delta\tau; \hat{H}(1))$. To sample correctly, need to start with $f(\mathbf{R}, \tau; \hat{H}(1))$ rather than original $f(\mathbf{R}, \tau; \hat{H}(0))$.

Use correlated sampling to reweight each walker

The propagation of walkers is governed by the DMC approximation for the Green's function which is a product of Gaussians describing diffusion and drift times a weight with an exponentiated local energy. Correlated sampling is used to correct the Green's function by a proper modification of the walker weights.

For each step, reweight each DMC walker by :

$$
\begin{aligned}
w(\mathbf{R}(\tau), t_{MD} = 1) &= \frac{G\left[\mathbf{R}(\tau) \leftarrow \mathbf{R}(\tau - \Delta\tau); \hat{H}(1)\right]}{G\left[\mathbf{R}(\tau) \leftarrow \mathbf{R}(\tau - \Delta\tau); \hat{H}(0)\right]} \\
&\cong \frac{\Psi_T^2\left[\mathbf{R}(\tau); \hat{H}(1)\right]}{\Psi_T^2\left[\mathbf{R}(\tau); \hat{H}(0)\right]} e^{-\Delta\tau\left(E_L\left[\hat{H}(1)\right] - E_L\left[\hat{H}(0)\right]\right)/2}
\end{aligned}
$$

Fix the algorithm II

Detailed balance

To account for MD steps, need to modify Green's function in the accept-reject step.

Orbital occupations

Possible that during dynamics the ordering of states is changed. Can detect orbital swapping before the first DMC step of a new MD step by computing the overlap of all orbitals with one another. Easy to estimate since same set of walkers can be used to perform integral.

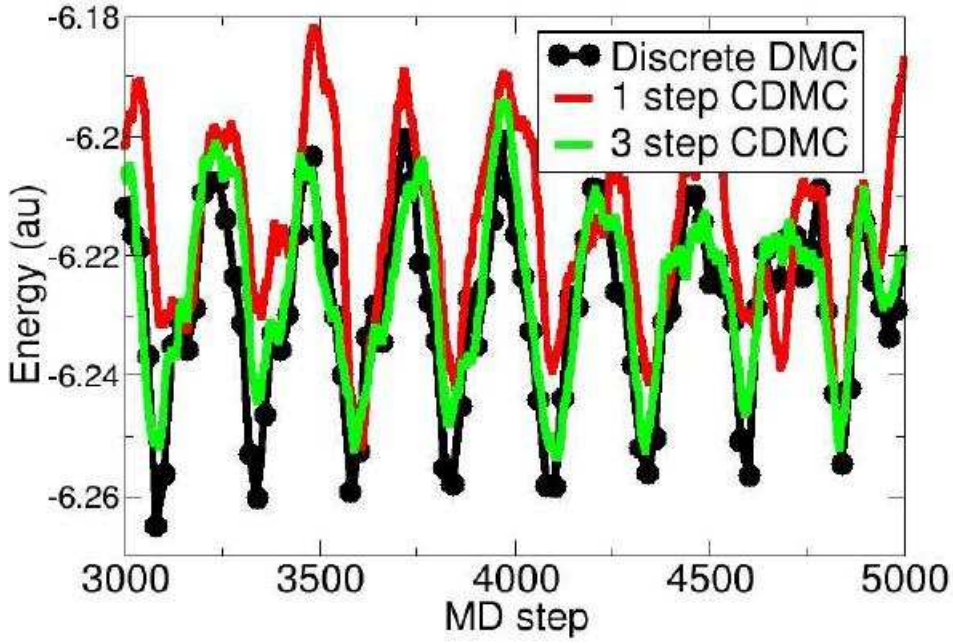
If unoccupied orbital swapped with occupied, simplest solution is to reconverge new set of walkers from scratch. 'Brute force' approach! However, this seems to happen quite rarely. For excited state of silane with near-degenerate orbitals accessible through fluctuations, happened only once in the simulation. For ground state, doesn't happen at all.

Node crossing

Eliminate walkers that cross nodes..

How many DMC steps needed per MD step?

Use large discrete sampled runs (1000 steps each) for comparison.



- As simulation progresses, 1-step CDMC energies begin to differ significantly from discrete DMC. Using 3 steps corrects 'time lag'. About an order of magnitude more efficient than discrete sampling.

Thermal Averages (over 1 ps)	
E(discrete DMC)	= - 6.228(2)
E(1 step continuous)	= - 6.220(2)
E(2 steps continuous)	= - 6.220(2)
E(3 steps continuous)	= - 6.226(2)
E(10 steps continuous)	= - 6.230(2)
E(20 steps continuous)	= - 6.228(2)

Thermal averages appear to be converged for $N \geq 3$.

Efficiency

Perform 3-step CDMC on Si_nH_n clusters :

Cluster	AIMD time	CDMC time	DMC steps	configs/node	nodes
SiH_4	12	10	3	300	16
Si_5H_{12}	35	45	3	300	32
$Si_{14}H_{20}$	106	56	3	50	128
$Si_{35}H_{36}$	310	290	3	35	256

Time for CDMC comparable to AIMD!

Grossman/Mitas also quote some interesting results:

- H_2O molecule dissociation path from constrained AIMD. CDMC energetics agree with discretely sampled result. Dissociation energy (127(2) kcal/mol) agrees with experiment (125.9 kcal/mol). Can describe bond-breaking..
- Calculated heat of vaporization of water (from full liquid simulations of 32 water molecules). Liquid binding energy (average CDMC energy for the 32 water system minus 32 times the DMC energy of a single water molecule) + 1.5 kcal/mol to account for the quantum motion of the H atoms. Results - LDA-PBE : $H_{vap} = 6.2(1)$, CDMC : $H_{vap} = 9.1(4)$, Expt. : $H_{vap} = 9.92$ kcal/mol.

Forces

So far :

- Accurate energies (at the level of fixed-node DMC).
- Can compute energy differences on the fly (couple two DMC electronic populations to the AIMD simulation e.g. one for the ground state and one for a given excited state).
- Good for calculations of excited states and thermodynamic averages.

However the ionic trajectories were all generated using DFT up to now.

Is it possible to incorporate QMC forces into the simulation?

Forces

So far :

- Accurate energies (at the level of fixed-node DMC).
- Can compute energy differences on the fly (couple two DMC electronic populations to the AIMD simulation e.g. one for the ground state and one for a given excited state).
- Good for calculations of excited states and thermodynamic averages.

However the ionic trajectories were all generated using DFT up to now.

Is it possible to incorporate QMC forces into the simulation?

- Using a standard method (basically finite differencing with some fancy tricks) Grossman and Mitas computed the DMC forces at each MD step for SiH_4 and compared results with LDA at around 1500K. Excellent agreement for Si and H.

Conclusions

- Grossman and Mitas have proposed a new method for coupling *ab initio* MD ionic steps with stochastic DMC electronic steps to provide accurate DMC energies on the fly.
- The technique exploits the slowness of MD evolution which enables the QMC sampling process to be updated very efficiently.
- Accurate for both thermal averages and description of energies along pathways.
- They have also carried out the first QMC/MD simulations using both forces and energies from QMC (albeit not very good forces). Presumably we could do this better now.

Running DMC-MD calculations with CASINO

Script `runqmcmd` used to automate DMC-MD using CASINO and the PWSCF DFT code (part of Quantum Espresso available for free at `www.quantum-espresso.org` - must use version 4.3+).

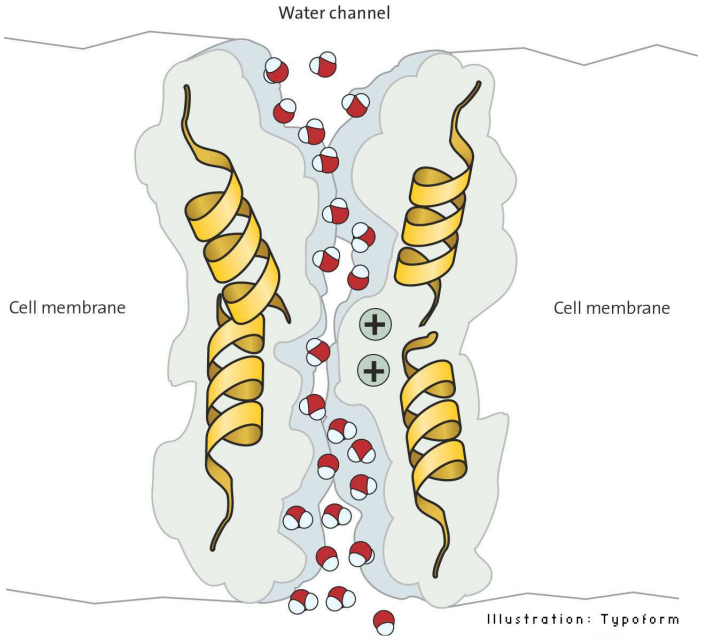
The script works by repeatedly calling CASINO run scripts `runpwscf` and `runqmc` which know how to run the two codes on any known machine. See instructions in `utils/runqmcmd/README`.

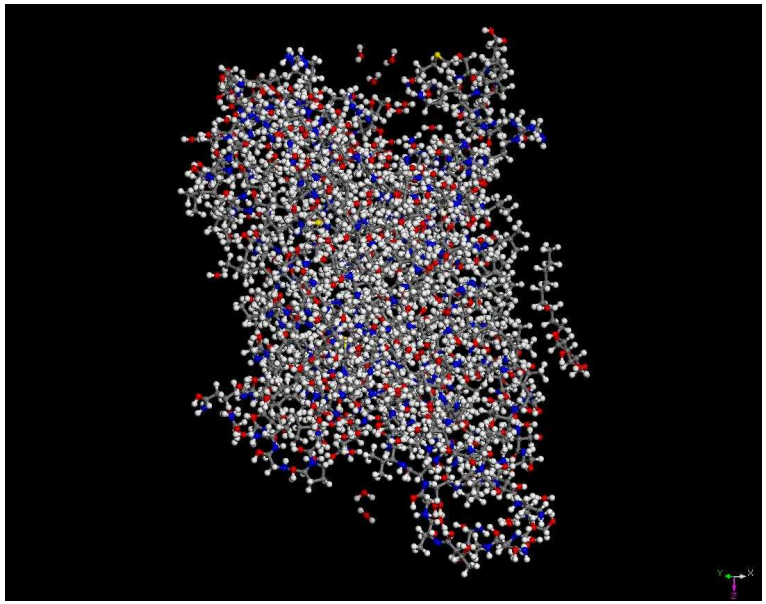
Setup the PWSCF input ('`in.pwscf`') and the CASINO input ('`input`' etc. but no wave function file) in the same directory. For the moment we assume you have an optimized Jastrow from somewhere (this will be automated later). Have the PWSCF setup as '**calculation = "md"**', and '**nstep = 100**' or whatever. The `runqmcmd` script will then run PWSCF once to generate 100 `xwfn.data` files, then it will run CASINO on each of the `xwfn.data`. The first will be a proper DMC run with full equilibration (using the values of **DMC_EQUIL_NSTEP**, **DMC_STATS_NSTEP** etc.). The second and subsequent steps (with slightly different nuclear positions) will be restarts from the previous converged config.in - each run will use new keywords **DMCMD_EQUIL_NSTEP** and **DMCMD_STATS_NSTEP** (with the number of blocks assumed to be 1. The latter values are used if new keyword **DMC_MD** is set to T, and they should be very small).

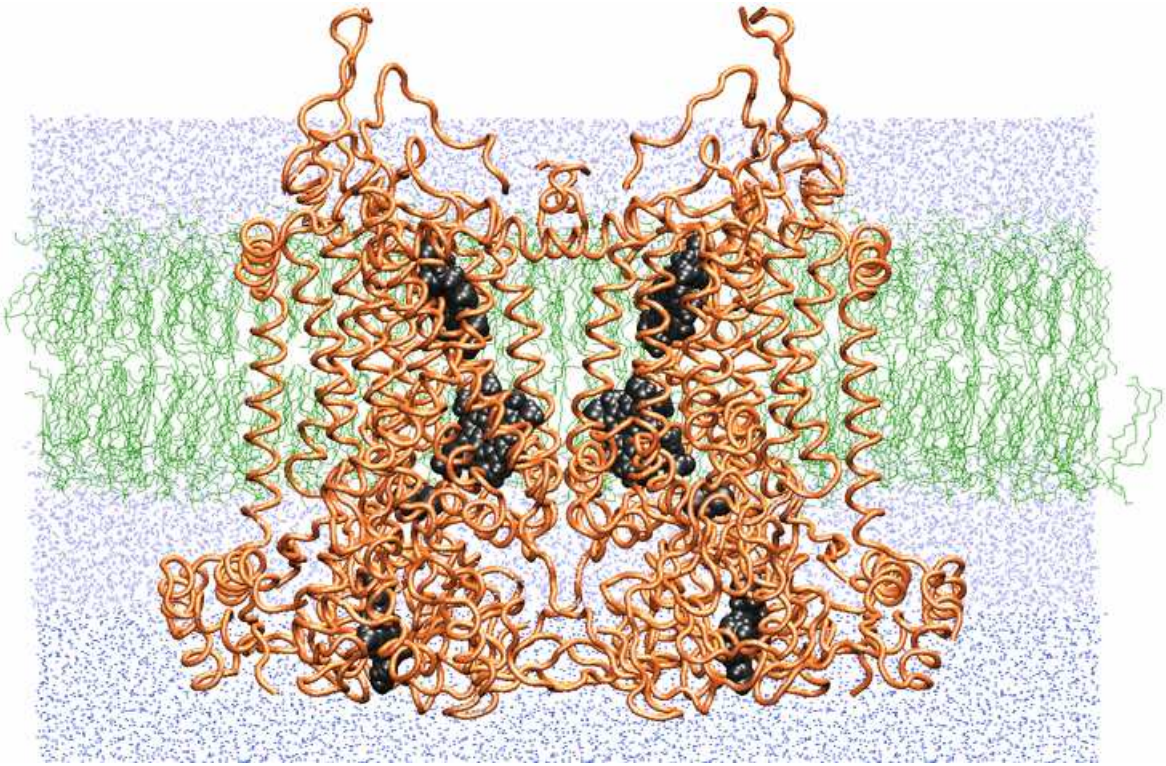
It is recommended that you set **DMC_SPACEWARPING** AND **DMC_REWEIGHT_CONF** to T in CASINO input when doing such calculations.

The calculation can be run through `pwfn.data`, `bwfn.data` or `bwfn.data.b1` formats as specified in the `pw2casino.dat` file (see CASINO and PWSCF documentation).

So obviously we'll be doing this sort of thing soon







But! What question am I trying to answer?

Dario asked me to "*show that GM-MD was* xxx *times faster than standard sampling*".

This question is based on the claims in the original Grossman-Mitas paper that:

- '*This continuous evolution of the QMC electrons results in highly accurate total energies for the full dynamical trajectory at a fraction of the cost of conventional, discrete sampling.*'
- '[It provides]*an improved, significantly more accurate total energy for the full dynamical trajectory*.'
- '*This approach provides the same energies as conventional, discrete QMC sampling and gives error bars comparable to separate, much longer QMC calculations.*'

A casual reading of this might lead to the conclusion that we get effectively the same results as conventional DMC at each point along the trajectory. As we are only doing e.g. 3 DMC stats accumulation per step, we thus appear to be getting 'something for nothing'. What are we missing?

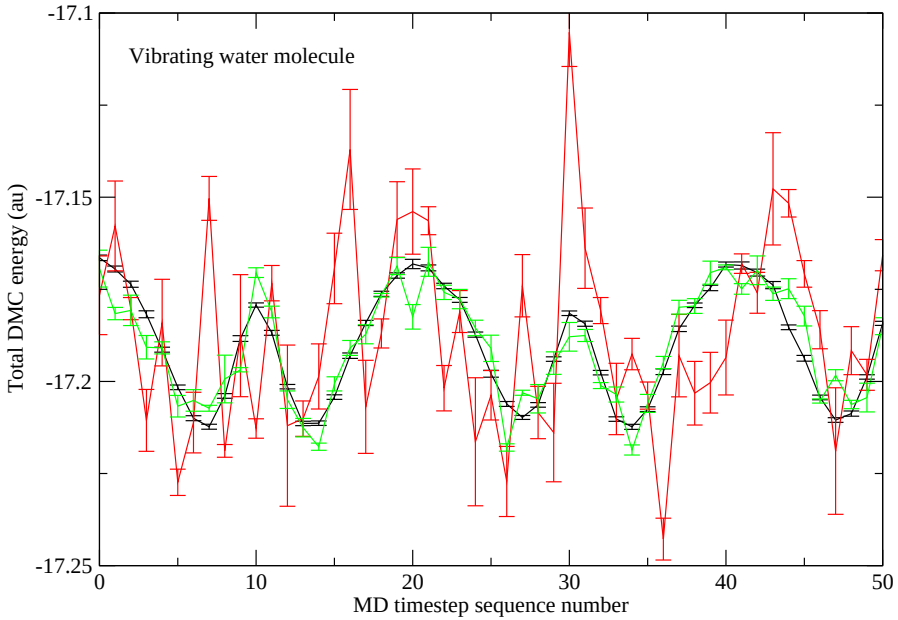
What are we missing?

(1) The point about GM-MD - and Lubos agrees with this - is that its true purpose is to calculate *thermodynamic averages* - an example being given in the paper of the heat of vaporization of H_2O. This is calculated as an average over the evolution path of the ions (at a given $T > 0$) and QMC and thermodynamic averages are done at the same time. Think of it as a method of Monte Carlo sampling a distribution in 3N-dimensional configuration space that is changing shape in time. Each nuclear configuration is sampled via far fewer (e.g. three moves times the number of walkers) electron configurations than usual. It is not intended that accurate energies for all nuclear configurations are calculated, only an accurate Monte-Carlo-sampled 'thermodynamic average' as the molecule (or whatever) vibrates or otherwise moves. This is not really made clear in the paper.

(2) If you want accurate answers and small error bars for the energies at each point along the MD trajectory then - done under normal conditions (number of DMC walkers etc.) - the method does not save you any time at all, apart from that spent doing DMC equilibration. It generally gives very noisy answers with huge error bars for the individual MD points - if you want proper DMC accuracy then you have to do the usual amount of statistical accumulation work.

(3) The only reason they are able to claim '*the same energies* [and error bars] *as conventional, discrete QMC sampling*' is because they use extremely large populations of walkers (this is difficult to scale to large systems). In the paper it is stated that '*we chose a number of walkers such that the statistical fluctuations in the GM-MD energies are one-tenth the size of the variation in the total energy as a function of the MD time* but it should be emphasized that this is considerably more than usual.

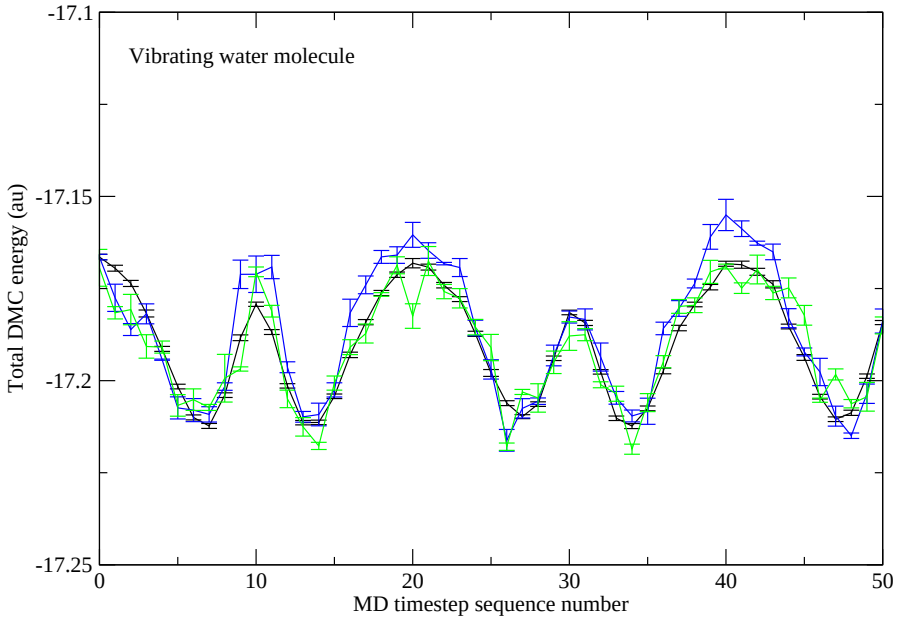
Some illustrative calculations (not GM-MD!)



Discrete DMC energies for vibrating water molecule. Black curve: converged DMC result with small error bar. Red curve: 1000 walkers (a 'typical' number), 1000 equil moves, 3 stats moves. Green curve: 14400 walkers, 1000 equil moves, 3 stats moves

- The red curve essentially does not follow the accurate curve. It can be made to do so by using a sufficiently large population of walkers - the green curve repeats the red calculations, but using 14400 walkers (14 times more) and the same number (3) of stats accumulation moves. This is effectively what GM do in order to claim '*the same energies as conventional discrete QMC sampling*' (though note the error bars in both the red and green curves are not accurate, there being insufficient data - only 3 moves worth - to calculate them properly).
- Because they do not mention the number of walkers used, it is possible to misconclude the nature of the speedup from their data e.g. if we normally do 3000 moves and now we're doing 3, we might think that a GM-MD calculation is 1000 times faster, but it's not because we're using 14 times more configs than usual.

Some GM-MD results



Discrete and GM-MD DMC energies for vibrating water molecule. Black curve: converged discrete DMC result with small error bar. Green curve: 14400 walkers, 1000 equil moves, 3 stats moves. Blue curve: 14400 walkers, no equil moves (each MD point restarted from previous), 3 stats moves.

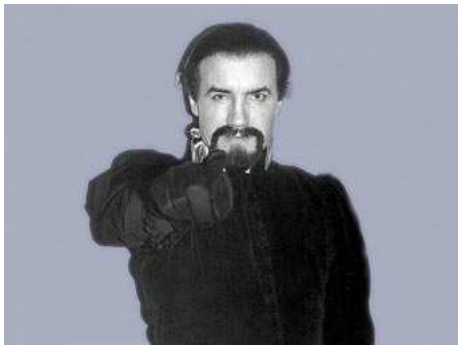
- A GM-MD restarted calculation is an *approximation* to the previous green curve (repeated above) (the approximation is done in order to save the time normally spent doing DMC equilibration).
- The new blue curve above is the GM-MD result with 14400 configs and 3 moves, but skipping the equilibration step for all MD steps after the first, and restarting from the equilbrated walkers from the previous step. Walker reweighting and spacewarping were both turned on.
- The GM-MD curve follows the green curve - where the equilibration is done explicitly - quite well (though its agreement with the full black curve is possibly not as good as that of the green curve, which is expected, as the reweighting is an additional approximation).
- Note that the approximation does not appear to 'get worse' over the course of MD time i.e. the error does not build-up.

Equilibration of walkers

Question: Can configs read at the restart be properly equilibrated for the new Ψ in so few moves?

- Can do 'extreme GM-MD' by just reweighting the configs when you read them in, recomputing the energy with the reweighted configs, and then not doing any DMC accumulation moves at all. The energies will be shifted compared to those of the original DFT calcs, but this is obtained almost in its entirety by the full DMC equilibration carried out at the initial nuclear configuration.
- Can reproduce demonstration calculation in GM paper this way. (where GM-MD total energies are shown tracking the discrete QMC energies for vibrating Si_xH_y molecules). The DMC energy curve simply runs parallel to the DFT energies. DMC does not demonstrate any new physics besides a shift in the total energy. This shift is obtained by the initial equilibration, then it is simply translated on along the trajectory. Nothing gained. Why do DMC-MD at all in this case?
- One might suspect that this issue should be important when calculating a trajectory where DMC shows a feature in the energy landscape that is not present in DFT. The GM approach ought simply to 'iron out' these features because the distribution does not have enough time to equilibrate into the special quantum-correlations that cause the energy differences.
- In the end, if you want to get any meaningful DMC energies along a DFT trajectory, it is essential to do some re-equilibration for each MD step to allow the population to properly respond to the new nuclear configuration. If this re-equilibration is too short, any DMC-specific features will be smeared out in the DMC-energy curve.
- Note equilibration is exponential process and timescale is determined by physics rather than initial distribution. The initial distribution largely determines the magnitude of the DMC energy error, not the rate at which it decays. Hence would think that if you want to do a good job of equilibrating (significantly improve your distribution) *you always need to equilibrate for at least time taken to diffuse across the longest physically relevant length scale*. TESTING REQUIRED.

GM-MD conclusions



- Rather than using 1000 times fewer stats accumulation moves than normal, and 10 times more configs than normal (say), would get same amount of statistical sampling using 100 times fewer stats accumulation moves with the usual number of configs. However, in the latter case have more propagation in imaginary time - expect wave function to adjust better to new nuclear configuration. So why not do this? Certainly it sounds more impressive to say 'you only need to do 2 moves' without mentioning larger number of configs required.
- Despite the fact that having a large number of walkers relative to the number of moves might facilitate parallelization, it remains the case that the total CPU time does not change when (number of walkers $\times$ number of DMC steps) is fixed.
- One might expect that it would be more efficient to sample widely differing nuclear configurations to get a thermodynamic average, rather than ones extremely close together as done in GM-MD.
- It would be nice to see - stated clearly and succinctly in a few lines, why we *expect* GM-MD to give us any speedup at all, even for thermodynamic averages.
- Problem: use of DFT forces! If LDA gives wrong structure then so does GM-MD. However DFT forces not intrinsic requirement; could use QMC forces in principle..

What to do next with this..? Hmmmm..

What have other people done?

Can just try direct '*DMC for electrons and nuclei*' by treating the nuclei in the same way as electrons [Ceperley and Alder, *PRB* **36**, 2092 (1987)]. However, no temperature effects and significant timescale separation problem, even for H.

- Proton is 1836 times more massive than the electron, it's diffusion with DMC (but not VMC, since the $1/m$ appears in the diffusion constant derived from the Green's function) is that much slower, and its root-mean-square displacement per step is hence 42 times smaller than that of the electron. Obviously much worse for higher Z nuclei!
- While the electronic distribution converges rapidly to its ground state, it is easy to find situations where the protonic distribution does not equilibrate in a reasonable amount of computer time. In principle the simulations for the electron-proton system several orders of magnitude longer than for a 1-component system, and this is not generally practical.

Obvious points

- Must thoroughly equilibrate initial ensemble by VMC so that for accurate trial functions the nuclear distribution will be close to the correct one.
- In crystal phases, the motion of the nuclei is severely limited in any case, so that the relevant equilibration time is much smaller (the inverse of the Debye temperature, in fact).
- Take care to use an accurate approximation to the nuclear wave function, if possible.

Can also use *restricted path-integral Monte Carlo* [Pierleoni, Ceperley, Bernu, Magro *PRL* **73**, 2145 (1994)] which uses the thermal density matrix to treat finite T electrons and nuclei. Very expensive, plus sampling problem at low temperature.

Coupled electron-ion Monte Carlo (CEIMC)

Born-Oppenheimer separation of time scales. Ground-state electrons, finite T nuclei. Samples the nuclear configuration space rather than trying to follow a dynamical trajectory.

- Metropolis Monte Carlo for the finite T nuclei. Generate Markov chain of nuclear configurations $\mathbf{R}$ according to the classical Boltzmann distribution $P(\mathbf{R}) = \exp[-\beta E_{BO}(\mathbf{R})]$. Propose move from $\mathbf{R}$ to $\mathbf{R}'$ and accept with probability

$$A(\mathbf{R} \longrightarrow \mathbf{R}') = \min\left[1, \frac{T(\mathbf{R}' \longrightarrow \mathbf{R}) e^{-\beta E_{BO}(\mathbf{R}')}}{T(\mathbf{R} \longrightarrow \mathbf{R}') e^{-\beta E_{BO}(\mathbf{R})}}\right]$$

  After a finite number of steps, the random walk will visit the states of the nuclear configuration space with a frequency proportional to their Boltzmann weight.
- Need to calculate full QMC energy E_{BO} for each nuclear configuration - expensive! Estimate of E_{BO} subject to noise - deal with this using *Penalty Method* - essentially requiring detailed balance to hold on average and not for any single energy configuration.
- Above requires evaluation of energy difference and the noise between two nuclear configurations (with all nuclei moved). Use *correlated sampling* to evaluate difference in 1 calculation rather than 2 separate ones.
- *Pre-reject* really stupid nuclear configurations using a classical potential model.
- Can apparently incorporate 'quantum nuclei' at little extra cost using a path-integral type thing.

See e.g. '*Computational methods in coupled electron-ion Monte Carlo simulations*', C. Pierleoni and D.M. Ceperley, ChemPhysChem **6**, 1872 (2005) and Pierleoni 2005 ESDG talk (see ESDG web page).

Other approaches: Attaccalite and Sorella

CEIMC very expensive since reasonable acceptance probability requires statistical error bars on energy of order kT, and amplitude of nuclear moves has to be decreased with increasing system size.

Attaccalite and Sorella have made an interesting proposal for an AIMD using noisy QMC forces, with a method that does not contain any Metropolis rejection scheme (at the expense of the usual MD time discretization error).. Finite T MD simulation requires some external noise on the forces, but you get this for free with QMC!

- Use a generalized Langevin dynamics, i.e. you add two extra force terms - a frictional one $\gamma\mathbf{v}$ (proportional to the nuclear velocity) and a random one η to Newton's equation, in order to approximate the effects of neglected degrees of freedom. Thus : $\dot{\mathbf{v}} = F(\mathbf{R})/m - \gamma(\mathbf{R})\mathbf{v} + \eta(t)$
- From *fluctuation-dissipation theorem* the friction matrix γ is related to the temperature T by

$$\gamma(\mathbf{R}) = \frac{1}{2T}\alpha(\mathbf{R})$$

  where α is a symmetric correlation matrix giving statistical correlation between force components.
- Whatever – ignoring details the practical upshot is that for a given noise on the forces, and a desired temperature, you can set the friction tensor so that the dynamics produces nuclear configurations distributed according to the classical Boltzmann distribution.

See e.g. '*Stable liquid hydrogen at high pressure by a novel ab initio molecular dynamics calculation*', C. Attaccalite and S. Sorella, *PRL* **100**, 114501 (2008).

Some simple things that could be done with CASINO

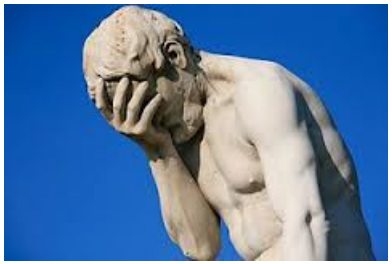
No reason we can't do DMC with light nuclei as quantum particles like in the Ceperley and Alder 1987 work mentioned earlier (computers are a bit faster these days!). Would be useful e.g. in hydrogen at high pressures - which is currently fashionable, see. e.g. RN/Chris Pickard work..

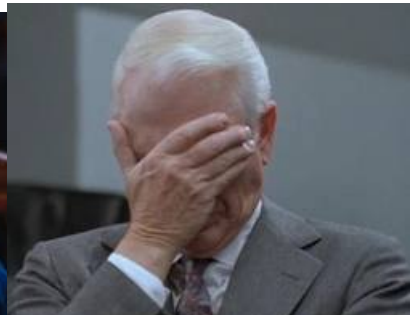
DMC calcs with quantum nuclei have been done more recently for hydrogen, see e.g. Natoli, Martin and Ceperley *PRL* **70**, 1952 (1993) and subsequent work (this would presumably be much easier now, and there is a greater knowledge of likely structures for high-pressure phases).

- Although DMC is a zero-temperature method, this would allow CASINO to e.g. calculate zero-point energies including anharmonicity (currently we have to get separate estimates from quasiharmonic DFT phonon calculations). Likely to be important to get this right.
- We would need to put some thought into representing the nuclear wave function. A good start (as Ceperley does) would be to use Gaussian orbitals centred on lattice sites with an optimizable width (no need to put them in a determinant, since they are effectively distinguishable), then stick some additional variational parameters describing nuclear-nuclear separations in the Jastrow (and possibly backflow) functions that form part of the standard Slater-Jastrow(-backflow) wave function used in CASINO.
- Apart from the extra wave function evaluation bits, need to include the masses in the relevant places in the DMC propagation routines, and a few more bits of administration. Relatively simple to code up!

Further though needs to be given to including temperature and to doing proper molecular dynamics trajectories to follow particular processes. It wouldn't be that difficult to code up something along the lines of CEIMC or Attacalite's work. Or is there something else we can do?

Let's use de Broglie-Bohm theory!

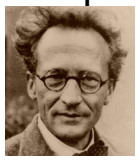




Classical atoms with *Newtonian* trajectories

Classical atoms are small and we cannot know their position with certainty, so we deal with a statistical ensemble in which only the *probability density* $\rho(\mathbf{x},t)$ is known.

- Probability must be conserved, i.e. $\int \rho d^3x = 1$ for each t. Therefore must satisfy continuity equation $\partial\rho/\partial t = -\nabla\cdot(\rho\mathbf{v})$ where $\mathbf{v}(\mathbf{x},t)$ is the velocity of the particle.
- Classical mechanics has various equivalent formulations. Choose the less well-known *Hamilton-Jacobi* version, where velocity $\mathbf{v}(\mathbf{x},t) = \frac{\nabla S(\mathbf{x},t)}{m}$ and $S(\mathbf{x},t)$ - related to the 'action' - is a solution of the Hamilton-Jacobi equation, $-\frac{\partial S}{\partial t} = \frac{(\nabla S)^2}{2m} + V$.
- Can write the two green *real* equations more elegantly as a single *complex* equation. To do this, introduce a complex function $\Psi = \sqrt{\rho}e^{\frac{iS}{\hbar}}$ where $\hbar$ is arbitrary constant giving dimensionless exponent. The two equations may then be rewritten as:



$$i\hbar\frac{\partial\Psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V - Q\right)\Psi \qquad \text{with} \qquad Q = -\frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}}.$$

This is the time-dependent Schrödinger equation (!) with an extra term Q. Note $|\psi(\mathbf{x},t)|^2$ has same interpretation as in QM: a probability density of particle positions. *So to recover classical mechanics from quantum mechanics we simply have to subtract out something that behaves exactly like a potential, thus implying that* **QM is just like classical statistical mechanics with a non-classical dynamics** *(due to an 'extra force' $-\nabla Q$ over and above the classical $-\nabla V$).*

First-order de Broglie-Bohm ('pilot-wave') theory

- Wave field evolution from Schrödinger equation $i\hbar\frac{\partial\Psi}{\partial t} = \sum_{i=1}^{N} -\frac{\hbar^2}{2m_i}\nabla_i^2\Psi + V\Psi$. Evolving quantum system behaves like 'probability fluid' of density $|\Psi|^2 = \Psi\Psi^*$ with an associated time-dependent quantum *probability current* $\mathbf{j} = \frac{\hbar}{m}\mathrm{Im}(\Psi^*\nabla\Psi)$.
- Suspect particle trajectories follow streamlines of current: velocity $\mathbf{v} = \frac{\hbar}{m}\mathrm{Im}\nabla\ln\Psi$ (current/density). Using complex polar form $\Psi = |\Psi|\exp[iS/\hbar]$, the wave function *phase* $S(\mathbf{x}_1,\ldots,\mathbf{x}_N,t)$ is given by $S = \hbar\mathrm{Im}\ln\Psi$ (similar to velocity expression). Thus deduce trajectories $\mathbf{x}_i(t)$ given by *de Broglie guidance equation* for velocity:

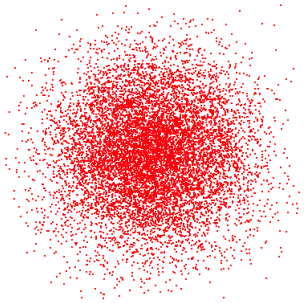
$$\mathbf{v}_i = \frac{d\mathbf{x}_i}{dt} = \frac{\nabla_i S}{m_i}$$

- Can write in 2nd-order '$F = ma$' form by taking t derivative: $m_i\ddot{\mathbf{x}}_i = -\nabla_i(V+Q)$, where $Q = -\sum_i \frac{\hbar^2}{2m_i}\frac{\nabla_i^2|\Psi|}{|\Psi|}$ (*quantum potential*). Extra 'quantum' force $-\nabla_i Q$ (big if large curvature in wave field). *Non-classical* dynamics since particles 'pushed along' by wave along trajectories perpendicular to surfaces of constant phase, as well as by classical force from other particles.
- Guidance equation *identical* to trajectory equation in Hamilton-Jacobi theory - a standard form of classical mechanics like Hamiltonian or Lagrangian dynamics. There S is indefinite integral of classical Lagrangian with respect to t (note the 'action' is the *definite* integral with fixed endpoints). Suggests immediately how to obtain the classical limit, i.e. when $Q = \nabla Q = 0$ the wave component of matter is passive and the particles follow classical trajectories (impossible in orthodox QM!). *This is what I did at the start when justifying semi-classical MD*.

Stochastic pilot-wave theories

To put DMC in deBB context, must first understand concept of *stochastic* pilot-wave theories. Additional random motion introduced in 1954 by Bohm and Vigier in context of why particles distributed as $|\Psi|^2$ (though no need - see MDT 2012). Imagine a deeper 'sub-quantum' level which imparts additional intrinsic randomness to particle motion (like in Brownian motion with pollen grains being hit by clouds of atoms). Velocity of individual particle is $\mathbf{v} = \frac{\nabla S}{m} + \xi(t)$ with $\xi(t)$ a chaotic contribution to the velocity fluctuating randomly with zero average. Usual $\frac{\nabla S}{m}$ trajectory produced by guiding equation thus *average* velocity not *actual* one.

- Assume - whatever its origin - stochastic process treatable as simple *diffusion*. With prob density P, diffusion constant D, there is diffusion current $\mathbf{j} = -D\nabla P$ and a conservation equation $\partial P/\partial t = -D\nabla^2 P$. Leads clearly to *uniform distribution* (change in P stops at zero density curvature, like ink drop spreading in water).



- If want *non-uniform* final distribution there must be another field giving rise to an *osmotic velocity*. **Example**: Einstein showed if add gravitational field in z-direction this velocity is $u = D\frac{mg}{kT}z$, the conservation equation becomes $\frac{\partial P}{\partial t} = -D\nabla\left[\frac{mg}{kT}zP + \nabla P\right]$. In equilibrium when $\frac{\partial P}{\partial t} = 0$ we have $\frac{\nabla P}{P} = \frac{mg}{kT}z + c$ or $P = A\exp(-\frac{mgz}{kT})$ - the Boltzmann factor.
- In stochastic pilot-wave theory require random diffusion process whose equilibrium state corresponds to prob density $P = |\Psi|^2 = \rho$ and mean current $\mathbf{j} = \rho\mathbf{v} = \rho(\frac{\nabla S}{m})$. Consistent possibility if $\Psi = \sqrt{\rho}\exp(\frac{iS}{\hbar})$ as this implies conservation equation $\frac{\partial \rho}{\partial t} + \nabla\cdot\mathbf{j} = 0$. Can be shown suitable osmotic velocity is $\mathbf{u} = \frac{D\nabla\rho}{\rho}$ - then follows there is an equilibrium state with $P = \rho$ in which osmotic velocity is balanced by diffusion current so the mean velocity is $\frac{\nabla S}{m}$.

DMC vs. stochastic pilot-wave theories

In the various theories at each timestep get change in particle position $\mathrm{d}\mathbf{r}$ from some combination of guided velocity, random diffusion and a drift (osmotic) velocity. The χ in the diffusion part is a random variable with zero mean and unit variance. Atomic units are dispensed with ($\hbar$ and m are back).

Standard pilot wave	$d\mathbf{r} = \frac{\nabla S}{m}\mathrm{d}t$
Stochastic pilot wave	$d\mathbf{r} = \frac{\nabla S}{m}\mathrm{d}t + \chi\sqrt{\frac{\hbar}{m}}\mathrm{d}t + \frac{\hbar}{2m}\frac{\nabla\lvert\Psi\rvert^2}{\lvert\Psi\rvert^2}\mathrm{d}t$
DMC	$d\mathbf{r} = \chi\sqrt{\frac{\hbar}{m}}\mathrm{d}t + \frac{\hbar}{m}\frac{\nabla\lvert\Phi_T\rvert}{\lvert\Phi_T\rvert}\mathrm{d}t$
DMC2†	$d\mathbf{r} = \frac{\nabla S}{m}\mathrm{d}t + \chi\sqrt{\frac{\hbar}{m}}\mathrm{d}t + \frac{\hbar}{m}\frac{\nabla\lvert\Phi_T\rvert}{\lvert\Phi_T\rvert}\mathrm{d}t$

† If use complex Φ_T and retain imaginary part of complex drift vector $\nabla\Phi_T/\Phi_T$ (since for $\Psi = R\mathrm{e}^{iS/\hbar}$ have $\frac{\hbar}{m}\frac{\nabla\Psi}{\Psi} = \frac{\hbar}{m}\nabla\ln\Psi = \frac{\hbar}{m}\frac{\nabla R}{R} + i\frac{\nabla S}{m}$). In this view, $\nabla S/m$ is that part of osmotic velocity accounting for target distribution changing shape in real time.

So methods have practically identical Langevin-type equations describing particle motion as result of drift and diffusion, and similar propagators K (one in real, one in imaginary time).

Notes

- In DMC complex Ψ hardly used: real arithmetic faster and real Ψ easier to map into probabilities.
- Where complex Ψ used one employs *fixed-phase approximation* instead of fixed-node i.e. say phase S is fixed and equal to phase of trial function Φ_T. DMC algorithm used to solve for modulus of Ψ.
- Note no-one ever does DMC for *time-dependent* wave functions - always stationary states.

Ideas

Question 1: Why does nobody do molecular dynamics using the first-order deBB theory i.e. calculate the trajectories directly from $m\mathbf{v} = \nabla S = \hbar \mathrm{Im} \frac{\nabla \Psi}{\Psi}$, instead of using its first time derivative $m\ddot{\mathbf{R}} = -\nabla V - \nabla Q$ or its classical limit.

- No need to calculate forces (in principle higher-order derivative than velocities)
- No integration step to get velocity from forces (what about T?)
- No need for BO approximation - full quantum effects without having to do path integrals.

Obviously because you need a proper wave function to take the gradient of, and because you presumably need to solve the time-dependent Schrödinger equation..?

Question 2: What's to stop you doing $\mathbf{v} = \frac{\nabla S}{m}$ directly in QMC?

- Need nuclear wave function, but not impossible - see earlier.
- No need to compute difficult forces? The $\frac{\nabla \Psi}{\Psi}$ is essentially the drift vector whose real part is already computed in order for DMC to work.

Question 3: If you insist on using the second order-form why does no-one just compute $-\nabla Q$ instead of doing path-integrals in order to get quantum trajectories?

Thought required!

And other idea: why not propagate in complex time?

Repeat DMC imaginary time analysis with complex time $\tau = t + it'$*:*

Choose constant offset E_T in TDSE to be ground-state energy E_0 then, as $\tau \to \infty$, Ψ comes to look more like ground state ϕ_0 (as before). Difference is that exponentially-decaying bit now has t-dependent moving nodal surface. (Recall that a linear combination of stationary TDSE solutions with different energies, each with its own t-dependent phase factor, gives overall t-dependence in $|\Psi|^2$.)

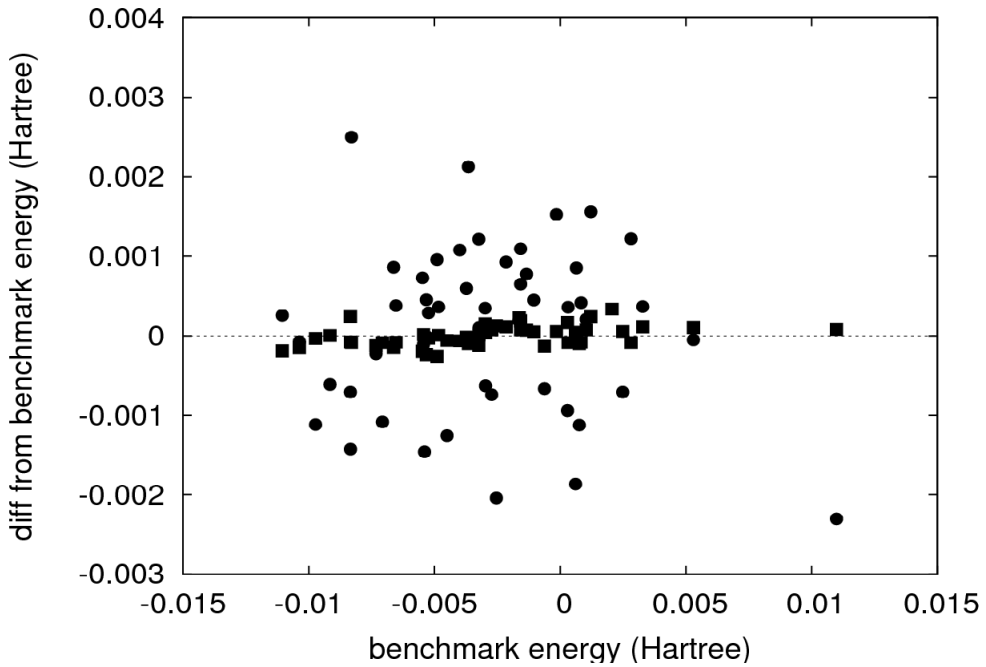
$$\Psi(\mathbf{x}, \tau) = c_0\phi_0 + \sum_{n=1}^{\infty} c_n\phi_n(\mathbf{x})\mathrm{e}^{i(E_n - E_0)t}\mathrm{e}^{-(E_n - E_0)t'}$$

- With t-dependent complex Ψ nodal surfaces dissolve into nodal lines where surfaces of real and imaginary functions intersect. Fewer barriers to motion of configurations?
- Simulations will show us how particles guided by the wave field with rapidly-moving nodes quickly became distributed according to $|\Psi|^2$. Also see that nodal lines moving through particle distribution acted as 'particle mixers'; trajectories become 'more chaotic' with more nodes.
- Might think that while imaginary time propagation improves Ψ, real time propagation allows nodal surface to relax and Ψ to be optimized more efficiently. As excited-state contributions die away and distribution approaches stationary state, $\frac{\nabla S}{m}$ (and hence guided particle velocity) tends to zero (only diffusion and real part of drift velocity remain for computing statistical data and expectation values).

This is very nebulous - proper thought required!

Give-up: do classical potential models

- Is QMC of any use in developing classical force fields?
- Can it help with Gabor's GAP potentials (which use arbitrary QM data as 'evidence' when generating interatomic potentials)? (Partay, ESDG May 2012)
- Can it help with hybrid classical/quantum molecular dynamics methods like Gabor's Learn On The Fly (LOTF)? (GC, ESDG August 2003, Feb 2010)



Comparison of DMC total energies (filled squares) with accurate quantum chemistry benchmarks at CCSD(T) level for a sample of 50 geometries of the H_2O trimer drawn from an MD simulation of liquid water. Horizontal axis shows CCSD(T) energy, vertical axis shows deviation of DMC energy from CCSD(T) energy. Filled circles show the same comparison for DFT(PBE).

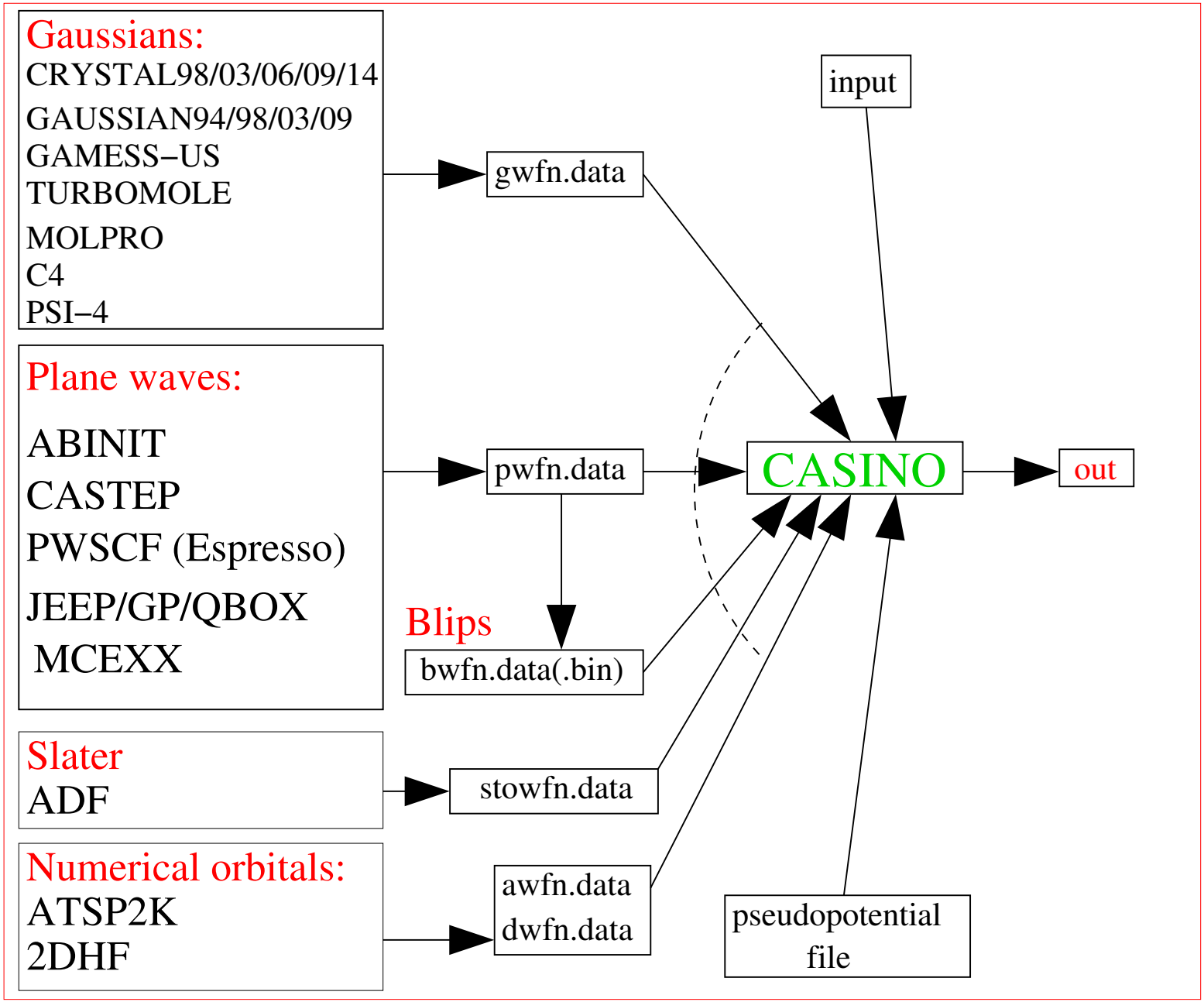
What expectation values can CASINO calculate?

- density
- spin density
- spin density matrix in non-collinear spin systems.
- pair correlation function (PCF)
- spherically-averaged pair correlation function (SPCF)
- localization tensor
- structure factor
- spherically-averaged structure factor
- one-particle density matrix (OBDM)
- two-particle density matrix (TBDM)
- condensate fraction estimator (unbiased TBDM, goes as TBDM-OBDM**2)
- dipole moment
- ionic populations (shortly to be generalized to charge, magnetic moment, and spin-resolved charge, and the spatial covariance of the charge, magnetic, and spin fluctuations)

CASINO input file (expectation value section)

```
# EXPECTATION VALUES
density             : F                #*! Accumulate density
spin_density        : F                #*! Accumulate spin densities
spin_density_mat    : F                #*! Accumulate spin density matrix
pair_corr           : F                #*! Accumulate recip. space PCF
pair_corr_sph       : F                #*! Accumulate spherical PCF
loc_tensor          : F                #*! Accumulate localization tensor
structure_factor    : F                #*! Accumulate structure factor
struc_factor_sph    : F                #*! Accumulate sph. struc. factor
onep_density_mat    : F                #*! Accumulate 1p density matrix
twop_density_mat    : F                #*! Accumulate 2p density matrix
cond_fraction       : F                #*! Accumulate cond fraction
dipole_moment       : F                #*! Accumulate elec. dipole moment
population          : F                #*! Calc ionic populations
expval_cutoff       : 30.d0 hartree    #*! G vector cutoff for expval
permit_den_symm     : F                #*! Symmetrize QMC charge data
qmc_density_mpc     : F                #*! Use QMC density in MPC int
int_sf              : F                #*! Calc ee int from strucfac
```

Interface to other packages



The `expval.data` file - header

```
START EXPVAL
Title
 Silicon 2x2x2
File version
 1
Number of particle types (e.g. 2=electrons, 4=electrons+holes)
 2
Number of each type of particle
 32 32
Dimensionality
 3
Periodicity
 3
Primitive translation vectors (au)
  0.000000000000000  5.129662155669400  5.129662155669400
  5.129662155669400  0.000000000000000  5.129662155669400
  5.129662155669400  5.129662155669400  0.000000000000000
Multiples of primitive translation vectors
 2 2 2
Volume of simulation cell
 2159.664411362648480
Wigner-Seitz radius of simulation cell
 7.254437790939672
Number of available G-vector sets
 1
```

The expval.data file - G vector sets

```
START GVECTOR SET 1
Energy cutoff (au) used to generate set
 30.000000000000000
Number of G-vectors in set
 725
Primitive reciprocal lattice vectors (au)
 -0.612436561756342  0.612436561756342  0.612436561756342
  0.612436561756342 -0.612436561756342  0.612436561756342
  0.612436561756342  0.612436561756342 -0.612436561756342
G-vector components Gx, Gy, Gz (au)
  0.000000000000000  0.000000000000000  0.000000000000000
  0.612436561756342 -0.612436561756342 -0.612436561756342
 -0.612436561756342  0.612436561756342  0.612436561756342
 -0.612436561756342 -0.612436561756342  0.612436561756342
  ....................................................
 -3.674619370538052 -3.674619370538052 -1.224873123512684
  3.674619370538052  3.674619370538052  1.224873123512684
END GVECTOR SET 1
```

The `expval.data` file - example data set

```
START DENSITY
Accumulation carried out using
 VMC
Use G-vector set
 1
Number of sets
 1
START SET 1
Particle type
 1
Total weight
   10.0000000000000
Complex charge-density coefficients (real part, imaginary part)
  8.00000000000000        0.
  1.67967897738036        0.
  ........................
  6.435364313328349E-004  0.
  6.435364313328349E-004  0.
END SET 1
END DENSITY
```

Continuing old runs

- Set **newrun** to F in the input file, and rename the `config.out` file - containing the final state of a previous VMC/DMC run - to `config.in`. Then CASINO will read in an old `expval.data` file before continuing the QMC calculation.
- This can be used either to continue the accumulation of an expectation value from the previous run, or to add a new expectation value to the `expval.data` file.

Accumulating things in reciprocal space

For example : the DENSITY

- After each single electron move from $\mathbf{r}' \longrightarrow \mathbf{r}$, accumulate density in reciprocal space as $\rho(\mathbf{G}) = \rho(\mathbf{G}) + \exp(i\mathbf{G} \cdot \mathbf{r})$ for each $\mathbf{G}$ vector.
- At the end of the simulation, divide by the total weight (e.g. the number of accumulation steps in VMC)to get the average of each $\rho(\mathbf{G})$.
- Fourier coefficients are normalized such that $\rho(\mathbf{G} = 0)$ is the number of electrons in the primitive cell. Get this by dividing the above average $\rho(\mathbf{G})$ by the total number of primitive cells in the simulation cell.
- Then the density in real space is given by

$$\rho(\mathbf{r}) = \sum_{i=1}^{n_{\mathbf{G}}} \rho(\mathbf{G}) \exp(i\mathbf{G} \cdot \mathbf{r})$$

Accumulating things in real space

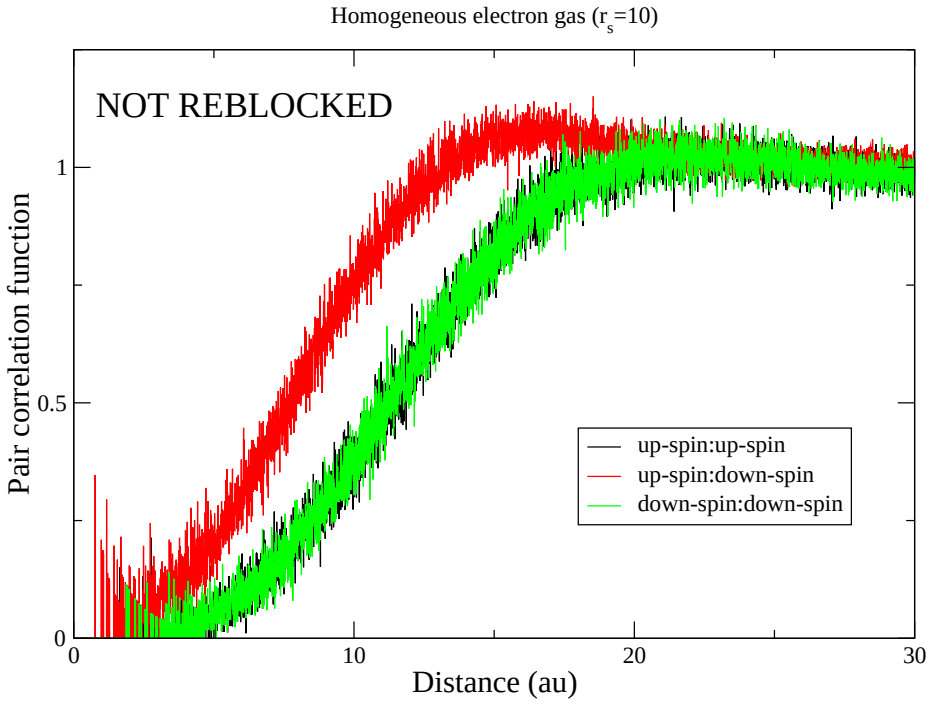
For example : pair correlation function (PCF) in 3D homogeneous, isotropic system

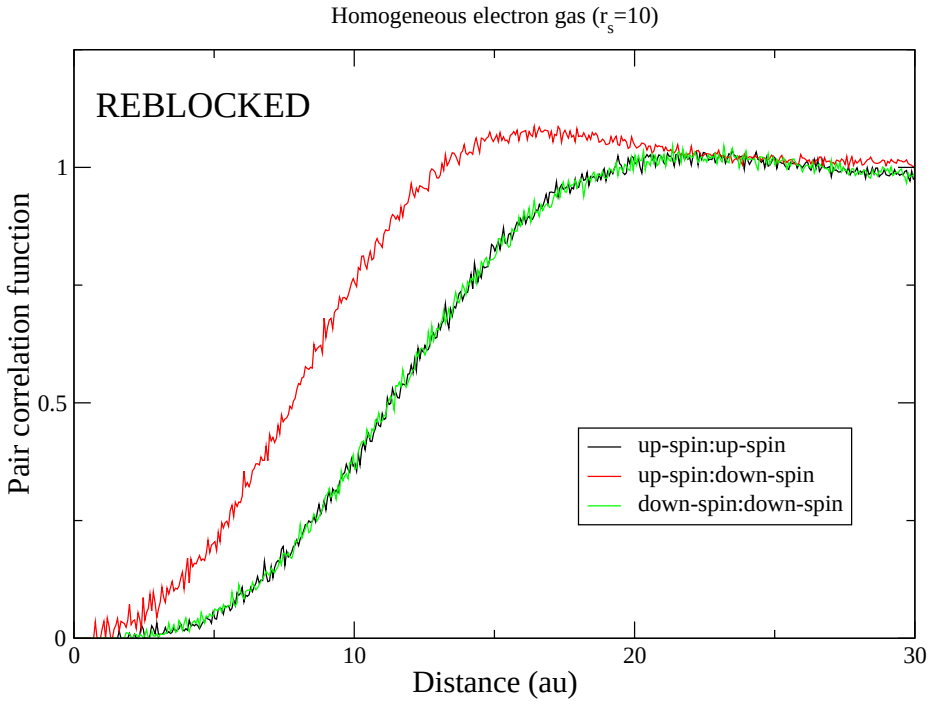
The PCF $g_{\alpha\beta}$ can be collected in bins of width Δ, giving

$$g_{\alpha\beta}(r_n) = \frac{\Omega}{\Omega_n} \left\langle \frac{N^n_{\alpha\beta}}{N_\alpha N_\beta} \right\rangle$$

where $N^n_{\alpha\beta}$ is the number of α, β-spin pairs whose separation falls within the nth bin and Ω_n is the volume of the nth bin. The volume of the nth bin is given by

$$\Omega_n = \frac{4}{3}\pi(n^3 - (n-1)^3)\Delta^3 = 4\pi(n^2 - n + \frac{1}{3})\Delta^3$$

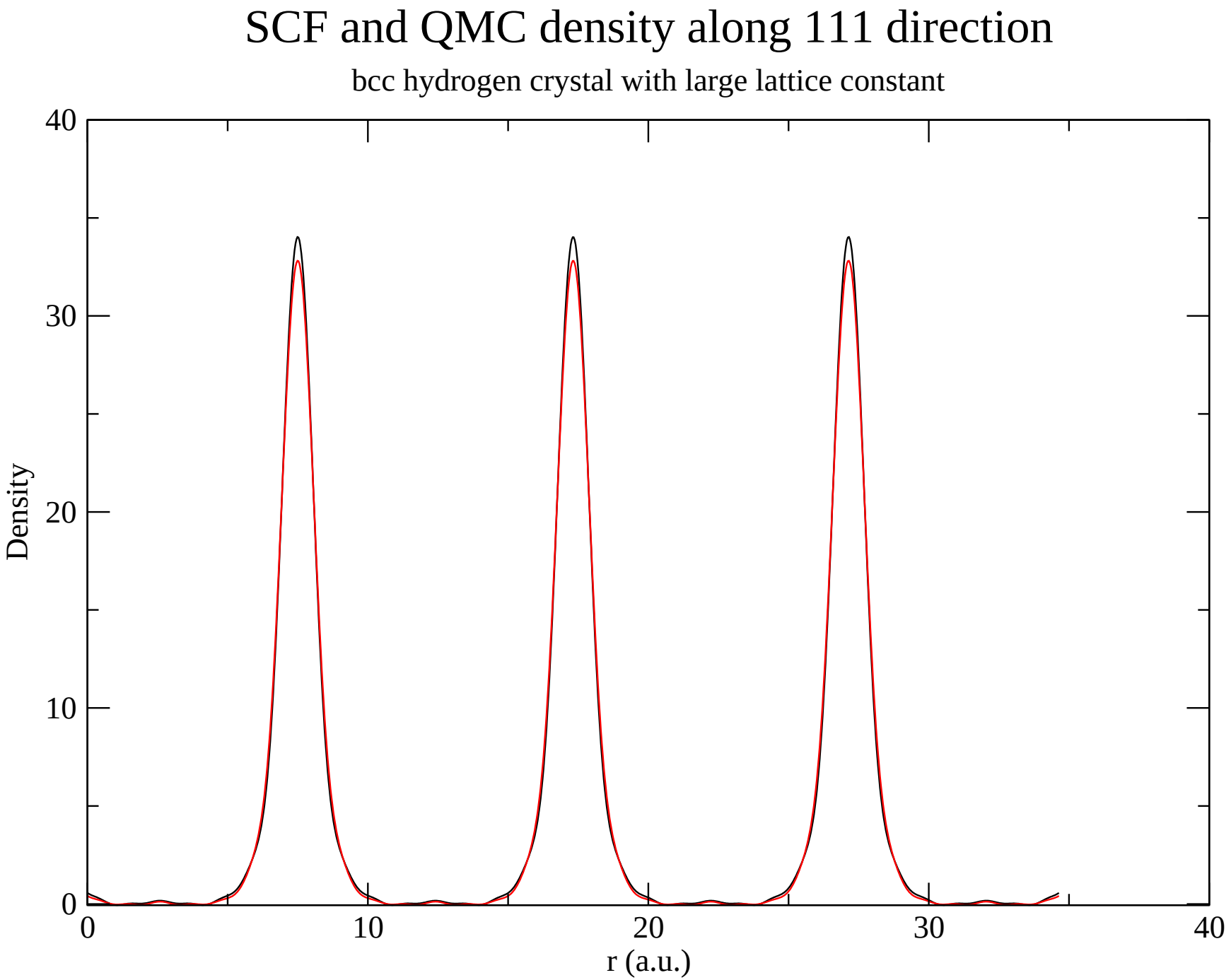




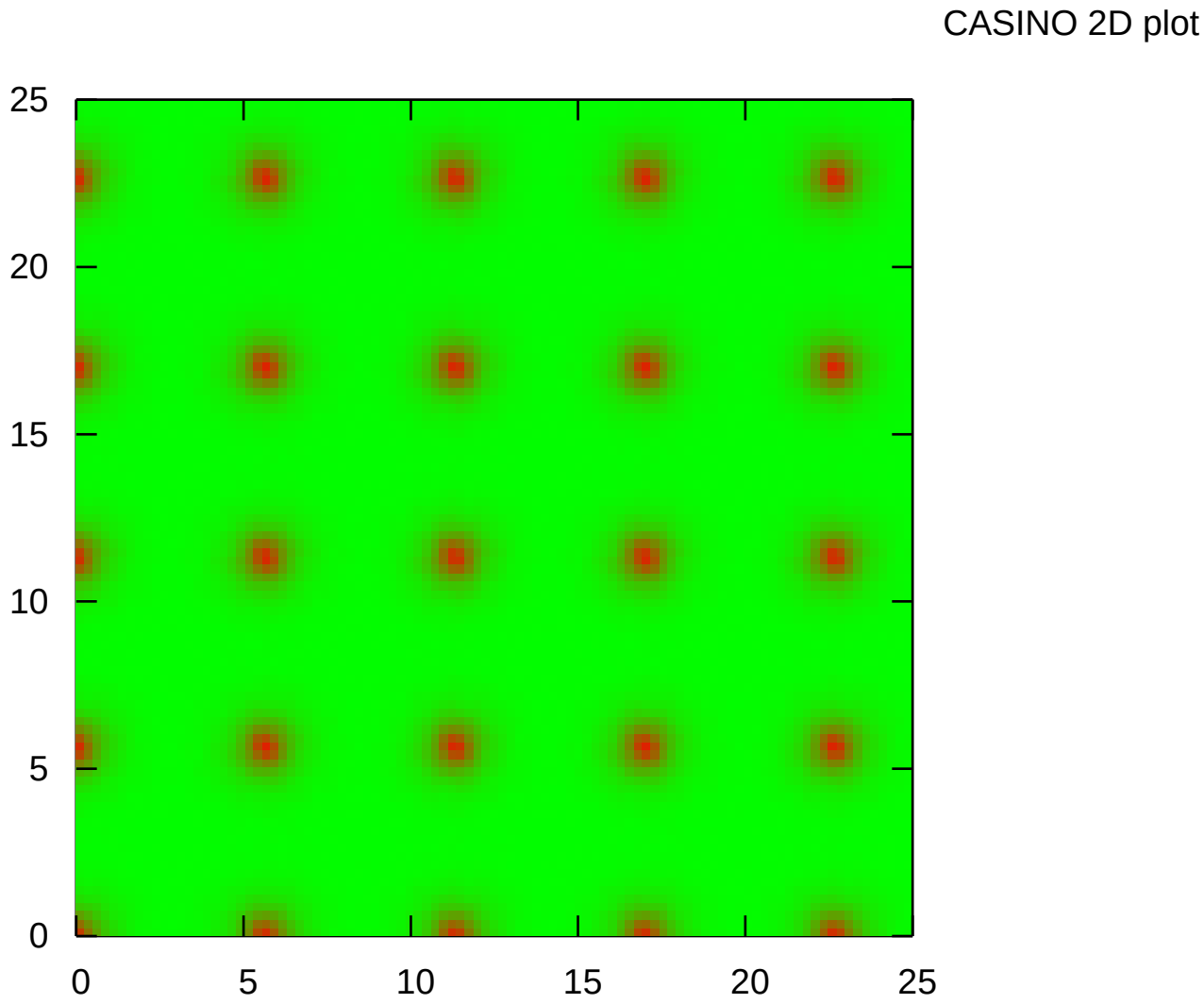
The `plot_expval` utility

- This utility reads the `expval.data` file produced by CASINO, performs appropriate conversions and transformations on the data therein, and writes the results to files visualizable by freely available standard packages such as XMGRACE, GNUPLOT etc.
- Output is placed in `lineplot.dat`, `2Dplot.dat`, or `3Dplot.dat` depending on the dimensionality of the plot.
- The utility takes its basic instructions from the **plot_expval** block in `input` which defines the geometrical region over which the data will be plotted (a line, plane or volume of appropriate size). Any additional options will be selected following direct appeal to the user at run time.

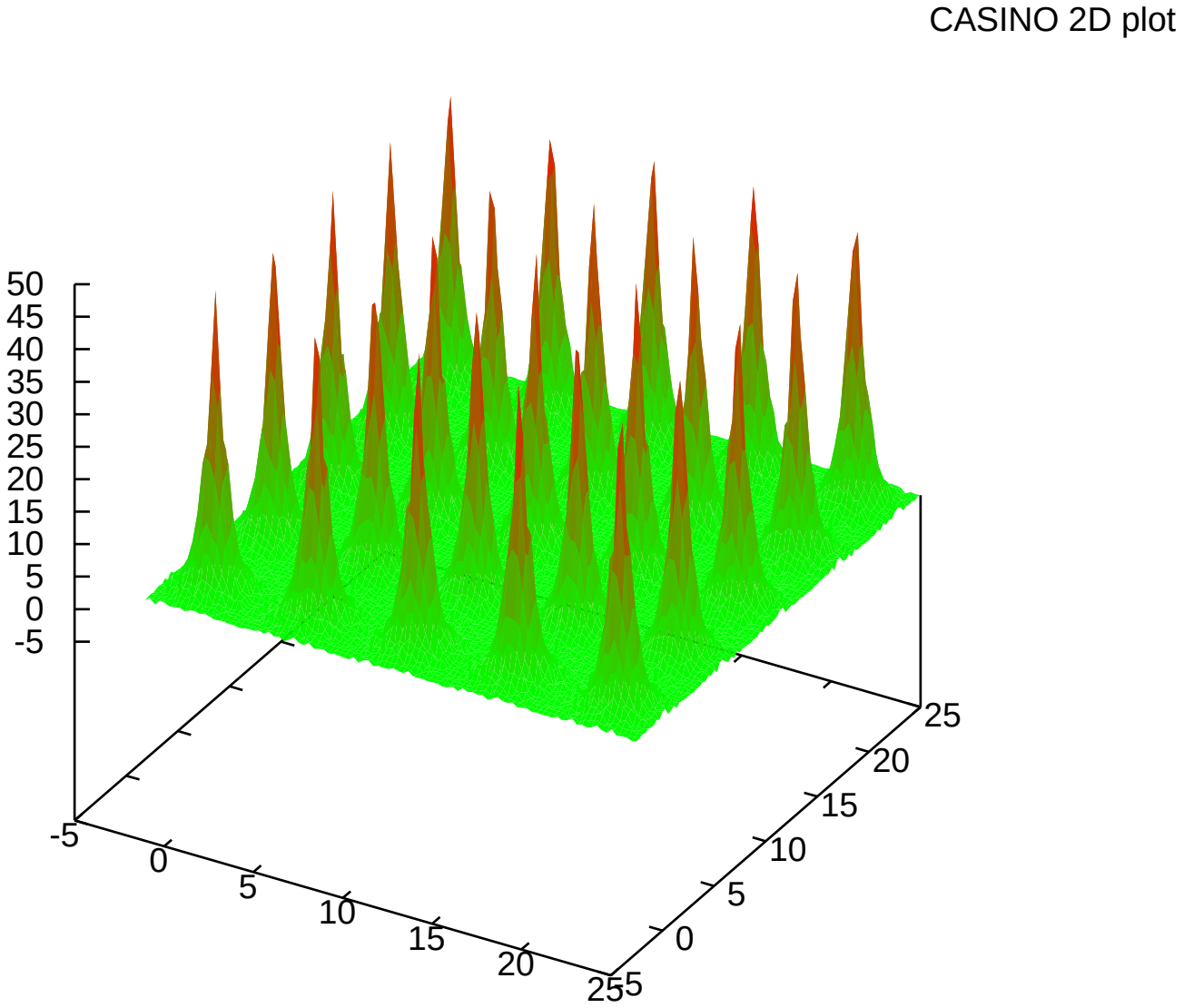
The XMGRACE package : 1D plots



The plot_2D utility



The plot_2D utility : -surf option



Mixed and pure estimators

Calculate energy from the mixed estimator:

$$E_{\mathrm{DMC}} = \frac{\int \Phi \hat{H} \Psi_T \, d\mathbf{R}}{\int \Phi \Psi_T \, d\mathbf{R}} = \frac{\int \Psi_T \Phi (\Psi_T^{-1} \hat{H} \Psi_T) \, d\mathbf{R}}{\int \Psi_T \Phi \, d\mathbf{R}}$$

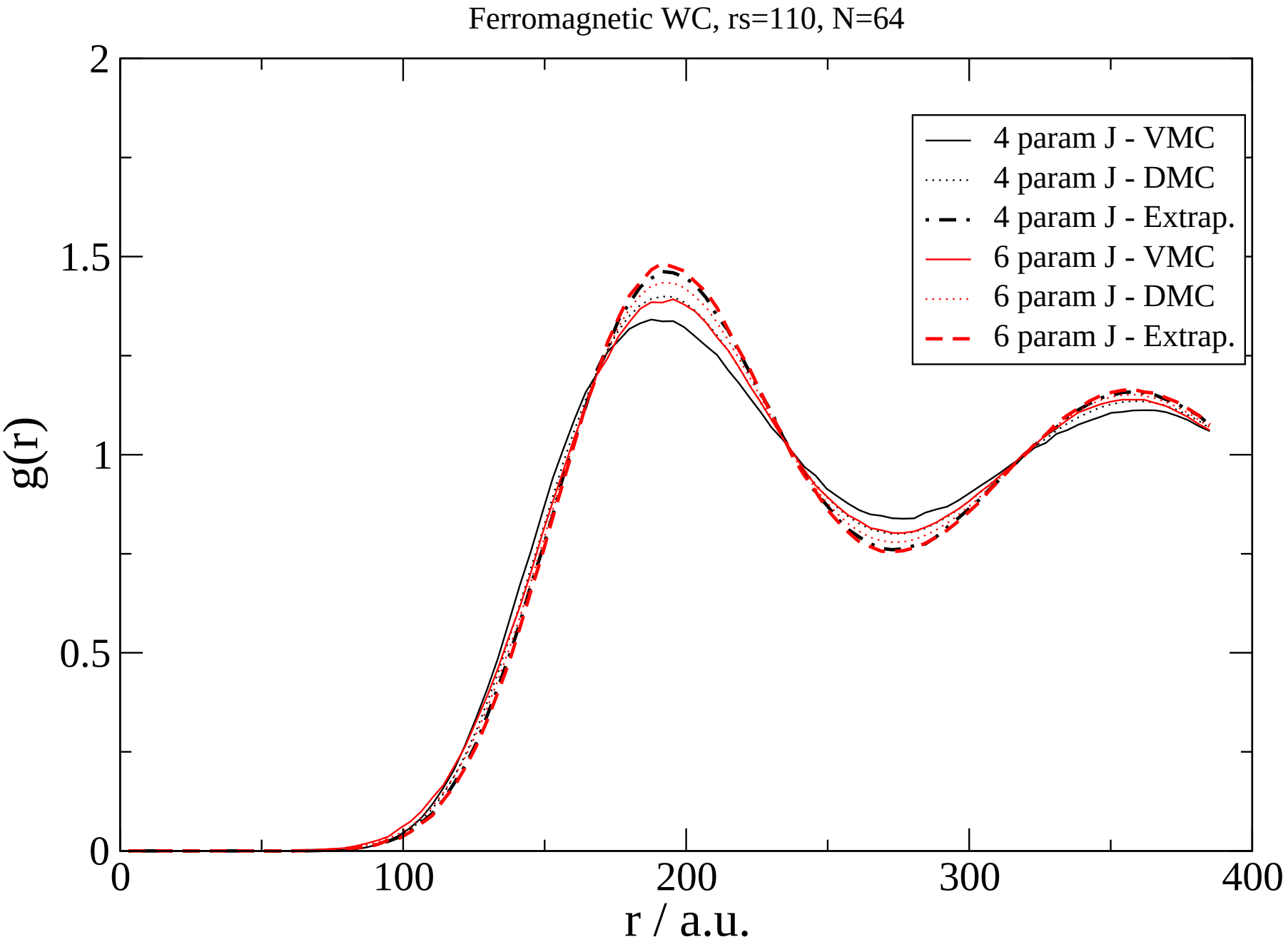
$$\simeq \frac{1}{M} \sum_i E_{\mathrm{L}}(\mathbf{R}_i)$$

- Crucial step in deriving this requires the equality of $\langle \Psi_T | \hat{H} | \Phi \rangle$ with $\langle \Phi | \hat{H} | \Psi_T \rangle$ together with the eigenvalue equation $\hat{H}\Phi = E\Phi$.
- For an operator that does not obey an eigenvalue equation with respect to Φ, i.e. one that does no commute with the Hamiltonian, the mixed estimator does not yield the correct expectation value.
- To use DMC to estimate the expectation value of such an operator, we need to sample from the "pure" distribution :

$$\langle A \rangle = \frac{\int \Phi \hat{A} \Phi \, d\mathbf{R}}{\int \Phi \Phi \, d\mathbf{R}}$$

Can obtain with e.g. extrapolated estimator, future walking, reptation MC.

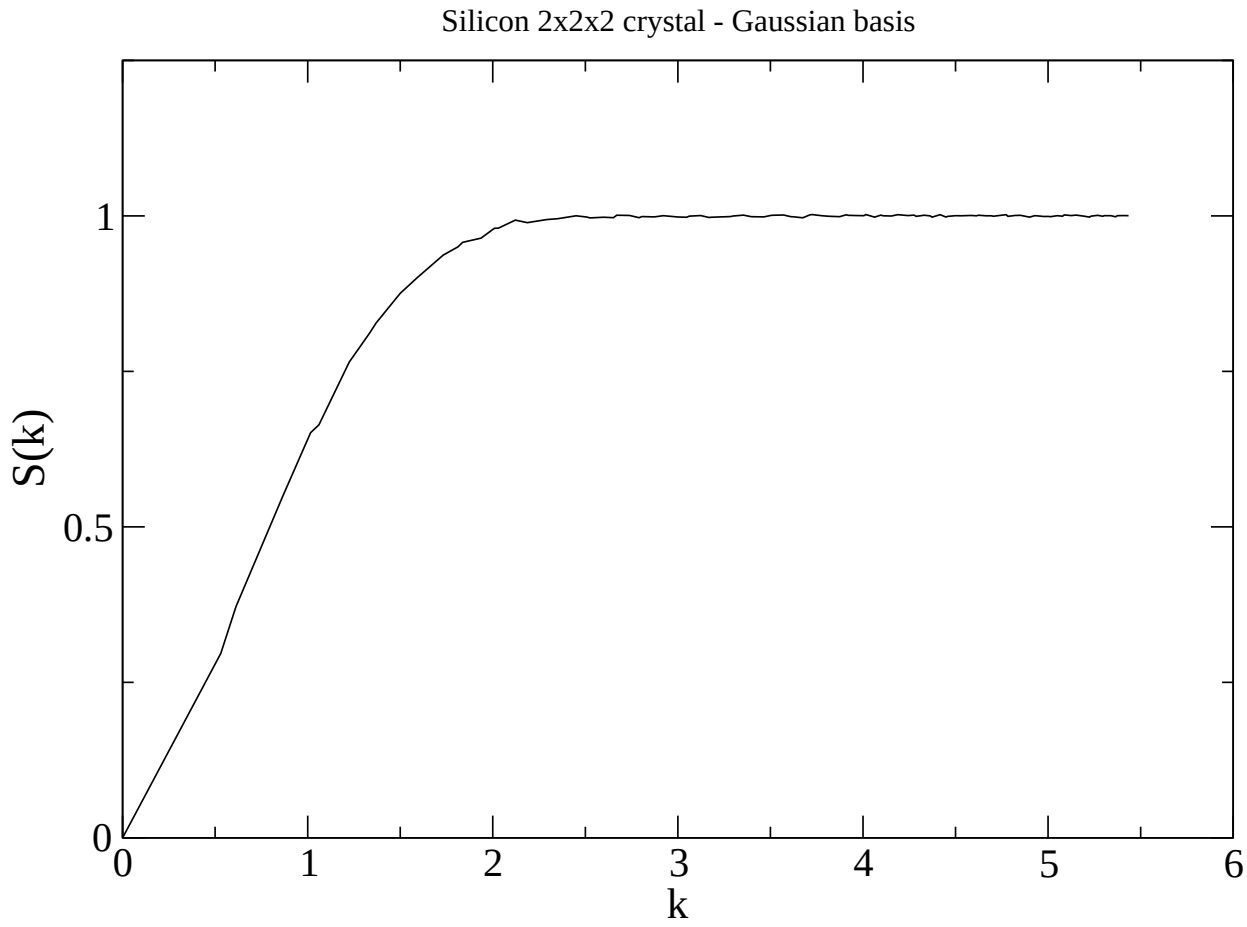
Extrapolation of pair correlation function



Structure factor

$$H\Psi = E\Psi$$

Potentially very important in new approaches to calculating electron-electron interactions with smaller finite-size effects in periodic systems.



Exchange-correlation hole and the adiabatic connection

Exact relationship between the exchange-correlation energy, E_{xc}, and the ground-state many-electron wave function Ψ_λ associated with different values of the Coulomb coupling constant, λ.

$$\bar{g}_{\alpha\beta}(\mathbf{r},\mathbf{r}') = \frac{N(N-1)}{n_\alpha(\mathbf{r})n_\beta(\mathbf{r}')}\int_0^1 d\lambda \int d\mathbf{x}_3\cdots d\mathbf{x}_N |\Psi_\lambda(\mathbf{r}\alpha,\mathbf{r}'\beta,\mathbf{x}_3,\ldots,\mathbf{x}_N|^2$$

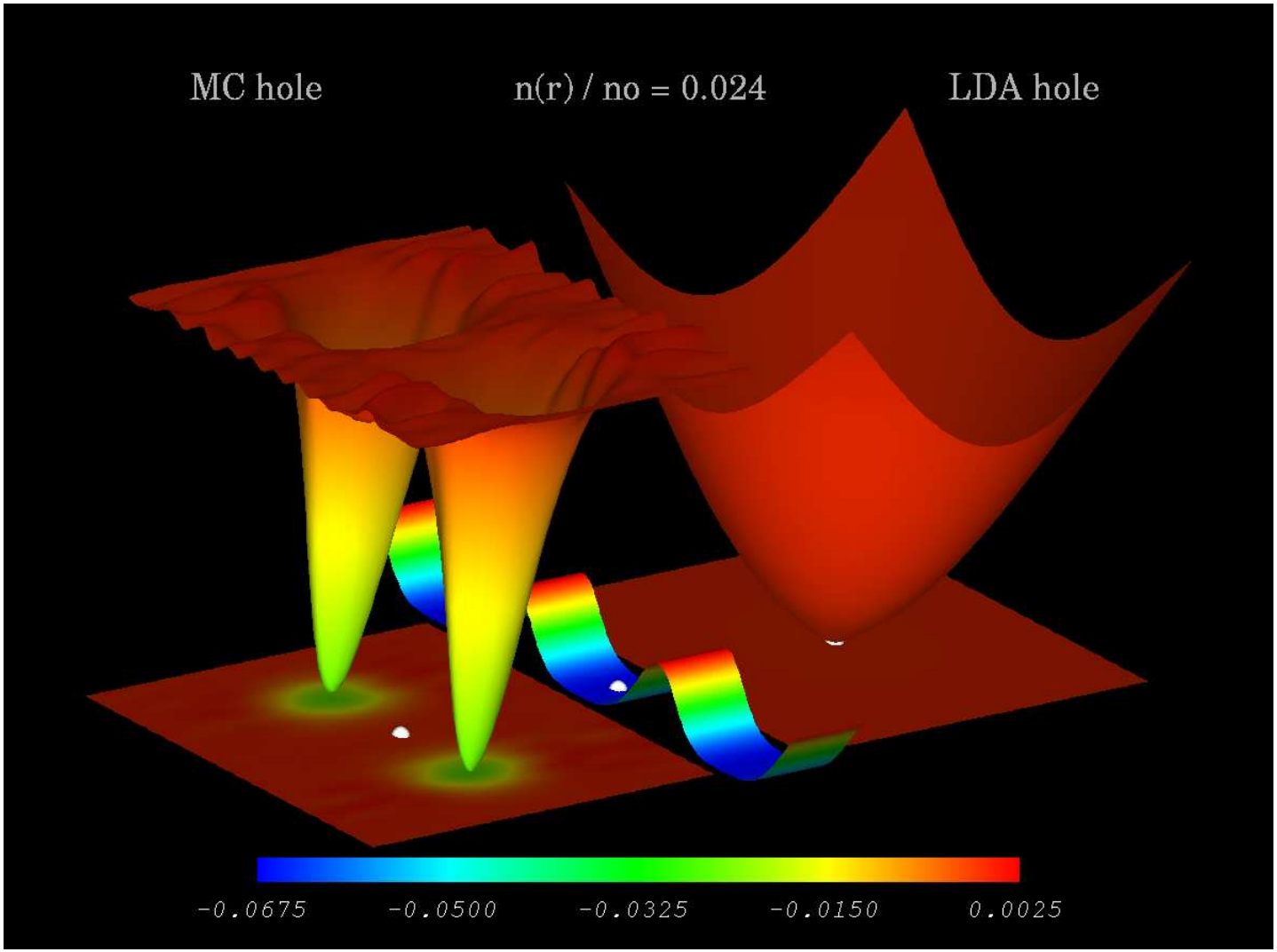
$$\bar{g}(\mathbf{r},\mathbf{r}') = \sum_{\alpha,\beta}\frac{n_\alpha(\mathbf{r})n_\beta(\mathbf{r}')}{n(\mathbf{r})n(\mathbf{r}')}\bar{g}_{\alpha\beta}(\mathbf{r},\mathbf{r}')$$

$$\rho_{xc}(\mathbf{r},\mathbf{r}') = n(\mathbf{r}')[\bar{g}_{\alpha\beta}(\mathbf{r},\mathbf{r}')-1]$$

$$e_{xc}(\mathbf{r}) = \frac{n(\mathbf{r})}{2}\int d\mathbf{r}'\frac{\rho_{xc}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

Integrate $e_{xc}(\mathbf{r})$ over all space to get exchange-correlation energy E_{xc}!

Exchange-correlation hole and the adiabatic connection



Some more interesting calculations

- Surface energy of materials
- Molecular adsorption on surfaces
- Water systems
- DMC-based statistical mechanics for surface problems.

Surface energy of materials

- When macroscopic sample of a material separated into two pieces, new surfaces are created. The work done in creating the new surfaces divided by their area is the surface formation energy, usually denoted by σ.
- For a crystal, σ depends on orientation of surface. Surface formation energies are important in fields as far apart as geology and fracture mechanics. Particularly important in nanoscience, because in particles a few nm across a significant fraction of the atoms are at or near the surface. One consequence is that the crystal structures of nano-particles are sometimes not those of the bulk material, because it may be advantageous to adopt a less stable bulk structure if the energies of the surfaces are lowered. For nano-particles supported on substrates, the equilibrium form of the nano-particles is often determined by the balance between surface and interfacial energies.
- DFT has quite serious problems in predicting σ. Known that predicted σ values can depend strongly on the exchange-correlation functional. For example, as a broad generalisation, it seems that GGA approximations tend to give σ values that are about 30 % less than LDA values. Even more surprising is that LDA values sometimes seem to agree better with experiment than their GGA counterparts. This is unexpected, because common sense suggests that the formation of a surface involves the breaking of bonds, and GGA is usually much better than LDA for bond formation energies.
- Problem is that σ values are not easy to measure experimentally, and those that have been measured are sometimes subject to large errors. In this unsatisfactory situation, there is an obvious need for accurate computed benchmarks for σ, which can be used to assess both DFT predictions and experimental values.

Surface energy of materials: example

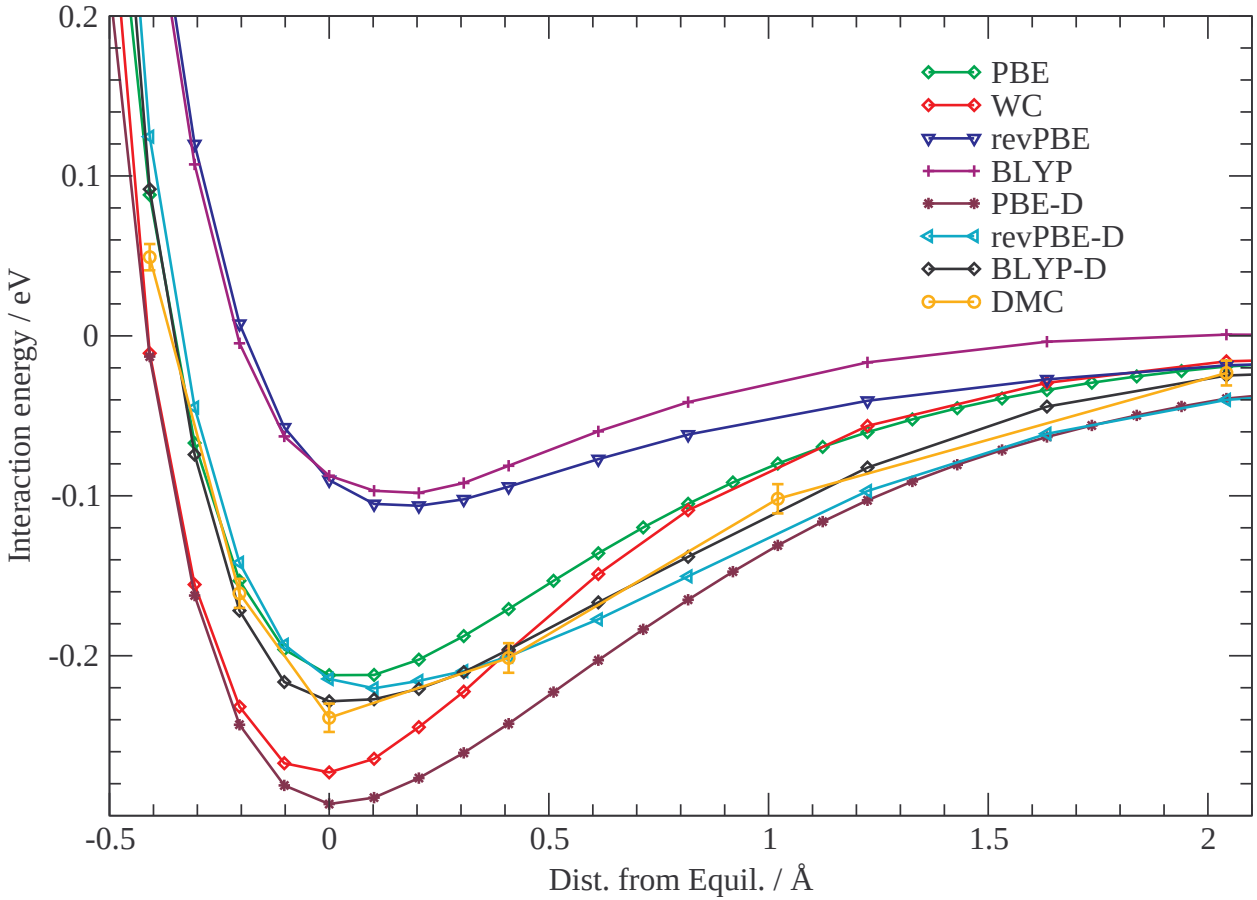
- Can now compare with accurate quantum chemistry calculations, using 'hierarchical method' that in some cases allows cohesive energy, equilibrium structure and elastic properties of perfect crystals and surface formation energies of crystals to be calculated using wave function-based quantum chemistry at the CCSD(T) level [S.J. Binnie et al., Phys. Rev. B **82**, 165431 (2010).]
- Calculations done for LiH, LiF and MgO. Disagree with DFT - who's right? Take LiH as example. Outstanding agreement with experiment given by all-electron DMC for LiH crystal. Equilibrium lattice parameter a_0 agrees to within 10^{-3} Å (both experiment and DMC give 4.061 Å at $T = 0$ K, with zero-point corrections). Cohesive energy agrees with experiment to within 20 meV per formula unit (comparable with experimental uncertainty). Hierarchical QC approach gives similar accuracy. No adjustable parameters in either method - gives us every reason to expect excellent values of surface formation energy.

Method	σ / J m^{-2}
DMC pseudo	0.373(3)
DMC all-elec	0.44(1)
Quantum chem (frozen core)	0.402
Quantum chem (with core)	0.434
DFT(LDA)	0.466
DFT(PBE)	0.337
DFT(rPBE)	0.272

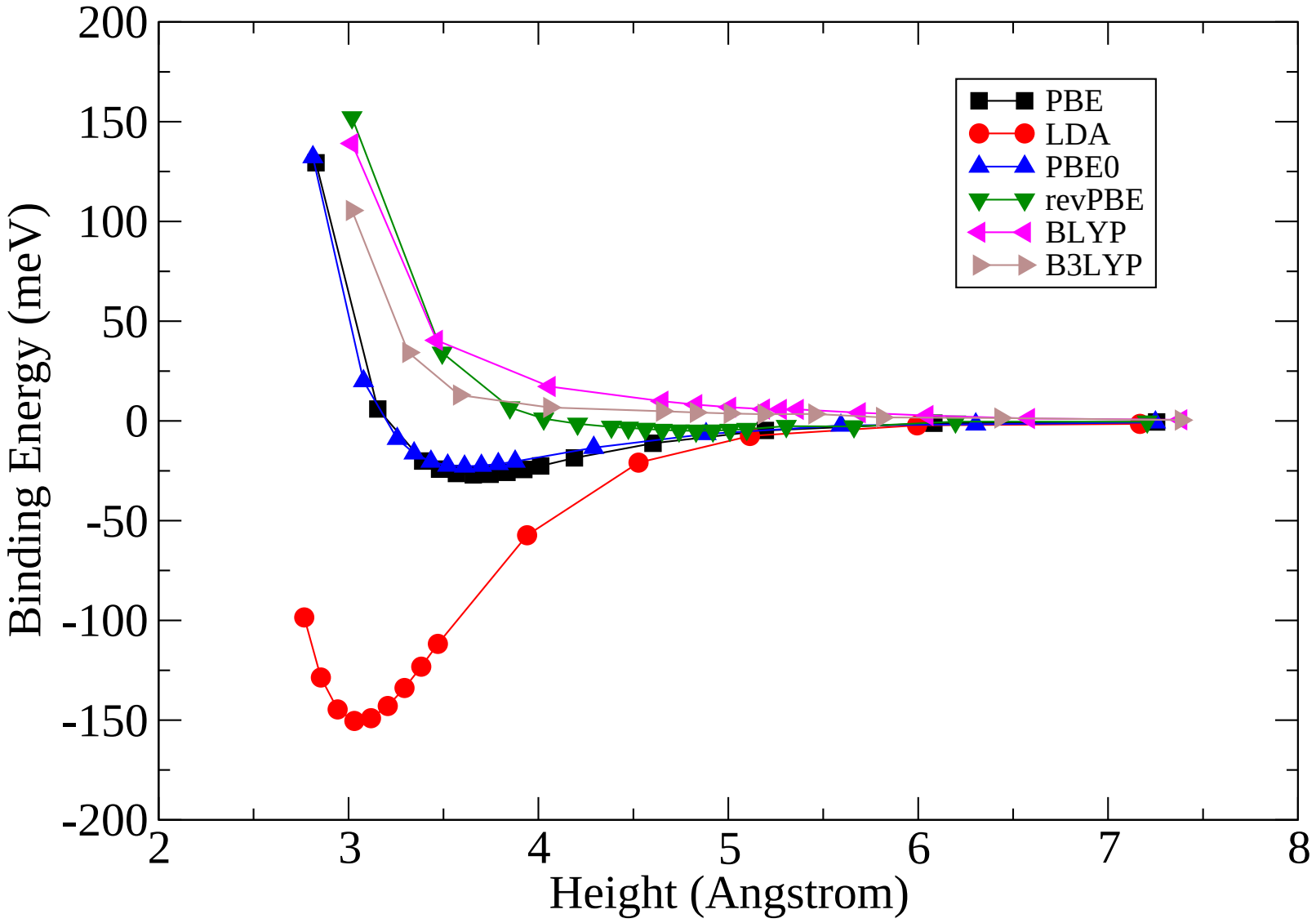
- DMC slab calculations on LiH went up to slab thicknesses of 6 ionic layers, with 18 ions per layer in the repeating cell (total number of ions per repeating cell thus went up to 108).
- Agreement between all-electron DMC and high-level quantum chemistry extremely close, as it should be, and the σ values give robust benchmarks against which to judge DFT. In LiH, LiF, and MgO, LDA performs rather well, the PBE value is about 30 % lower and is considerably less good, while the revised PBE value (rPBE) is ~ 40 % below correct value.

Molecular adsorption on surfaces

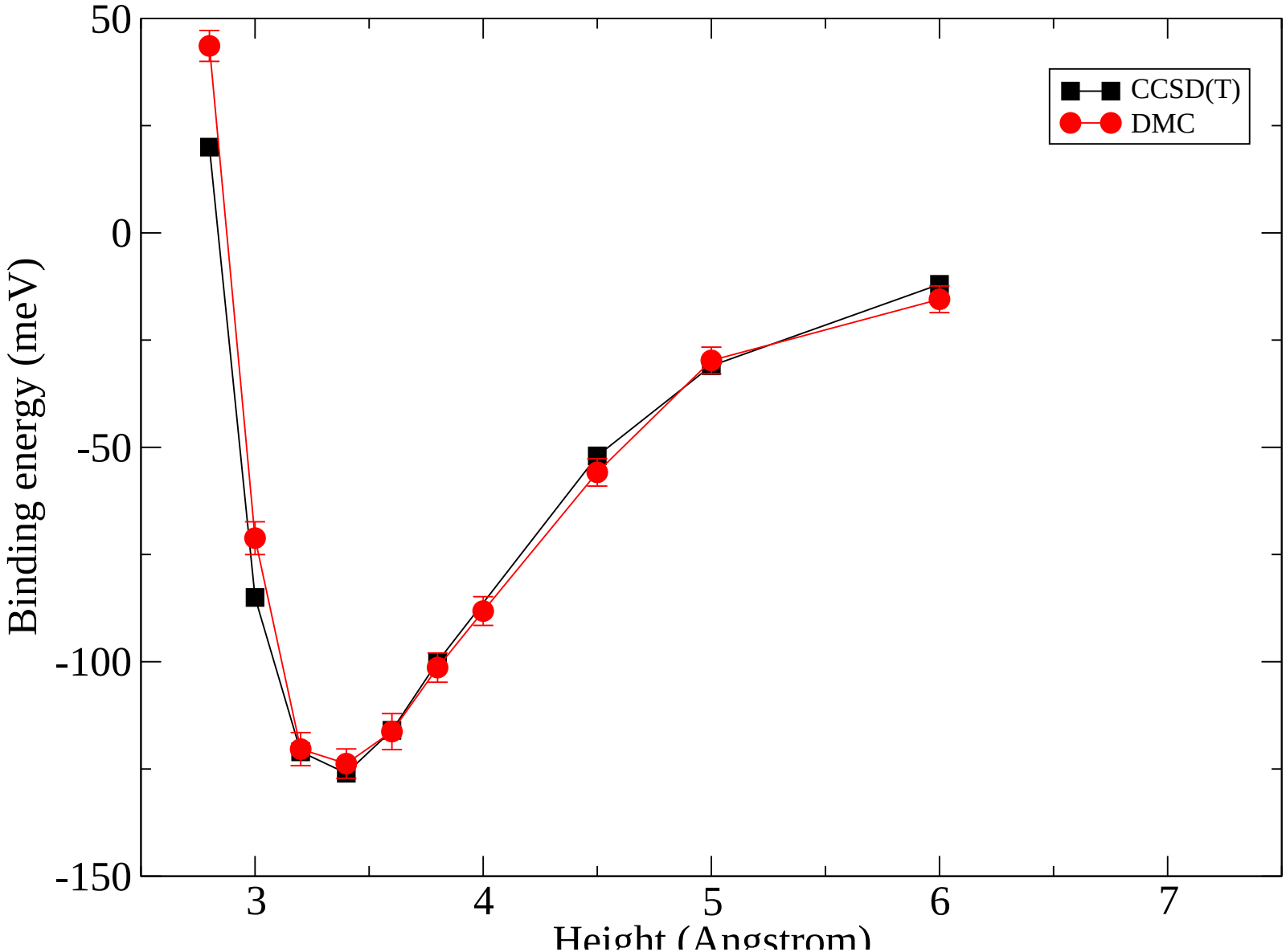
- Adsorption energies of molecules on surfaces important for many reasons (gas sensing, catalysis, corrosion, gas purification, chemical reactions in interstellar space, atmospheric processss).
- Our quantitative knowledge of adsorption energies remains poor, largely because DFT has rather poor predictive power for the energetics of adsorption (especially when van der Waals etc. important).



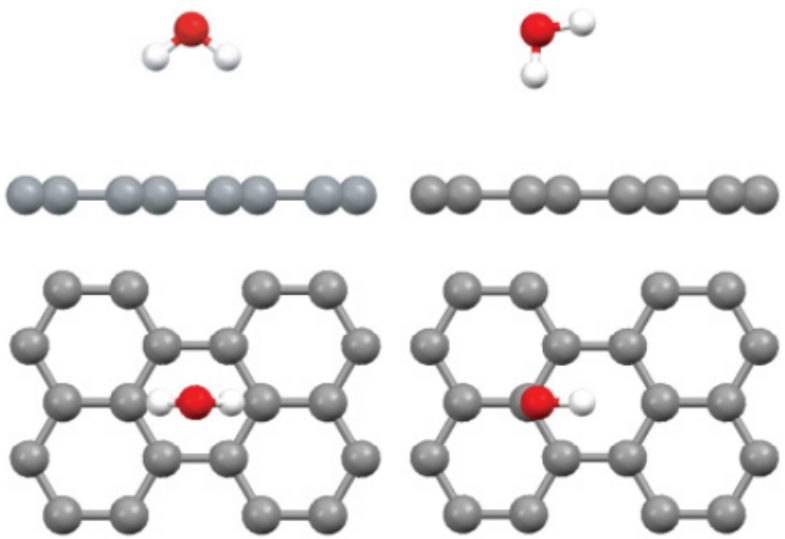
Water on graphene: DFT is completely useless



Water on benzene: DMC compared with CCSD(T)



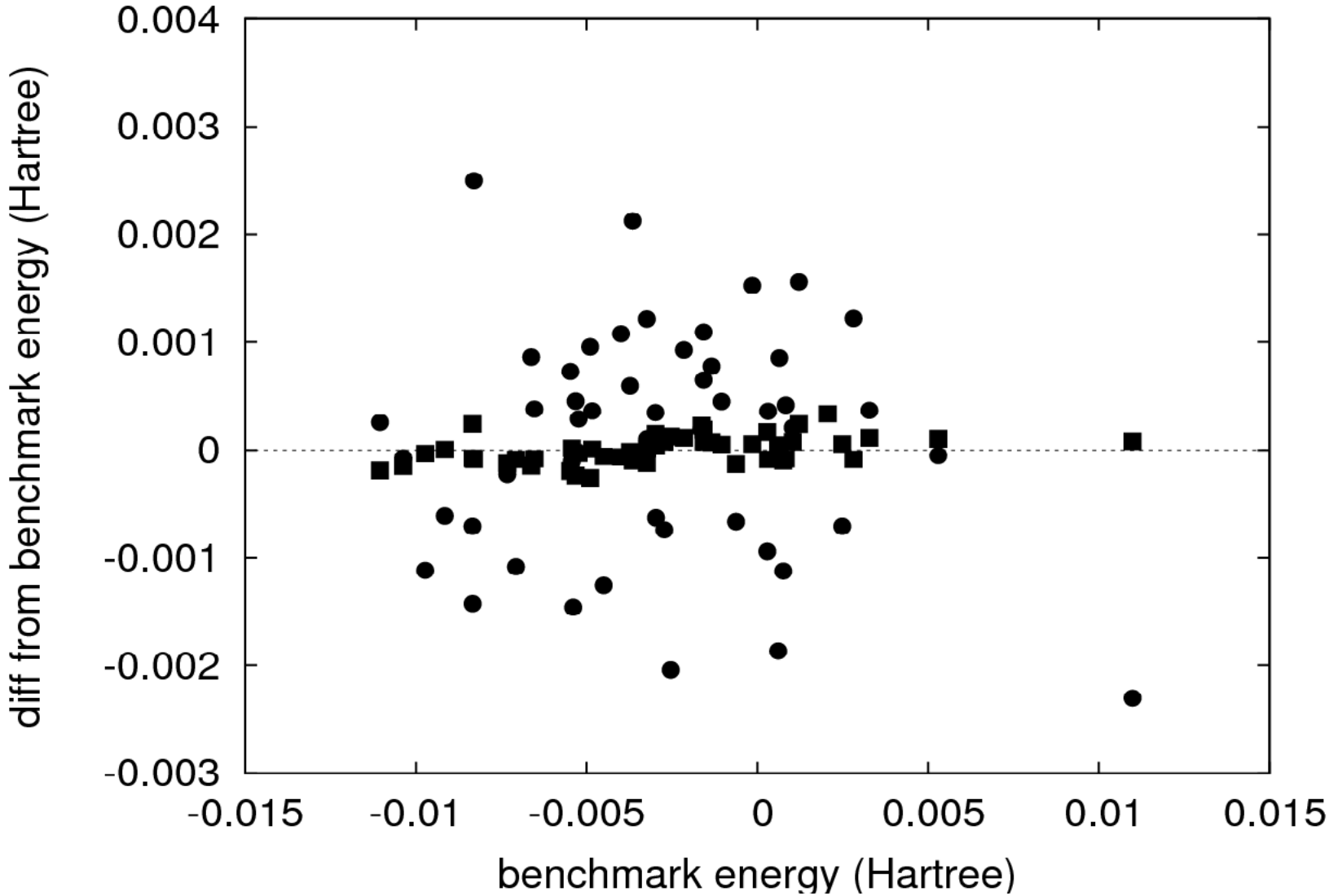
Water on graphene: DMC vs. DFT



	Two leg		One leg	
Approach	E_{ads} (meV)	Height (Å)	E_{ads} (meV)	Height (Å)
DMC	-70 ± 10	3.4-4.0	-70 ± 10	3.4-4.0
RPA	-98 (-77)	3.42	-82 (-81)	3.55
LDA	-151	3.04	-139	3.15
PBE	-27	3.65	-31	3.65
PBE0	-23	3.62	-27	3.66
revPBE	-4	4.66	-7	4.42
PBE-D	-90	3.35	-87	3.45
BLYP-D	-90	3.35	-87	3.47

J. Ma, A. Michaelides, D. Alf, L. Schimka, G. Kresse, E. Wang, *Adsorption and diffusion of water on graphene from first principles*, Physical Review B **84**, 033402 1-4 (2011).

Water systems



Comparison of DMC total energies (filled squares) with accurate quantum chemistry benchmarks at CCSD(T) level for a sample of 50 geometries of the H_2O trimer drawn from an MD simulation of liquid water. Horizontal axis shows CCSD(T) energy, vertical axis shows deviation of DMC energy from CCSD(T) energy. Filled circles show the same comparison for DFT(PBE).

DMC-based statistical mechanics for surface problems

Combination of DMC with statistical mechanics is feasible, because Alfè *et al.* have recently shown it in action for liquid and solid iron at high temperatures. [E. Sola and D. Alfè, Phys. Rev. Lett. **103**, 078501 (2009).]

Alfè group now addressing new challenge of doing *thermally disordered* interacting molecules on surfaces, including interactions between adsorbed molecules and thermal disorder/entropic effects. Need to combine DMC with other techniques.

Calculate e.g. :

- adsorption isotherms (coverage θ as a function of gas pressure)
- thermal desorption rates as a function of θ and temperature T.

The outcome of all this will be the first ever computations of thermodynamic functions of a dissociating adsorbate with full inclusion of entropic effects - at the chemical accuracy given by DMC.

General advice

Given developments underway, we believe it is now very timely for more research groups to consider becoming involved in the QMC enterprise. But what is the right way to do this? Some suggestions.

- Start small and work upwards: Clearly, one should gain experience first with small problems (for example, problems involving small molecules), which can easily be run on modest local resources. Work up from there to more ambitious problems that need large machines.
- Use standard codes: A large development effort has gone into CASINO and other QMC codes, and it makes sense to work with a code that already has a substantial publication record.
- Collaborate: Even more than DFT, there is much that you need to know about QMC before trying calculations. No QMC code can (yet) be treated as a black box, and it is wise to learn from experienced practitioners. Attending this school is thus a good idea.
- Choose your problem well: Just like DFT, QMC cannot solve all the world's problems, and it is important to play to the strengths of the techniques and to be aware of their weaknesses (fixed-node error/pseudopotential non-locality potential weaknesses of DMC in its present form).

Our hope for the future is that more researchers will discover for themselves the possibilities offered by QMC on machines ranging all the way from local clusters to the national and international petascale services now becoming available.

The end of Mike's contributions to the summer school