Continuum Quantum Monte Carlo Methods Introduction, Recent Developments, Successes, Failures, and Prospects

W.M.C. Foulkes

Department of Physics Imperial College London

From Basic Concepts to Real Materials, KITP, 2009



Outline



- 2 Example Applications
- The Good, The Bad and The Ugly
- 4 Recent Advances and Prospects

Outline



- 2 Example Applications
- 3 The Good, The Bad and The Ugly
- 4 Recent Advances and Prospects

Bias

- I am an electronic structure/materials theorist.
 - Want real numbers for real materials.
 - More interested in solids than molecules.
 - More interested in ground states than excitations.
 - DFT is wonderful but not enough.
 - Lattice models are not enough.

Bias

• I am an electronic structure/materials theorist.

- Want real numbers for real materials.
- More interested in solids than molecules.
- More interested in ground states than excitations.
- DFT is wonderful but not enough.
- Lattice models are not enough.
- Methods that interest me ...
 - Get the electronic structure right first.
 - Describe correlation better than DFT.

The Earth's Core

- What is it?
- How hot is it?

About 5500K at inner/outer core boundary

How viscous it is?

Estimates span 12 orders of magnitude



Cytochrome P450 Family of Enzymes

 $RH + O_2 + 2H^+ + 2e^- \longrightarrow ROH + H_2O$





Continuum QMC

Continuum QMC methods are

- Primarily for ground states/total energies.
- Are not perturbative (variational instead).
- Are systematically improvable.
- Useful in large systems.

Continuum QMC

Continuum QMC methods are

- Primarily for ground states/total energies.
- Are not perturbative (variational instead).
- Are systematically improvable.
- Useful in large systems.

Tackle the many-electron problem head-on!

[The many-electron problem is (NP-)hard!]



The Many-Electron Schrödinger Equation

Textbook Notation

$$\left(\sum_{i=1}^{N} \left\{-\frac{\hbar^2}{2m} \nabla_i^2 + v_{\text{ext}}(\mathbf{r}_i)\right\} + \sum_{i>j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}\right) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

The Many-Electron Schrödinger Equation

Textbook Notation

$$\left(\sum_{i=1}^{N} \left\{-\frac{\hbar^2}{2m} \nabla_i^2 + v_{\text{ext}}(\mathbf{r}_i)\right\} + \sum_{i>j} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}\right) \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

Simpler Notation

$$\left\{-\frac{1}{2}
abla_{\mathbf{R}}^{2}+V(\mathbf{R})
ight\}\Psi(\mathbf{R}) = E\Psi(\mathbf{R})$$

 Atomic units: $e = \hbar = m = 4\pi\epsilon_0 = 1$ Unit of length = $a_0 = 0.529177$ Å Unit of energy = Hartree = 27.211 eV

Energy Scales

Total electronic Chemical bond Chemical reaction Room temperature High T_c superconductivity $\begin{array}{l} > 10^2 \, eV \qquad (> 10^3 \, eV) \\ few \, eV \\ < 10^{-1} \, eV \\ 2.5 \times 10^{-2} \, eV \\ < 10^{-2} \, eV \end{array}$

C atom

C pseudo-atom

$$\begin{array}{rcl} {\sf E}_{total} & = & -150 \mbox{ eV} \\ {\sf E}_{HF} & = & 98.2\% \mbox{ E}_{total} & = & -147.3 \mbox{ eV} \\ {\sf E}_{corr} & = & 1.8\% \mbox{ E}_{total} & = & -2.7 \mbox{ eV} \end{array}$$

Continuum QMC

The

- variational
- diffusion
- reptation, path-integral, phaseless auxiliary-field, ...

QMC methods take the HF approximation as a starting point and add a reasonable description of correlation.

- Not accurate enough to describe subtle correlation physics (< 10⁻² eV).
- $\bullet\,$ Can attain chemical accuracy (1 kcal mole^{-1} \sim 0.04 eV) in molecules.
- Can answer most questions about chemical bonding.
- About 10 times (?) better than DFT (but \gg 10 times the effort).

- N^3 (or better) scaling with system size.
- Can study systems of 1000+ electrons.



576 valence electrons (without vacancy)



The VMC Algorithm

• Guess $\Psi_T(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{10^{23}}) = \Psi_T(\mathbf{R}).$

Evaluate

$$E_T = \int \Psi_{\mathrm{T}}^*(\mathbf{R}) \hat{H} \Psi_{\mathrm{T}}(\mathbf{R}) d\mathbf{R}$$
$$= \int \left(\frac{\hat{H} \Psi_{\mathrm{T}}(\mathbf{R})}{\Psi_{\mathrm{T}}(\mathbf{R})} \right) |\Psi_{\mathrm{T}}(\mathbf{R})|^2 d\mathbf{R}$$

using Monte Carlo integration.

• Probability density $= |\Psi(\mathbf{R})|^2$; score $= E_L(\mathbf{R}) = \frac{\hat{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}$.



Slater-Jastrow Trial Functions

Strategy: first get mean-field right, then add a reasonable description of correlation:

$$\Psi(\mathbf{R}) = \exp\left[-\sum_{i>j} u(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \chi(\mathbf{r}_i)\right] \sum_{\alpha} c_{\alpha} D_{\alpha}(\mathbf{R}) ,$$

where $\alpha = (i_1, i_2, \dots, i_N)$ and

$$D_{\alpha}(\mathbf{R}) = \begin{vmatrix} \psi_{i_1}(\mathbf{r}_1) & \psi_{i_1}(\mathbf{r}_2) & \dots & \psi_{i_1}(\mathbf{r}_N) \\ \psi_{i_2}(\mathbf{r}_1) & \psi_{i_2}(\mathbf{r}_2) & \dots & \psi_{i_2}(\mathbf{r}_N) \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \psi_{i_N}(\mathbf{r}_1) & \psi_{i_N}(\mathbf{r}_2) & \dots & \dots & \psi_{i_N}(\mathbf{r}_N) \end{vmatrix}$$

Slater-Jastrow Trial Functions

- Orbitals $\psi_i(\mathbf{r})$ from DFT or Hartree-Fock.
- *u*(**r**_i, **r**_j) increases when **r**_i → **r**_j. Helps keep electrons apart (correlation). χ(**r**_i) counteracts smearing effect of *u*.
- u and χ given flexible variational representations.
- Expansion coefficients c_{α} treated as variational parameters.
- Orbitals in determinants may be optimized too.

HF Density



Ge : HF density



W.M.C. Foulkes (Imperial College London)

Continuum QM0

EXCITCM

HFJCHI Density

Ge: HFJ + CHI density





W.M.C. Foulkes (Imperial College London)



Effect of Jastrow Factor in Silane Molecule





The Cusp Conditions

As any two electrons approach one another, both terms in

$$\textit{E}_{L}(\textbf{R}) = \frac{-\frac{1}{2} \nabla_{\textbf{R}}^{2} \Psi(\textbf{R})}{\Psi(\textbf{R})} + \textit{V}(\textbf{R})$$

diverge.

- In exact ground state, divergences cancel exactly.
- For Slater determinant, KE term does not diverge.
- ► For SJ trial function, divergences cancel if u(r_{ij}) obeys the cusp conditions:

$$\left. \frac{\partial u(r_{ij})}{\partial r_{ij}} \right|_{r_{ij}=0} = \begin{cases} -\frac{1}{2} & \text{parallel spins} \\ -\frac{1}{4} & \text{antiparallel spins} \end{cases}$$

Multiple Determinants

In molecules with strong static correlation, it is common to use trial functions including hundreds (but not millions) of Slater determinants.

In solids, this is less common:

- Suppose you have N electrons and use 2N basis functions in your mean-field calculation.
- Could construct ${}^{2N}C_N \approx e^{(2 \ln 2)N}$ determinants!
- Which to choose?

Fortunately, one-determinant SJ trial functions work surprisingly well in weakly correlated solids. Typically account for 85%+ of correlation energy.

Advantages and Disadvantages

Advantages:

- Surprising that such a simple approach returns 85% of E_c.
- Relatively quick (only 10 times DFT!)

Disadvantages:

- 85% of E_c is not enough.
- Hard to ensure error cancellation.
- Wave function optimization is a pain.
- Trial functions for La₂CuO₄?
- Garbage in garbage out!

Diffusion QMC

- The DMC method overcomes many of the problems of VMC.
- Uses a stochastic projection technique to improve an initial trial function (often taken from a prior VMC calculation).
- Typically returns 95%+ of correlation energy.
- DMC and affiliates are much the most accurate total energy methods available for large systems.

The Imaginary-Time Schrödinger Equation

$$\frac{\partial \Psi(t)}{\partial t} = -\hat{H}\Psi(t)$$
$$\frac{\partial \Psi(\mathbf{R}, t)}{\partial t} = \left[\frac{1}{2}\nabla_{\mathbf{R}}^2 - V(\mathbf{R})\right]\Psi(\mathbf{R}, t)$$

Cross between a 3N-dimensional diffusion equation

$$\frac{\partial n(\mathbf{R},t)}{\partial t} = D \nabla_{\mathbf{R}}^2 n(\mathbf{R},t)$$

and an exponential decay equation

$$\frac{\partial \mathbf{N}}{\partial t} = -\kappa \mathbf{N}$$

with a position-dependent decay rate.

Simple DMC Algorithm

Simple DMC Algorithm

- Scatter walkers **R** through configuration space. Each walker has initial "weight" w = 1.
- In time Δt , walker at **R**:
 - moves to $\mathbf{R}' = \mathbf{R} + \boldsymbol{\xi}\sqrt{2D\Delta t} = \mathbf{R} + \boldsymbol{\xi}\sqrt{\Delta t}$,
 - changes weight from w to $w' = we^{-V(\mathbf{R})\Delta t}$.
- To prevent walkers accumulating wild weights, allow low-weight walkers to die out and high-weight walkers to split up, keeping expected weight at every point unchanged.

DMC Evolution





Since $\psi(\mathbf{R})$ is antisymmetric, it cannot be a probability density!

DMC algorithm works, but yields the nodeless many-boson ground state of \hat{H} .

Can we circumvent this problem?

First Excited State of a Particle in a Box

• At t = 0, write $\Psi = \Psi_+ - \Psi_-$, where

$$\Psi_+ = rac{1}{2}(|\Psi|+\Psi) \ , \ \ \Psi_- = rac{1}{2}(|\Psi|-\Psi) \, .$$

- Evolve positive distributions $\Psi_+(t)$ and $\Psi_-(t)$ separately.
- Antisymmetric ground state = $\lim_{t\to\infty} [\Psi_+(t) - \Psi_-(t)].$



Failure!

- Unfortunately, both Ψ₊ and Ψ₋ tend to the symmetric ground state as t → ∞.
- The antisymmetric components decrease like $e^{-(E_1-E_0)t}$ in comparison.
- The subtraction required to extract the antisymmetric component becomes more and more difficult as *t* increases and the result is soon swamped by the statistical noise.

The Fixed-Node Approximation for the First Excited State





The Fixed-Node Approximation for the Many-Electron Ground State

The Fixed-Node DMC Algorithm

- Take an antisymmetric trial wavefunction (from VMC).
- Scatter walkers as normal and start DMC dynamics.
- Put impenetrable barriers on the nodal surface.
- Delete any walker attempting a trial move for which

 $\Psi_{T}(\mathbf{R}_{new})/\Psi_{T}(\mathbf{R}_{old}) < 0$.

- The fixed-node approximation is variational and normally very accurate (> 95% of correlation energy).
- We know of nothing better to do in large systems.

Nodal Surface



A 2D slice through the 321-dimensional nodal surface of a gas of 161 spin-up electrons. The positions of 160 electrons are fixed and the nodes plotted as a function of the position of the 161st.

Outline

1 Continuum QMC Methods

2 Example Applications

- 3 The Good, The Bad and The Ugly
- 4 Recent Advances and Prospects

The G1 Set

- Simplest plausible approach:
 - one determinant
 - no orbital optimization
- Mean absolute deviation: 2.9 kcal/mol (0.13 eV). (Comparable to B3LYP and CCSD(T) with similar basis.)
- Maximum deviation (SO₂): 14 kcal/mol (0.61 eV).

Not great, but remember that DMC scales

[J.G. Grossman, J. Chem. Phys. 117, 1434 (2002)]



Can We Do Better?

- Well depths for first-row homonumclear diatomic molecules.
- Fully optimized $J \times FVCAS$ trial function with up to \sim 150 CSFs.
- Typical accuracy 1 kcal/mol = 0.04 eV; largest error (N₂) 0.1 eV.
- Be₂ and Ne₂ OK.



[J. Toulouse and C.J. Umrigar, J. Chem. Phys. 128, 174101 (2008).]

W.M.C. Foulkes (Imperial College London)

Continuum QMC

EXCITCM

Cohesive Energies

Method	Si	Ge	С	BN	Al	
Experiment	4.62	3.85	7.37	12.9	3.43	
DFT	5.28	4.59	8.61	15.07	4.21	
VMC	4.48	3.80	7.36	12.85	3.23	
DMC	4.63	3.85	7.46		3.47	

in eV per atom statistical errors typically ± 0.04 eV/atom

Band Structure of Si



[A.J. Williamson, R.Q. Hood, R.J. Needs, and G. Rajagopal, Phys. Rev. B 57, 12140 (1998).]

Finding out about Exchange and Correlation

The Adiabatic Connection Formula

$$E_{\rm xc}[n] = \frac{1}{2} \iint \frac{n(r)\bar{n}_{\rm xc}(r,r')}{|r-r'|} \, dr \, dr'$$

where

$$\bar{n}_{xc}(r, r') = \int_{\lambda=0}^{1} n_{xc}^{\lambda}(r, r') \, d\lambda$$

$$\hat{H}^{\lambda}\psi^{\lambda} = \left(\hat{T} + \lambda \hat{V}_{ee} + \hat{V}^{\lambda}\right)\psi^{\lambda} = E^{\lambda}\psi^{\lambda}$$

$$n^{\lambda}(r) = n(r)$$

VMC Realisation of the Adiabatic Connection

$$E_{\rm xc} = \frac{1}{2} \iint \frac{n(r)\bar{n}_{\rm xc}(r,r')}{|r-r'|} \, dr' \, dr = \int e_{\rm xc}(r,[n]) \, dr$$

- Slater-Jastrow ansatz for many-body wavefunction.
- Integral over λ discretised.
- Variational parameters in ψ^λ and Fourier components of V^λ determined at every λ.
- Statistical errors small.
- Tried very hard to account for all the systematic errors (particularly finite-size errors).

The Exchange-Correlation Hole in Si



[R.Q. Hood, M.-Y. Chou, A.J. Williamson, G. Rajagopal, and R.J. Needs, Phys. Rev. B 57, 8972 (1998).]

The Exchange-Correlation Hole in Si



W.M.C. Foulkes (Imperial College London)



Exchange-Correlation Energy Density in Silicon



W.M.C. Foulkes (Imperial College London)



Error in LDA Exchange-Correlation Energy Density in Silicon





Point Defects in Alumina

Same mineral, different defects: Corundum, Ruby and Sapphire





Cr doped

Ti/Fe doped



N.D.M. Hine, K. Frensch, M.W. Finnis, W.M.C. Foulkes, and A.H. Heuer, Phys. Rev. B **79**, 024112 (2009) and tbp.

Formation Energies

• For defects in alumina of charge $q = -\Delta n_e$,

$$E_{f} = G_{def}^{q} - G_{perf} - \Delta n_{AI}\mu_{AI} - \Delta n_{O}\mu_{O} + q\mu_{e}$$
.

- *G*^{*q*}_{*def*} and *G*_{*perf*} approximated by total energies plus vibrational contributions.
- μ_O , μ_{Al} fixed by annealing conditions.

Why DMC?

Table: Formation (all in eV). ΔH_0^{Al} and ΔH_0^O are the formation enthalpies per atom of Al and O atoms in the gas phase. ΔH_0^{AlO} is the formation energy of an AlO molecule.

Method	ΔH_0^{A1}	$\Delta H_0^{\rm O}$	ΔH_0^{AlO}
LDA-USP	4.05	3.62	0.91
LDA-DF	4.10	3.67	1.13
GGA-USP	3.41	2.82	0.74
DMC	3.47(1)	2.54(1)	0.68(1)
Experiment	3.42	2.58	0.69

Why DMC?

In case this is not yet convincing ...

Occupied defect states deriving from conduction band states are too low in energy because of DFT gap underestimation.



Correction $m \times \Delta E_g$ is normally applied to DFT formation energies.



DMC Formation Energies

If no bandgap correction is applied, DMC results agree well with DFT except for correcting self-interaction error of localised states.



DFT (including bandgap correction) appears to be significantly overbinding. Real cost to break bonds is lower.

DMC Formation Energies



Oxygen Interstitial Formation Energies

Interstitial is consistently easier to form in DFT, also suggesting DFT overbinds.

Other Applications

- Surface energies and reactions at surfaces.
- Clusters.
- Static linear response.
- Fermi liquid parameters.
- Hydrogen bonding and van der Waals bonding.
- Transition metal oxides.
- Metal-insulator transitions.
- Phase stability.
- . . .

Outline



- Example Applications
- The Good, The Bad and The Ugly
 - 4 Recent Advances and Prospects

The Good

- DMC works well stable and nearly automatic once it is going.
- DMC algorithm is "trivially parallel".
- More accurate than DFT. Normally much more accurate.
- The only (?) way to benchmark DFT in large systems.
- System-size scaling no worse than DFT.
- Provides direct information about correlation.
- Excellent total energies; often lower than full CI.



The Bad

- DFT is 10³-10⁴ (?) times faster. DMC is currently too slow for serious quantum molecular dynamics.
- Excellent total energies do not imply similarly excellent energy differences: fixed-node errors do not cancel as well as you might hope.
- Limited ideas about how to improve nodes in large systems.
- Very limited information about excited states.
- Force calculations are hard and slow.

The Ugly

- Practicality relies on pseudopotentials:
 - Not a many-electron concept.
 - Treatment of non-locality requires additional approximation.
- Trial function construction and optimization is a pain.
- Finite-size errors are a pain.

Outline



- 2 Example Applications
- 3 The Good, The Bad and The Ugly
- 4 Recent Advances and Prospects

... are the largest errors in most QMC calculations of solids.



... are the largest errors in most QMC calculations of solids.



... are the largest errors in most QMC calculations of solids.



... are the largest errors in most QMC calculations of solids.



... are the largest errors in most QMC calculations of solids.



Trial Wavefunction Optimization

- Filippi, Nightingale, Sorella, Toulouse, Trail, Umrigar, ..., have enormously improved wavefunction optimization techniques over the past few years.
 - Automatic optimization (almost) a reality.
 - No more need to sacrifice graduate students.

This is *much* more important than it sounds

Better Trial Functions for Large Systems

Two main ideas are being investigated.

Geminals and Pfaffians

Attacalite, Bajdich, Casula, Mitas, Schmidt, Sorella, Wagner.

Replace determinant by antisymmtrized product of pairing functions:

$$\hat{\mathcal{A}}\phi_1(\mathbf{x}_1,\mathbf{x}_2)\phi_2(\mathbf{x}_3,\mathbf{x}_4)\ldots\phi_{N/2}(\mathbf{x}_{N-1},\mathbf{x}_N).$$

- If all φ functions are identical singlets, get BCS wavefunction.
- If $\phi(\mathbf{x}_1, \mathbf{x}_2) = \sum_{k=1}^{N} \phi_k(\mathbf{x}_1) \phi_k(\mathbf{x}_2)$, the BCS wavefunction reduces to a Slater determinant.
- Can combine singlet pairing, triplet pairing and unpaired electrons.

Backflow

Ceperley, Esler, Drummond, Holzmann, Kwon, Lin, López-Ríos, Ma, Martin, Needs, Pierleoni, Towler, Zong.

Electron positions in determinants are replaced by collective coordinates:

$$D(\mathbf{R}) \rightarrow D(\tilde{\mathbf{R}}),$$

where $\tilde{\mathbf{r}}_i = \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R})$.

Forces

- Difficult because the nodal constraint (which is in effect a boundary condition on the wave function) depends on nuclear positions. Leads to hard-to-evaluate Pulay-like terms.
- Combining recent advances in low-variance estimators (Assaraf, Caffarel, Filippi, Umrigar, ...) with the Hellmann-Feynman theorem allowed Badinski, Needs and co-workers to evaluate forces accurately (although approximately). Effort remains large.



QMC MD



FIG. 1 (color online). DMC total ground state energies vs time for 5 fs time intervals of an AIMD simulation of SiH_4 , Si_5H_{12} , and $Si_{14}H_{20}$ at 1000 K. The circles (lines) correspond to discretely sampled (continuous) DMC calculations. The error bars of the discretely sampled data are smaller than the symbol size.

[J.C. Grossman and L. Mitas, Phys. Rev. Lett. 94, 056403 (2005).]

The Melting Curve of Iron

$$\Delta F = \int_0^1 d\lambda \langle \Delta U \rangle_{\lambda} \approx \langle \Delta U \rangle_{\lambda=0} - \frac{1}{2k_BT} \left\langle (\Delta U - \langle \Delta U \rangle_{\lambda=0})^2 \right\rangle_{\lambda=0}$$



Solid is blue Liquid is red

 $P=330 {
m GPa}$ $T_{
m DFT}^{
m m}=6350\pm300 {
m K}$ $\Delta T^{
m m}=550\pm230 {
m K}$

[E. Sola and D. Alfè, Phys. Rev. Lett. 103, 078501 (2009).]

New QMC Algorithms

Phaseless Auxiliary-Field QMC

S. Zhang, in *Theoretical Methods for Strongly Correlated Electron Systems*, D. Senechal and A.-M. Tremblay, Eds. (Springer-Verlag, 2003).

FCI-QMC

G.H. Booth, A.J.W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

Challenges

- "Real" QMC MD.
- "Real" excitations.
- Stronger correlations.
- Systematic improvement of nodal surface.
- Larger systems and more real applications (no more dimers!).

Summary

- QMC is useful but field is small and relatively underdeveloped.
 - Still room for fundamental methods development.
 - Still room to make a big splash.
- QMC is perfect for next generation hardware. QMC MD is on the way.
- QMC is not the one true answer: you cannot calculate everything and you have to work hard to get results.
- After a slow patch, techniques have advanced substantially during the past few years, leaving me optimistic about the future.