

Introduction to the Variational and Diffusion Monte Carlo Methods

UCSB, 12 noon, Wednesday 7th August 2002

W.M.C. Foulkes
M. Nekovee

Imperial College

R.J. Needs
G. Rajagopal
M.D. Towler

Cavendish Laboratory

Outline

1. The many-electron problem
2. VMC basics
3. DMC basics
4. The simulation cell Hamiltonian
5. What can VMC/DMC do?
 - (a) What can they calculate?
 - (b) Where is the effort?
 - (c) Excitations
 - (d) Strong correlations
6. Conclusions

1. The Many-Electron Problem

Why compute?

- I can't do the mathematics ...
- I can't do the experiment ...

A good computational method is one that

- Produces good data
- Is general and robust enough to yield surprises
- Is simple enough to trust
- Can be systematically improved

The many-electron problem is easy because

- Linear, 2nd order, Hermitian
- Screening
- Pauli principle

The many-electron problem is hard because

- $3 \times 6 \times 10^{23}$ dimensions
- Pauli principle
- Wide range of energy/time/length scales
 - LDA accounts for $> 99\%$ of total energy \Rightarrow separate energy scales or get the chemistry *right!*
 - Precision matters: basis sets must be very accurate or systematically improvable.

Where we stand now

- Ground-state properties of weakly interacting systems: DFT, QMC
- Excitations of weakly interacting systems: *GW*, TDDFT, (QMC)
- Ground-state properties of strongly interacting systems: DFT, QMC?, DMFT?
- Excitations (on 1 eV scale) of strongly interacting systems: DMFT, TDDFT?, (QMC?)

Where we would like to stand

- Ground state and excitations on scale $\leq kT$

2. VMC Basics

- Guess $\psi_T(r_1, r_2, \dots, r_{10^{23}}) = \psi_T(R)$.
- Evaluate

$$\begin{aligned} E_T &= \int \psi_T^*(R) \hat{H} \psi_T(R) dR \\ &= \int \left(\frac{\hat{H} \psi_T(R)}{\psi_T(R)} \right) |\psi_T(R)|^2 dR \\ &= \int E_L(R) |\psi_T(R)|^2 dR \end{aligned}$$

using Monte Carlo integration.

- Adjust ψ_T to minimise E_T .

But isn't guessing ψ_T impossible? Not always:

Laughlin: FQHE

BCS: superconductivity

Slater-Jastrow: weakly correlated Fermi liquids

Slater-Jastrow trial wavefunction

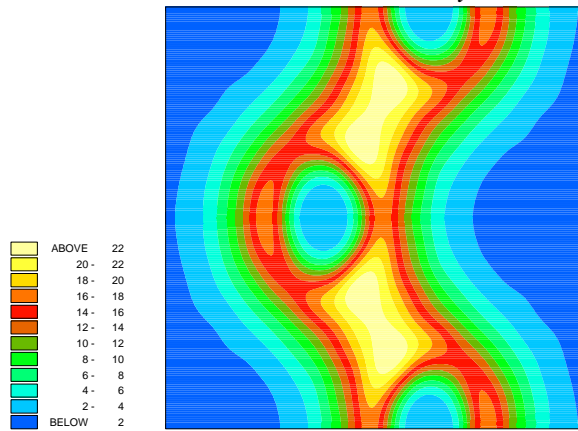
$$\psi_T = \exp\left(-\sum_i \chi(r_i)\right) \exp\left(-\sum_{i>j} u(r_{ij})\right) D$$

where

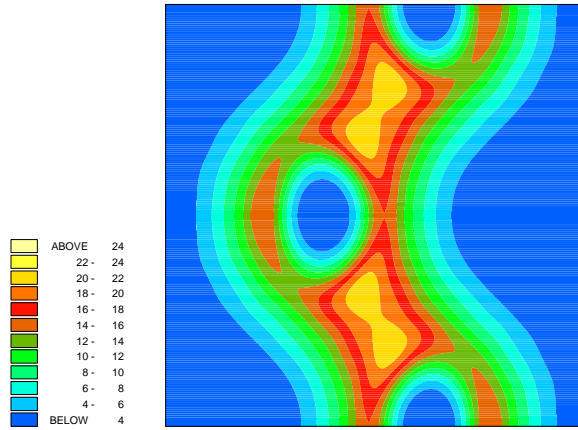
$$D = \begin{vmatrix} \phi_1(r_1) & \phi_1(r_2) & \dots & \dots & \phi_1(r_N) \\ \phi_2(r_1) & \phi_2(r_2) & \dots & \dots & \phi_2(r_N) \\ \dots & \dots & & & \dots \\ \dots & \dots & & & \dots \\ \phi_N(r_1) & \phi_N(r_2) & \dots & \dots & \phi_N(r_N) \end{vmatrix}$$

- The $\exp(-\sum u)$ term decreases the chance of two electrons coming close together but spoils the accurate LDA electron density. The $\exp(-\sum \chi)$ term cures this problem.
- $u(r_{ij})$ incorporates the e-e cusp conditions at small r_{ij} and approaches the Bohm-Pines RPA limit at large r_{ij} .
- u and χ (and ϕ_i) treated variationally.

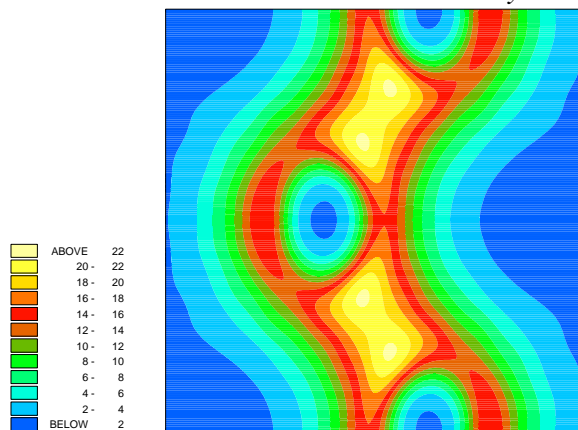
Ge : HF density



Ge : HFJ density



Ge : HFJ + CHI density



Other expectation values

$$\begin{aligned} O_T &= \int \left(\frac{\hat{O}\psi_T(R)}{\psi_T(R)} \right) |\psi_T(R)|^2 dR \\ &= \frac{1}{M} \sum_{i=1}^M \frac{\hat{O}\psi_T(R_i)}{\psi_T(R_i)} \end{aligned}$$

If $\hat{O} = \hat{H}$, variance tends to zero as $\psi_T \rightarrow \psi_0$.

Zero Variance Principle

Problems with VMC

VMC seems to be remarkably accurate (10 times more accurate than LDA for cohesive energies) in weakly correlated solids such as jellium, Ge, Si, C, Al, but:

- Get out what you put in (GIGO).
- ψ_T for interesting materials?

Are there any better methods?

3. DMC Basics

The imaginary-time Schrödinger equation

$$\begin{aligned}\frac{\partial \psi}{\partial t} &= -\hat{H}\psi \\ \psi(t) &= e^{-\hat{H}t}\psi(0)\end{aligned}$$

If $\psi(0) = \sum c_i \psi_i$,

$$\psi(t) = \sum c_i e^{-E_i t} \psi_i$$

becomes proportional to ψ_0 as $t \rightarrow \infty$.

Simplest one-particle algorithm

$$\frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \frac{1}{2} \nabla^2 \psi(\mathbf{r}, t) - U(\mathbf{r}) \psi(\mathbf{r}, t)$$

Distribute walkers (*not* particles) randomly

For each time step {

For each walker {

Move by step $\Delta \mathbf{r}$ chosen from Gaussian of variance Δt in each dimension

Calculate $w = e^{-U(\mathbf{r})\Delta t}$

Replace walker at \mathbf{r} by $\text{INT}(w + \eta)$ copies, where η is chosen from $U(0, 1)$

}

}

In limit as $t \rightarrow \infty$, probability density is $\propto \psi_0(\mathbf{r})$.

Importance sampling in theory

The simplest algorithm is disastrous in realistic cases. Must deal with singularities in $U(\mathbf{r})$.

Assume that $\psi_0(\mathbf{r})$ is known and consider

$$f(\mathbf{r}, t) = \psi_0(\mathbf{r})\psi(\mathbf{r}, t)$$

Its equation of motion is

$$\begin{aligned}\frac{\partial f}{\partial t} &= \psi_0 \left(\frac{1}{2} \nabla^2 \psi - U \psi \right) \\ &= \frac{1}{2} \nabla^2 f - \nabla \cdot (\mathbf{v} f) - E_0 f\end{aligned}$$

where $\mathbf{v} = \nabla \psi_0 / \psi_0$.

- Choose zero of energy such that $E_0 = 0$
- Fokker-Planck equation conserves walker number
- $\mathbf{v}(\mathbf{r})$ is well behaved ($\psi_0 > 0$)
- If you know ψ_0 , you can sample ψ_0^2 and find E_0 !

Importance sampling in practice

In practice, of course, we don't know ψ_0 , but we can try the same manoeuvre with ψ_T .

Define

$$f(\mathbf{r}, t) = \psi_T(\mathbf{r})\psi(\mathbf{r}, t)$$

Its equation of motion is

$$\frac{\partial f}{\partial t} = \frac{1}{2}\nabla^2 f - \nabla \cdot (\mathbf{v}f) - E_L f$$

where

$$\mathbf{v}(\mathbf{r}) = \frac{\nabla\psi_T(\mathbf{r})}{\psi_T(\mathbf{r})} \quad E_L(\mathbf{r}) = \frac{\hat{H}\psi_T(\mathbf{r})}{\psi_T(\mathbf{r})}$$

- Choose zero of energy such that walker number is conserved on average
- $\mathbf{v}(\mathbf{r})$ is well behaved ($\psi_T > 0$)
- $E_L(\mathbf{r})$ is well behaved ($\psi_T > 0$ and obeys cusp conditions)
- As $\psi_T \rightarrow \psi_0$, $E_L(\mathbf{r}) \rightarrow E_0$, population fluctuations vanish

Better one-particle algorithm

$$\frac{\partial f}{\partial t} = \frac{1}{2}\nabla^2 f - \nabla \cdot (\mathbf{v}f) - E_L f$$

Distribute walkers (*not* particles) randomly

For each time step {

For each walker {

Move by step $\Delta\mathbf{r} = \mathbf{v}(\mathbf{r})\Delta t + \boldsymbol{\xi}$, where $\boldsymbol{\xi}$ is a random vector with each component chosen from Gaussian of variance Δt .

Calculate $w = e^{-E_L(\mathbf{r})\Delta t}$

Replace walker by $\text{INT}(w + \eta)$ copies, where η is chosen from $U(0, 1)$

}

}

In limit as $t \rightarrow \infty$, probability density is $\propto \psi_T(\mathbf{r})\psi_0(\mathbf{r})$.

Expectation values

Ground-state energy is easy:

$$\begin{aligned} E_0 &= \frac{\int \psi_0 \hat{H} \psi_T d^3r}{\int \psi_0 \psi_T d^3r} \\ &= \int \left(\frac{\hat{H} \psi_T}{\psi_T} \right) \frac{\psi_T \psi_0}{\int \psi_T \psi_0 d^3r} d^3r \\ &= \langle E_L \rangle \end{aligned}$$

zero variance principle applies

Other expectation values are more troublesome:

$$\int \left(\frac{\hat{O} \psi_T}{\psi_T} \right) \frac{\psi_T \psi_0}{\int \psi_T \psi_0 d^3r} d^3r = \frac{\int \psi_0 \hat{O} \psi_T d^3r}{\int \psi_0 \psi_T d^3r}$$

One approach is extrapolated estimation:

$$\begin{aligned} O_0^{\text{extrapolated}} &= 2O_0^{\text{DMC}} - O_0^{\text{VMC}} \\ &= O_0 + \mathcal{O}[(\psi_T - \psi_0)^2] \end{aligned}$$

DMC with N particles

$$\begin{aligned}\hat{H} &= \sum_i -\frac{1}{2}\nabla_i^2 + \sum_i V_{\text{ext}}(r_i) + \sum_{j>i} \frac{1}{|r_i - r_j|} \\ &= -\frac{1}{2}\nabla_R^2 + U(R)\end{aligned}$$

has same form as one-particle \hat{H} (with r replaced by R), so everything works just as before.

- Walkers drift/diffuse/branch in $3N$ -dimensional configuration space.
- In limit as $t \rightarrow \infty$, walker density samples $\psi_T(R)\psi_0(R)$.
- Calculation of expectation values exactly as for one electron.

The sign problem (without importance sampling)

Walker density $\psi(R, t) \geq 0$, so simple DMC algorithm cannot produce fermion ground state. In fact, it gives the totally symmetric boson ground state.

How can we impose antisymmetry?

1. Build it into the propagator. Instead of defining

$$\begin{aligned} G(R', R, \Delta t) &= \langle R' | \exp(-\Delta t \hat{H}) | R \rangle \\ &= \frac{1}{(2\pi\Delta t)^{3N/2}} e^{-(R'-R)^2/2\Delta t} e^{-U(R)\Delta t} \end{aligned}$$

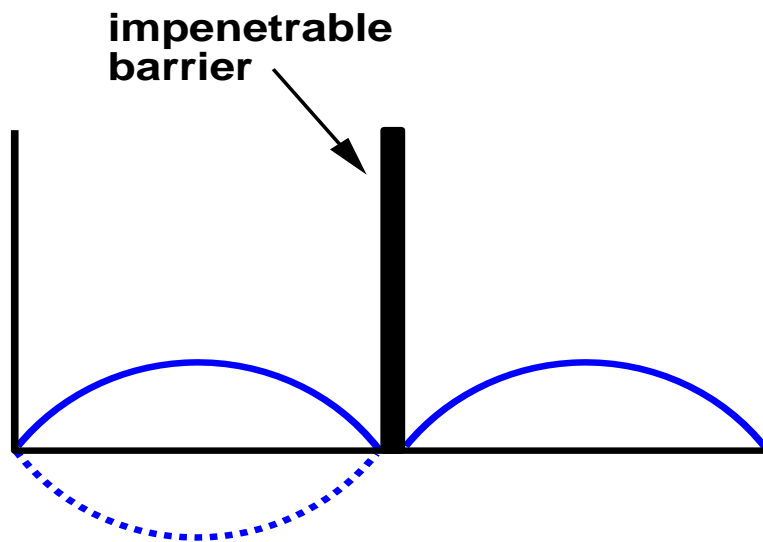
with unsymmetrised states $|R\rangle$, use antisymmetrised states.

2. Build it into the starting state. If $\psi(t=0)$ is anti-symmetric, so is $\exp(-t\hat{H})\psi(t=0)$... in principle.

The fixed node approximation

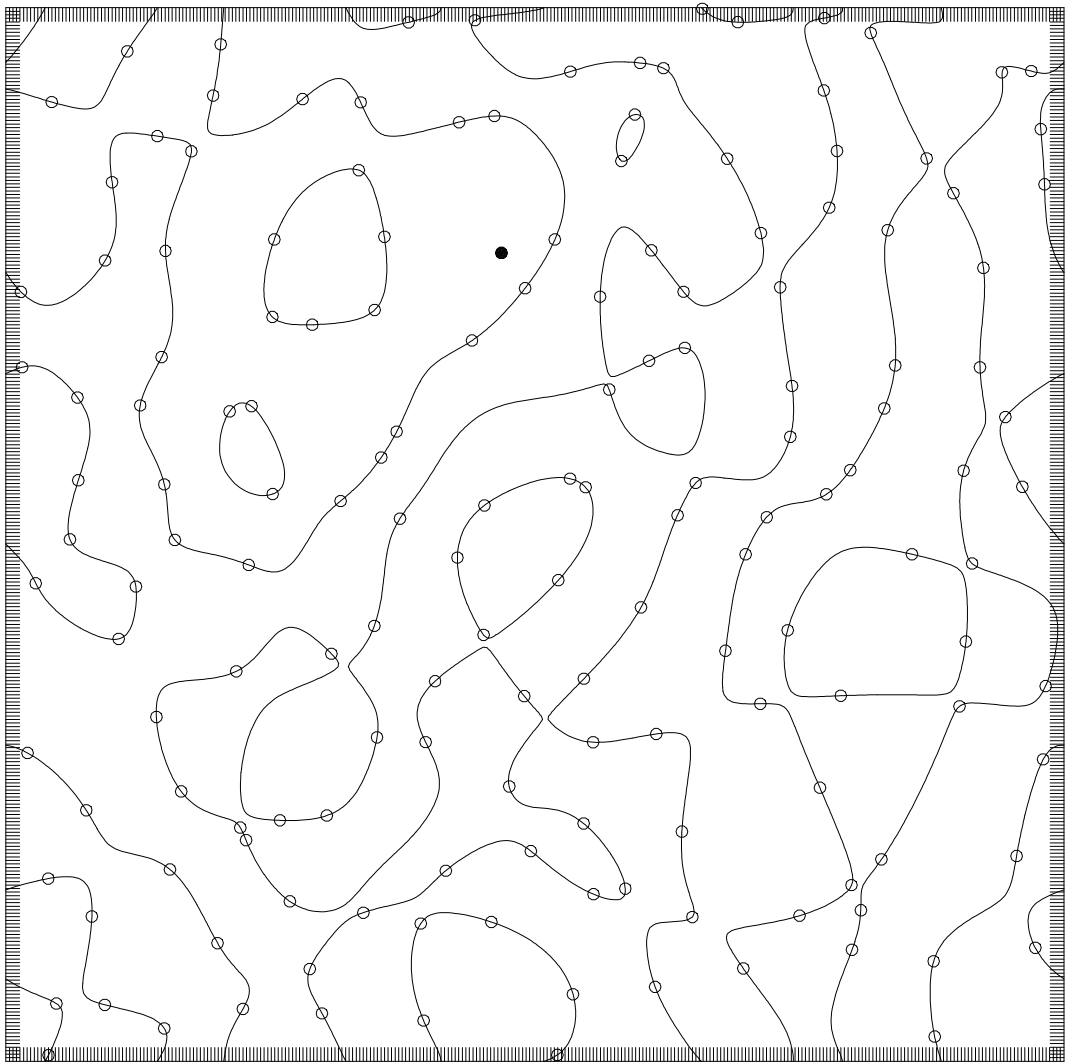
In large systems, only practical solution is the *fixed node approximation*:

- Impose nodes as a boundary condition and solve Schrödinger equation separately in each nodal pocket.



- If nodes are right, so is result; if nodes are wrong, result is variational.
- In solids, normally use LDA nodal surface. Results are more accurate than can be obtained using any other known ground-state method.

Exact nodal surfaces are weird, so the fixed-node variational principle is important.

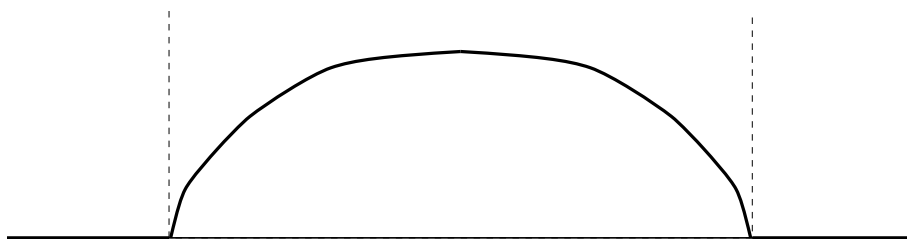


A two-dimensional slice through the 321-dimensional nodal surface of a two-dimensional electron gas containing 161 spin-up electrons subject to periodic boundary conditions (from Ceperley, *J. Stat. Phys.* **63**, 1237, 1991).

The fixed-node variational principle

DMC algorithm in nodal pocket v_α produces state $\phi_\alpha(\mathbf{R})$ satisfying:

$$\begin{aligned}\hat{H}\phi_\alpha(\mathbf{R}) &= \epsilon_\alpha\phi_\alpha(\mathbf{R}) + \delta_\alpha && \text{when } \mathbf{R} \in v_\alpha \\ \phi_\alpha(\mathbf{R}) &= 0 && \text{when } \mathbf{R} \notin v_\alpha\end{aligned}$$



Construct state Φ_α by antisymmetrising

$$\Phi_\alpha(\mathbf{R}) = \frac{1}{N!} \sum_P (-1)^P \phi_\alpha(P\mathbf{R}) \equiv \hat{\mathcal{A}}\phi_\alpha(\mathbf{R})$$

This state cannot be zero.

The standard QM variational principle then gives

$$\begin{aligned} E_0 &\leq \frac{\langle \Phi_\alpha | \hat{H} | \Phi_\alpha \rangle}{\langle \Phi_\alpha | \Phi_\alpha \rangle} = \frac{\langle \Phi_\alpha | \hat{H} \hat{\mathcal{A}} | \phi_\alpha \rangle}{\langle \Phi_\alpha | \hat{\mathcal{A}} | \phi_\alpha \rangle} \\ &= \frac{\langle \Phi_\alpha | \hat{H} | \phi_\alpha \rangle}{\langle \Phi_\alpha | \phi_\alpha \rangle} = \epsilon_\alpha \end{aligned}$$

- $\hat{\mathcal{A}}$ commutes with \hat{H} .
- $\hat{\mathcal{A}}$ is self-adjoint and idempotent.
- The delta functions in $\hat{H}\phi_\alpha$ don't contribute because they appear on the nodal surface where $\Phi_\alpha(\mathbf{R})$ is zero.

The tiling theorem

All nodal pockets of the fermion ground state Ψ_0 are equivalent by symmetry

- Take one nodal pocket of Ψ_0 and colour it blue.
- Each permutation P maps the blue pocket into itself or into another pocket equivalent by symmetry. This is also coloured blue.
- Do the blue pockets fill configuration space? Assume not and consider the state Φ defined by:

$$\begin{aligned}\Phi(\mathbf{R}) &= \Psi_0(\mathbf{R}) && \text{in blue pockets} \\ \Phi(\mathbf{R}) &= 0 && \text{elsewhere}\end{aligned}$$

- $\Phi(\mathbf{R})$ satisfies

$$\hat{H}\Phi(\mathbf{R}) = E_0\Phi(\mathbf{R}) + \delta$$

where E_0 is the exact ground-state energy.

- Hence

$$E_0 = \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle}$$

- But Φ has gradient discontinuities and *can't* be the ground state. Hence the blue region must fill configuration space.

4. The Simulation Cell Hamiltonian

Pseudopotentials

In small systems we can use the full Hamiltonian,

$$\sum_i -\frac{1}{2}\nabla_i^2 + \sum_{i>j} \frac{e^2}{|r_i - r_j|} + \sum_{i,\alpha} \frac{eZ_\alpha}{|r_i - d_\alpha|} + \sum_{\alpha>\beta} \frac{Z_\alpha Z_\beta}{|d_\alpha - d_\beta|} ,$$

but the required computer time scales like Z^{5-6} !

We're stuck with first row atoms unless we use *pseudopotentials* (warning: a "one-electron" idea).

VMC: use best non-local norm-conserving pseudopotentials

DMC: technical problems with non-local potentials only partly solved

Pseudopotential problems are main limitation on fun applications of DMC

Periodic boundary conditions

In small systems we can use the full Hamiltonian, but in solids or liquids we must:

- Replace infinite system by finite *model system* with $N \leq 1000$.
- Apply periodic (*not* Bloch) boundary conditions.
- Use a model electron-electron interaction energy.

The model electron-electron interaction should

model the forces in the real solid;

but depend only on the positions of the electrons in the simulation cell.

We use the energy per simulation cell of an infinite periodic lattice of identical copies of the simulation cell, calculated using the Ewald method.

5. What Can VMC/DMC Do?

(a) What can they calculate?

- Total (pseudo)energies; energy barriers; energy differences.
 - Accurate (0.1 eV per atom)
 - Comparatively efficient (zero-variance principle)
 - Exact bar fixed-node approximation — typical error $\sim 5\%$ of correlation energy
- Forces on atoms — are just energy derivatives. (Not yet efficient enough for molecular dynamics.)
- Other static ground-state expectation values (including many-particle operators).
 - Not variational; no zero-variance principle; DMC only yields mixed estimator
 - Nevertheless practical
 - Better than DFT?
- Excitation spectra?

(b) Where is the effort?

Time to move all the electrons in one walker:

- Recalculate one-electron orbitals: N^2 (?)
- Update Slater inverse N^3 (?)
- Jastrow factor: N^2 (?)
- Coulomb energy: N^2 (?)

Variance of local energy: $\sigma_{E_L}^2 \sim N$

Variance of energy per electron: $\sigma_e^2 \sim 1/N$

To converge energy per electron to Δ requires M walker moves, where $\Delta \sim \sigma_e / \sqrt{M}$.

Hence $M \sim \sigma_e^2 \sim 1/N$. Total effort $\sim N^2$.

To converge total energy to Δ requires $M \sim \sigma_{E_L}^2 \sim N$ moves. Total effort $\sim N^4$.

(c) Excitations

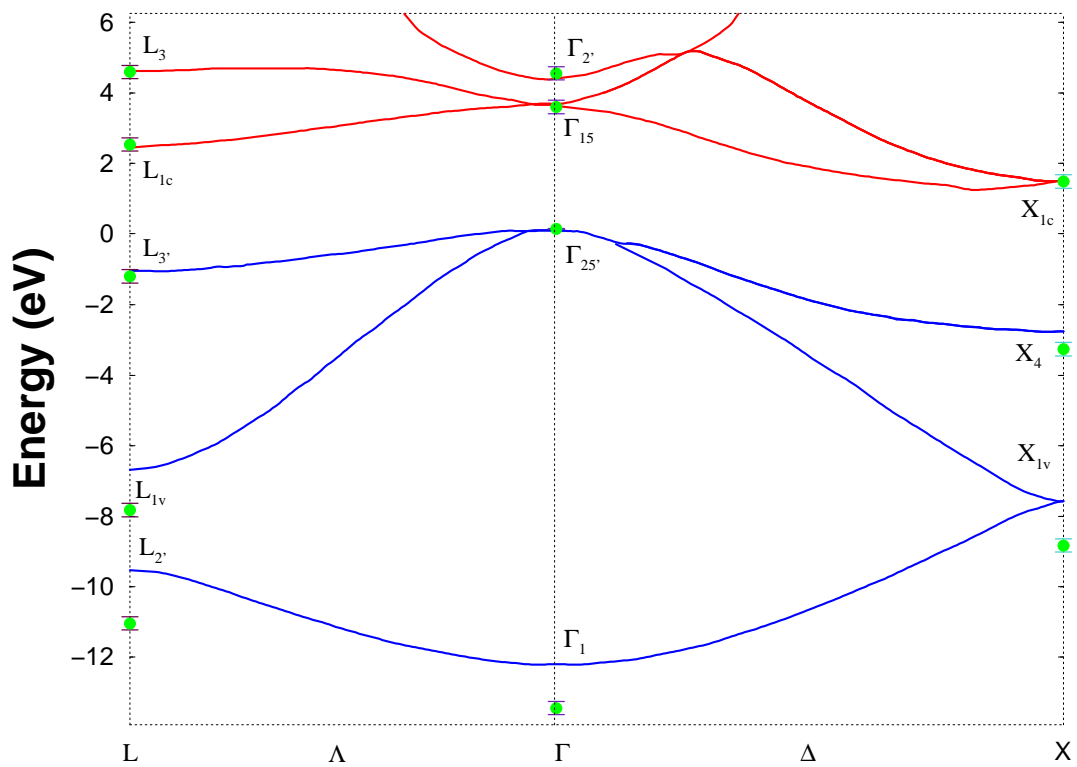
Limited information is available.

Band Gaps

$$E_g = (E_0^{N+1} - E_0^N) - (E_0^N - E_0^{N-1})$$

Band Structures

By addition, subtraction, promotion



Not well justified

Useful only when have sharp quasiparticle peaks

Better approaches?

Would like, e.g.,

$$A_+(k, \omega) = \sum_i \left| \langle \psi_i^{N+1} | \hat{c}_k^\dagger | \psi_0^N \rangle \right|^2 \delta(\omega - E_i^{N+1} - E_0^N)$$

Formally, DMC has no problem calculating

$$C_{AB}(t) = \langle \psi_T | \hat{A}^\dagger e^{-(\hat{H} - E_T)t} \hat{B} | \psi_T \rangle$$

which is Laplace transform of

$$C_{AB}(\omega) = \sum_i \langle \psi_T | \hat{A}^\dagger | \psi_i \rangle \langle \psi_i | \hat{B} | \psi_T \rangle \delta(\omega - E_i + E_T)$$

This looks promising, but the folk wisdom is that it doesn't work — inconsistent with fixed-node approximation. (I'm not convinced.)

(Relaxing FNA \Rightarrow exponentially growing noise \Rightarrow Laplace inversion even worse than usual.)

Silly idea

Combine QMC and Lanczos?

(d) Strong correlations

Why not?

6. Conclusions

- QMC calculations in real solids are practical; there's lots to do.
- Slater-Jastrow trial functions work very well in all (weakly correlated) cases investigated so far.
- Cohesive energies are an order of magnitude more accurate than those obtained from LDA calculations.
- Pair-correlation functions/xc holes.
- No LDA band-gap problem.

But:

- Need a lot of computer time.
- Pseudopotential problems in DMC.
- Finite size errors.
- More interesting excitations.
- Strongly correlated systems?