Some mathematical questions related to quantum Monte Carlo approaches for molecules

Michel Caffarel

Laboratoire de Physique Quantique, IRSAMC Université Paul Sabatier, Toulouse

QMC ?

Quantum Monte Carlo (QMC) = generic name for a variety of methods “solving stochastically” the stationary Schrödinger equation.

More precisely: Finding the lowest eigenvalue E_0 of:

Continuous systems: $H = -\frac{1}{2}\nabla^2 + V(x) \quad x \in R^d \; d = 3N$

Discrete systems: H= very large **sparse** matrix (linear size much greater than 10^{10})

QMC community

Four pertinent parameters:

- Temperature: $T = 0$ or $T \neq 0$
- Statistics: F (Fermions), B (Bosons), Bz (Boltzmannons: distinguishable particles)
- Nature of configuration space: D (discrete), C (continuous)
- Extension of the system: infinite (solids,liquids, the thermodynamical limit has to be taken: $N \to \infty$, finite (molecules, cluster). M(macro) ou m(micro)

$\Rightarrow$ 2x3x2x2= potentially: 24 communities...

Physics:

(0,C,B,M), (T,C,B,M): Bosonic liquids He_4, Superfluidity

(0,C,F,M),(T,C,B,M): Fermi liquids He_3 Very rich phase diagram

(0,D,F,M),(0,D,F,M): Theoretical condensed-matter physics Hubbard model, High-T_c superconductivity

(0,C,F,M): Solid-state physics Silicium

(0,C,F,m): Nuclear physics Tritium nucleus

Chemistry:

(0,C,Bz,m): Ro-vibrational spectroscopy Water IR spectrum

(0,C,F,m): Electronic Structure of molecules H_2?, rather Li...

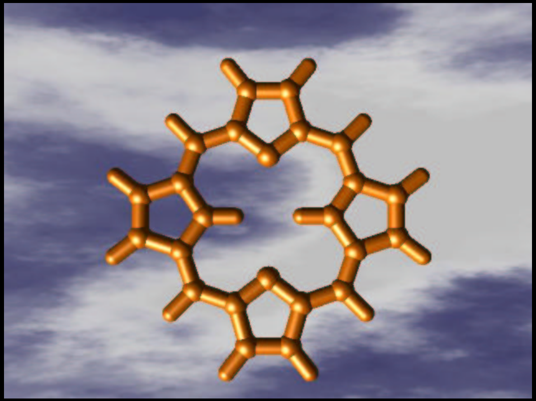
Warning!

Numerous QMC algorithms with various acronyms:

VMC, DMC, PDMC, GFMC, PIMC, projector MC, Worldline MC, etc....

Here: Electronic Sructure for quantum Chemistry: T=0 , continuous space, fermions, finite system.

Recent Advances: Porphyrin



- Excitations (eV) of porphyrin (162 electrons)
- Largest all-electron DMC calculation to date

Method / Excitation	Adiabatic	Vertical
	$S_0 \rightarrow T_1$	$S_0 \rightarrow S_2$
CASPT2	-	2.26
CIS	-	2.66
SAC-CI	-	2.25
TD-DFT	-	2.39
DFT-MRCI	-	2.38
MR-SD CI	-	2.40
EOM-CCSD	-	2.76
STEOM-CC	-	2.61
DMC *	**1.60(1)**	**2.45(08)**
Exp.	1.58	2.42,2.46

A. Aspuru-Guzik O. El Akramine, J. C. Grossman, and WAL J. Chem. Phys. February 15, 2004

Fundamental idea

Quantum Monte Carlo (QMC) = “stochastic” power method

Let $|u_0>$ be an arbitrary vector:

Power method: $H^n|u_0>$

$\Rightarrow$ Extraction of the eigenvector associated with the largest eigenvalue [$(E-H)^n|u_0>$: lowest eigenvalue]

Fundamental step: Multiplication of H by |u>; $\sum_{j=1}^{N} H_{ij}u_j$

- Full multiplication = usual deterministic power method

- Importance sampled multiplication = QMC (only the most significant contributions to the sum are computed and accumulated)

In practice

$G(H)$: $G(H)|u_0>$

Lattice: $G(H) = 1 - \tau(H - E_T)$

Continuous: $G(H) = \exp[-\tau(H - E_T)]$

Matrix elements of $G(H)$ are written as a product of a probability transition (stochastic matrix) times a weight:

$$\boxed{G_{ij} = p(i \to j) w_{ij}}$$

Example:

$< x, \exp[-\tau(H - E_T), y > \sim_{small\ \tau} < x, \exp(\tau/2\nabla^2, y >$ $\exp[-\tau(V(x) - E_T)]$

with $p(x \to y) = (\frac{1}{2\pi\tau})^d \exp(-\frac{(x-y)^2}{2\tau})$

Feynman-Kac type formulae

$p(i \rightarrow j)$ allow to generate series of “states” (walkers, particles,...) in time (iteration number n)

$\rightarrow$ notion of “trajectories”

Quantum averages can be written as average over the set of trajectories (path integral formalism)

$\rightarrow$ Feynman-Kac point of view

Example:

$\phi_0(x) \sim < e^{-\int_0^{+\infty} V(x(s))ds} >$

$< \cdots >$= average over the set of all free brownian trajectories starting at x; (this formula is just the rewritting of $[e^{-\tau H}]^n$ applied to $\delta(x-y)$

Importance sampling

As in any Monte Carlo scheme, importance sampling has to be introduced.

Here, importance sampling = introduction of some good guess of the unknown eigenvector, $|u^T>$ or ψ_T

$$p^T(i \rightarrow j) = u_j^T / u_i^T p(i \rightarrow j)$$

Continuous systems:

$$p^T(x \rightarrow y) = (\frac{1}{2\pi\tau})^d e^{-\frac{(y-x-b(x)\tau)^2}{2\tau}}$$

$b(x) = \nabla \psi_T / \psi_T$ drift vector

$\rightarrow$ “drifted” random walks instead of free random walks

Importance sampling II

The usual Feynman-Kac weight:

$\exp[-\tau(V(x) - E_T)]$

is “renormalized” and replaced by:

$\exp[-\tau(E_L - E_T)]$ $E_L = H\psi_T/\psi_T$, local energy

"Sign" problem

Considered as one of the most important problem in computational physics.

Electrons are fermions: From a mathematical point of view, two types of electrons (α,β)

$$\phi_0(\vec{r}_1, ..., \vec{r}_{N_\alpha} | \vec{r}_{N_\alpha+1}, ..., \vec{r}_{N_\alpha+N_\beta})$$

ϕ_0 must be antisymmetric under the exchange of two α or two β electrons.

"Sign" problem II

$\psi_T = \psi_T^+ - \psi_T^-$ fermionic

Power method: $e^{-nH}\psi_T = e^{-nH}\psi_T^+ - e^{-nH}\psi_T^-$

At large n (to extract the ground-state)

$$e^{-nH}\psi_T \sim e^{-nE_B}\phi_0^{Bose} - e^{-nE_B}\phi_0^{Bose} + \ldots$$

All bosonic terms vanish and the fermionic contribution, $e^{-nE_F}\phi_0^{Fermi}$, is exponentially small with respect to the leading bosonic terms.

→ Exponentially bad signal-to-noise ratio

Fixed-node Approach

FN approach = runs the algorithm as it is!

$$p(x \rightarrow y) = (\frac{1}{2\pi\tau})^d e^{-\frac{(y-x-b(x)\tau)^2}{2\tau}}$$

$b(x) = \nabla\psi_T/\psi_T$ drift vector

ψ_T antisymmetric.

What happens?

Walkers are trapped into the nodal pockets of the trial wavefunction (the zeroes of ψ_T act as infinitely repulsive barriers)

Fixed-node Approach II

Walkers are trapped into the nodal pockets of the trial wavefunction (the zeroes of ψ_T act as infinitely repulsive barriers)

Mathematically: we are extracting the lowest antisymmetric eigenvector of H with the additional constraints that $\phi_0 = 0$ whenever $\psi_T = 0$

$\Rightarrow$ fixed-node approximation

the algorithm is stable but biased

the error is variational: the FN energy is an upper bound of the exact one.

"Nodal-Release" Approaches

$$p(x \rightarrow y) = (\frac{1}{2\pi\tau})^d e^{-\frac{(y-x-b(x)\tau)^2}{2\tau}}$$

$b(x) = \nabla\psi_T/\psi_T$ drift vector

ψ_T symmetric nodeless vector

the antisymmetry is introduced into averages

$\Rightarrow$ no bias but unstable: explosion of the noise with respect to the signal.

Some mathematical questions

Mostly all aspects of QMC deserve some serious mathematical analysis

- Many different variants of QMC.
  What is the "best" algorithm?
  For a given algorithm, what is the best numerical implementation?
- Fixed-node approach.
  Validity of the fixed-node approach?
  Optimal numerical implementation?
  Study of the properties of the nodal hypersurfaces?
- Beyond the fixed-node approximation
  Is it possible to define a stable and unbiased algorithm?
- Improved estimators
- ....

Best QMC algorithm?

Many different variants of QMC depending on the way of treating weights.

- “Pure” Feynman Kac approaches: the elementary weights are factorized along the trajectories, the total weights become either zero or infinite in the large-n limit, (Pure Diffusion Monte Carlo, method of P. Del Moral and M. Rousset)
- Weights can be treated via a birth-death process at each step (branching process). Walkers are (locally) multiplied or killed according to the elementary weight. Fluctuating population of walkers: Some population control step is necessary. (Diffusion Monte Carlo, the most standard scheme)

Best QMC algorithm II

- It is also possible to work at constant number of walkers. Use of a so-called reconfiguration step (Hetherington, Sorella et al., Assaraf et al.)
- A mixture of the various alternatives is possible...

What is the best strategy?

Numerical implementation

Control of the various errors:

- finite time-step
- finite number of walkers
- population control error
- finite simulation time
- finite projection time “n”

How to propagate the walkers in the vicinity of the nodes (infinite drift)?

etc.

Cermics group (ENPC, Paris): E. Cancès, M. El Makrini, T. Lelièvre, B. Jourdain, A. Scemama

Fixed-Node Approach

Validity of the FN approach:

Does a FN-QMC simulation give the miminum of the energy $< \Psi|H|\Psi >$ over the set of the antisymmetric functions Ψ vanishing at the nodes of ψ_T?

Up to now considered as correct by physicists/chemists.

Mathematical proof recently presented by: E.Cancès, B. Jourdain, and T. Lelièvre “Quantum Monte Carlo simulations of fermions. A mathematical analysis of the fixed-node approximation, Preprint Cermics 2004-270.

Nodal properties

Very little is known about the properties of nodal hypersurfaces of fermionic wavefunctions.

- Tiling property (Ceperley, J.Stat. Phys 63,1237 (1991)) : By applying all possible permutations to an arbitrary nodal cell of a ground-state wavefunction, one covers the complete configuration space.

  In practice, domain of validity ?
- Quite remarkably nodes appear to have a much higher symmetry than the exact wavefunction.

Nodal properties

Examples: He, Li+,etc. ^3S state: $\vec{r_1} = \vec{r_2}$: “exchange” nodes

$|\vec{r_1}| = |\vec{r_2}|$ exact nodes : independent of r_{12} and Z

For the lithium ground-state (3 electrons system), with great numerical accuracy: $|\vec{r_1}| = |\vec{r_2}|$ (two elecrons of same spin) seems also to be the exact nodes (Bressanini, Ceperley, Reynolds (2000).

More generally, there is some hope that knowledge about nodes is less difficult to obtain than knowledge about the wavefunction (Bressanini, Reynolds (2005))

Beyond Fixed-Node Approximation

Fixed-Node Approximation is very accurate but still not sufficient!

Example: F_2 molecule at equilibrium geometry (coll. B.Braida, Paris)

Binding energy calculated within the fixed-node approximation:

D_e = 24(1) kcal with SCF nodes

D_e = 34(1) kcal with “Breathing Orbitals Valence Bond” nodes

D_e(Exp.)= 36.9 Kcal

Dissociation barrier of O_4

Work in collaboration with A. Scemama (Paris) and A.Ramirez-Solis (Mexico)

2 $O_2 \rightarrow O_4$ = 114. +/- 1.9 kcal

$O_4 \rightarrow O_4$TS = 12.7 +/- 3.7 kcal with MCSCF nodes

$O_4 \rightarrow O_4$TS = 26.2 +/- 2.9 kcal with SCF nodes

Expt. Barrier (?) $>$ 10 kcal

Fermion Monte Carlo (Kalos et al., 2000)

A recent proposal presented as an exact and stable QMC approach for fermions (FMC algorithm)

Ingredients:

- Two types of walkers, positive and negative. Two positive guiding wavefunctions $\Psi_G^+(R)$ and $\Psi_G^-(R)$ such that $\Psi_G^+(PR)$=$\Psi_G^-(R)$, P permutation of electrons
- Correlation between walkers: they are correlated such that a pair of walkers (+,-) meet with nonzero probability in a given time interval
- Walkers are cancelled when they meet.

Fermion Monte Carlo II

We have shown (Assaraf et al., 2005) that:

- the method is exact (unbiased) for an infinite number of walkers

- stable for a finite number M of walkers

However:

- the stability at fixed number of walkers is obtained at the price of introducing a bias $\epsilon(M)$

Fermion Monte Carlo III

We have shown that $\boxed{\epsilon(M) = 1/M^{\gamma}}$

with $\gamma = \Delta_{Fermi}/(\Delta_{Bose-Fermi} + \Delta_{Fermi})$

$$\Delta_{Fermi} \equiv E_1^{Fermi} - E_0^{Fermi}$$

$$\Delta_{Bose-Fermi} \equiv E_0^{Fermi} - E_0^{Bose}$$

The key idea of FMC, the cancellation process, leads only to an effective decrease of $\Delta_{Bose-Fermi}$.

The sign is solved only when $\Delta_{Bose-Fermi}$=0, which is not the case, the sign problem is still there.

Preprint: R. Assaraf, M.C.,and A. Khelif "The Fermion Monte Carlo revisited"