Quantum Monte Carlo methods

Lubos Mitas
North Carolina State University

Urbana, August 2006
The Theory of Everything

\[ i \hbar \frac{\partial \Psi}{\partial t} = \mathcal{H} \Psi \]

\[ \mathcal{H} = - \sum_{j} \frac{k^2}{2m} \nabla_j^2 - \sum_{\alpha} \frac{k^2}{2m_{\alpha}} \nabla_{\alpha}^2 - \sum_{j} \sum_{\alpha} \frac{Ze_{\alpha}}{|r_j - r_{\alpha}|} + \sum_{j<k} \frac{e^2}{|r_j - r_k|} + \sum_{\alpha<\beta} \frac{Z_{\alpha}Z_{\beta}e^2}{|R_{\alpha} - R_{\beta}|} \]

* Air  * Steel  * Paper  * Vitamins
* Water  * Plastic  * Dynamite  * Ham Sandwiches
* Fire  * Glass  * Antifreeze  * Ebola Virus
* Rocks  * Wood  * Glue  * Economists
* Cement  * Asphalt  * Dyes  * ...
Electronic structure and properties of materials

Hamiltonian of interacting electrons and ions

\[ H = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{i,I} \frac{Z_I}{r_{II}} + \sum_{i<j} \frac{1}{r_{ij}} + E_{\text{ion-ion}} \]

\[ H \psi(r_1, r_2, ...) = E \psi(r_1, r_2, ...) \]

- ground states
- excited states, optical properties
- responses to external fields
- T>0, etc ...
Density Functional Theory:
- based on one-particle density
  \[ E_{tot} = \int F[\rho(r)] \, dr \]
- exact functional \( F \) is unknown
- various approximations for \( F \):
  LDA (local density approx.)
  GGA (generalized grad. approx.)

Problem: efficient but inaccurate
(need accuracy 0.1 eV or higher)

Hartree-Fock:
- wavefunction as Slater determinant (antisymmetry)
  of one-particle orbitals
  \[ \psi_{HF}(r_1, r_2, \ldots) = \text{Det}\left[\left\{ \phi_i(r_j) \right\}\right] \]

Post-Hartree-Fock: expansion
in excitations
  \[ \psi_{corr}(r_1, r_2, \ldots) = \sum_n d_n \text{Det}_n\left[\left\{ \phi_i(r_j) \right\}\right] \]

More advanced: CC, MBPT etc

Problem: accurate but inefficient

Alternatives? Quantum Monte Carlo (QMC)
Accuracies which we need

- cohesion, optical excitations, barriers : $\sim 0.1 - 0.01\text{eV} \sim 1000 - 100\text{K}$

- magnetism, superconductivity, spintronics : $\sim 0.001\text{eV} \sim 10\text{K}$

- QED (important) : $\sim 0.000001\text{eV}$

- recent calculations of sixth order QED corrections for He atom:
  12 digit accuracy

Nature employs energy, length, etc scales as a composer employs various orchestral instruments
Density Functional Theory:
- based on one-particle density
\[ E_{tot} = \int F[\rho(r)] \, dr \]
- exact functional \( F \) is unknown
- various approximations for \( F \):
  - LDA (local density approx.)
  - GGA (generalized grad. approx.)

Problem: efficient but inaccurate (need accuracy 0.1 eV or higher)

Hartree-Fock: wavefunction as Slater determinant (antisymmetry)
of one-particle orbitals
\[ \psi_{HF}(r_1, r_2, \ldots) = \text{Det} \left[ \{ \phi_i(r_j) \} \right] \]

Post-Hartree-Fock: expansion in excitations
\[ \psi_{corr}(r_1, r_2, \ldots) = \sum_n d_n \text{Det}_n \left[ \{ \phi_i(r_j) \} \right] \]

Problem: accurate but inefficient

More advanced: CC, MBPT etc

Alternatives? Quantum Monte Carlo (QMC)

Lubos_Mitas@ncsu.edu
Traditional electronic structure methods and quantum Monte Carlo

Post-HF methods:
- convergence in one-particle basis sets is slow, inefficient
description of many-body effects
- need to explicitly evaluate integrals restricts functional forms
  which can be used

DFT approaches:
- difficult systematical improvement (the fundamental proof
  is not constructive)

QMC:
- use stochastic methods to map the many-body problem onto
  a sampling/simulation problem
- focus on many-body effects and efficiency of their description
- many ideas applicable to other systems/models: lattices, etc.
- in many ways, complimentary to traditional approaches
Variational Monte Carlo (VMC): Slater-Jastrow wavefunction

"Zero order" wave function: Slater determinant(s)

$$\psi_0(R) = \sum_n d_n \text{Det}_n[\phi_\alpha] \text{Det}_n[\phi_\beta]$$

one-particle orbitals from - Hartree-Fock, MCSCF
- Density Functional Theory (DFT)
- ...

Correlate the Slater wave function explicitly: e-e, e-e-l terms, ...

$$\psi_T(R) = \sum_n d_n \text{Det}_n[\phi_\alpha] \text{Det}_n[\phi_\beta] \exp(U_{\text{corr}})$$

$$U_{\text{corr}} = \sum_{i,j} \sum_{l} \sum_{k,m} c_{kml} a_k(r_{ij}) b_l(r_{il}) b_k(r_{jl})$$

electrons ions correlation basis expansion

Other functional forms: pairing BCS wavefunctions, pfaffians, ...
Variational Monte Carlo (VMC): stochastic methods for multi-D integrals

- evaluate the expectation values by sampling $3N$ electron coordinates
- no other choice! For example, variational energy is given by

$$E_{VMC} = \frac{\int \psi_T^2 [H \psi_T / \psi_T] dR}{\int \psi^2 dR} = \lim_{M \to \infty} \frac{1}{M} \sum_{sample} \frac{H \psi(R_{sample})}{\psi(R_{sample})}$$

where samples are distributed as $\psi_T^2(R)$ and $R = (r_1, r_2, ..., r_N)$

- optimize any variational parameters $\{c_{klm}\}$, eg, minimize the energy or local energy variance (Umrigar et al,'88)

$$\sigma_{VMC}^2 \approx \frac{1}{M} \sum_{sample} \left[ \frac{H \psi(R_{sample})}{\psi(R_{sample})} - E_{VMC} \right]^2 = \text{min}$$
Beyond VMC

**VMC** is straightforward, fast, efficient, but not accurate enough: **variational bias**

**Eliminate the VMC bias by** projecting out the (ground) state in imaginary time

\[
\psi(R, t) = \exp(-tH)\psi_T(R)
\]

\[
-\partial_t \psi(R, t) = H \psi(R, t)
\]

\[
\psi(R, t+\tau) = \int G(R, R', \tau) \psi(R', t) dR'
\]

Wave function can be sampled and the equation solved by interpreting the Green's functions as a transition probability density: simulate by a stochastic process
Toy model: 1D harmonic oscillator

\[ H = T + V(x) \]

**Propagator**

\[ G(x, x', \tau) \]

\[ Ce^{-\frac{(x-x')^2}{4\tau}} \cdot e^{-\left(V(x) - E_T\right)\tau} \]

**diffusion**  
**renorm**

\[ V(x) = x^2 \]

\[ \psi_{\text{init}}(x) \]

\[ \psi_{\text{ground}}(x) \]

Lubos_Mitas@ncsu.edu
Fermion sign problem

Naïve approach for fermionic wave functions: decompose to + and -

\[ \psi_T(R) = \psi_T^+(R) - \psi_T^-(R) \]

\[ -\partial_t \psi^+(R, t) = H \psi^+(R, t) \]

\[ -\partial_t \psi^-(R, t) = H \psi^-(R, t) \]

Unfortunately, + and - components converge independently to the lowest energy solution (which is bosonic) because Schr. eq. is linear!

\[ \lim_{t \to \infty} \psi^+(R, t) - \lim_{t \to \infty} \psi^-(R, t) \propto \exp[-(E_{Fermi} - E_{Boson}) t] \]

Fermion "signal" decays exponentially quickly into a bosonic "noise"
Importance sampling and fixed-node diffusion Monte Carlo (FNDMC)

\[
f(R, t + \tau) = \int G^*(R, R', \tau) f(R', t) dR'
\]

\[
f(R, t) = \psi_T(R) \phi(R, t), \quad \psi_T = \psi_{HF} e^{U_{corr}} = \det\{\phi_\alpha\} \det\{\phi_\beta\} e^{U_{corr}}
\]

\[
f(R, t \to \infty) \propto \psi_T(R) \phi_{ground}(R) \quad G^*(R, R', \tau) = \frac{\langle R | \exp(-\tau H) | R' \rangle}{\psi_T(R') \psi_T^{-1}(R)}
\]

Fermion node: (3N-1)-dimen. hypersurface defined as \[\phi(r_1, r_2, ..., r_N) = 0\]

Fixed-node approximation: \[f(R, t) > 0\]

- antisymmetry (nonlocal) replaced by a boundary (local)
- exact node implies recovering exact energy (in polynomial time)

Accuracy quite high: energy differences within a few % of experiment
Fermion node toy model: excited state of harmonic oscillator

\[ H = T + V(x) \]

Propagator
\[ G(x, x', \tau) \]

\[ C e^{-\frac{(x-x')^2}{4\tau}} e^{-\frac{(V(x)-E_T)t}{\tau}} \]

+ boundary condition (evaluate trial function)
Fixed-node approximation

- bosonization of the fermionic problem

- important shift (and a leap of faith):
  - antisymmetry $\rightarrow$ boundary condition
    - (nonlocal) $\rightarrow$ (local)

- fermion node is (3N-1)-dim hyper-surface:
  - easy to enforce (check the sign of the determinant)
  - difficult to parametrize with arbitrary accuracy

Green surface: 3D cut of 59-dimensional fermion node hypersurface
The key paradigm shift in QMC

- the many-body (Schrodinger) differential equation mapped onto equivalent stochastic process which is easy to simulate

density of walkers in 3N-dim space <=> value of the wavefunction
The key directions of (my) QMC research

- applications of QMC to challenging materials and to materials design: strongly correlated systems, competing many-body effects, new materials

- development of fundamental understanding of quantum many-body systems and methods capable of solving such challenging problems: fermion nodes, treatment of magnetic effects, etc

- coupling of QMC with other approaches for multi-scale application: coupling of QMC with molecular dynamics, T>0, etc
Example of application: which is the lowest energy isomer???

C20 challenge: ring                                bowl                                  cage

QMC was the first method to predict the energy order of isomers

Calculate correlation energy: from a single atom to molecule to clusters all the way to the solid
Performance of the fixed-node DMC method using a few-determinant trial functions

\[ \text{E}_{\text{tot}} = \text{E}_{\text{HF}} + \text{E}_{\text{corr}} \]

~3% \( \text{E}_{\text{tot}} \) -> \( \text{E}_{\text{corr}} \) -> Fixed-node error ~5% \( \text{E}_{\text{corr}} \)

Fixed-node DMC gets about 95% of \( \text{E}_{\text{corr}} \)

Cohesion, excitations, barriers, ...
Solving the stationary Schrödinger equation: key ingredients

The many-body problem splits into two parts

- **fermion nodes**, ie, boundary conditions
  - often one or few determinants are very reasonable
  - "non-dynamical" correlation, multi-reference wavefunctions
    - can be obtained within mean-field/basis set methods

- **"bosonic" correlations** within the fermion node domains
  - e-e cusps, multiple collision points, etc
  - "dynamical" correlation, very difficult for other methods
    - easy within QMC

The combined strategy “corners” the error in the treatment
Summary on application and performance of QMC methods

- fixed-node DMC typically recovers about 95% of the valence correlation energy for up to hundreds of electrons

- energy differences agree with experiments within a few percent

- method scales like a $N^3$ where N is the number of valence electrons (core electrons eliminated by pseudopotentials)

- applied to a number of systems, eg, electron gas and quantum liquids, atoms, molecules, solids etc; often the results became benchmarks for other methods

- about two orders of magnitude slower than mean-field methods but very efficient (perfectly scalable) on parallel architectures
Computational advantages

- scaling in the number of particles: $N^a$, $a = 1 \sim 3$

- sampling walkers independent -> natural parallelism and scalability, slow networks OK, heterogeneous clusters, ...

- cycle intensive, less memory intensive

- robust

- enable to focus in interesting physics
Impact of QMC methods

Breakthroughs and benchmark calculations:

- homogeneous electron gas in '80 (Ceperley & Alder)
- quantum liquids and solids
- barrier of H+H2 -> H2+H with 0.001 eV accuracy (J. B. Anderson)
- calculations of solids, clusters, etc (last 15 years)

Relative accuracy for s-p systems
- energy differences typically within 1-3% of experiment

Sizes of systems and timing
- 100-200 valence electron systems becoming routine;
  (1000 or more doable at the current level of development)
- typical run: 32 processors for a day
- about 100 times slower than mean-field approaches
Current limitations of QMC methods

Fundamentally, fixed-node approximation is still present:
- sometimes the physical effect of interest is in the last few percent of the correlation

Quantities beyond energies:
- we need better methods for getting correlation functions, forces, responses to external fields, etc
- small energy scales (ie, magnetism) is buried in the error bars

Need inputs (orbitals) from other approaches:
- often that is the most time consuming part :-(
- many of currently available DFT/HF codes are not built for the purpose to provide the best possible inputs to QMC
One- and two-particle quantities vs many-particle wave functions

Traditional approaches are based on one- or two-particle quantities:
- one-particle density (matrix), orbitals
- one-particle Green's function
- pair-correlation functions, etc
since these contain just the right amount of information necessary to describe the physics.

We think and "understand" in terms of reduced quantities.

QMC, however, is based on wave functions: ... despite that
- in general, wave functions are intractable and contain too much of information most of which is irrelevant
- limited to small number of particles, otherwise intractable
- difficult to understand the physics

QMC shows that maybe the traditional reductionism might have reached its limits ...
QMC and wavefunctions: only an accurate method or a paradigm shift?

- in many cases it is more efficient to carry out the many-particle calculations rather than to (re)build a better mean-field

- let the machine to worry about reducing the unnecessary information

- combination of analytical insights and stochastic techniques a key for getting high accuracy solutions of Schr. eq.

- in working with wave functions one is closest to the many-body physics and understanding (perhaps even the most efficient way how to capture the correlation effects)

- history of high accuracy benchmark/reference calculations