united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

SMR.1348 - 12

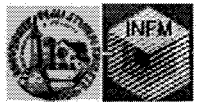
SECOND EUROPEAN SUMMER SCHOOL on
MICROSCOPIC QUANTUM MANY-BODY THEORIES
and their APPLICATIONS

(3 - 14 September 2001)

QUANTUM SIMULATIONS
A partial overview of Quantum Monte Carlo methods and applications on the continuum

Part I

Gaetano SENATORE
Dipartimento di Fisica Teorica
Universita' di Trieste e INFM
Strada Costiera 11
Trieste, I-34014
ITALY



Quantum Simulations

A partial overview of Quantum Monte Carlo methods and applications on the continuum

Gaetano Senatore*

senatore@ts.infn.it

Università di Trieste and INFM, Italy

The consequences of the physical law

Laplace, 1820

An intelligent being who, at a given moment, knows all the forces that cause nature to move and the positions of the objects that it is made from, if also it is powerful enough to analyze this data, would have described in the same formula the movements of the largest bodies of the universe and those of the lightest atoms. Although scientific research steadily approaches the abilities of this intelligent being, complete prediction will always remain infinitely far away.

Dirac, 1929

The general theory of quantum mechanics is now almost complete. The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

Many–body and Computational Physics

- The equations that Dirac alluded to are complicated because of the many-body nature of the systems that they describe.
- One may try with approximate analytic (many-body) theories.
- Or one may turn to computers (Metropolis *et al.*, 1953; Alder and Wainright, 1957; . . .).
- Computational physics has turned into a third way of doing physics, together with theory and experiments:
  - It is often able to give essentially exact predictions where experiments are not possible, difficult, expensive.
  - It allows for a check of approximate analytic theories.
  - It offers, when feasible, a direct approach to problems.

Plan of the lectures

- Lecture 1: Generalities. Multidimensional integrals and the Monte Carlo method. Random numbers and generators. Markov chains and random walks. Estimate of the errors.
- Lecture 2: Variational Monte Carlo. Optimization of the wavefunction. Correlated sampling ad reweighting. Variance minimization.
- Lecture 3: Diffusion Monte Carlo. Imaginary time evolution. Small time Green function. Importance sampling. Fermions, nodes, and the sign problem. Other ground state methods properties. Path Integral Monte Carlo and finite temperature properties.
- Lecture 4: Selected applications – I.
- Lecture 5: Selected applications – II.

A partial list of references

- David Ceperley lecture notes on Quantum Monte Carlo (QMC): http://archive.ncsa.uiuc.edu/Apps/CMP/papers/cep96b/lnotes.ps
- A recent school on QMC: *NATO Advanced Study Institute: Quantum Monte Carlo Methods in Physics and Chemistry*, ed. P. Nightingale and C. Umrigar (Kluwer, Dordrecht, 1999).
- A book on QMC: *Monte Carlo Methods in Ab Initio Quantum Chemistry*, by B.L Hammond, W.A. Lester, Jr., and P.J. Reynolds (World Scientific, Singapore, 1994).
- A book and a long article on Monte Carlo methods: *Monte Carlo Methods Volume I*, M.H. Kalos and P.A. Whitlock (Wiley, New York, 1981).
  Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms, Alan D. Sokal (Cargèse Summer School on " Functional Integration: Basics and Applications, 1996; also Cours de Troisième Cycle de la Physique en Suisse Romand, Lausanne, Switzerland, 1989)

Notation

- The N-body Hamiltonian (for a one–component system!) is

$$H = -D\nabla^2 + V(R),$$

with

$$D = \frac{\hbar^2}{2m}, \quad R = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N),$$

and

$$\nabla \equiv \nabla_R = (\nabla_1, \nabla_2, \ldots, \nabla_N).$$

- R is a d·N–dimensional vector
- $\mathcal{O}$ is a generic Hermitian operator (I, H, $\hat{n}(\mathbf{r})$, $\hat{n}(\mathbf{r}, \mathbf{r}')$, $\ldots$)
- τ is the imaginary time ($\tau = it/\hbar$).

Quantum Averages I

Quantum averages involve multidimensional integrations (here $\int$) :

- (Ground) state average

$$\langle \mathcal{O} \rangle_\phi = \frac{\langle \phi | \mathcal{O} | \phi \rangle}{\langle \phi | \phi \rangle} = \frac{\int dR\, \phi(R)^* \mathcal{O} \phi(R)}{\int dR |\phi(R)|^2}$$

$$= \int dR \left[\frac{|\phi(\mathbf{R})|^2}{\langle \phi | \phi \rangle} \right] \left[\frac{\mathcal{O}\phi(\mathbf{R})}{\phi(\mathbf{R})} \right] \equiv \int dR\, \pi(\mathbf{R}) \mathcal{O}_\mathbf{L}(\mathbf{R}).$$

- The method of choice is Monte Carlo (MC) Integration.

$$\int dR\, \pi(\mathbf{R}) \mathcal{O}_L(R) \simeq \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}_L(R_i) \quad \left[\pm \frac{\mathbf{cost}}{\sqrt{\mathbf{M}}} \right], \quad M \text{ large.}$$

if the *walkers* $\{R_i\}$ are distributed with the probability $\pi(R)$.

Quantum Averages II

Temperature averages also involve multidimensional integrations:

- For classical systems

$$\langle \mathcal{O} \rangle_\beta = \frac{\int dR e^{-\beta V(R)} \mathcal{O}(R)}{Z}, \quad Z = \int dR e^{-\beta V(R)}.$$

- For quantum systems

$$\langle \mathcal{O} \rangle_\beta = \frac{\int dR dR' \rho(R, R', \beta) \langle R|\mathcal{O}|R'\rangle}{Z}, \quad Z = \int dR \rho(R, R, \beta),$$

and

$$\rho(R, R', \beta) = \langle R|e^{-\beta H}|R'\rangle$$

is the temperature density matrix.

- Above, Z is the canonical partition function.

Multidimensional Integration

- Quadrature (Simpson–like) schemes are unfeasible!
  - A regular grid with 10 mesh points per axis would require $10^{d \cdot N}$ evaluations of the integrand, for N particles in d dimensions, i.e., 10^{30} operations for 10 particles in 3 dimensions!
  - A simple operation takes say about 10^{-9}s on a present computer.
  - A year is about 3×10^{7}s.
  - Integration by quadrature even for 10 particles would take too many years! About 3×10^{13} years!

Error scaling for quadrature

- If we integrate over a hypercube of side L, with a mesh of size h, the number of grid points is $M = (L/h)^{d \cdot N}$, i.e. $h \propto M^{-1/(d \cdot N)}$.
- Assume that the error $\propto h^l$. Hence

$$\text{error} \propto 1/M^{l/(d \cdot N)}$$

Since l is of order unity, the error decays exceedingly slowly with M. In fact, the larger is $d \cdot N$ the slower decays the error.

- For $N = 20$, $d = 2$, $l = 4$ (Simpson rule), halving the error of an evaluation with M points requires going to

$$2^{d \cdot N/l} \cdot M = 2^{2 \cdot 20/4} \cdot M = 1024 \cdot M$$

points; to reduce it by a factor 4 requires $10^6 \cdot M$ points, and so on!

Monte Carlo Integration

- Monte Carlo Integration is the only choice:

$$\int dR\,\pi(R)\mathcal{O}(R) \simeq \frac{1}{M}\sum_{i=1}^{M}\mathcal{O}(R_i), \quad M \text{ large}$$

with an

$$\text{error} \propto 1/\sqrt{M},$$

provided that the configurations or walkers $\{R_i\}$ are distributed with the probability $\pi(R)$.

- To halve the error only $4 \cdot M$ points are required; $16 \cdot M$ points are sufficient to reduce the error by a factor 4; and so on. Also, there is no dependence on the dimensionality of the configuration space.

Generating the random walk

Configurations distributed with a given probability can be generated with a variety of algorithms:

- (generalized) Metropolis algorithm,
- Molecular Dynamics,
- Langevin Dynamics,
- combination of the above,
- other.

In the following we shall restrict to the (generalized) Metropolis algorithm. Metropolis method generates configurations resorting to random walks, for which random numbers need to be generated.

Random numbers

- Random sequence of numbers drawn from an assigned probability density, say $u(x)$:

  $u(x)dx =$ probability that x falls between x and $x+dx$, $\int u(x)dx = 1$

- Uniform variates:

$$\begin{aligned} u(x) &= 1/(\beta-\alpha), \quad \alpha < x < \beta, \quad \text{or} \\ u(x) &= 1, \quad 0 < x < 1. \end{aligned}$$

- Other variates

$$\begin{aligned} u(y) &= e^{-y}, \quad 0 < y < \infty, \\ u(y) &= e^{-y^2}/\sqrt{\pi}, \quad -\infty < y < \infty. \end{aligned}$$

Generation of pseudo numbers

- Let's concentrate on the uniform distribution ($0 < x < 1$):
  - a proper generator will produce values of x placed at random in the given interval;
  - for large generation numbers the x values will be uniformly distributed in $0 < x < 1$;
  - for large generation numbers both the average and variance calculated on the generated values will reproduce those of the assigned uniform distribution.
- In practice pseudo random numbers are generate on computers with deterministic rules [sequences are perfectly reproducible!].
- In the following we shall assume that a good generator of uniform variates is provided, disregarding the issue of how to device it.

Random variates: the inversion method

How we can generate non-uniform variates from uniform ones, $w(y)$ from $u(x) = 1,\ 0 < x < 1$? Let's consider the inversion method.

- Look for $y = f(x)$ such that if x's are distributed according to $u(x)$, then $y = f(x)$ are distributed according to $w(y)$.
- Start from

$$u(x)dx = w(y)dy = u(f^{-1}(y))df^{-1}(y),$$

and use $u(f^{-1}(y)) = 1$ to get

$$f^{-1}(y) = \int_{-\infty}^{y} dy' w(y') = W(y).$$

If $W(y)$ and its inverse are known, $y = f(x) = W^{-1}(x)$

Random variates: the rejection method

When the inverse of $W(y) = \int_{-\infty}^{y} dy' w(y')$ is not known [$w(y)$ non-uniform], one may resort to the rejection methods

- Look for $f(y) \geq w(y)$. Here we choose $f(y) = w_{max}$.
- Generate a uniform random number y
- Generate a second uniform random number ξ

(1) $w(y)/w_{max} \geq \xi,$ accept y

(2) $w(y)/w_{max} < \xi,$ reject y

It easily seen that y is distributed according to $w(y)$.

- Note: the normalization of $w(y)$ is not necessary!
- Metropolis method is a particular rejection method.

Random Walks (Markov Chains)

A Markov chain is fully specified by the initial distribution and by the transition probability $p(s, s')$. Markov chains provide a convenient way to sample multidimensional probability distributions.

The state (or configuration) s of the system is changed randomly according to the transition probability $p(s, s') = p(s \to s')$ satisfying

$$\sum_{s'} p(s, s') = 1 \quad \text{and} \quad p(s, s') \geq 0,$$

thus generating a random walk (or sample) $(s_0, s_1, s_2, \ldots)$.

If $p(s, s')$ is ergodic there exists a (unique) probability measure $\pi(s)$ satisfying at equilibrium the stationarity condition:

$$\sum_{s} \pi(s) p(s, s') = \pi(s').$$

Moreover if $p^n(s, s')$ is the probability to reach s' from s in n steps then

$$\lim_{n \to \infty} p^n(s, s') = \pi(s') :$$

the random walk converges to the equilibrium distribution irrespective of the initial distribution.

The transition $p(s, s')$ is ergodic if the following conditions are verified:

- Irreducibility: for each (s, s') there exists an $n \geq 0$ such that $p^n(s, s') > 0$;
- Aperiodicity: $p(s, s) > 0$;
- The average return time is finite: it exists $N_{s,s'} < \infty$ such that, for $n > N_{s,s'}$, $p^n(s, s') > 0$.

Detailed balance

A sufficient condition to obtain $\pi(s)$ as stationary distribution is to chose the transition probability to satisfy

$$\pi(s)p(s,s') = \pi(s')p(s',s).$$

In fact summing the above over s one gets

$$\sum_s \pi(s)p(s,s') = \pi(s') \sum_s p(s',s) = \pi(s').$$

Generalized Metropolis Algorithm

The transition probability may be conveniently decomposed into the product of an irreducible proposal or sampling matrix $T(s, s')$ and an acceptance matrix $A(s, s')$

$$p(s, s') = T(s, s')A(s, s').$$

Imposing the detailed balance yields

$$\frac{A(s, s')}{A(s', s)} = \frac{\pi(s')T(s', s)}{\pi(s)T(s, s')} \equiv q(s, s'),$$

which can be satisfied quite generally by choosing

$$A(s, s') = F[q(s, s')],$$

where the function $F : [0, \infty] \rightarrow [0, 1]$ satisfies

$$\frac{F[z]}{F[1/z]} = z, \quad \text{for all z.}$$

- Metropolis choice:

$$F[z] = min[1, z]$$

- An alternative choice could be:

$$F[z] = \frac{z}{1+z}$$

Monte Carlo Estimates and Averages

One would like to evaluate the *true* mean

$$\langle \mathcal{O} \rangle = \int ds\, \pi(s) \mathcal{O}(s),$$

whereas MC yield a sample $(s_1, s_2, ..., s_M)$ of length $\simeq M$ of states distributed according to $\pi(s)$. Evidently, one can define a sample mean

$$\overline{\mathcal{O}} = \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}_i,$$

with $\mathcal{O}_i \equiv \mathcal{O}(s_i)$.

The sample mean is an unbiased estimator of the true mean, i.e., $\langle \overline{\mathcal{O}} \rangle = \langle \mathcal{O} \rangle$ independently of M.Also, it is possible to prove:

- the law of large numbers, $\lim_{M\to\infty} \overline{\mathcal{O}} = \langle \mathcal{O} \rangle$;
- The central limit theorem, which states that $\overline{\mathcal{O}}$ is *normally* distributed around $\langle \mathcal{O} \rangle$.

Therefore we need to evaluate the variance

$$\sigma^2(\overline{\mathcal{O}}) = \langle (\overline{\mathcal{O}} - \langle \mathcal{O} \rangle)^2 \rangle,$$

whose root we may interpret as statistical error on $\overline{\mathcal{O}}$.

The statistical error

Using $\overline{\mathcal{O}} = (1/M)\sum_{i=1}^{M}\mathcal{O}_i$, one obtains for the variance

$$\sigma^2(\overline{\mathcal{O}}) = \frac{1}{M}\sum_{t=-(M-1)}^{t=M-1}\left(1-\frac{|t|}{M}\right)C(t) \approx \frac{\tau}{M}C(0) = \frac{\tau}{M}\sigma^2(\mathcal{O}).$$

Here

$$C(t) = \langle \mathcal{O}_s \mathcal{O}_{s+t}\rangle - \langle \mathcal{O}\rangle^2$$

is the normalized *time* autocorrelation function, which evidently reduces to the variance of $\mathcal{O}$ at time 0, $C(0) = \sigma^2(\mathcal{O})$, and the integrated correlation time

$$\tau = 1 + 2\sum_{t=1}^{\infty}\frac{C(t)}{C(0)},$$

accounts for the correlation existing between walkers in the Markov chain.

In general $\tau > 1$.

A sample estimate of $C(t)$, with a bias of order $1/M$ is given by

$$\tilde{C}(t) = \frac{1}{M - |t|} \sum_{i=1}^{M-|t|} (\mathcal{O}_i - \overline{\mathcal{O}})(\mathcal{O}_{i+|t|} - \overline{\mathcal{O}}).$$

Thus one has an estimate for $\sigma^2(\mathcal{O}) = C(0) \approx \tilde{C}(0)$,

$$\tilde{\sigma}^2(\mathcal{O}) = \frac{1}{M} \sum_{i=1}^{M} (\mathcal{O}_i - \overline{\mathcal{O}})^2,$$

and the correlation time can also be calculated from $\tilde{C}(t)$.

Blocking and estimate of errors

The precise estimate of the error bar requires the calculation of time correlation functions, which one would rather avoid.

An alternative is provided by the blocking procedure. The sample is broken in a number of blocks $M = N_b n_b$, with N_B the number of blocks and and n_b the length of each block. New variable are constructed as block averages

$$\mathcal{O}_{b,I} = \frac{1}{n_b} \sum_{i=1}^{n_b} \mathcal{O}_{(I-1)n_b+i},$$

and clearly have a mean equal to the run mean $\overline{\mathcal{O}}$. Intuitively, if $n_b \gg \tau$, this new variables should become statistically independent and therefore have a variance around their mean $\overline{\mathcal{O}}$ given by

$$\sigma^2(\mathcal{O}_b) = \frac{1}{N_b(N_b-1)} \sum_{I=1}^{N_b} (\mathcal{O}_{b,I} - \overline{\mathcal{O}})^2$$

One can indeed show that provided $n_b \gg \tau$ and yet $n_b \ll M$ or equivalently N_b large

$$\sigma^2(\mathcal{O}_b) \simeq \sigma^2(\overline{\mathcal{O}}).$$

A plot of $\sigma^2(\mathcal{O}_b)$ versus n_b will reveal a plateau, where in fact the above relation holds, and therefore it also yields an estimate of the correlation time.

Implementation of Metropolis algorithm

Given a probability $\pi(s)$ to sample (here, s is the state of the system and $\pi(s)$ may be not known in closed form, see, e.g., DMC, GFMC):

- Chose the proposal matrix $T(s, s')$;
- Initialize the system in the state s_0;
- To advance from s_n to s_{n+1}:
  - sample s' from $T(s_n, s')$,
  - calculate

$$q(s_n, s') = \frac{\pi(s')T(s', s_n)}{\pi(s_n)T(s_n, s')},$$

- generate a random number r_n and compare it with $q(s_n, s')$:
  - if $q(s_n, s') > r_n$: $s_{n+1} = s'$
  - else $s_{n+1} = s_n$.
- Throw away the first k *states* as being out of equilibrium;
- Collect averages using the configurations with $n > k$ and block them to calculate error bars (???).
- Example: T a constant in a cube, $\pi(s) \propto exp(-\beta V(s))$.

Some facts about Metropolis:

- The normalization of the probability, $\int ds\, \pi(s)$, is never needed and in fact cannot be calculated (... easily).
- Particles can be moved one at time (hard spheres!);
- For the generalized algorithm ($T(s, s')$ is not a constant) one has to sample both forward and reverse transition;
- An optimal acceptance is

$$\mathcal{A} = \frac{\text{moves accepted}}{\text{total moves}} \simeq 1/2.$$

In fact the overall efficiency may dictate different choices (see, e.g., DMC).

- The length of the necessary *thermalization* (deciding the number k of initial moves to discard) can be investigate monitoring cumulative averages of physically relevant quantities (energy, density profile, ...).

Plan of lecture III and IV

- Variational Monte Carlo
- Variance minimization.
- Optimization of the wavefunction
- Correlated sampling and reweighting
- Size extrapolation
- Diffusion Monte Carlo
- The algorithm
- Fermion and Fixed-node Approximation

Variational Monte Carlo

- If the wavefunction is given, MC provides a straightforward mean of evaluating state averages.
- Assume that we have a trial wavefunction $\Psi(R;\mathbf{a})$ depending on a set of parameters $\mathbf{a} = (a_1, a_2, \ldots, a_p)$. The variational theorem states that

$$E(\mathbf{a}) = \frac{\int dR\, \Psi(R;\mathbf{a})^* H \Psi(R;\mathbf{a})}{\int dR |\Psi(R;\mathbf{a})|^2} \geq E_0,$$

  with E_0 the exact ground state energy.
- We know how to generate configurations (say M) distributed with the probability $\pi(R;\mathbf{a}) \propto |\Psi(R,\mathbf{a})|^2$. Hence we can estimate $E(\mathbf{a})$,

$$E(\mathbf{a}) = \int dR \left[\frac{|\Psi(\mathbf{R};\mathbf{a})|^2}{\langle\Psi_\mathbf{a}|\Psi_\mathbf{a}\rangle}\right]\left[\frac{\mathbf{H}\Psi(\mathbf{R};\mathbf{a})}{\Psi(\mathbf{R};\mathbf{a})}\right]$$

$$\equiv \int dR\, \pi(\mathbf{R};,\mathbf{a})\mathbf{E_L}(\mathbf{R};\mathbf{a}) = \frac{1}{M}\sum_{i=1}^{M} E_L(R_i;\mathbf{a}).$$

We can optimize $\Psi(R;\mathbf{a})$ by minimizing $E(\mathbf{a})$, to obtain the best upper bound to E_0.

- Note: as $\Psi_\mathbf{a} \rightarrow \Phi_n$ (an exact eigenstate), $E_L(R);\mathbf{a})$ approaches a costant everywhere, $E_L(R);\mathbf{a}) \rightarrow E_n$.
- The above implies the zero variance property: as $\Psi_\mathbf{a}$ approaches an eigenstate the MC estimate of the energy converges more rapidly with the number M of steps in the random walk.

Variance

The optimization of $\Psi(R;\mathbf{a})$ can be also achieved with other techniques.

- Another quantity enjoying a minimum property is the variance

$$\sigma^2(\mathbf{a}) = \frac{\int dR\, \Psi(R;\mathbf{a}))^*(H - E_{ref})^2\Psi(R;\mathbf{a})}{\int dR|\Psi(R;\mathbf{a})|^2} \geq 0.$$

  The minimum value attainable by σ^2 is 0, which is achieved whenever $\Psi(R,\mathbf{a})$ coincides with an exact eigenstate of H, say Φ_n with eigenvalue E_n, and E_{ref} is set equal to the E_n.

- In principle one could judge on the quality of a minimization by looking at the size of σ^2. In practice, variance minimization has a number of bonuses:

- It can be used to study excited states (by a proper choice of the constant E_{ref}).
- Being a sum of squares, σ^2 can be efficiently minimized using efficient algorithms like that of Levenberg and Marquand.
- It requires a smaller number of configurations as compared with the energy minimization.
- The only way $\sigma^2(\mathbf{a})$ can be made small is by having $E_L(R;\mathbf{a})$ smooth and close to an eigenvalue, whereas the energy minimization can be biased by configurations with $E_L(R;\mathbf{a})$ to low.

Weinstein criterion

- The variance

$$\sigma_\psi^2 = \frac{\langle\psi|(H - E_\psi)^2|\psi\rangle}{\langle\psi|\psi\rangle}, \quad E_\psi = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle},$$

  is a measure of the distance of E_ψ from exact eigenvalues. In fact, σ_ψ only vanishes at an energy eigenstate.

- It is easily shown that in the range

$$[E_\psi - \sigma_\psi, E_\psi - \sigma_\psi]$$

  there is always at least one exact energy eigenvalue.

- There is always an energy eigenvalue whose distance from E_ψ is at most σ_ψ. Hence, σ_ψ can be employed to determine the (energy) accuracy of a minimization.

The Trial Wavefunction

A good trial wavefunction, apart from being flexible enough should satisfy a minimum number of basic requirements:

- Ψ and $\nabla\Psi$ should be continuous for finite potential $V(R)$.
- Not only $\int \Psi^2$ and $\int \Psi^* H \Psi$ should exists, but also $\int \Psi^* H^2 \Psi$, in order that the variance exists and statistical errors are finite.
- The wavefunction should embody all the know exact behaviors, so as to make the local energy $E_L(R)$ as smooth as possible.

Symmetry

Ψ should have the correct symmetry property, i.e., for any particle permutation P:

- for Bosons

$$\Psi(PR) = \Psi(R)$$

- and for Fermions

$$\Psi(PR, P\Sigma) = (-)^P \Psi(R, \Sigma).$$

Above

$$\Sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)$$

denotes the set of discrete spin projection variables.

Symmetry – Bosons

- For an uniform Bose fluid the simplest trial function is of the Bijl–Jastrow type

$$\Psi_2(R) = \exp[-\sum_{i<j} u(r_{ij})] \equiv J(R).$$

- More refined wavefunctions include three–body correlations to read

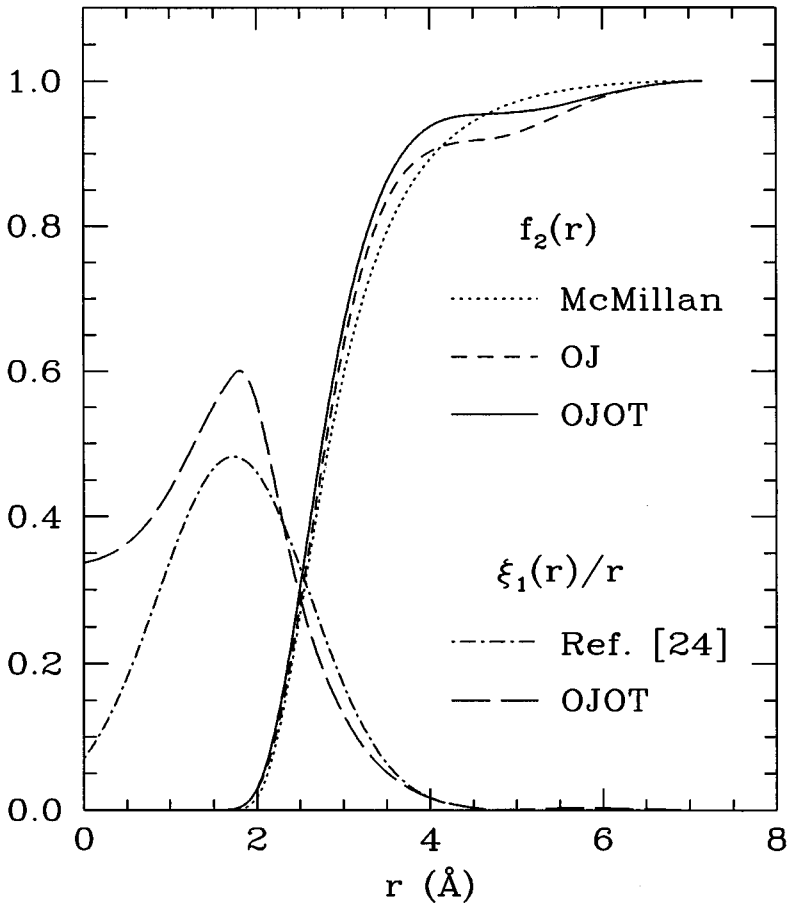
$$\Psi_3(R) = J(R) \cdot \exp[-\sum_{i<j<k} u_3(r_{ij}, r_{ik}, cos(\hat{r}_{ij} \cdot \hat{r}_{ik}))].$$

- The Feynman ansatz $u_3^F = \sum_{cyc} \xi_1(r_{ij})\xi_1(r_{ik})\hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{r}}_{ik}$ yields

$$\Psi_F(R) = J(R) \cdot \exp\left[-\sum_i \left[\sum_j \xi_1(r_{ij})\hat{\mathbf{r}}_{ij}\right]^2\right]$$

Pseudopotentials for He_4

- $f_2(r) = \exp[-u(r)]$
- Mc Millan: $u(r) \propto 1/r^5$.
- OJ: optimized $u(r)$.
- OT: optimized $u_3(r)$.



Variational calculations for He_4

Table 1: Energy E of He_4 with various wavefunctions

	E	T	$(E-E_0)/(2T)$
M	-5.702(5)	14.712(50)	5.1%
OJ	-6.001(16)	14.709(20)	4.1%
OJT^a	-6.862(16)		0.99%
OJOT	-6.901(4)	14.233(8)	0.86%
DMC	-7.143(4)	14.049(18)	0.0%

Symmetry – Fermions

- A typical Fermion wavefunction is obtained augmenting the symmetric Jastrow function by a Slater determinant ensuring antisymmetry,

$$\Psi(R,\Sigma) = J(R) \cdot det[\varphi_i(r_k,\sigma_k)],$$

with $\varphi_i(r_k,\sigma_k)$ the i-th orbital. One needs, for N particle N distinct spin orbitals to get a non-vanishing determinant.

- More refined wavefunctions are obtained
  - with resorting to triplet *pseudopotentials*: $\Psi(R,\Sigma) = \Psi_3(R) \cdot det[\varphi_i(r_k,\sigma_k)]$.
  - By including backflow in the the Slater determinant. [Particle coordinates are replaced by suitable collective coordinates].

Smoothness of $E_L(R)$ – small r_{ij}

Look at the dominant terms in $E_L(R)$ when two particles come close.

- For He atoms

$$E_L(R) = v(r) + 2D\nabla^2 u(r) - 2D(\nabla u(r))^2 + \cdots,$$

with $v(r)$ the pair interaction. So for LJ interaction one is led to McMillan $u(r) \propto 1/r^5, \quad r \to 0$. Here $r \equiv r_{ij}$.

- For electrons one would chose the $\varphi_i(r)$ as exact solution of the independent particle problem; then with an analysis similar to the one above one would get $u(r) = a_{\sigma,\sigma'} r, \quad \to 0$, with
  - $a_{\uparrow\downarrow} = -e^2/(4D)$, and
  - $a_{\downarrow\downarrow} = -e^2/(8D)$.

Smoothness of $E_L(R)$ – large r_{ij}

The study of $u(r)$ at large r is most easily accomplished by rewriting the variational energy in reciprocal space in terms of the collective coordinates

$$\rho_{\mathbf{k}} = \sum_i exp(i\mathbf{k} \cdot \mathbf{r}_i),$$

with

$$S(k) = (1/N)\langle \rho_{\mathbf{k}} \rho_{-\mathbf{k}} \rangle$$

the static structure factor.

Using the RPA approximation, one obtains for the energy

$$E_V = T_0 + \sum_{\mathbf{k}} \left(S(k) \cdot (Dk^2 u^2(k) + \frac{1}{2} v(k)) - \frac{v(k)}{2} \right)$$

and for the structure factor

$$S(k) = \frac{S_0(k)}{1 + 2u(k)S_0(k)}.$$

Variation with respect to $u(k)$ immediately yields

$$2u(k) = -\frac{1}{S_0(k)} + \sqrt{\frac{1}{S_0^2(k)} + \frac{2v(k)}{Dk^2}}.$$

- For charged Fermions in 3 dimensions this implies, $u(r) \propto 1/r, \quad \to \infty.$
- For short ranged potentials, like Helium, replacing $u(k)$ with a constant for small k one obtains $u(r) \propto 1/r^2, \quad \to \infty.$

Reweighting

Optimization techniques require taking derivatives of MC estimates with respect to the variational parameters.

In principle one would obtain a sample $\{R_i\}$ from $\pi(R;\mathbf{a})$ and a sample $\{R'_i\}$ from $\pi(R;\mathbf{a}')$ to calculate for example

$$E(\mathbf{a}) = \frac{1}{M}\sum_{i=1}^{M} E_L(R_i;\mathbf{a}),$$

and

$$E(\mathbf{a}') = \frac{1}{M}\sum_{i=1}^{M} E_L(R'_i;\mathbf{a}'),$$

and from these the derivative of $E(\mathbf{a})$. This procedure, however, turns out to be unstable due to the independent statistical errors on the two estimates of the energy.

If the two parameters sets are close enough to each other a winning strategy is to use the same sample, say $\{R_i\}$ for both evaluations. In other words,

$$E(\mathbf{a}') = \frac{1}{M} \frac{\sum_{i=1}^{M} E_L(R_i'; \mathbf{a}') w(R_i)}{\sum_i w(R_i)},$$

with

$$w(R) = |\Psi(R; \mathbf{a}')|^2 / |\Psi(R; \mathbf{a})|^2.$$

Ewald sums and N extrapolation

- Properties in the long-wavelength limit depend on long–range behaviour of the trial function as well as on the modeling of the system under study.
- Especially for Coulomb systems it is crucial to consider periodic replicas of the simulation cell, and to sum interactions with all the replicas (Ewald sums). It is also important to Ewald sum the pseudopotential.
- Even allowing for Ewald sums Coulomb systems have residual size effects due to the size of the cell (finite number of particles).
- Important size effects are present in Fermion systems in the Fermi liquid regime due to levels shell structure.

- To study systems in the thermodynamic limit, one can:
  - Study the system at various N and then try to extrapolate to $N = \infty$. This is usually the case with VMC. Some times This is also done for DMC simulations (see following lectures).
  - Assume that the N dependence does not depend much on the details of the simulation and borrow the number dependence of VMC. This is usually done with DMC and GFMC.

Some filtering (projection) techniques

Suitable evolution to evolve a initial guess or trial wavefunction $\Psi_T(R)$ into the sought ground state $\Phi_0(R)$. The evolution can be either in imaginary time (DMC, PIGS), or time integrated (GFMC).

- GFMC

$$\Psi_{n+1}(R) = (E_T + V_0) \int dR'\, g(R, R')\Psi_n(R'),$$

$$g(R, R') = \langle R| \frac{1}{H + V_0} |R'\rangle,$$

with $H + V_0$ positive definite and E_T close to the ground state energy E_0.

Expanding Ψ_T in eigenfunctions of H, $\Psi_0 = \Psi_T = \sum_i c_i \Phi_i$, one immediately gets

$$\Psi_n(R) = \sum_i \left[\frac{E_T + V_0}{E_i + V_0}\right]^n c_i \Phi_i(R),$$

which for large n implies

$$\Psi_n \propto \left[\frac{E_T + V_0}{E_o + V_0}\right]^n c_o \Phi_0,$$

provided that $c_0 \neq 0$.

Evidently, higher energy components, compared with the funtamental one, die out exponentially as

$$\left[\frac{E_0 + V_0}{E_i + V_0}\right]^n = exp\left[-n \log\left(\frac{E_i + V_0}{E_0 + V_0}\right)\right].$$

Knowledge (construction) of $g(R, R')$ is necessary.

- DMC: $|\Psi_n\rangle = e^{-\tau(H-E_T)}|\Psi_{n-1}\rangle$, $\Psi_0 = \Psi_T$, or

$$\Psi(R;(n+1)\tau) = \int dR'\, G(R,R';\tau)\Psi(R';n\tau),$$

$$G(R,R';\tau) = \langle R|\exp[-\tau(H-E_T)]|R'\rangle,$$

and E_T close to the ground state. Again for large n one projects out the ground state according to

$$\Psi_n \propto \exp(-n\tau(E_0 - E_T))c_0\Phi_0.$$

- Note that

$$g(R,R') = \int_0^\infty d\tau G(R,R';\tau)$$

with $E_T = -V_0$.

- Both DMC and GFMC involve $d \cdot N$ dimensional integrations, which require MC techniques and can be dealt with by resorting to random walks.
- One has to sample the appropriate Green's functions: this can be arranged with a small time expansion for DMC and other techniques for GFMC.
- Both DMC and GFMC are implemented as power methods, i.e., in an iterative manner.

- PIGS

$$\frac{\langle \Psi_T | e^{-\tau H/2} H e^{-\tau H/2} | \Psi_T \rangle}{\langle \Psi_T | e^{-\tau H} | \Psi_T \rangle}$$

It implies classical simulation of interacting polymers. It allows for pure stimates and is easily extended at finite temperature.

Diffusion Monte Carlo (DMC)

- From a trial wavefunction to the exact ground state: $\Psi_T \longrightarrow \Phi_0$,
- Imaginary time evolution ($t \longrightarrow -i\tau\hbar$, $D \equiv \hbar^2/2m$)

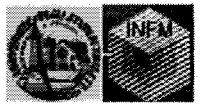
$$-\frac{\partial}{\partial\tau}\Phi(R,\tau) = (H - E_T)\Phi = -D\nabla^2\Phi + [V(R) - E_T]\Phi,$$

$$\Phi(R,0) = \Psi_T(R) = \sum_{n=0}^{\infty} c_n \Phi_n(R),$$

filters out high energy components:

$$\Phi(R,\tau) = \sum_{n=0}^{\infty} c_n e^{-(E_n - E_T)\tau}\Phi_n(R) \longrightarrow c_0 e^{-(E_0 - E_T)\tau}\Phi_0(R)$$

for large enough times τ.



- If $\Phi(R,\tau) \geq 0$, then it can be regarded as a probability and the equation above can be exactly integrated by numerical means.
- It is convenient to resort to Green function of the equation, $G(R,R';\tau) = \langle R|\exp[-\tau(H-E_T)]|R'\rangle$, and put it in integral form and use random walks.
- The equation contains:
  - a diffusion term, describing brownian motion, and
  - a rate term describing death and birth processes
- In practice the equation is not easy to deal with because usually the potential is not boundend: is infinite for certain configurations.
- The cure to the above problem is importance sampling.

Importance sampling

- A smooth evolution equation is obtained evolving the mixed distribution

$$f(R,\tau) = \Phi(R;\tau)\Psi_T(R),$$

with the initial condition $f(R;0) = \Psi_T^2(R)$. The resulting equation is:

$$-\frac{\partial f(R,t)}{\partial t} = -D\nabla^2 f + [E_L(R) - E_T]f + D\nabla\cdot[f\mathbf{F_Q}(\mathbf{R})],$$

$$\mathbf{F}_Q(R) = \nabla \log|\Psi_T(R)|^2.$$

$\mathbf{F}_Q(R)$ is called the quantum force.

- Note: here and in the following we assume real wavenfunctions, which is always possible for the ground state without magnetic field.

Small time Green Function

- The evolution of $f(R,\tau) = \Phi(R;\tau)\Psi_T(R)$, involves a modified Green's function,

$$G(R,R';\tau) \rightarrow K(R,R';\tau) = \Psi_T(R)G(R,R';\tau)\Psi_T^{-1}(R'),$$

- which, for small time $\Delta\tau$ reads

$$K(R',R;\Delta\tau) = K_1 \times K_2 + O(\Delta\tau^2),$$

$$K_1 = \frac{1}{(4\pi D\Delta\tau)^{\frac{Nd}{2}}} \times \exp\left(\frac{-[R'-R-D\Delta\tau\mathbf{F}_Q(R)]^2}{4D\Delta\tau}\right),$$

$$K_2 = \exp\left(-\left(\frac{1}{2}[E_L(R')+E_L(R)]-E_T\right)\Delta\tau\right)$$

Processed described by the $K(R, R'; \tau)$

It can be easily shown that the small time Green function describes to distinct processes:

- K_1 is the small time Green function of

$$-\frac{\partial f(R,t)}{\partial t} = -D\nabla^2 f + D\nabla \cdot [f\mathbf{F_Q}(\mathbf{R})],$$

  which has the steady-state solution $f = \Psi_T^2$.

- K_2 is the small time Green function of

$$-\frac{\partial f(R,t)}{\partial t} = [E_L(R) - E_T]f,$$

  and describes rate processes, or branching.

The algorithm to sample $f(R,\tau)$

- To perform the filtering and sampling use

$$f(R,(n+1)\Delta\tau) = \int dR'\, K(R,R';\Delta\tau) f(R';n\Delta\tau).$$

Start from

$$\begin{aligned} f(R,0) &= \sum_{i=1}^{N_W} \delta(R-R_{i,0}), \\ |\Psi_T(R)|^2 &\longrightarrow \{R_{i,0}\}. \end{aligned}$$

- Evolve $f(R,0)$ by one step to get

$$\begin{aligned} f(R,\Delta\tau) &= \sum_i \int dR'\, K_2(R,R_{i,0};\Delta\tau) \times K_1(R,R_{i,0};\Delta\tau) \\ &\equiv \sum_i w_i \times K_1(R,R_{i,0};\Delta\tau) \end{aligned}$$

The algorithm: Diffusion with a drift

- Sample $f(R, \Delta\tau)$ to get a new generation of walkers

$$f(R, \Delta\tau) = \sum_{i=1}^{N'_W} \delta(R - R_{i,1}),$$

($N'_W \neq N_W$). as follows:

- K_1: $R_{i,0} \longrightarrow R_{i,0} + D\Delta\tau F_Q(R_{i,0}) + \chi$
  - χ normally distributed, variance $2D\Delta\tau$,
  - accept the move with probability

$$p = \min\left[\frac{|\Psi_T(R')|^2 K_1(R, R'; \Delta\tau)}{|\Psi_T(R)|^2 K_1(R', R; \Delta\tau)}, 1\right],$$

    to eliminate time step errors in this part of the evolution.
  - Note that the exact G is symmetric ($\longrightarrow p = 1$)!.

The algorithm: Branching

- Of each walker take m (integer!) copies, according
  - $\mathrm{m} = w_i + \chi' = e^{-(E_L(R)+E_L(R')/2-E_T)\Delta\tau} + \chi'$
  - χ' flat in $]0,1[$.
- Periodically adjust E_T to keep the walker population $\simeq N_W$.

- ITERATE, discard the initial transient (filtering!); accumulate averages.

Importance sampling: *mixed estimates*

- With importance sampling it is easy to evaluate mixed estimates

$$< \mathcal{O} >_{mix} = \frac{\int dR\, \Phi_0(R) \mathcal{O} \Psi_T(R)}{\int dR\, \Phi_0(R) \Psi_T(R)} = \frac{\int dR\, f(R) \mathcal{O}_L(R)}{\int dR\, f(R)}$$

$$\simeq \frac{1}{M} \sum_{i=1}^{M} \mathcal{O}_L(R_i),$$

- For the Hamiltonian and for operators that commute with it, the mixed estimate coincides with the pure estimate,

$$< H >_{mix} = E_0, \quad < \mathcal{O} >_{mix} = < \mathcal{O} >_0 .$$

- For other operators $< \mathcal{O} >_{mix}$ is different from the pure ground state expectation value $< \mathcal{O} >_0 = < \Phi_o | \mathcal{O} | \Phi_0 > ..$

Extrapolated estimates

- For operators that do not commute with H, assume that Ψ_T is close enough to Φ_0

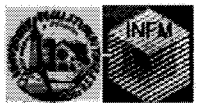
$$|\Phi_0 >= |\Psi_T > +\delta|\Psi > .$$

Then a reasonable approximatation to the pure average is given by the extrapolated estimate

$$\begin{aligned} < \mathcal{O} > &= 2 < \mathcal{O} >_{mix} - < \mathcal{O} >_{VMC} + O(\delta^2) \\ &\equiv < \mathcal{O} >_{\mathbf{extr}} + O(\delta^2) \end{aligned}$$

- Another possibility is to resort to forward walking, whereby keeping memory for a while of the walkers evolution, one becomes able to evaluate pure estimates.

Fermions and DMC: the nodes



- Nodal surfaces

$$H = -D\nabla^2 + V(R),$$

$$H\Phi_n(R) = E_n\Phi_n(R).$$

Statistics $\leftrightarrow$ symmetry of $\Phi(R)$

- Bosons: $\Phi_0^B = \Phi_0$
- Fermions: $\Phi_0^F = \Phi_n, n > 0$ if $N > 2$.
  - An excited state of H ($\Phi_n, n > 0$) has nodes

- $\Phi_0^F = \Phi_n, n > 0$
  - Absence of exact and stable algorithms: because of roundoff errors even if Ψ_T is antisymmetric, : $\Psi_T \longrightarrow \Phi_0 = \Phi_0^B$.

Fixed-Node Approximation (Fermions)

- $\Phi_0^F = \Phi_n, n > 0$
  - $\Phi_n(R) \gtrless 0$ is not a probability
- However, the nodal regions of the lowest state of given discrete symmetry possess a *tiling property*: given a nodal region all the others are obtained by the first by permutational symmetry.
  - Thus It is possible to consider one of the equivalent *nodal region*, and get a Bosonic problem with homogeneous boundary conditions.
  - The exact nodal surface *unknown* in general: fix the nodes to those of $\Psi_T(R)$, which implies $f(R) \geq 0$ everywhere.
  - Fixed–node is variational and stable.

Fixing the nodes on the walkers

- Constrain the random walk to a given nodal region, rejecting moves whereby a walker crosses in a different nodal pocket.
- In practice, monitor the sign of Ψ_T. If the proposed move is such that $\Psi_T(R')/\Psi_T(R) < 0$ reject the move.

Plan of last lecture

- Helium–4
- Electrons in low dimension
  - Magnetization transition in the 2D electron gas (egas)
  - A model quantum wire
  - Electron–hole bilayer

He_4 by DMC

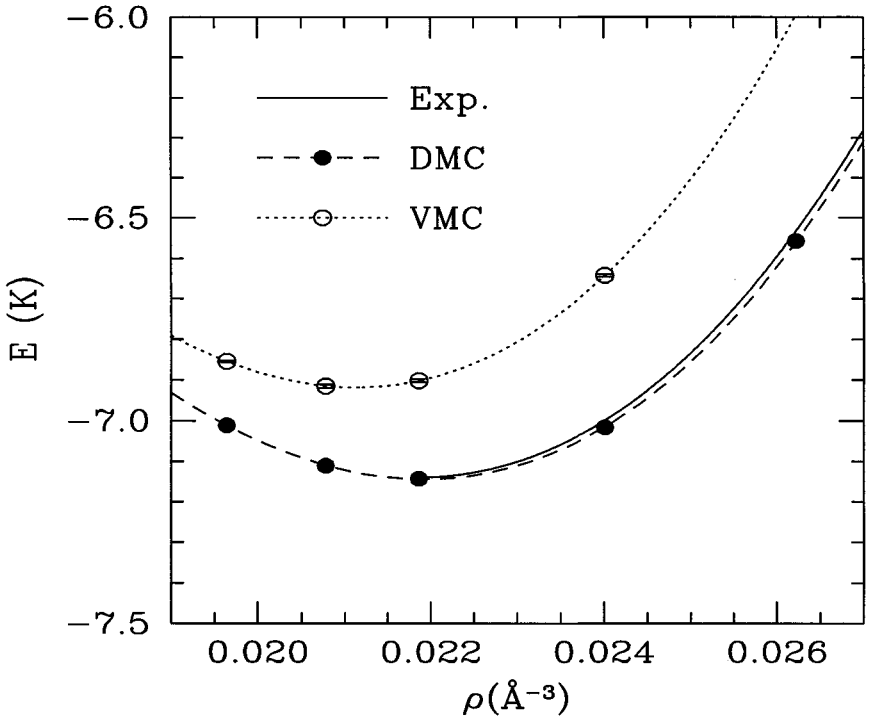
In the following as a first illustration of the method to Bosons, we show selected results of DMC calculations for He_4, after Moroni *et al* (PRB 52, 13547, 95).

In particular we shall give results for

- The equation of state
- Structural properties
- The condensate
- Note: the HFDHE2 of Aziz *et al* (JCP 70, 4330, 1979) was used .

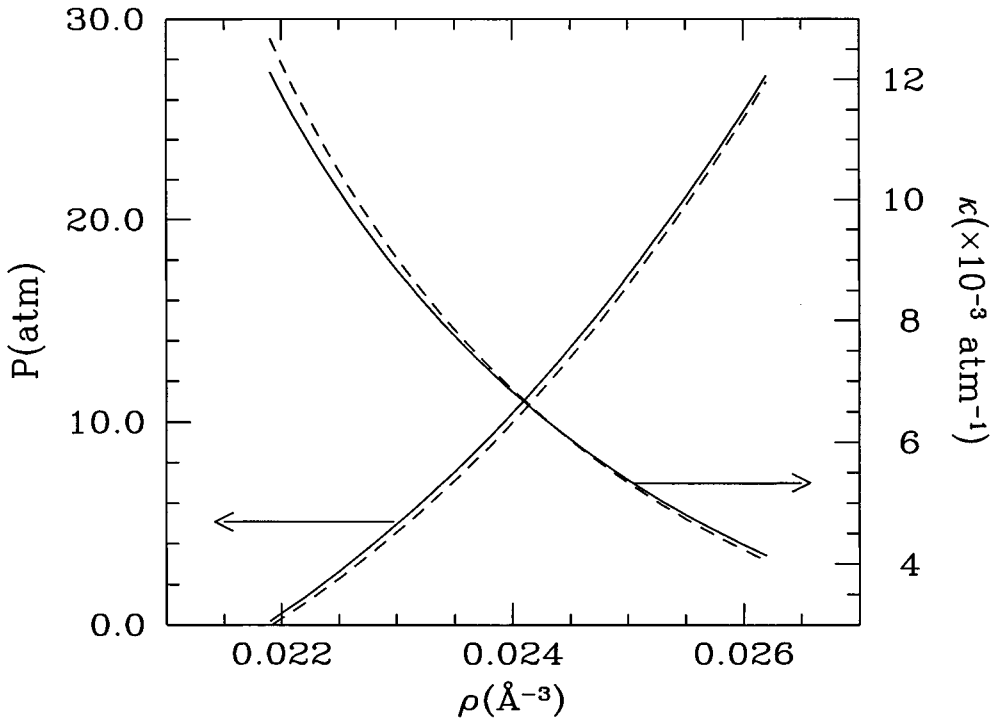
Equation of state of He_4

- Dashed and dotted lines are fits to the MC results.
- Exp: Roach *et al*, (PRA 2, 543, 1970).
- VMC: OJOT



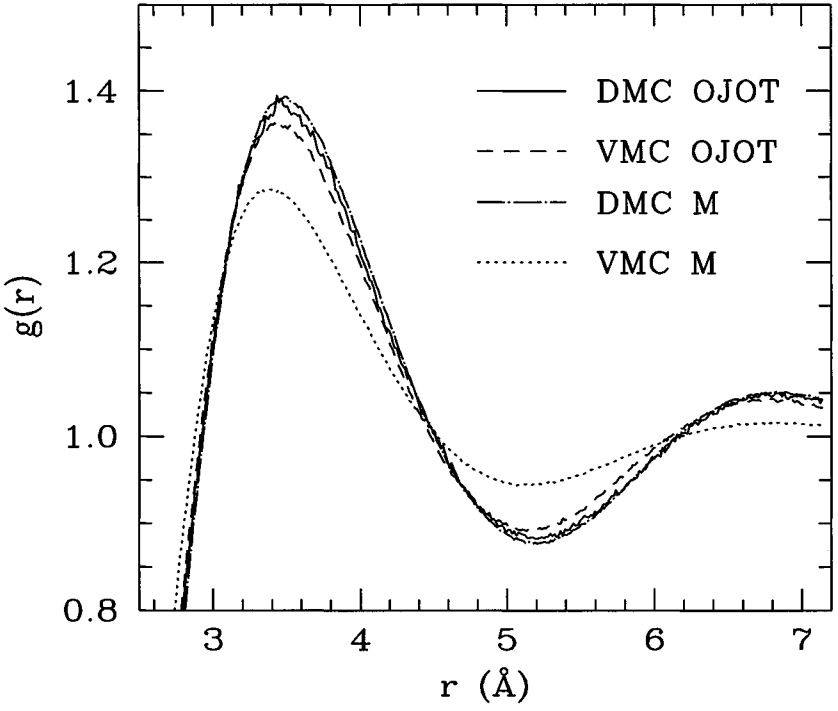
Pressure and compressibility of He_4

- Dashed lines are fits to the DMC results.
- Exp: Roach *et al*, (PRA 2, 543, 1970).

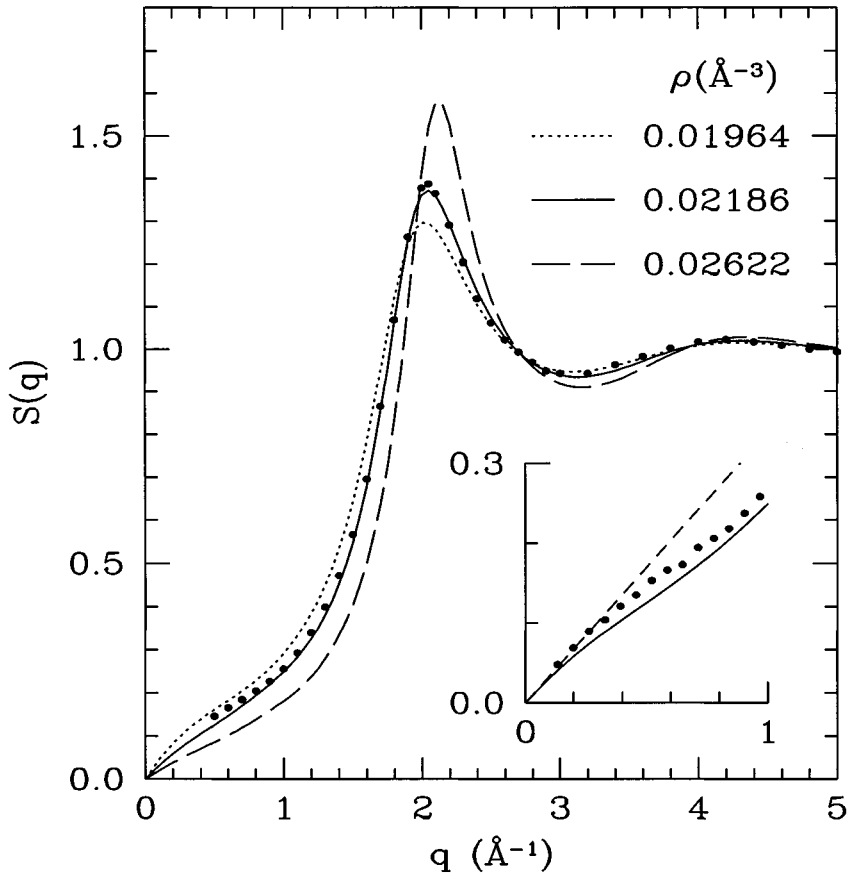


The g(r) of He_4 at the equilibrium density

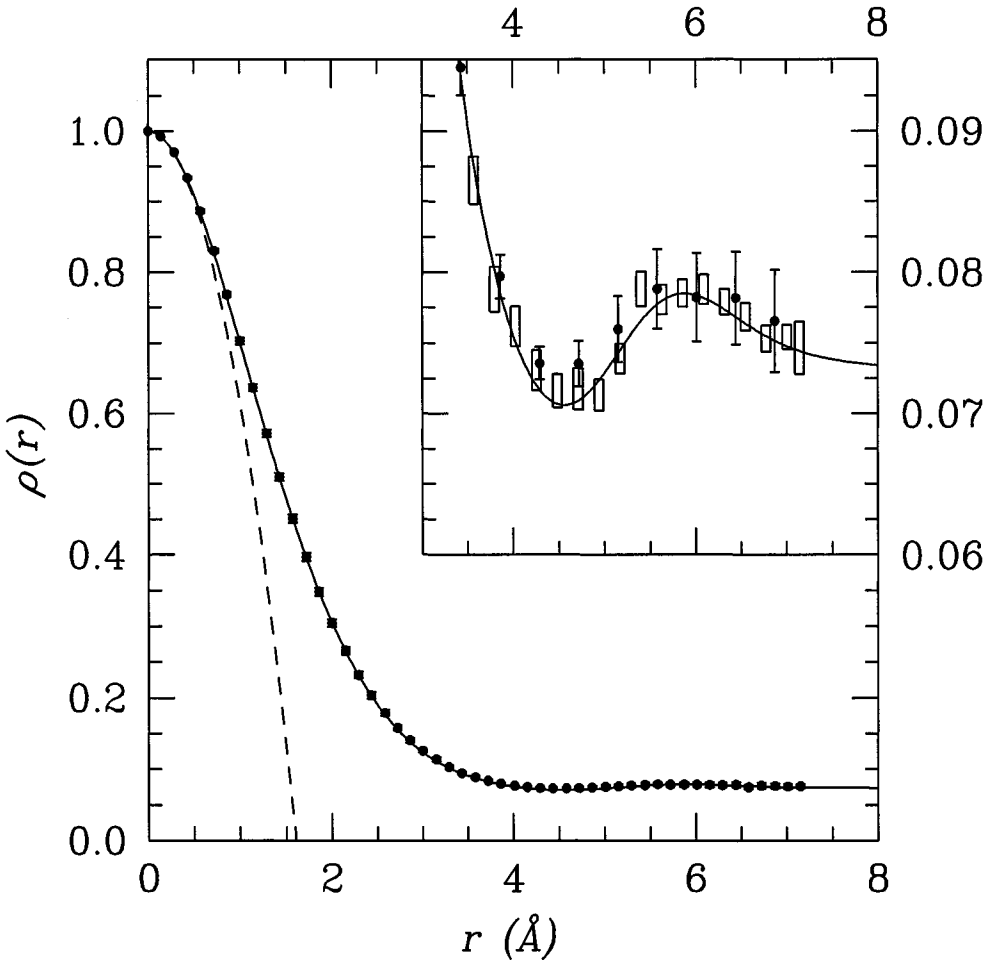
- Note the improvement from the VMC - M to the OJOT.



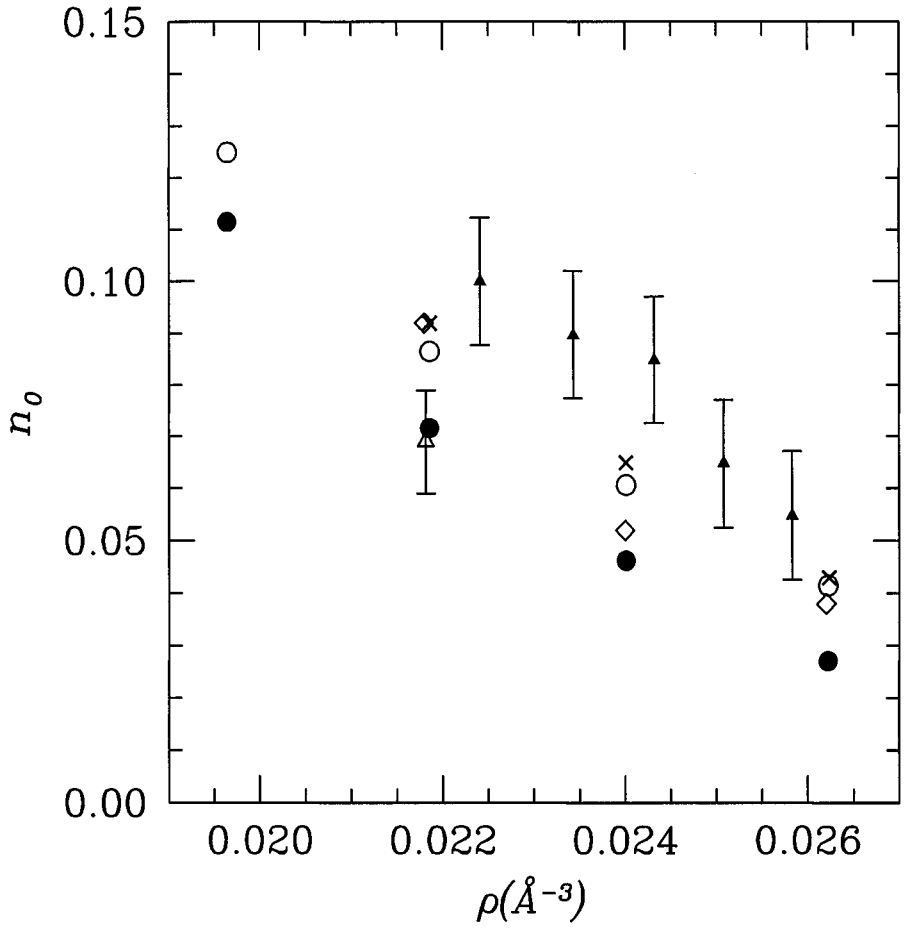
The S(q) of He_4



The one–body density matrix of He_4



The condensate fraction in He_4



Condensate fraction in He_4

Table 2: Condensate fraction n_0 in ^{4}He. DMC, GFMC(Whitlock et al, 87), and HNC(Manousakis et al, 85) predictions are at $T = 0$. The PIMC(Ceperley, Pollock, 86) result is at $T = 1.18^o$K. The density is in $\mathring{A}^{-3}$.

ρ	0.01964	0.02186	0.02401	0.02622
DMC	0.112(1)	0.0717(5)	0.0462(6)	0.02.71(6)
PIMC		0.069(10)		
GFMC		0.092(1)	0.052(1)	0.037(2)
HNC		0.092	0.065	0.043

Electrons in low dimension

- Quasi 2D and 1D electron (hole) systems are nowadays routinely realized in the labs.
- Very low densities are becoming achievable: exchange and correlation (!) should have measurable consequences in these systems. however, their balance is extremely delicate.
- Coupled electron (hole) layers [Quantum Wells (QW)] should bring about addictional correlations effects and phases.
- High accuracy is crucial for reliable predictions.
- Accurate calculations should clarify some of the points above and may be guide experiments.