The (negative) sign problem in Full Configuration Interaction Quantum Monte Carlo and other short stories

James Spencer\textsuperscript{1,2} and Matthew Foulkes\textsuperscript{2}

\textsuperscript{1}Thomas Young Centre, Dept. of Materials, Imperial College London

\textsuperscript{2}Condensed Matter Theory, Dept. of Physics, Imperial College London

July 30, 2012
HANDE-QMC code

Highly Accurate N-Determinant Quantum Monte Carlo

- Systems:
  - Hubbard model (local and Bloch orbitals)
  - Uniform electron gas
  - Heisenberg model
  - Molecular systems via precomputed integrals

- Methods:
  - Full Configuration Interaction
  - Full Configuration Interaction Quantum Monte Carlo
  - Coupled Cluster Monte Carlo
  - Initiator approximation
  - Folded spectrum FCIQMC
  - Density Matrix Quantum Monte Carlo

- Much more to come...

Available to collaborators. Open-source release in the next year-ish.
Acknowledgements

- UROP students (funding: EPSRC via CDT on Theory and Simulation of Materials)
  - Joe Weston (Imperial)
  - Nick Blunt (Imperial)
  - Will Handley (Cambridge)
- MSci students
  - Nick Blunt
  - Tom Rogers
- Alex Thom
- Richard Needs
Essentially exploit the power method for finding the eigenstate, \( c_0 \) with the largest absolute eigenvalue of a matrix, \( M \):

1. Take a starting vector, \( n(t = 0) \) with a non-zero overlap with \( c_0 \); \( n(0) = \sum_i x_i c_i \).
2. Let \( n_i(t + \Delta \tau) = n_i(t) + \sum_j M_{ij} n_j(t) \Delta \tau \).
3. Contribution from eigenstate \( c_i \) decays as \( ((1 + \Delta \tau \lambda_i)/(1 + \Delta \tau \lambda_0))^t/\Delta \tau \).
4. \( n(t \to \infty) \propto c_0 \).

\( \rightarrow \) Can easily be performed stochastically by sampling the action of \( M \) on \( n^1 \).

Win if memory demands are less than two vectors the size of the Hilbert space!

\[^1\text{G.H. Booth, A.J.W. Thom and Ali Alavi, JCP 131, 054106 (2009)}\]
Imaginary-time Schrödinger equation

\[ n(\tau = k\Delta\tau) = (I - H\Delta\tau)^k n(0) \]  \hspace{1cm} (1)

is a first-order approximation to

\[ n(\tau) = e^{-H\tau} n(0) \]  \hspace{1cm} (2)

which is the solution to the imaginary-time Schrödinger equation:

\[ \frac{d n_i}{d\tau} = - \sum_j H_{ij} n_j. \]  \hspace{1cm} (3)

FCIQMC appears to be particularly efficient for (some) quantum systems.
FCIQMC algorithm

for each occupied site $i$

for each psip, sign $s_i$, on $i$

energy contribution: $\frac{H_{i0} - S\delta_{i0}}{n_0} s_i$

annihilate parent and child psips

select a random site, $j$

spawn new psip on $j$ with probability

$$\frac{|H_{ij} - S\delta_{ij}| \Delta \tau}{p(j|i)},$$

$$-\text{sign} \left( (H_{ij} - S\delta_{ij}) s_i \right)$$
Example: CN

UHF single-particle basis; cc-pVDZ; CAS (9,12); 98476 determinants.
FCIQMC: successes and failures

✓ Exact (within finite basis results) for wide variety of atoms and molecules
✓ Benchmark results for ionisation and electron affinity energies
✓ Largest calculation done: $> \mathcal{O}(10^{15})$ [largest FCI: $\mathcal{O}(10^{10})$]
× Methane is ‘hard’!
× Hubbard model is a disaster...
Hubbard model plateau

\[ \hat{H} = -t \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle, \sigma} \hat{c}_{\mathbf{r}, \sigma}^{\dagger} \hat{c}_{\mathbf{r}', \sigma} + U \sum_{\mathbf{r}} \hat{n}_{\mathbf{r}, \uparrow} \hat{n}_{\mathbf{r}, \downarrow} \]  

(4)

\[ \hat{H} = \sum_{\mathbf{k}, \sigma} \epsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}, \sigma}^{\dagger} \hat{c}_{\mathbf{k}, \sigma} + \frac{U}{M} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3} \hat{c}_{\mathbf{k}_1, \uparrow}^{\dagger} \hat{c}_{\mathbf{k}_2, \downarrow}^{\dagger} \hat{c}_{\mathbf{k}_3, \downarrow} \hat{c}_{\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3, \uparrow} \]  

(5)

18 site 2D Hubbard model at \( \mathbf{k} = (0, 0) \):
Annihilation is crucial
FCIQMC without annihilation

(Let $T = -(H - SI) = T^+ - T^-$.)

Separate, but coupled, populations of positive and negative psips$^2$:

\[
\frac{dn_i^+}{d\tau} = \sum_j \left( T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right),
\]

\[
\frac{dn_i^-}{d\tau} = \sum_j \left( T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right).
\]

(6)

Can combine in-phase and out-of-phase:

\[
\frac{d(n_i^+ + n_i^-)}{d\tau} = \sum_j \left( T_{ij}^+ + T_{ij}^- \right) \left( n_j^+ + n_j^- \right),
\]

\[
\frac{d(n_i^+ - n_i^-)}{d\tau} = \sum_j \left( T_{ij}^+ - T_{ij}^- \right) \left( n_j^+ - n_j^- \right).
\]

(7)

$^2$ JSS, N.S. Blunt, WMCF, JCP 136 054110 (2012)
Convergence to $H^+ + H^-$
Effect of annihilation

\[
\frac{dn_i^+}{d\tau} = \sum_j \left( T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right)
\]

\[
\frac{dn_i^-}{d\tau} = \sum_j \left( T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right)
\]  

(8)
Effect of annihilation

\[
\frac{dn_i^+}{d\tau} = \sum_j \left( T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right) - 2\kappa n_i^+ n_i^- \\
\frac{dn_i^-}{d\tau} = \sum_j \left( T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right) - 2\kappa n_i^+ n_i^-
\]

(8)

Destabilises in-phase state \( n^+ + n^- \).
Leaves true solution, \( n^+ - n^- \), unchanged.
Sign-problem-free systems

If $T^+ + T^-$ and $T^+ - T^-$ are related by a unitary transform then:

- identical set of eigenvalues;
- identical growth rates;
- no annihilation events;
- no sign problem in FCIQMC $\Rightarrow$ sample FCI ground state with arbitrary number of psips.

Sign-problem-free systems: 1D Hubbard model in a local orbital basis; Heisenberg bipartite lattices.

Example: 18-site, 18-electron 1D Hubbard model at $U = t$:

<table>
<thead>
<tr>
<th>basis</th>
<th>Hilbert space</th>
<th>plateau height</th>
<th># psips</th>
<th>energy $(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bloch</td>
<td>$1.31 \times 10^8$</td>
<td>$6.9 \times 10^6$</td>
<td>$2.3 \times 10^7$</td>
<td>$-18.84248(8)$</td>
</tr>
<tr>
<td>local</td>
<td>$2.36 \times 10^9$</td>
<td>n/a</td>
<td>$2.8 \times 10^5$</td>
<td>$-18.8423(3)$</td>
</tr>
</tbody>
</table>
Population dynamics

\( \text{Let } p = n^+ + n^- \text{ and } n = n^+ - n^- \).

\[
\begin{align*}
\frac{dp_i}{d \tau} &= \sum_j \left( T_{ij}^+ + T_{ij}^- \right) p_j - \kappa (p_i^2 - n_i^2) \\
\frac{dn_i}{d \tau} &= \sum_j \left( T_{ij}^+ - T_{ij}^- \right) n_j.
\end{align*}
\] (9)

As \( \tau \to \infty \), \( n(\tau) \) tends to ground-state wavefunction, \( n_0 \):

\[
\frac{dp_i}{d \tau} \approx \sum_j \left( T_{ij}^+ + T_{ij}^- \right) p_j - \kappa p_i^2 + \kappa \alpha^2 e^{2T_{\max} \tau} n_0^2. \] (10)

\( \Rightarrow \) Initial exponential growth followed by a plateau followed by a second (slower) exponential growth.
One-component analogue

\[
\frac{dp_i}{d\tau} \approx \sum_j \left( T_{ij}^+ + T_{ij}^- \right) p_j - \kappa p_i^2 + \kappa \alpha^2 e^{2T_{\text{max}} \tau} n_{0i}^2.
\]  \hspace{1cm} (11)

One-component analogue of population ODE:

\[
\frac{dp}{d\tau} = V_{\text{max}} p - \kappa p^2 + \kappa \left( n_0 e^{T_{\text{max}} \tau} \right)^2.
\]  \hspace{1cm} (12)
One-component analogue

One-component analogue of population ODE:

\[
\frac{dp}{d\tau} = V_{\text{max}}p - \kappa p^2 + \kappa \left(n_0 e^{T_{\text{max}} \tau}\right)^2.
\] (11)

Riccati differential equations can be solved:

\[
p(\tau) = \frac{1}{\kappa u} \frac{du}{d\tau},
\] (12)

\[
u(\tau) = c_1 \cdot {}_0F_1 \left( ; 1 - \frac{V_{\text{max}}}{2T_{\text{max}}}; z \right)
+ c_2 z^{V_{\text{max}}/2T_{\text{max}}} \cdot {}_0F_1 \left( ; 1 + \frac{V_{\text{max}}}{2T_{\text{max}}}; z \right),
\] (13)

\[
z = \frac{\kappa^2 n_0^2 e^{2T_{\text{max}} \tau}}{4T_{\text{max}}^2}.
\] (14)
Model population dynamics

\[ \frac{p(\tau)}{n_0} \]

\[ V_{\text{max}}/\kappa n_0 = 25.0; \quad V_{\text{max}}/T_{\text{max}} = 30.0 \]

\[ V_{\text{max}}/\kappa n_0 = 12.5; \quad V_{\text{max}}/T_{\text{max}} = 15.0 \]

\[ V_{\text{max}}/\kappa n_0 = 12.5; \quad V_{\text{max}}/T_{\text{max}} = 30.0 \]
Hubbard model: plateau height $\propto U/t$

Kinetic contributions ($\propto t$) to the Hamiltonian matrix are diagonal in the Bloch basis.

\[
\frac{dp_i}{d\tau} = \sum_j \left( T_{ij}^+ + T_{ij}^- \right) p_j - \kappa p_i^2 + \kappa \alpha^2 e^{2T_{\text{max}}\tau} n_{0i}^2 \quad (15)
\]

\[
\downarrow
\]

\[
U^2 \sum_{ij} \left( T_{ij}^{+\prime} + T_{ij}^{-\prime} \right) p_j' \approx \kappa U^2 \sum_i p_i'^2. \quad (16)
\]

Total population psip population at the plateau:

\[
\sum_i p_i = U \sum_i p_i' \quad (17)
\]
Propagator \((I - H \Delta \tau)\) only gives access to the maximal eigenstate of \(H\).

Use folded spectrum method:

\[
M = (H - \varepsilon I) \tag{18}
\]

and solve for \(M^2\):

\[
n_i(\tau + \Delta \tau) = n_i(\tau) + \sum_j \sum_k (M_{ij} M_{jk} - S \delta_{ij} \delta_{jk}) \Delta \tau n_k(\tau) \tag{19}
\]

Sample action of \((H - \varepsilon I)^2 - SI\) rather than action of \(H - SI\).
Excitation generation

\[
\begin{align*}
\frac{(M_{ii}M_{ii} - S) \Delta \tau}{p(\circlearrowleft)} & \quad \frac{M_{jj}M_{ji} \Delta \tau}{p(\rightleftharpoons)p(i|j)} & \quad \frac{M_{ji}M_{ii} \Delta \tau}{p(\rightleftharpoons)p(i|j)} \\
\frac{M_{jk}M_{ki} \Delta \tau}{p(\左上rightarrow \rightleftharpoons \leftdownarrow \rightarrow)} & \quad & \\
\end{align*}
\]
Preliminary results: $3 \times 3$ Hubbard model, $U = t$, 8 electrons
Conclusions

▶ Negative sign problem in FCIQMC is due to instability to a non-physical state.
▶ Annihilation ensures convergence to the true ground state of the Hamiltonian.
▶ Characteristic population dynamics is due to the interplay between the instability, annihilation and the true ground state.
▶ Severity of the sign problem is dependent upon the underlying basis.
▶ Excited states accessible via the folded spectrum approach.
Bonus slides
Heisenberg spin model (Nick Blunt)

\[ \hat{H} = J \sum_{<ij>} \hat{S}_i \cdot \hat{S}_j \quad (20) \]

5 × 5 anti-ferromagnetic \((J > 0)\) triangular lattice with periodic boundary conditions.
Time-step error (in limit $S \to E_0$)

Exact propagator $e^{-(H-SI)\Delta \tau}$: $\lambda_0 = 1$.

Approximate propagator $I - (H - SI)\Delta \tau$:

$\lambda_{\text{max}} = 1 - (E_0 - S)\Delta \tau = 1$ or

$\lambda_{\text{max}} = 1 - (E_{\text{max}} - S)\Delta \tau = 1 - (E_{\text{max}} - E_0)\Delta \tau$.

Disaster occurs if

$$\Delta \tau > \frac{2}{E_{\text{max}} - E_0}$$ (21)
Time step and the sign problem

Hamiltonian matrices are (often) diagonally dominant.

Sign problem is actually not so bad if a psip cannot create more than one psip of the opposite sign on its own basis function.

Example: uniform electron gas $r_s = 1.0$, $n = 4^3$.

<table>
<thead>
<tr>
<th>$M_{ij}$</th>
<th>Lowest eigenvalue (a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle D_i</td>
<td>\hat{H}</td>
</tr>
<tr>
<td>$-</td>
<td>\langle D_i</td>
</tr>
<tr>
<td>$-</td>
<td>\langle D_i</td>
</tr>
</tbody>
</table>

\[3\text{http://github.com/jsspencer/toy_fci}\]
Convergence to the ground state

\[
\frac{dp_i}{d\tau} = \sum_j \left( T_{ij}^+ + T_{ij}^- \right) p_j - \kappa(p_i^2 - n_i^2) \tag{22}
\]

After the plateau the shift is adjusted to the ground state energy:

\[
0 = \frac{dp_i}{d\tau} \approx \sum_j \kappa(n_i^2 - p_i^2) \tag{23}
\]

⇒ basis functions occupied by positive or negative psips.

\textbf{n}: stochastic representation of the ground-state wavefunction

\textbf{\textit{n}}: psip population