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

James Spencer[1,2] and Matthew Foulkes[2]

[1]Thomas Young Centre, Dept. of Materials, Imperial College London

[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

Stochastic diagonalisation

Essentially exploit the power method for finding the eigenstate, $\mathbf{c}_0$ with the largest absolute eigenvalue of a matrix, $\mathbf{M}$:

1. Take a starting vector, $\mathbf{n}(t=0)$ with a non-zero overlap with $\mathbf{c}_0$; $\mathbf{n}(0) = \sum_i x_i \mathbf{c}_i$.
2. Let $n_i(t+\Delta\tau) = n_i(t) + \sum_j M_{ij} n_j(t) \Delta\tau$.
3. Contribution from eigenstate $\mathbf{c}_i$ decays as $((1+\Delta\tau\lambda_i)/(1+\Delta\tau\lambda_0))^{t/\Delta\tau}$.
4. $\mathbf{n}(t \to \infty) \propto \mathbf{c}_0$.

$\rightarrow$ Can easily be performed stochastically by sampling the action of $\mathbf{M}$ on $\mathbf{n}$[1].

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

[1] G.H. Booth, A.J.W. Thom and Ali Alavi, JCP 131 054106 (2009)

Imaginary-time Schrödinger equation

$$\mathbf{n}(\tau = k\Delta\tau) = (\mathbf{I} - \mathbf{H}\Delta\tau)^k \mathbf{n}(0) \tag{1}$$

is a first-order approximation to

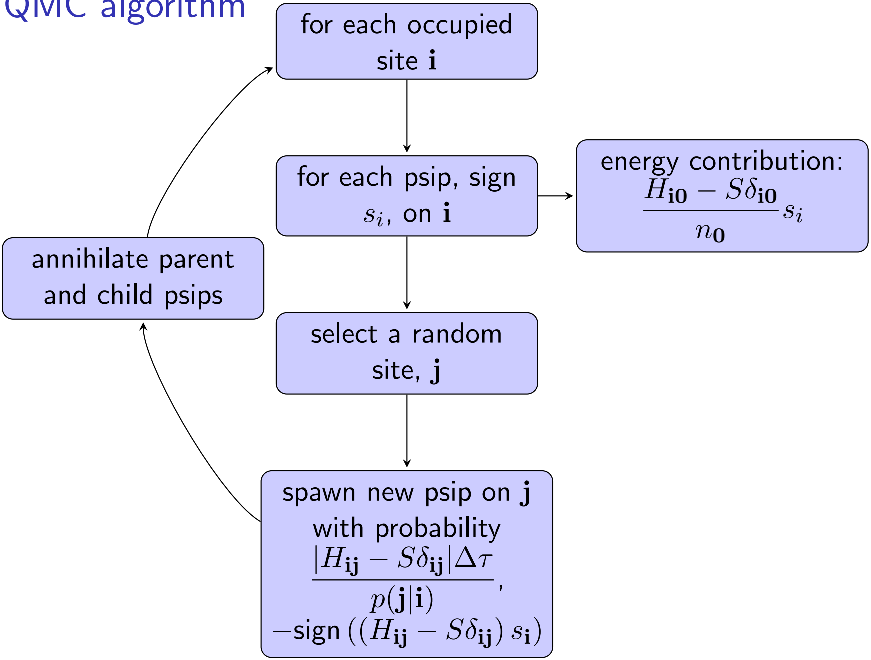
$$\mathbf{n}(\tau) = e^{-\mathbf{H}\tau}\mathbf{n}(0) \tag{2}$$

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

$$\frac{dn_{\mathbf{i}}}{d\tau} = -\sum_{\mathbf{j}} H_{\mathbf{ij}} n_{\mathbf{j}}. \tag{3}$$

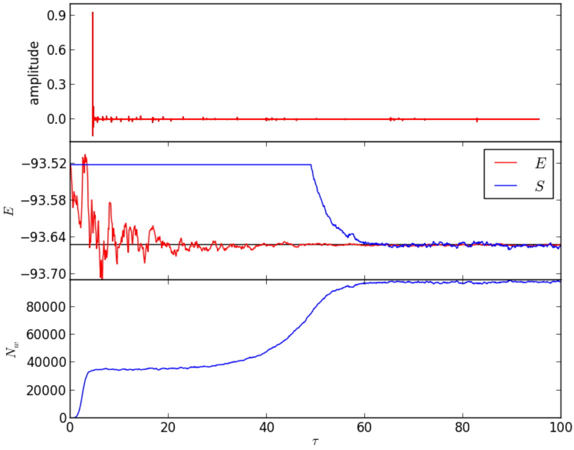
FCIQMC appears to be particularly efficient for (some) quantum systems.

FCIQMC algorithm



Example: CN

UHF single-particle basis; cc-pVDZ; CAS (9,12); 98476 determinants.



FCIQMC: successes and failures

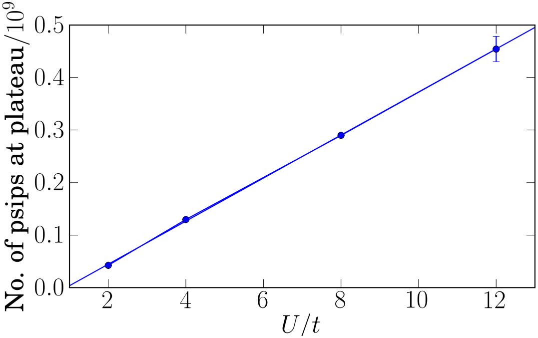
- ✓ Exact (within finite basis results) for wide variety of atoms and molecules
- ✓ Benchmark results for ionisation and electron affinity energies
- ✓ Largest calculation done: $> \mathcal{O}(10^{15})$ [largest FCI: $\mathcal{O}(10^{10})$]
- × Methane is 'hard'!
- × Hubbard model is a disaster...

Hubbard model plateau

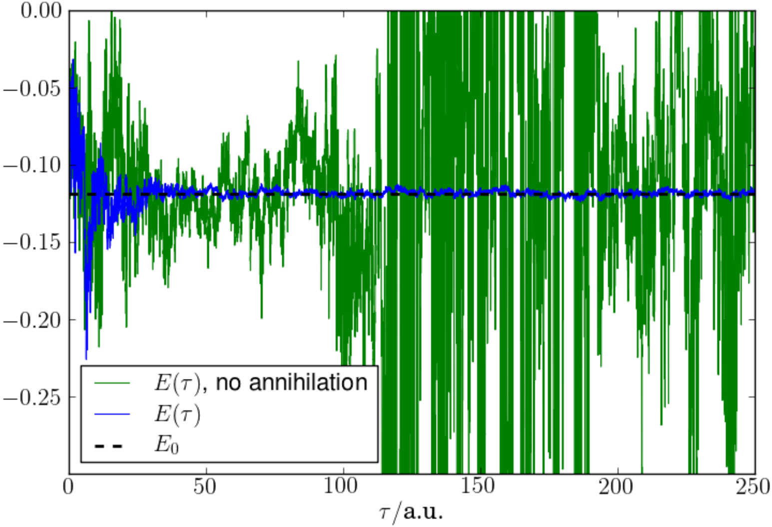
$$\hat{H} = -t \sum_{\langle \mathbf{r},\mathbf{r}' \rangle,\sigma} \hat{c}^{\dagger}_{\mathbf{r},\sigma} \hat{c}_{\mathbf{r}',\sigma} + U \sum_{\mathbf{r}} \hat{n}_{\mathbf{r},\uparrow} \hat{n}_{\mathbf{r},\downarrow} \tag{4}$$

$$\hat{H} = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} \hat{c}^{\dagger}_{\mathbf{k},\sigma} \hat{c}_{\mathbf{k},\sigma} + \frac{U}{M} \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3} \hat{c}^{\dagger}_{\mathbf{k}_1,\uparrow} \hat{c}^{\dagger}_{\mathbf{k}_2,\downarrow} \hat{c}_{\mathbf{k}_3,\downarrow} \hat{c}_{\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3,\uparrow}, \tag{5}$$

18 site 2D Hubbard model at $\mathbf{k} = (0,0)$:



Annihilation is crucial



FCIQMC without annihilation

(Let $\mathbf{T} = -(\mathbf{H} - S\mathbf{I}) = \mathbf{T}^{+} - \mathbf{T}^{-}$.)
Separate, but coupled, populations of positive and negative psips[2]:

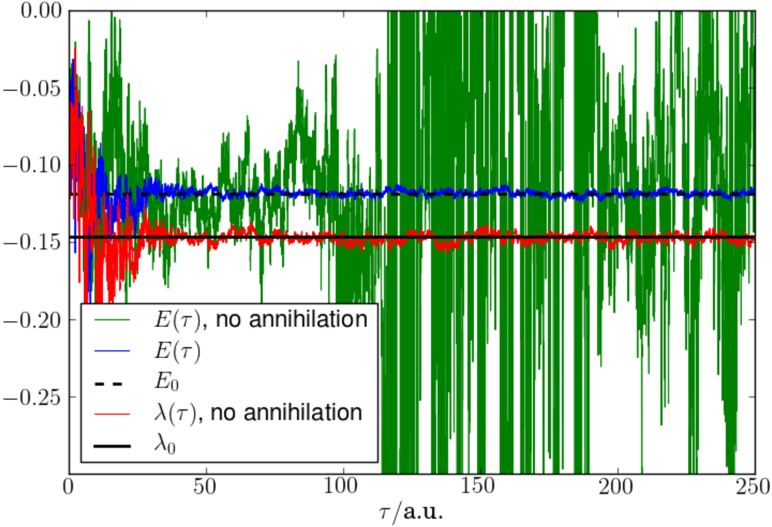
$$\begin{aligned} \frac{dn_{\mathbf{i}}^{+}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{+} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{-} \right), \\ \frac{dn_{\mathbf{i}}^{-}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{-} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{+} \right). \end{aligned} \tag{6}$$

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

$$\begin{aligned} \frac{d\left(n_{\mathbf{i}}^{+} + n_{\mathbf{i}}^{-}\right)}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} + T_{\mathbf{ij}}^{-} \right) \left(n_{\mathbf{j}}^{+} + n_{\mathbf{j}}^{-} \right), \\ \frac{d\left(n_{\mathbf{i}}^{+} - n_{\mathbf{i}}^{-}\right)}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} - T_{\mathbf{ij}}^{-} \right) \left(n_{\mathbf{j}}^{+} - n_{\mathbf{j}}^{-} \right). \end{aligned} \tag{7}$$

[2] JSS, N.S. Blunt, WMCF, JCP 136 054110 (2012)

Convergence to $\mathbf{H}^+ + \mathbf{H}^-$



Effect of annihilation

$$
\begin{aligned}
\frac{dn_{\mathbf{i}}^{+}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{+} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{-} \right) \\
\frac{dn_{\mathbf{i}}^{-}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{-} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{+} \right)
\end{aligned} \tag{8}
$$

Effect of annihilation

$$\begin{aligned}\frac{dn_{\mathbf{i}}^{+}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{+} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{-} \right) - 2\kappa n_{\mathbf{i}}^{+} n_{\mathbf{i}}^{-} \\ \frac{dn_{\mathbf{i}}^{-}}{d\tau} &= \sum_{\mathbf{j}} \left(T_{\mathbf{ij}}^{+} n_{\mathbf{j}}^{-} + T_{\mathbf{ij}}^{-} n_{\mathbf{j}}^{+} \right) - 2\kappa n_{\mathbf{i}}^{+} n_{\mathbf{i}}^{-}\end{aligned} \tag{8}$$

Destabilises in-phase state $\mathbf{n}^{+} + \mathbf{n}^{-}$.
Leaves true solution, $\mathbf{n}^{+} - \mathbf{n}^{-}$, unchanged.

Sign-problem-free systems

If $\mathbf{T}^+ + \mathbf{T}^-$ and $\mathbf{T}^+ - \mathbf{T}^-$ 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 $U = t$:

basis	Hilbert space	plateau height	# psips	energy (t)
Bloch	1.31×10^8	6.9×10^6	2.3×10^7	$-18.84248(8)$
local	2.36×10^9	n/a	2.8×10^5	$-18.8423(3)$

Population dynamics

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

$$\begin{aligned} \frac{dp_{\mathbf{i}}}{d\tau} &= \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} + T^{-}_{\mathbf{ij}}\right) p_{\mathbf{j}} - \kappa(p_{\mathbf{i}}^{2} - n_{\mathbf{i}}^{2}) \\ \frac{dn_{\mathbf{i}}}{d\tau} &= \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} - T^{-}_{\mathbf{ij}}\right) n_{\mathbf{j}}. \end{aligned} \tag{9}$$

As $\tau \to \infty$, $\mathbf{n}(\tau)$ tends to ground-state wavefunction, $\mathbf{n}_0$:

$$\frac{dp_{\mathbf{i}}}{d\tau} \approx \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} + T^{-}_{\mathbf{ij}}\right) p_{\mathbf{j}} - \kappa p_{\mathbf{i}}^{2} + \kappa \alpha^{2} e^{2T_{\mathsf{max}}\tau} n_{0\mathbf{i}}^{2}. \tag{10}$$

$\Rightarrow$ Initial exponential growth followed by a plateau followed by a second (slower) exponential growth.

One-component analogue

$$\frac{dp_{\mathbf{i}}}{d\tau} \approx \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} + T^{-}_{\mathbf{ij}} \right) p_{\mathbf{j}} - \kappa p_{\mathbf{i}}^2 + \kappa \alpha^2 e^{2T_{\mathsf{max}}\tau} n_{0\mathbf{i}}^2. \tag{11}$$

One-component analogue of population ODE:

$$\frac{dp}{d\tau} = V_{\mathsf{max}} p - \kappa p^2 + \kappa \left(n_0 e^{T_{\mathsf{max}}\tau} \right)^2. \tag{12}$$

One-component analogue

One-component analogue of population ODE:

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

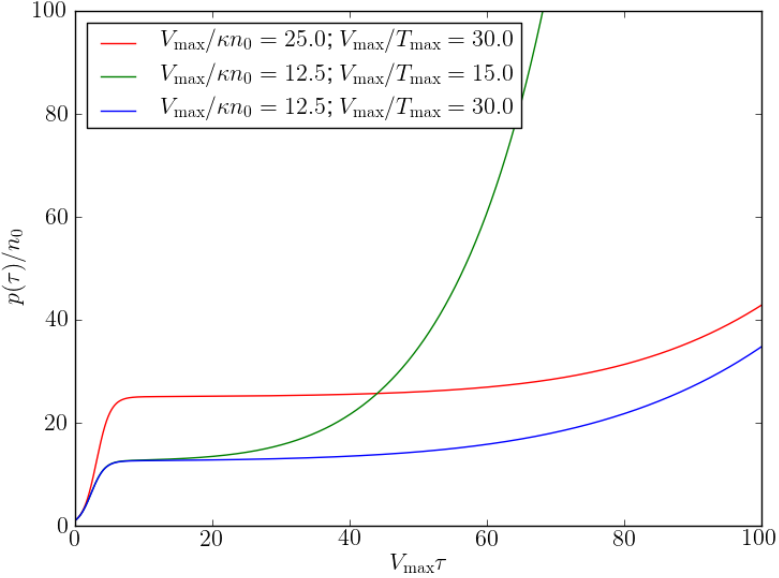
Riccati differential equations can be solved:

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

$$\begin{aligned} u(\tau) = & c_1 \cdot {}_0F_1 \left(; 1 - \frac{V_{\text{max}}}{2T_{\text{max}}} ; z \right) \\ & + c_2 z^{V_{\text{max}}/2T_{\text{max}}} \cdot {}_0F_1 \left(; 1 + \frac{V_{\text{max}}}{2T_{\text{max}}} ; z \right) , \end{aligned} \tag{13}$$

$$z = \frac{\kappa^2 n_0^2 e^{2T_{\text{max}} \tau}}{4T_{\text{max}}^2} . \tag{14}$$

Model population dynamics



Hubbard model: plateau height $\propto U/t$

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

$$\frac{dp_{\mathbf{i}}}{d\tau} = \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} + T^{-}_{\mathbf{ij}} \right) p_{\mathbf{j}} - \kappa p_{\mathbf{i}}^2 + \kappa \alpha^2 e^{2T_{\mathsf{max}}\tau} n_{0\mathbf{i}}^2 \tag{15}$$

$$\Downarrow$$

$$U^2 \sum_{\mathbf{ij}} \left(T^{+\prime}_{\mathbf{ij}} + T^{-\prime}_{\mathbf{ij}} \right) p'_{\mathbf{j}} \approx \kappa U^2 \sum_{\mathbf{i}} {p'_{\mathbf{i}}}^2. \tag{16}$$

Total population psip population at the plateau:

$$\sum_{\mathbf{i}} p_{\mathbf{i}} = U \sum_{\mathbf{i}} p'_{\mathbf{i}} \tag{17}$$

Accessing excited states in FCIQMC (Will Handley)

Propagator $(\mathbf{I} - \mathbf{H}\Delta\tau)$ only gives access to the maximal eigenstate of $\mathbf{H}$.

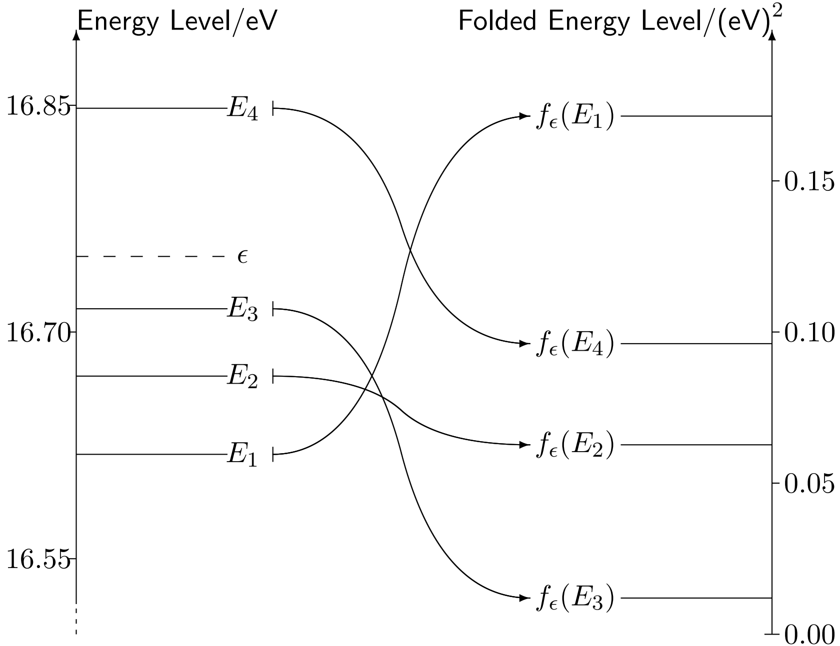
Use folded spectrum method:

$$\mathbf{M} = (\mathbf{H} - \varepsilon\mathbf{I}) \tag{18}$$

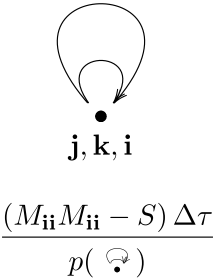
and solve for $\mathbf{M}^2$:

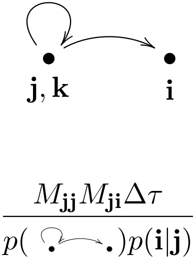
$$n_{\mathbf{i}}(\tau + \Delta\tau) = n_{\mathbf{i}}(\tau) + \sum_{\mathbf{j}} \sum_{\mathbf{k}} \left(M_{\mathbf{ij}} M_{\mathbf{jk}} - S\delta_{\mathbf{ij}}\delta_{\mathbf{jk}}\right) \Delta\tau n_{\mathbf{k}}(\tau) \tag{19}$$

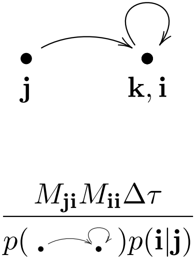
Sample action of $(\mathbf{H} - \varepsilon\mathbf{I})^2 - S\mathbf{I}$ rather than action of $\mathbf{H} - S\mathbf{I}$.

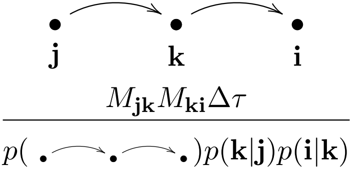

Energy Level/eV
Folded Energy Level/(eV)^2
16.85
16.70
16.55
E_4
ϵ
E_3
E_2
E_1
$f_\epsilon(E_1)$
$f_\epsilon(E_4)$
$f_\epsilon(E_2)$
$f_\epsilon(E_3)$
0.15
0.10
0.05
0.00

Excitation generation

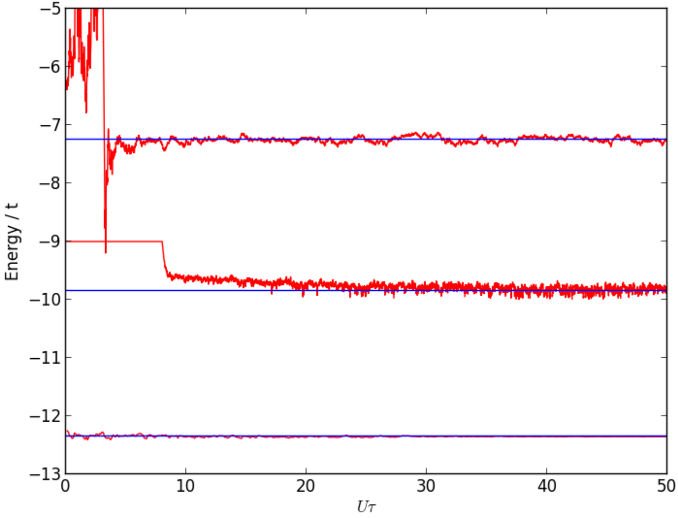








Preliminary results: 3×3 Hubbard model, $U = t$, 8 electrons



Conclusions

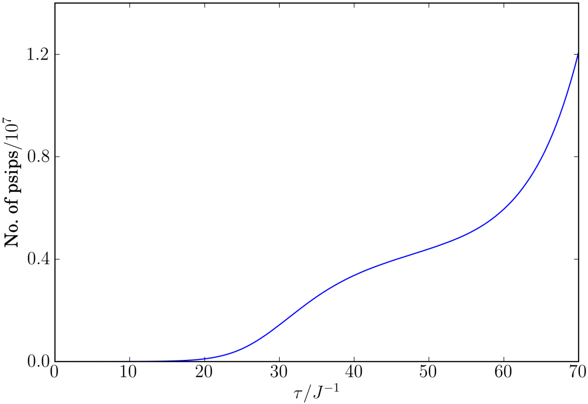
- Negative sign problem in FCIQMC is due to instability to a non-physical state.
- Annihilation ensures convergence to the true ground state of the Hamiltonian.
- Characteristic population dynamics is due to the interplay between the instability, annihilation and the true ground state.
- Severity of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

Bonus slides

Heisenberg spin model (Nick Blunt)

$$\hat{H} = J \sum_{<ij>} \hat{S}_i . \hat{S}_j \quad (20)$$

5×5 anti-ferromagnetic ($J > 0$) triangular lattice with periodic boundary conditions.



Time-step error (in limit $S \to E_0$)

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

Approximate propagator $\mathbf{I} - (\mathbf{H} - S\mathbf{I})\Delta\tau$:
$\lambda_{\text{max}} = 1 - (E_0 - S)\Delta\tau = 1$ or
$\lambda_{\text{max}} = 1 - (E_{\text{max}} - S)\Delta\tau = 1 - (E_{\text{max}} - E_0)\Delta\tau$.

Disaster occurs if

$$\Delta\tau > \frac{2}{E_{\text{max}} - E_0} \tag{21}$$

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 $r_s = 1.0, n = 4^3$.

$M_{\mathbf{ij}}$	Lowest eigenvalue (a.u.)
$\langle D_{\mathbf{i}} \vert \hat{H} \vert D_{\mathbf{j}} \rangle$	5.631330
$-\vert\langle D_{\mathbf{i}} \vert \hat{H} \vert D_{\mathbf{j}} \rangle\vert$	-25.032719
$-\vert\langle D_{\mathbf{i}} \vert \hat{H} \vert D_{\mathbf{j}} \rangle\vert, \mathbf{i} \neq \mathbf{j}$	5.349003

[3] http://github.com/jsspencer/toy_fci

Convergence to the ground state

$$\frac{dp_{\mathbf{i}}}{d\tau} = \sum_{\mathbf{j}} \left(T^{+}_{\mathbf{ij}} + T^{-}_{\mathbf{ij}} \right) p_{\mathbf{j}} - \kappa (p^2_{\mathbf{i}} - n^2_{\mathbf{i}}) \qquad (22)$$

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

$$0 = \frac{dp_{\mathbf{i}}}{d\tau} \approx \sum_{\mathbf{j}} \kappa (n^2_{\mathbf{i}} - p^2_{\mathbf{i}}) \qquad (23)$$

$\Rightarrow$ basis functions occupied by positive or negative psips.
$\mathbf{n}$: stochastic representation of the ground-state wavefunction
$|\mathbf{n}|$: psip population