Experiments with Alavi's FCI-QMC Method

J.S. Spencer[1,2] W.M.C. Foulkes[1]

[1]Department of Physics
Imperial College London

[2]Department of Materials
Imperial College London

Quantum Monte Carlo in the Apuan Alps VI

26th July 2010

Outline

Outline

The Full Configuration Interaction Method

- Choose a basis of single-particle orbitals ϕ_n, $n = 1, 2, \ldots, M$.
- Construct all possible N-electron determinants from the corresponding set of $2M$ spin-orbitals.

$$
\begin{aligned}
D_{\boldsymbol{i}} &= D_{i_1, i_2, \ldots, i_N} = \frac{1}{\sqrt{N!}} \left| \phi_{i_1} \phi_{i_2} \ldots \phi_{i_N} \right| \\
&= \frac{1}{\sqrt{N!}} \begin{vmatrix}
\phi_{i_1}(1) & \phi_{i_1}(2) & \ldots & \ldots & \phi_{i_1}(N) \\
\phi_{i_2}(2) & \phi_{i_2}(2) & \ldots & \ldots & \phi_{i_2}(N) \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
\phi_{i_N}(1) & \phi_{i_N}(2) & \ldots & \ldots & \phi_{i_N}(N)
\end{vmatrix}
\end{aligned}
$$

The Full Configuration Interaction Method (cont.)

- Use the determinants as a basis:

$$|\Psi\rangle = \sum_i c_i |D_i\rangle$$

- The coefficients c_i that minimise

$$E(\{c_i\}) = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$$

satisfy the matrix eigevalue problem

$$\sum_j H_{ij} c_j = E_0^{\mathrm{FCI}} c_i$$

where $H_{ij} = \langle D_i|\hat{H}|D_j\rangle$.

Advantages

- Accurate and (in principle) systematically improvable.
- Total energies often higher than fixed-node QMC energies, but cancellation of errors seems to be more reliable.

Disadvantages

- The total number of determinants is huge and rises exponentially with system size.
- For a 2D Hubbard model with 18 sites and 18 electrons, the total number of determinants is 1.3×10^8.
- Diagonalizing $10^8 \times 10^8$ matrices is hard!

Monte Carlo methods are good at high-dimensional problems

Outline

The Full-CI QMC Algorithm

- The lowest eigenvector of the full-CI matrix eigenproblem

$$\sum_j H_{ij} c_j = E_0^{\mathrm{FCI}} c_i$$

may be obtained by solving the imaginary-time Schrödinger equation

$$\begin{aligned} \frac{dc_i(t)}{dt} &= -\sum_j \left(H_{ij} - (E_{\mathrm{HF}} + S)\delta_{ij}\right) c_j(t) \\ &= -\sum_j \left(K_{ij} - S\delta_{ij}\right) c_j(t) \end{aligned}$$

- Since $H_{ii} = \langle D_i | \hat{H} | D_i \rangle \geq E_{\mathrm{HF}}$, it follows that $K_{ii} \geq 0$.

Sampling Algorithm

Imagine a population of signed walkers moving through the configuration space, time step by time step. Suppose that generation n includes a walker on D_i. In generation $n+1$, this walker gives rise to:

1. An expected number of walkers $(1-(K_{ii}-S)\Delta t)$ on D_i. These have the same sign as the parent.
2. An expected number of walkers $|K_{ji}\,dt|$ on D_j.
   - If $-K_{ji} > 0$, these walker(s) have the same sign as their parent
   - If $-K_{j,i} < 0$, they have the opposite sign

Then

$$\langle q_i \rangle_{n+1} \;=\; (1-(K_{ii}-S)\Delta t)\langle q_i \rangle_n - \sum_{j \neq i} (K_{ij}\Delta t)\langle q_j \rangle_n$$

Sampling Algorithm (cont.)

Rearranging gives

$$\langle q_i \rangle_{n+1} = \langle q_i \rangle_n - \Delta t \sum_j (K_{ij} - S\delta_{i,j}) \langle q_j \rangle_n$$

which is a first-order approximation to the imaginary-time Schrödinger equation

$$\frac{dc_i(t)}{dt} = -\sum_j (K_{ij} - S\delta_{ij}) c_j(t)$$

Thus, this walker dynamics solves the imaginary-time Schrödinger equation. It is analogous to DMC without importance sampling or the fixed-node approximation.

Time-Step Errors

- We are iterating with

$$\tilde{\mathbf{G}} = \mathbf{I} - (\mathbf{H} - S\mathbf{I})\Delta t$$

  instead of

$$\mathbf{G} = \exp\left(-(\mathbf{H} - S\mathbf{I})\Delta t\right)$$

- Fortunately, as long as $\Delta t \leq 2/(E_{\max} - S)$, both methods yield the exact ground state.

There is no time-step error

The Local Energy Estimator

- FCIQMC has an analogue of the local energy estimator

$$\begin{aligned} E(t) &= \frac{\langle e^{-t\hat{H}} D_0 | \hat{H} | D_0 \rangle}{\langle e^{-t\hat{H}} D_0 | D_0 \rangle} \\ &= E_{\mathrm{HF}} + \sum_{j \neq 0} \frac{c_j(t) \langle D_j | \hat{H} | D_0 \rangle}{c_0(t)} \\ &= E_{\mathrm{HF}} + \sum_{j \neq 0} \frac{\langle q_j(t) \rangle H_{j0}}{\langle q_0(t) \rangle} \end{aligned}$$

- Note that it is important to average q_j and q_0 separately, taking the ratio afterwards:

$$\left\langle \frac{q_j}{q_0} \right\rangle \neq \frac{\langle q_j \rangle}{\langle q_0 \rangle}$$

Walker Cancellation

As with all fermion QMC methods, there is a sign problem. Walkers of both signs appear on the same configurations and the positive and negative populations almost cancel.

- To help control this problem, positive and negative walkers on the same determinant at the same time are cancelled out.
- Similar walker cancellation algorithms have been tried many times in continuum DMC. They do not work very well.
- The surprise in Alavi's work was that for FCI spaces walker cancellation works **much better**.

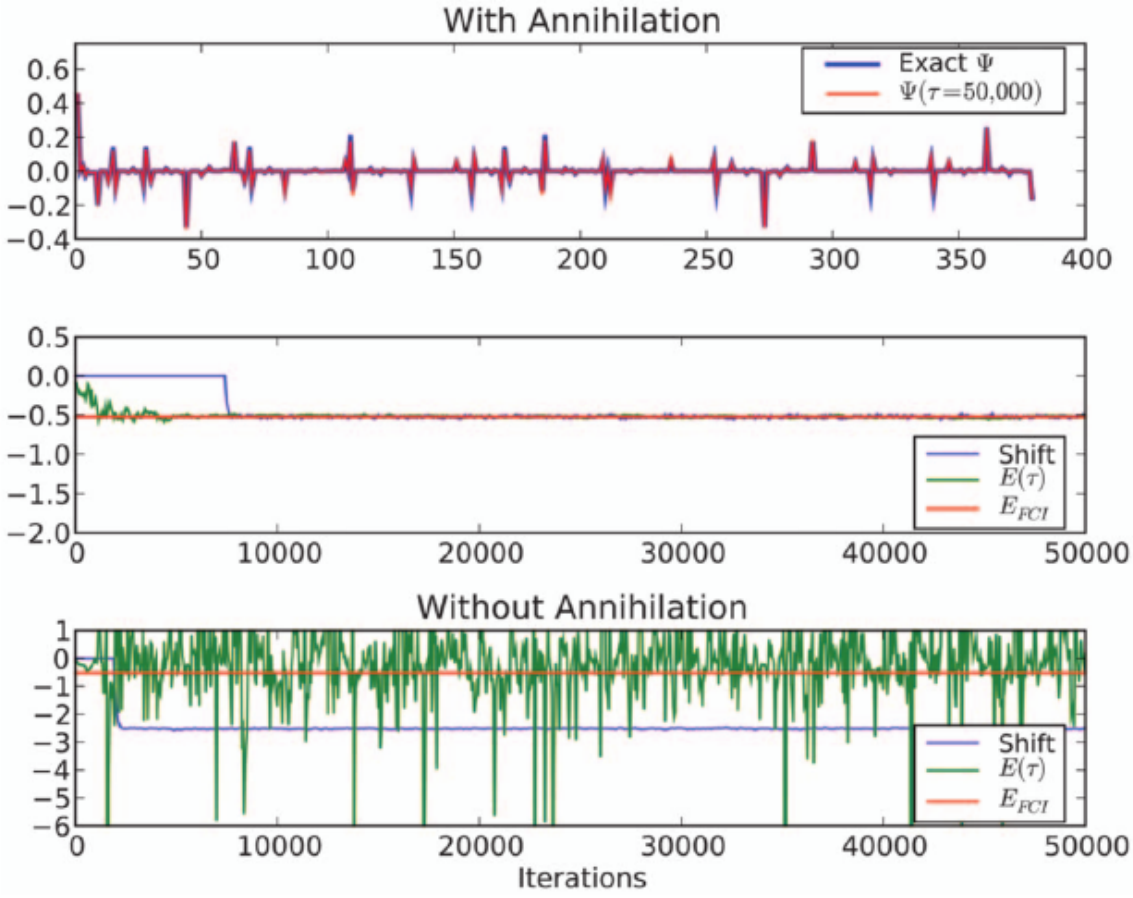


FIG. 1. Comparison of sampling of determinants compared to full diagonalization result for a stretched N_2 molecule in a space of 379 determinants. Ψ indicates the normalized walker number on each of the determinants after 50 000 iterations, comparing it to the exact wave function. The shift and E_0 values for each iteration are shown in the lower plot.

(Booth, Thom, Alavi, J. Chem. Phys. **131**, 054106 (2009))

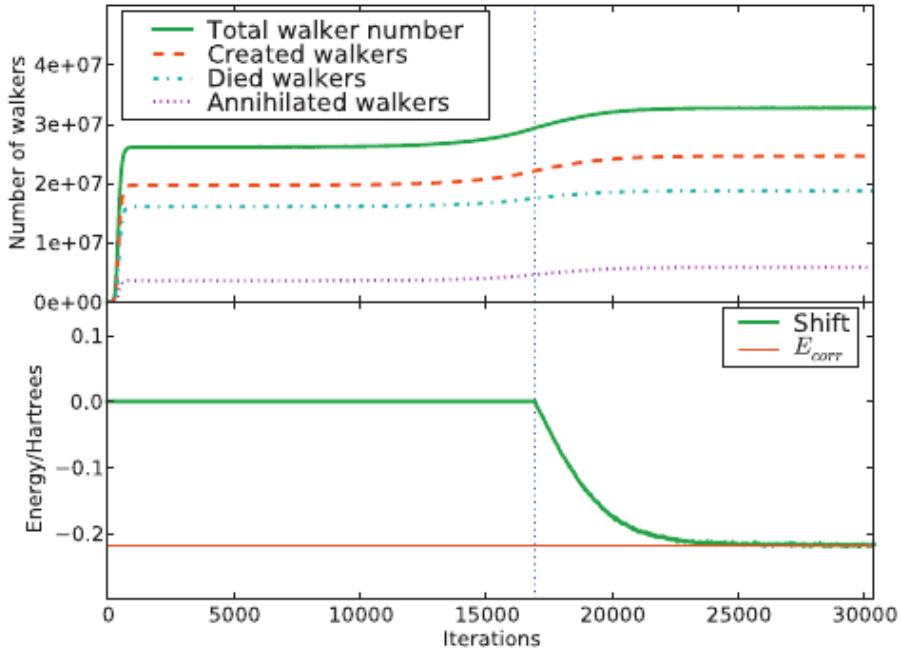


FIG. 2. A typical trend of walker growth for H_2O in a cc-pVDZ basis set in constant shift mode ($S=0$) until iteration 16 970, showing the appearance of a well-defined plateau in N_w, in this case at $N_c=26\times10^6$ walkers. For this system, $N_{FCI}=451\times10^6$. The number of walkers at the plateau is therefore about 6% of the FCI space. Also shown on the plot are the numbers of walkers created, died, and annihilated per $A=10$ iterations. It is evident that at the plateau, the combined rate of death and annihilation matches the birth rate from the spawning. The plateau gradually gives way to a growth phase in walker number, which increases exponentially. Once in the growth phase, the shift is then allowed to vary according to Eq. (17), (in this example on iteration number 16 970). The number of walkers then rapidly stabilizes. The lower plot shows that the value for the shift which stabilizes the walker growth is exactly the correlation energy for the system.

(Booth, Thom, Alavi, J. Chem. Phys. **131**, 054106 (2009))

The Sign-Coherence Phase Transition

- Setting $S = 0$, the population grows exponentially until the numbers of positive and negative walkers on important configurations become large enough for cancellation (which $\propto$ the square of the walker population) to kick in.
- The simulation then enters a plateau region, during which the total number of walkers remains roughly fixed. The "sign-coherence" of the population grows steadily and the population begins to sample the ground state.
- The population then begins to grow again, at which point the shift is adjusted to prevent a population explosion.
- As long as you reach the plateau before running out of memory, you get the right ground-state energy.

Successes and Limitations of FCIQMC

Working on a small parallel machine over a period of a few months, Alavi was able to reproduce essentially all the exact FCI calculations ever done, to full precision. He was also able to do some FCI calculations that had been impossible before.

But what if the critical number of walkers is too large?

Outline

Initiator FCIQMC

i-FCIQMC

is a systematically improvable approximate approach that works very well in many small molecules.

- Divide determinants into two classes:
  - Initiator determinants are allowed to create new walkers on unoccupied determinants.
  - Ordinary determinants are only allowed to create new walkers on occupied determinants.
  - The fixed set of initiators is normally chosen according to a CAS criterion.
  - If an ordinary determinant acquires a population $\geq n_{\text{add}}$, it becomes an initiator and stays an initiator until its population drops below n_{add}.

Properties of i-FCIQMC

- Helps the walker population choose a global sign and stabilises the FCIQMC algorithm at lower walker number.
- Becomes exact as:
  - the number of walkers $\rightarrow \infty$.
  - the set of initiator determinants is enlarged.
- Otherwise, i-FCIQMC is biased.
- In m<any cases, the bias appears to be small.

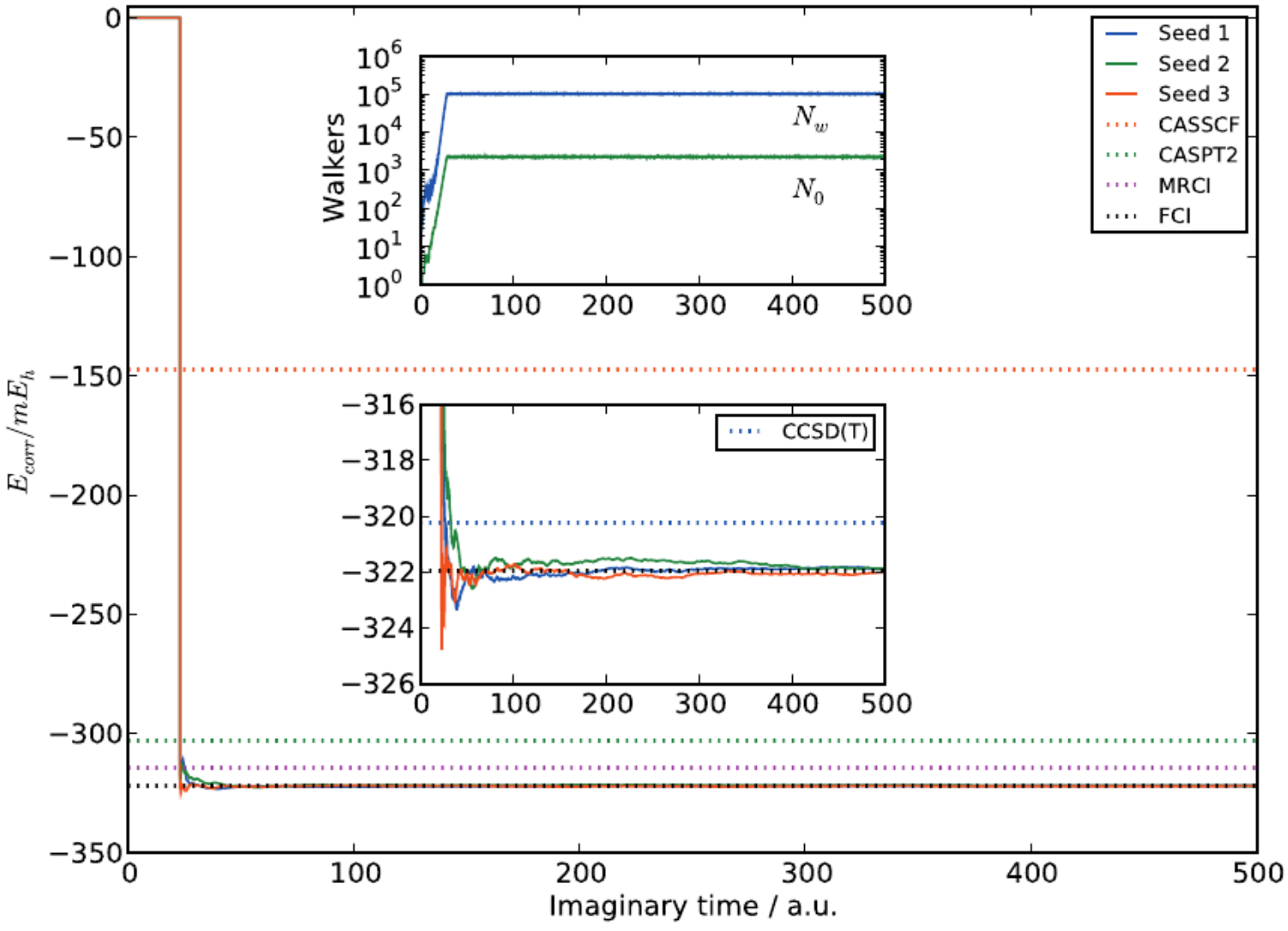
N_2 Example

N_2 molecule, cc-pVDZ basis

- FCIQMC requires 2.7×10^8 walkers — about half the FCI space.
- i-FCIQMC with a 396-determinant fixed initiator space and $n_{\text{add}} = 3$ requires only $\sim 10^5$ walkers.

(Cleland, Booth and Alavi, J. Chem. Phys. **132**, 041103 (2010))

N_2 Example



(Cleland, Booth and Alavi, J. Chem. Phys. **132**, 041103 (2010))

Our Aims

- System-size scaling?
- Phase transition?
- Better algorithms?

Our Aims

By studying the simplest system we can think of — the Hubbard model — we plan to investigate these questions as quickly and easily as possible.

(Can we do the world's biggest Hubbard model?)

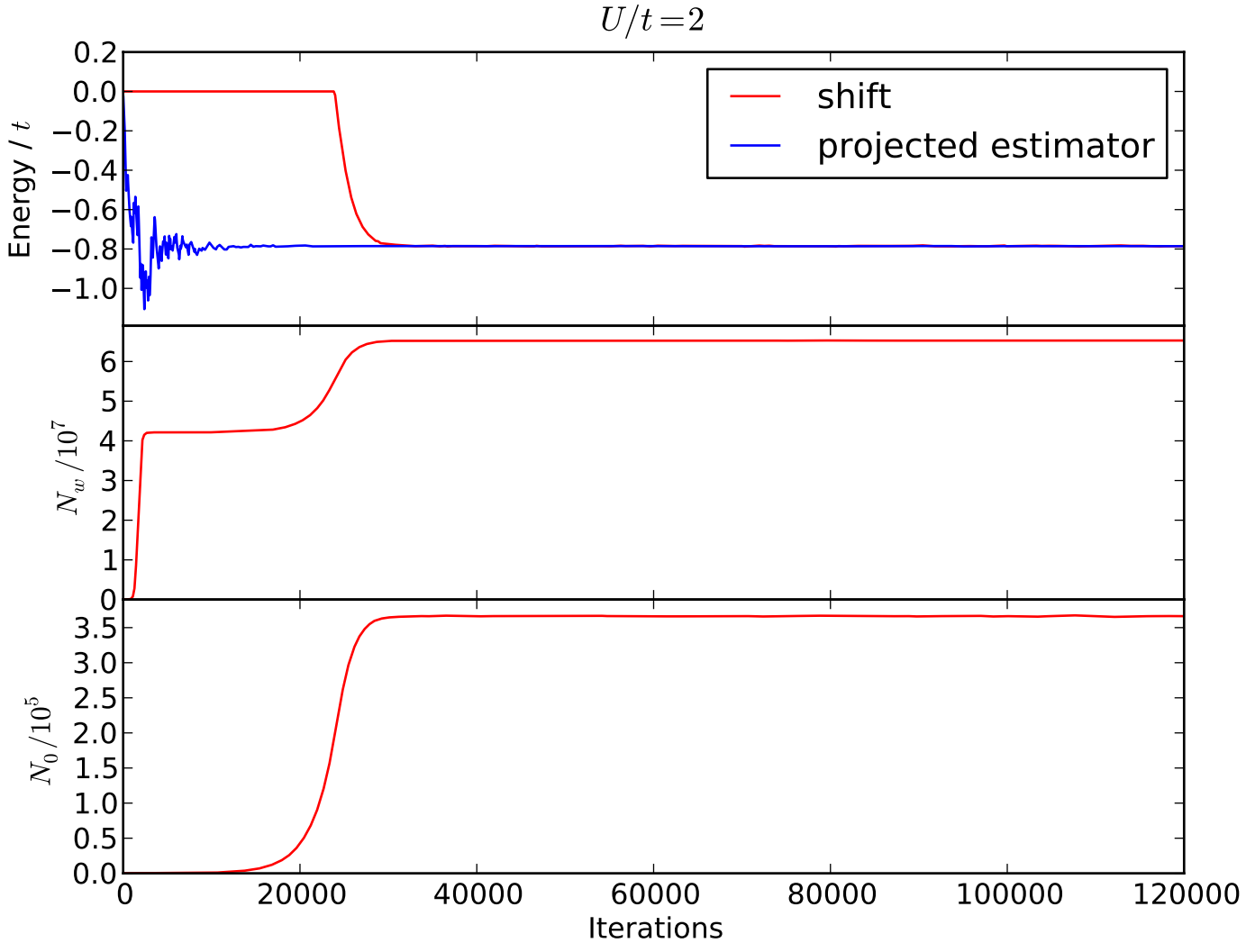
Outline

Results

So far . . .

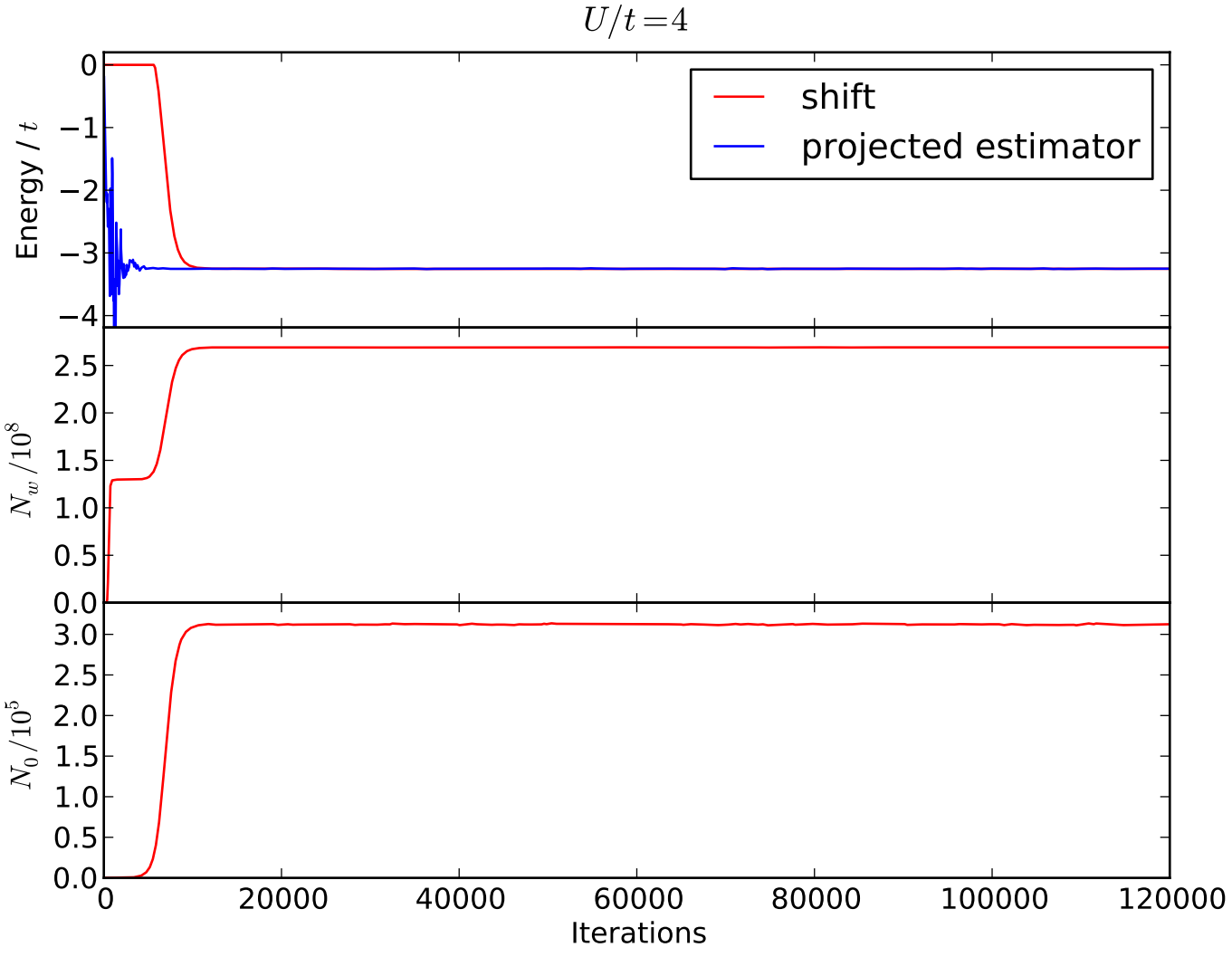
- We have written an (i-)FCIQMC code for the Hubbard model from scratch.
- We have tested that it works and observe similar behaviour to Alavi.
- . . . except that the Hubbard model seems much less favourable than molecules.

FCIQMC: 18-Site 18-Electron 2D Hubbard Model with $U/t = 2$



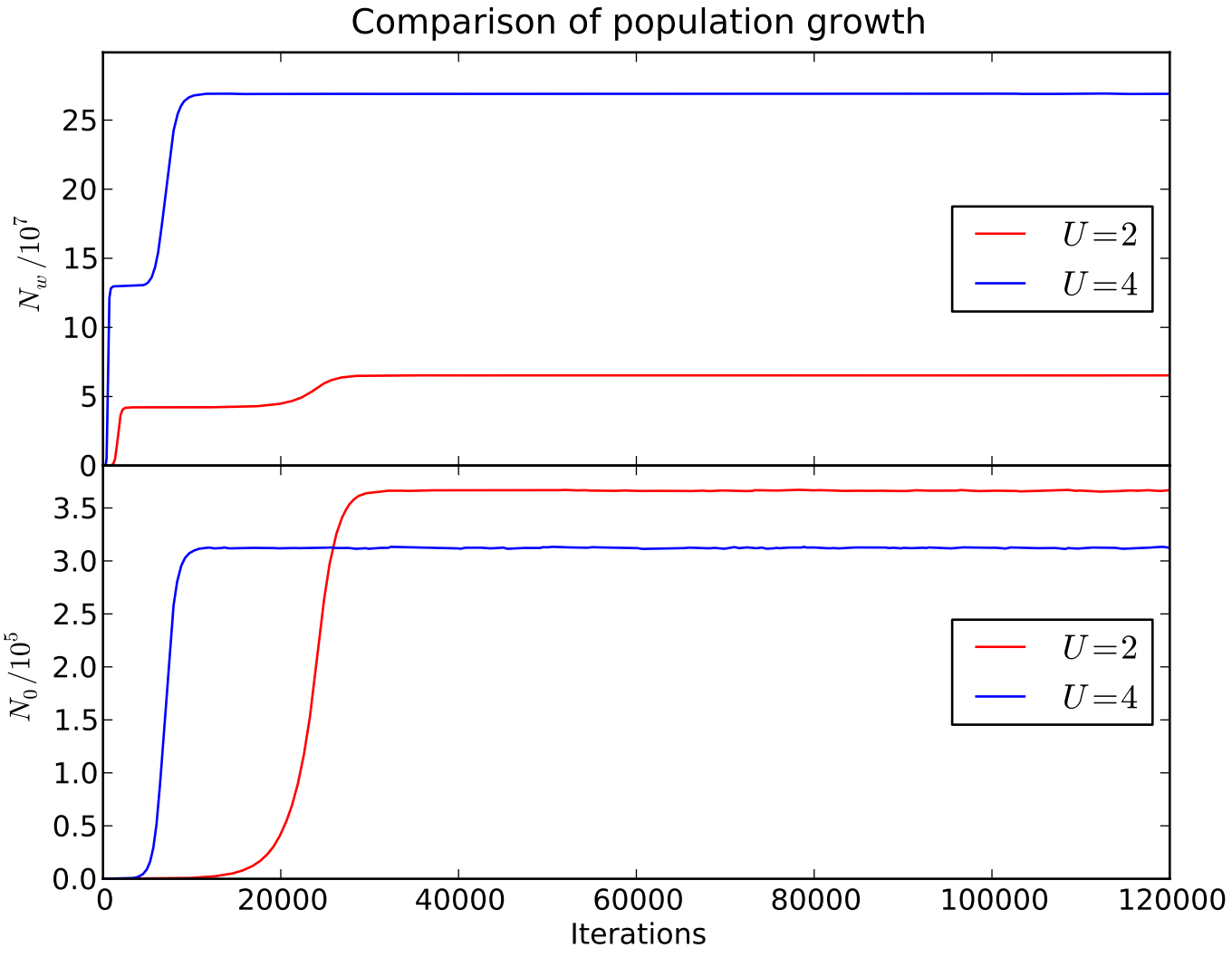
(Hilbert space: 1.3×10^8)

FCIQMC: 18-Site 18-Electron 2D Hubbard Model with $U/t = 4$



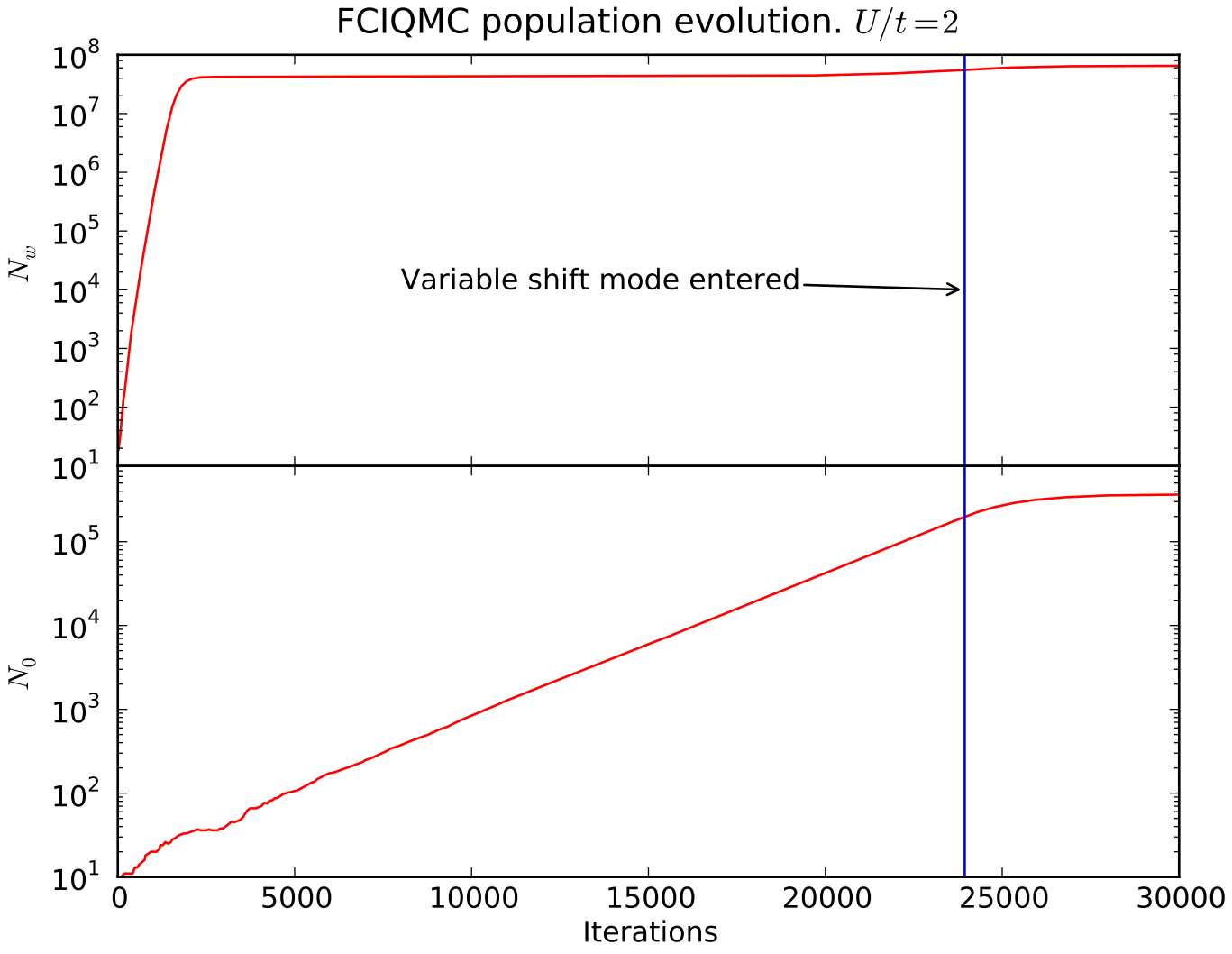
(Hilbert space: 1.3×10^8)

FCIQMC: Comparsion of $U/t = 2$ and $U/t = 4$ Populations



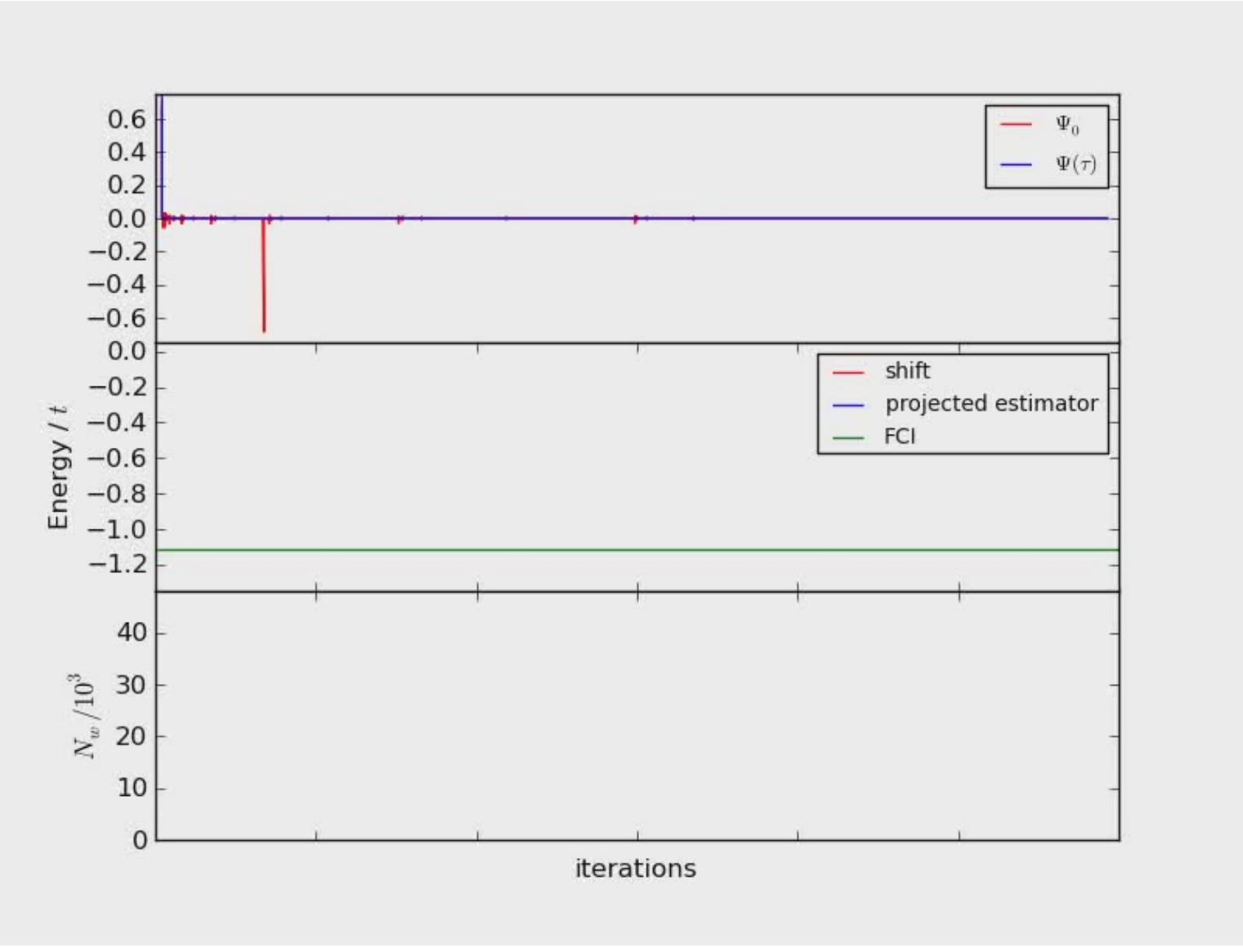
(Hilbert space: 1.3×10^8)

FCIQMC: The Plateau Region



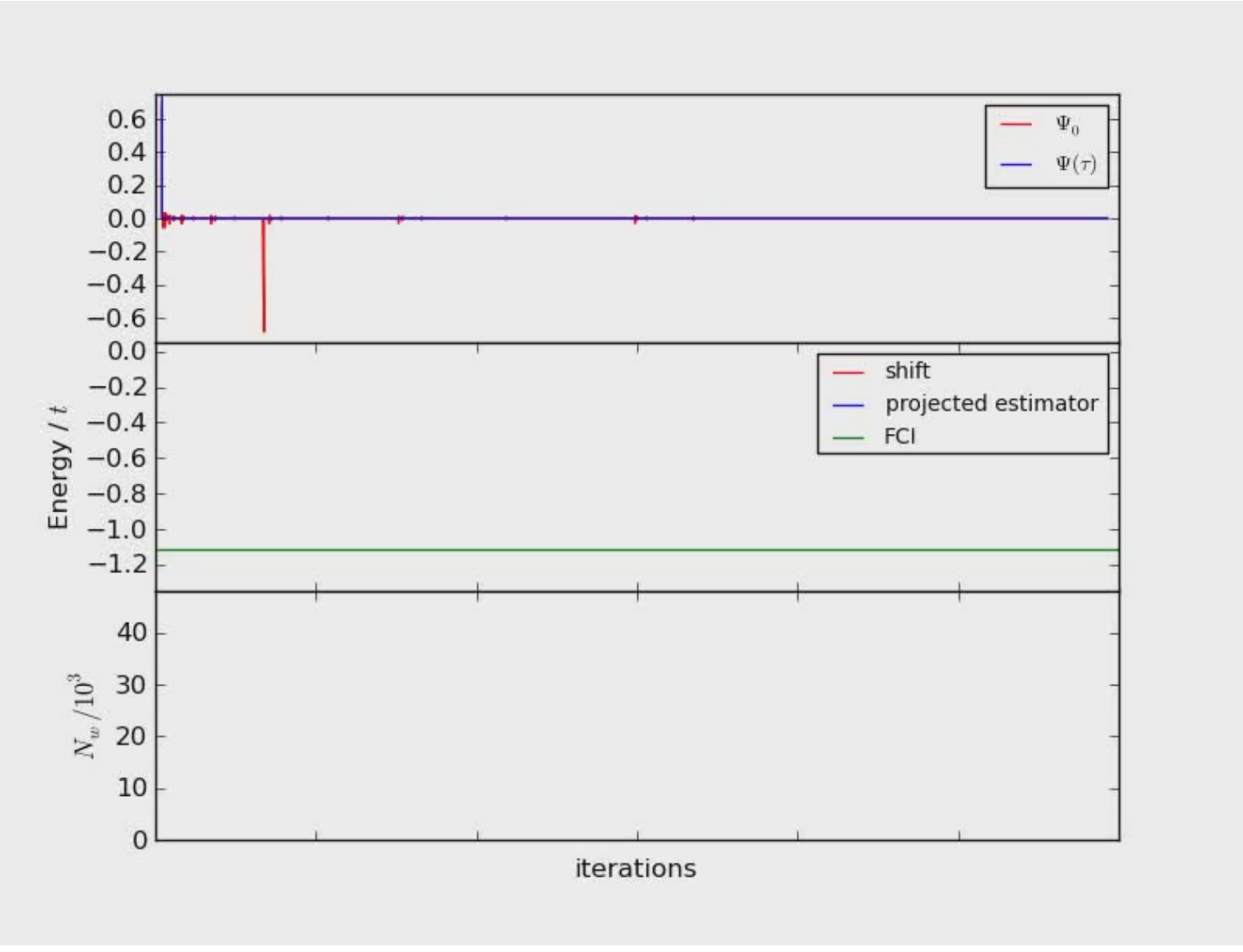
(Hilbert space: 1.3×10^8)

FCIQMC: 12-Site 12-Electron 1D Hubbard Model with $U/t = 2$



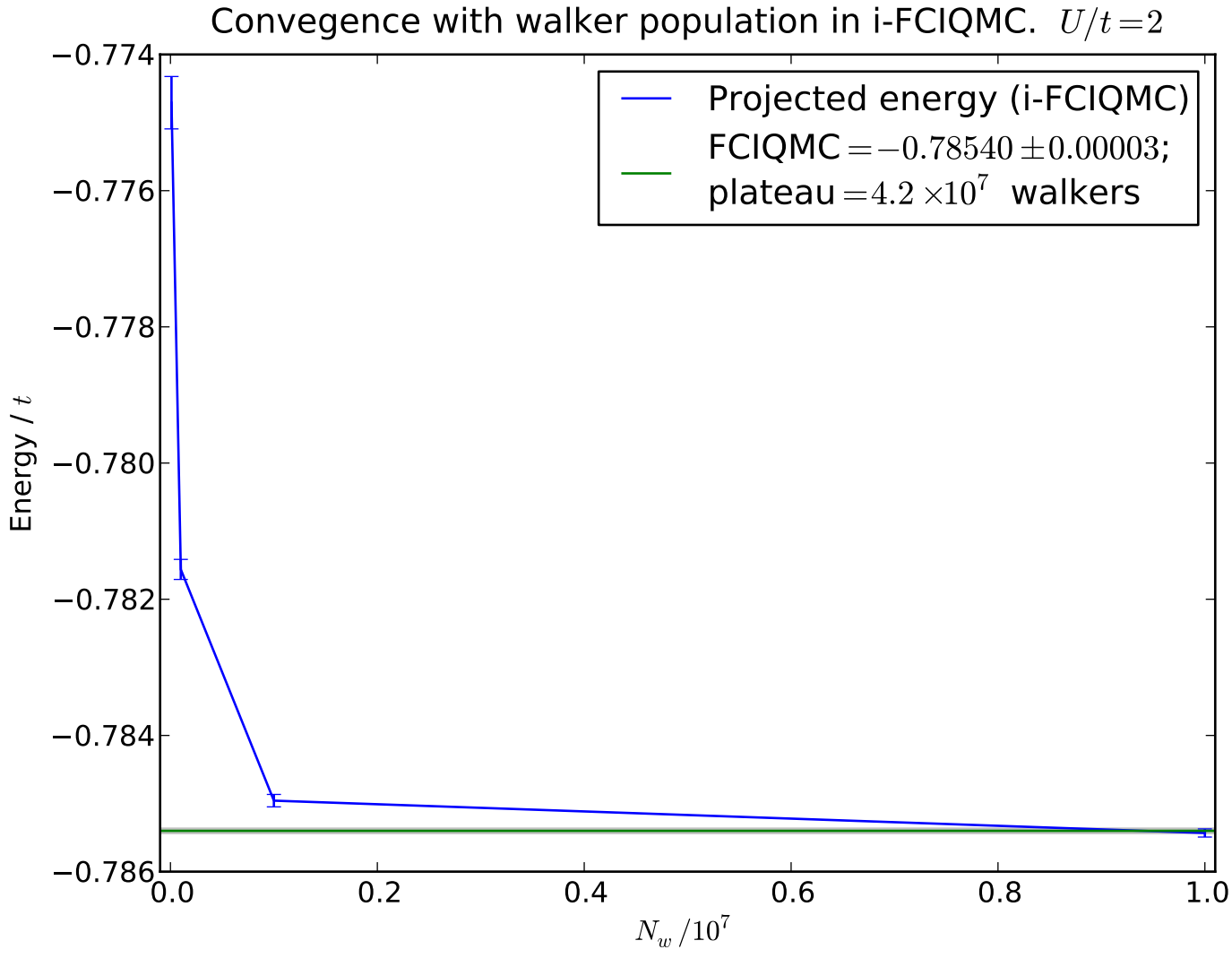
(Hilbert space: 71100, $k = \pi/2$)

FCIQMC: 12-Site 12-Electron 1D Hubbard Model with $U/t = 2$



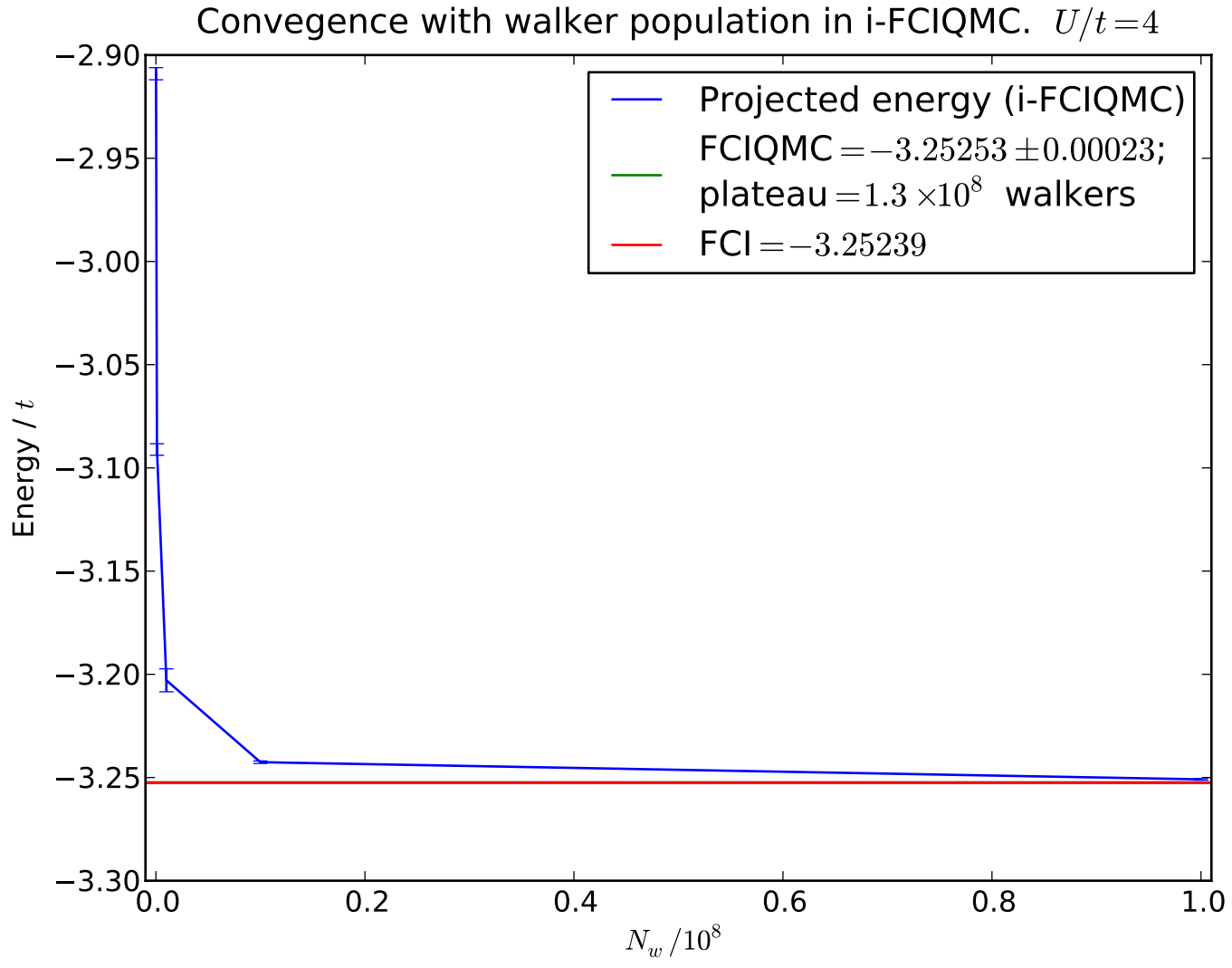
(Hilbert space: 71100, $k = \pi/2$)

i-FCIQMC: 18-Site 18-Electron 2D Hubbard Model with $U/t = 2$



(Hilbert space: 1.3×10^8)

i-FCIQMC: 18-Site 18-Electron 2D Hubbard Model with $U/t = 4$



(Hilbert space: 1.3×10^8)

Outline

Current Work

Expectation Values

$$\begin{aligned}\langle\Psi_0|\hat{O}|\Psi_0\rangle &= \frac{\lim_{t\to\infty}\sum_{i,j} c_i(t) O_{ij} c_j(t)}{\lim_{t\to\infty}\sum_i c_i c_i} \\ &= \frac{\lim_{t\to\infty}\sum_{i,j}\langle q_i\rangle O_{ij}\langle q_j\rangle}{\lim_{t\to\infty}\sum_i\langle q_i\rangle\langle q_i\rangle}\end{aligned}$$

- Unfortunately, we do not have enough memory to accumulate and store all the $\langle q_i\rangle$.
- On-the-fly averaging is no good because

$$\langle q_i\rangle O_{ij}\langle q_j\rangle \neq \langle q_i O_{ij} q_j\rangle$$

Current Work

Hellmann-Feynman Sampling

- We are investigating an adaptation of Gaudoin's Hellmann-Feynman sampling algorithm (R. Gaudoin and J.M. Pitarke, Phys. Rev. Lett. **99**, 126406 (2007)).
- FCIQMC version is not restricted to diagonal operators (although they are easier).
- We think it will work.

Current Work

Other Basis Sets

- FCIQMC is not restricted to the basis of determinants of HF orbitals we have been using so far.
- We can also use a basis of determinants of local orbitals — closer in spirit to DMC and surely more efficient in the strongly correlated limit?
- Currently very inefficient, but might combine well with the next idea . . .

Current Work

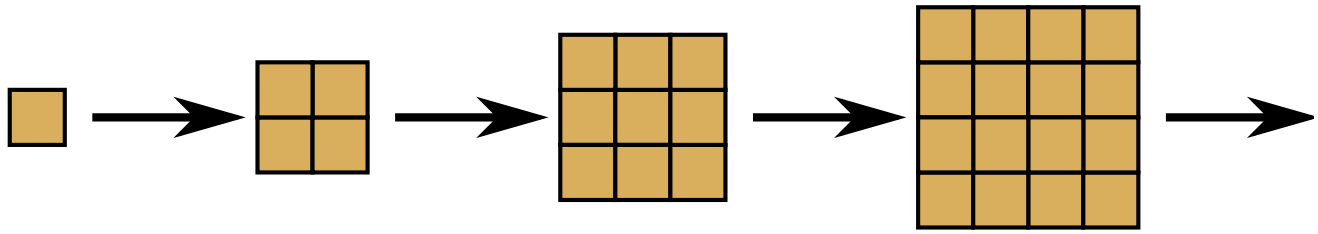
Importance Sampling

- $\Psi(\mathbf{R}, t)\Psi^T(\mathbf{R}) \rightarrow c_{\boldsymbol{i}}(t)c_{\boldsymbol{i}}^T$.
- Carrying a trial wavefunction around with each walker is problematic if you have 10^8 walkers. Importance sampling might not be as useful as in DMC.
- But it might help establish sign coherence. If $\Psi^T = \Psi_0$, then $\lim_{t\rightarrow\infty} c_{\boldsymbol{i}}(t)c_{\boldsymbol{i}}^T$ has the same sign on every configuration.

Current Work

System-Size Scaling

- How does the critical number of walkers scale with system size?



Current Work

Better Sign Cancellation Algorithms

- Can we adapt continuum sign-cancellation techniques developed by Kalos, Anderson, and others to improve sign cancellation in FCIQMC?

Continuous Time Algorithms

- Since set of allowed moves is finite, can sample from Poisson distribution and move straight to next event.

Summary

- FCIQMC contains almost no completely new ideas but works much better than most of us would have guessed.
- In many cases, it is already the best full-CI method available.
- Still early days; may get considerably better.

- Useful for solids? Some people think so; I am sceptical.
- We can't solve the largest Hubbard model ever.