Density Matrix Quantum Monte Carlo

N.S. Blunt[1] T.W. Rogers[1] 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 VII
30th July 2012

Outline

Outline

The Full Configuration Interaction Method

- Expand the wave function in a basis of Slater determinants

$$|\Psi\rangle = \sum_{i} c_i |D_i\rangle$$

- 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

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

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

The FCIQMC Algorithm

The lowest eigenvector of the full-CI matrix eigenproblem

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

can be obtained by solving the imaginary-time Schrödinger equation

$$\frac{\partial c_i(t)}{\partial t} = -\left(H_{ij} - S\delta_{ij}\right) c_j(t) = T_{ij}\, c_j(t)$$

Sampling Algorithm

$$\frac{\partial c_i(t)}{\partial t} == T_{ij}\, c_j(t)$$

- Imagine a population of signed "psips" scattered through the configuration space.
- In one time step, a psip on D_j may spawn children on any configuration D_i for which T_{ij} is non-zero.
- The *expected* number of children spawned on D_i is $|T_{ij}\Delta t|$.
  - If $T_{ij} > 0$, the children have the same sign as their parent.
  - If $T_{ij} < 0$, they have the opposite sign.

Then

$$\langle q_i \rangle_{n+1} \;=\; \langle q_i \rangle_n + (T_{ij}\Delta t)\langle q_j \rangle_n$$

Sampling Algorithm (cont.)

$$\langle q_i \rangle_{n+1} = \langle q_i \rangle_n + (T_{ij} \Delta t) \langle q_j \rangle_n$$

This is a first-order finite-difference approximation to the imaginary-time Schrödinger equation

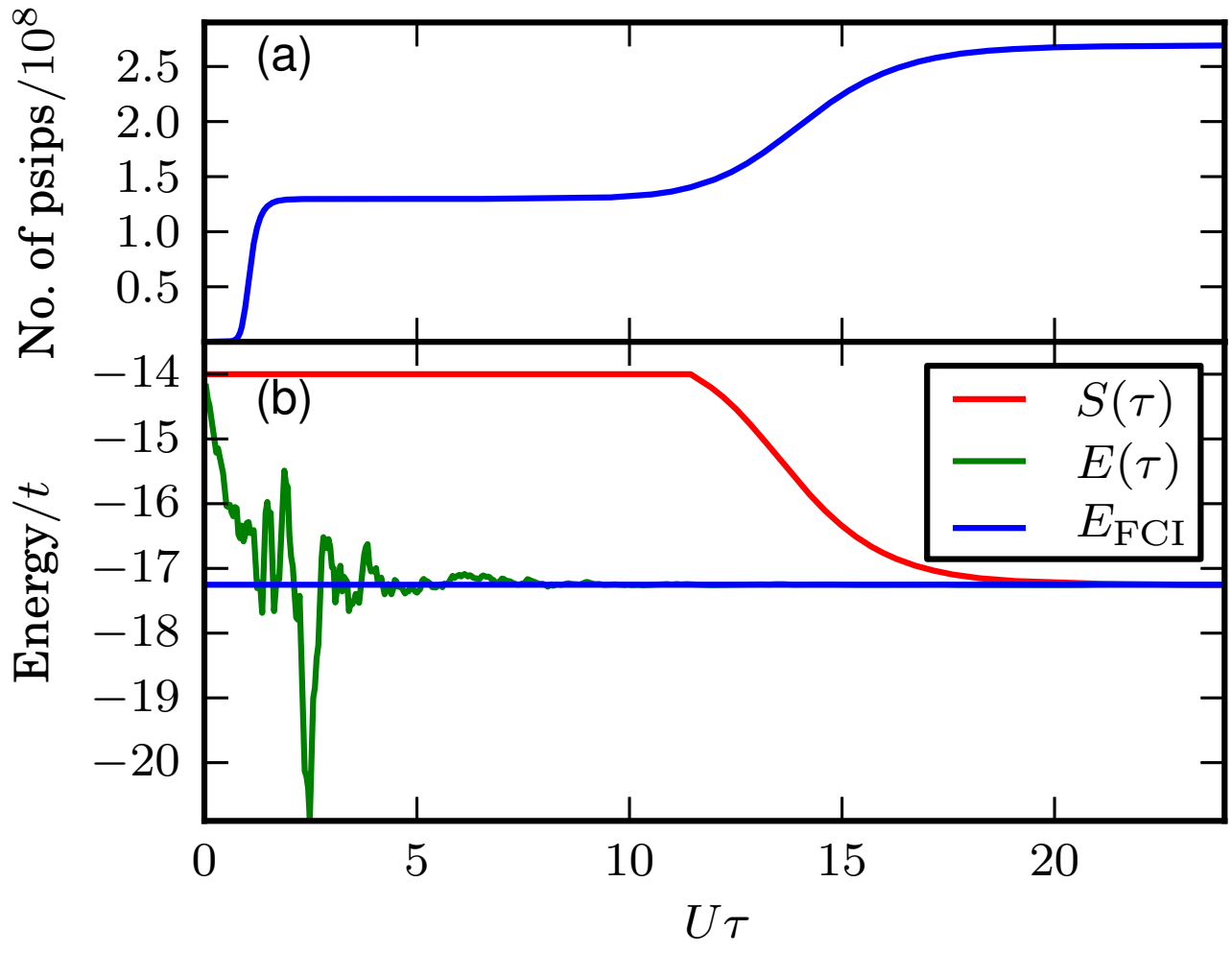
$$\frac{\partial c_i(t)}{\partial t} = T_{ij}\, c_j(t)$$

The FCIQMC psip dynamics solves the imaginary-time Schrödinger equation.

Psip Cancellation

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

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



2D 18-site Hubbard model at half filling, $U = 4t$.

(Spencer, Blunt, and Foulkes, J. Chem. Phys. **136**, 054110 (2012))

Outline

DMQMC

- The density operator $\hat{\rho}(\beta) = \exp[-\beta(\hat{H} - S\hat{I})]$ satisfies the Bloch equation

$$\frac{\partial \hat{\rho}}{\partial \beta} = -(\hat{H} - S\hat{I})\hat{\rho} = \hat{T}\hat{\rho}$$

DMQMC

- The density operator $\hat{\rho}(\beta) = \exp[-\beta(\hat{H} - S\hat{I})]$ satisfies the Bloch equation

$$\frac{\partial \hat{\rho}}{\partial \beta} = -(\hat{H} - S\hat{I})\hat{\rho} = \hat{T}\hat{\rho}$$

$$\frac{\partial \rho_{ij}}{\partial \beta} = T_{ik}\,\rho_{kj}$$

DMQMC

- The density operator $\hat{\rho}(\beta) = \exp[-\beta(\hat{H} - S\hat{I})]$ satisfies the Bloch equation

$$\frac{\partial \hat{\rho}}{\partial \beta} = -(\hat{H} - S\hat{I})\hat{\rho} = \hat{T}\hat{\rho}$$

$$\frac{\partial \rho_{ij}}{\partial \beta} = T_{ik}\,\rho_{kj}$$

- This looks rather like the imaginary-time Schrödinger equation.

$$\frac{\partial c_i}{\partial \beta} = T_{ij}\,c_j$$

Idea

Can we solve the Bloch equation using an FCIQMC-like technique that works in the space of operators/matrices instead of in the space of states/vectors?

Sampling Algorithm

$$\frac{\partial \rho_{ij}}{\partial \beta} = T_{ik}\rho_{kj}$$

- Imagine a population of *two-index* psips scattered through the configuration space of *operators* $|k\rangle\langle j|$.
- In one time step, a psip on $|k\rangle\langle j|$ may spawn children on any operator $|i\rangle\langle j|$ for which T_{ik} is non-zero.
- The expected number of children spawned on $|i\rangle\langle j|$ is $|T_{ik}\Delta t|$.
  - If $T_{ik} > 0$, the children have the same sign as their parent.
  - If $T_{ik} < 0$, they have the opposite sign.
- Psips of opposite sign on the same operator cancel.

Sampling Algorithm (cont.)

$$\frac{\partial \rho_{ij}}{\partial \beta} = T_{ik}\rho_{kj}$$

- Psips spawn along columns of the density matrix only. However, since $[\hat{\rho}, \hat{H}] = 0$, we can equally well solve the symmetrized Bloch equation

$$\frac{\partial \rho_{ij}}{\partial \beta} = \frac{1}{2}\left(T_{ik}\rho_{kj} + \rho_{ik}T_{kj}\right)$$

  Now psips spawn along row and columns.
- $\rho_{kj}(\beta = 0) = \delta_{kj}$, so psips are initially scattered at random along the diagonal of the density matrix.

FCIQMC and DMQMC Compared

The DMQMC equation of motion

$$\frac{\partial \rho_{ij}}{\partial t} = \frac{1}{2}\left(T_{ik}\rho_{kj} + \rho_{il}T_{lj}\right)$$

can be rewritten in the form

$$\frac{\partial \rho_{ij}}{\partial t} = L_{ij,kl}\rho_{kl}$$

where $L_{ij,kl} = T_{ik}\delta_{jl} + T_{lj}\delta_{ik}$.

- FCIQMC finds the large eigenvalue of the matrix T_{ij}.
- DMQMC finds the largest eigenvalue of the matrix $L_{ij,kl}$, with ij and kl regarded as composite indices.

Advantages of DMQMC

- Finite-temperature properties accessible.
- Expectation values of operators that do not commute with $\hat{H}$ easily obtainable:

$$\langle \hat{O} \rangle = \frac{O_{ij}\rho_{ji}}{\rho_{kk}}$$

Disadvantages of DMQMC

Problem 1

The dimension of the space of operators is the *square* of the dimension of the space of states.

But . . .

- The number of determinants in the FCI space rises exponentially with the number of particles/sites n, and

$$N_D \propto e^{\alpha n} \qquad \Rightarrow \qquad N_D^2 \propto (e^{\alpha n})^2 = e^{\alpha(2n)}$$

- FCIQMC for an n-site Hubbard model $\sim$ DMQMC for an $n/2$-site Hubbard model. Not so bad.

Problem 2

The simulation goes straight past every inverse temperature β: you cannot stop to accumulate statistics. Necessary to run many β-loops and average.

But . . .

- Every β-loop provides data at all inverse temperatures from $\beta = 0$ ($T = \infty$) to β large ($T \to 0$).
- Independence of β-loops makes statistical analysis easy.

Outline

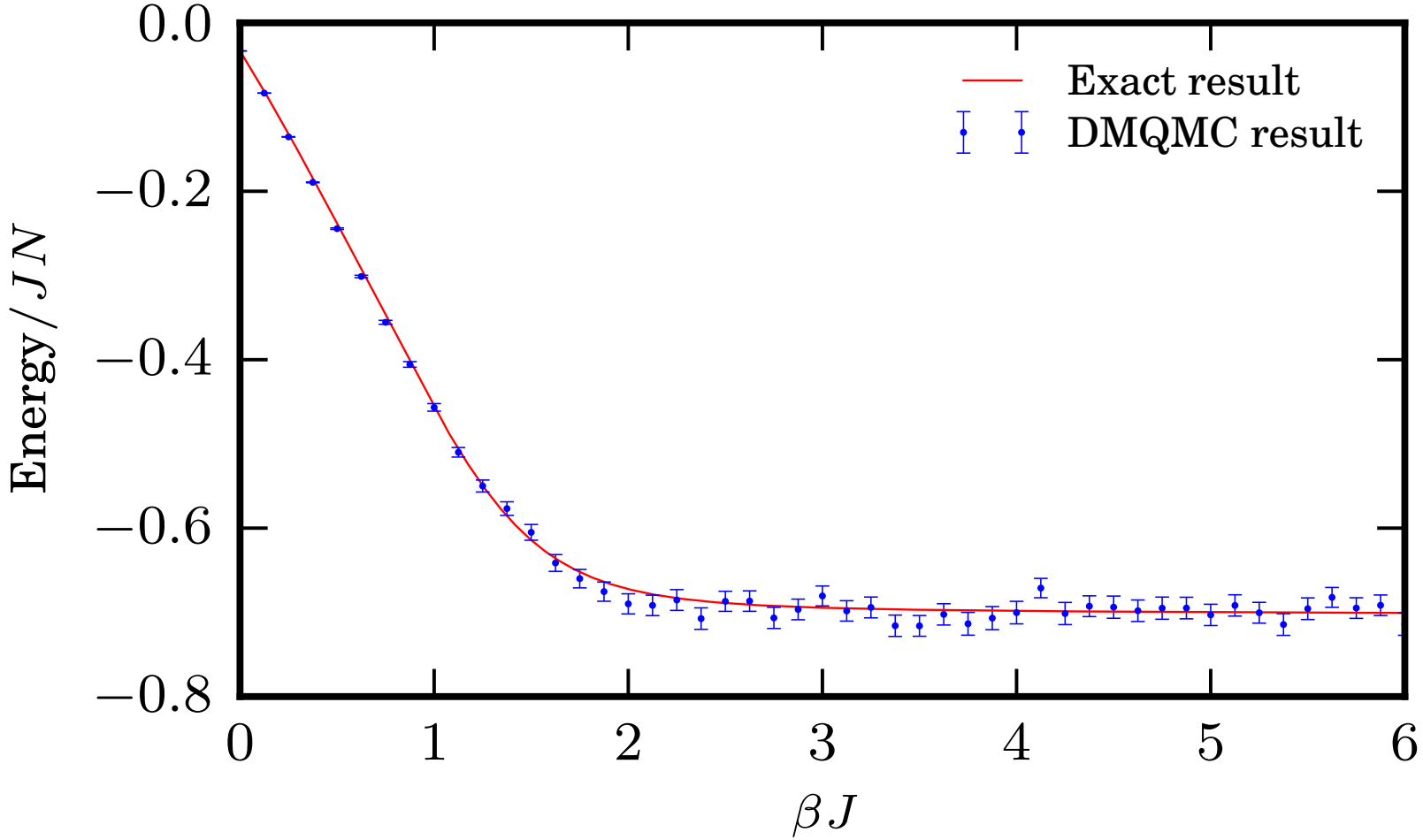
The 2D Heisenberg Model

Antiferromagnetic Heisenberg Hamiltonian ($J > 0$)

$$\hat{H} = J \sum_{\langle ij \rangle} \hat{\boldsymbol{S}}_i \cdot \hat{\boldsymbol{S}}_j = J \sum_{\langle ij \rangle} \left[\hat{S}_{iz} \hat{S}_{jz} + \frac{1}{2} (\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) \right]$$

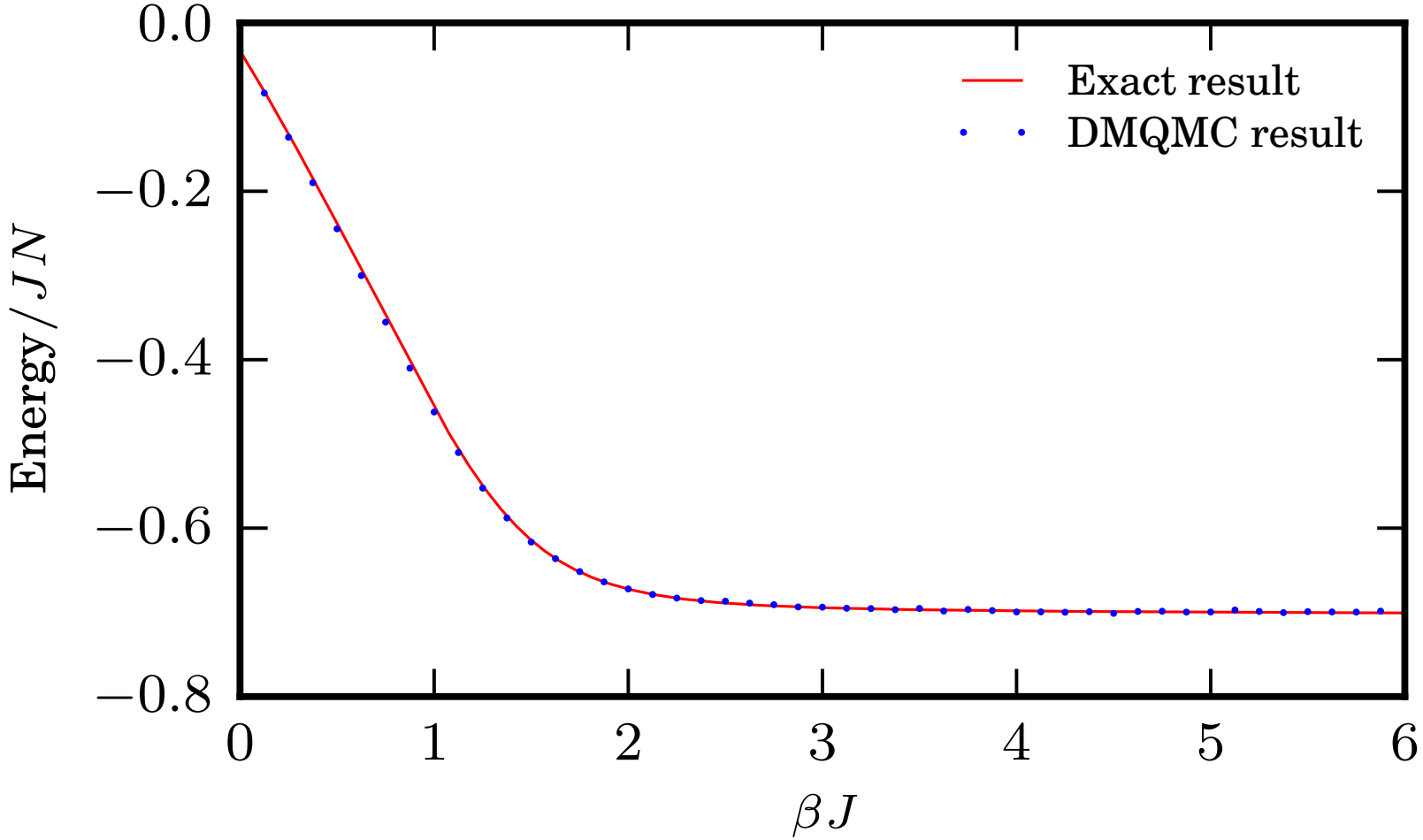
- 2D square lattice is bipartite: no sign problem; no annihilation.
- Ground state *and our simulations* have $M_s = 0$.
- For 4×4 lattice, $M_s = 0$ Hilbert space has dimension $N_C = {}^{16}C_8 = 12870$. Direct diagonalisation possible.
- For 6×6 lattice, $N_C = 9.08 \times 10^9$.

$\langle \hat{H} \rangle$



4×4; 10^3 psips; 10^3 β-loops.

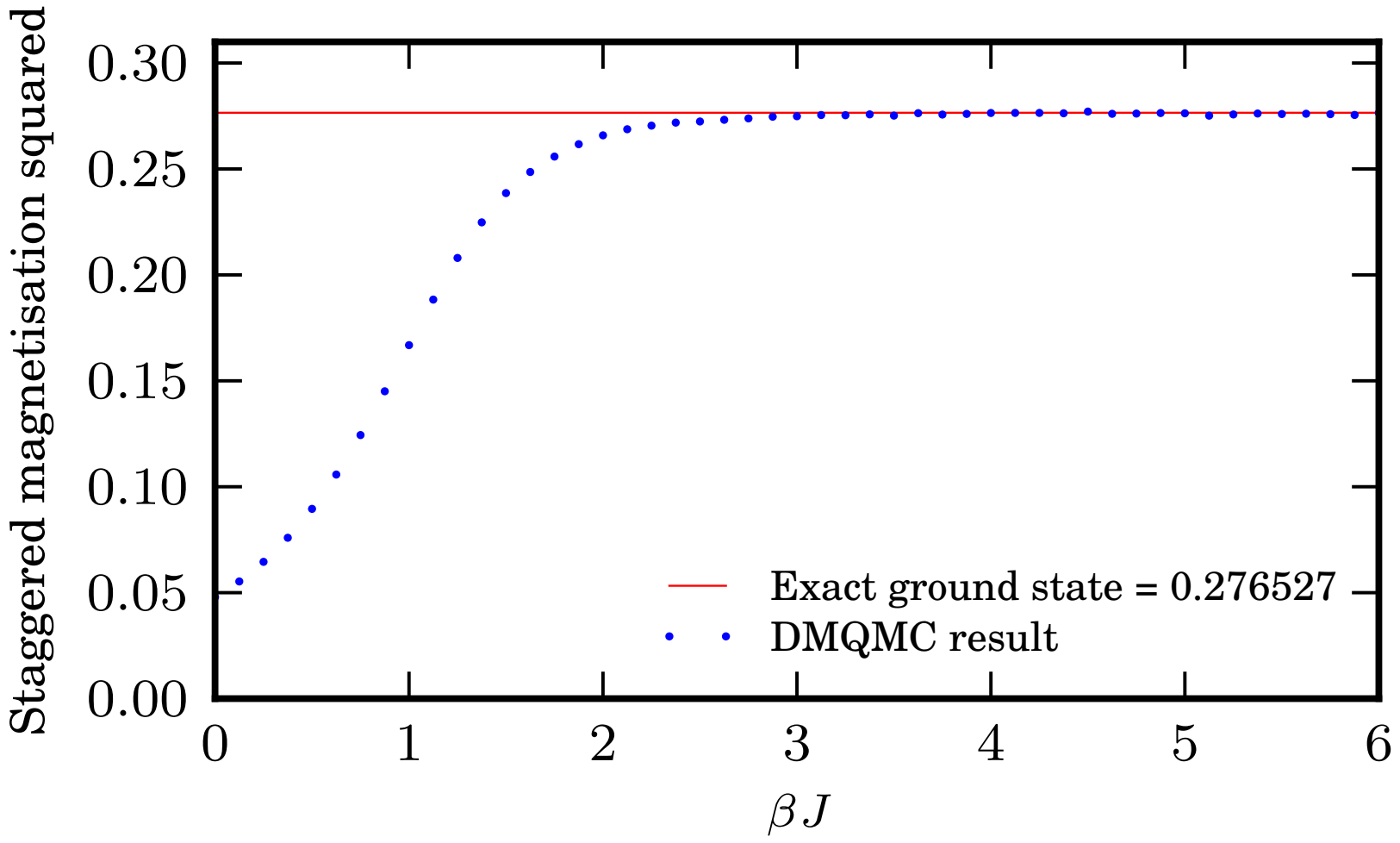
$\langle \hat{H} \rangle$ (cont.)



4×4; 10^5 psips; 10^3 β-loops.

Staggered Magnetization

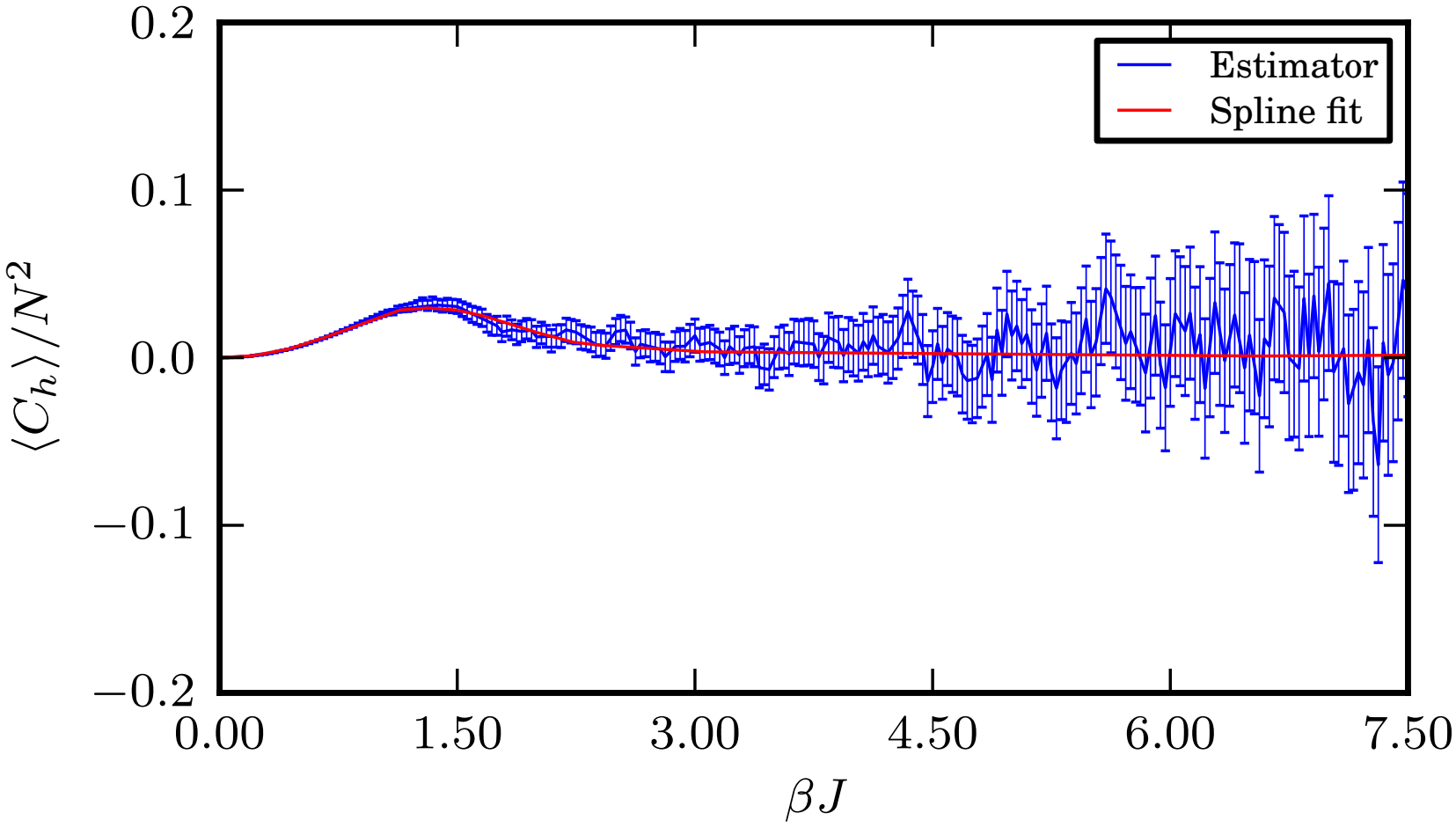
$\langle \hat{M}^2 \rangle$, where $\hat{\boldsymbol{M}} = \frac{1}{N} \sum_i (-1)^{x_i + y_i} \hat{\boldsymbol{S}}_i$



4×4; 10^5 psips; 10^3 β-loops.

Heat Capacity

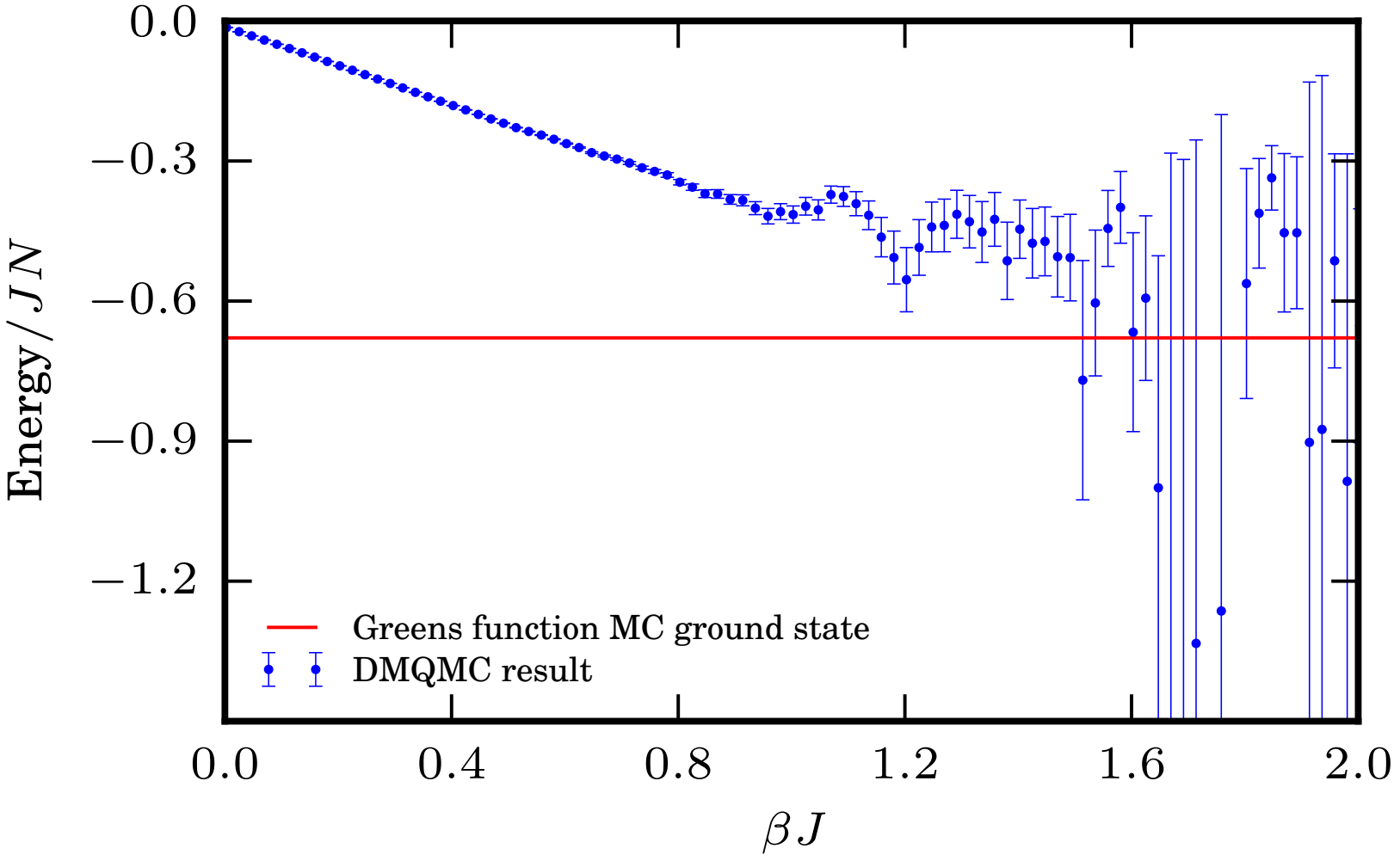
$$C_h = -\beta^2 d\langle \hat{H} \rangle / d\beta = \beta^2 (\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2)$$



4×4; 10^5 psips; 10^3 β-loops.

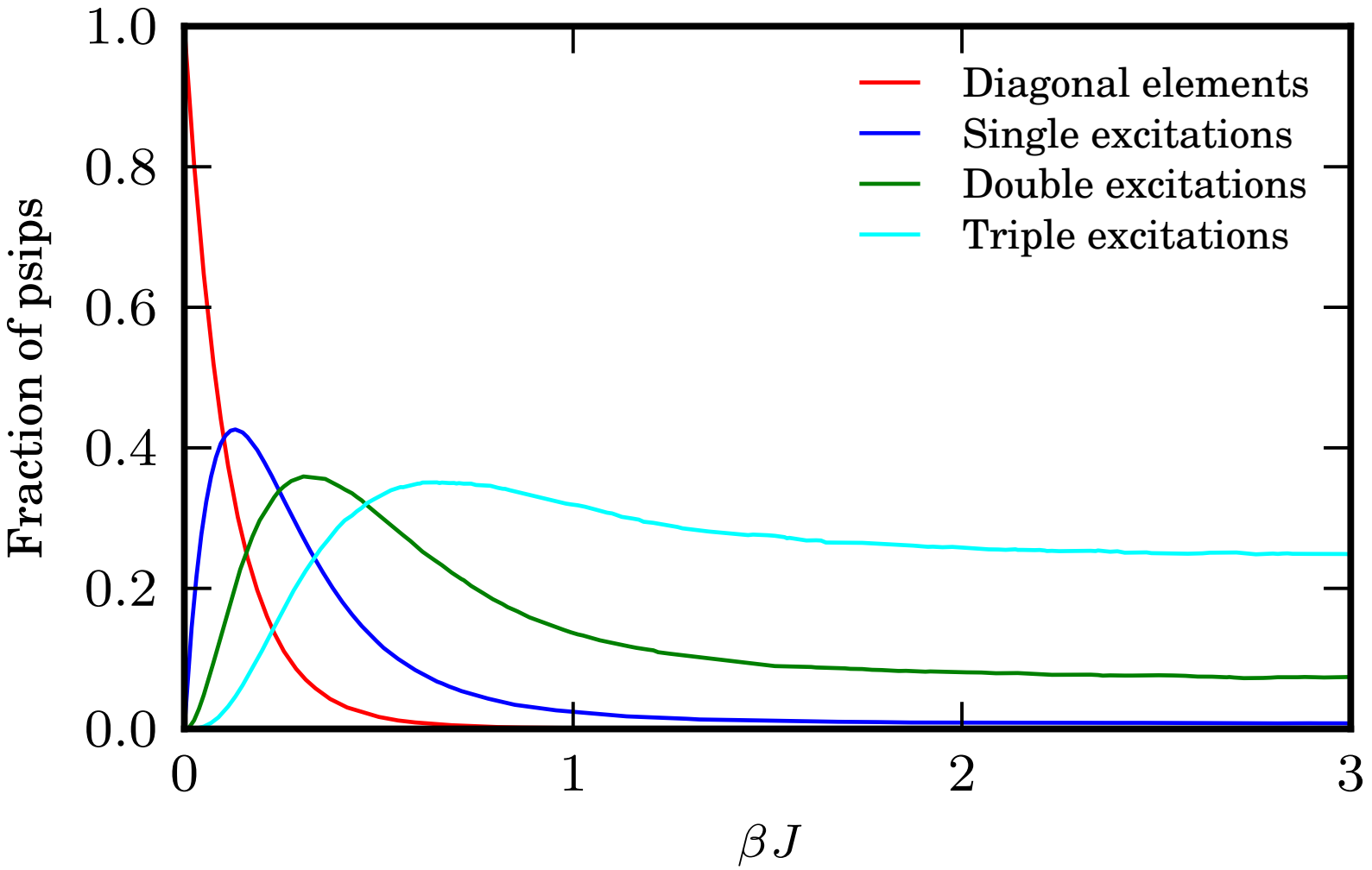
Larger Systems

4 $\times$ **4**: $N_C = 12870$ **6** $\times$ **6**: $N_C = 9.08 \times 10^9$



6 $\times$ 6; 10^5 psips; 5×10^3 β-loops.

Psip Spreading



Psip Spreading (cont.)

- As β increases, psips appear further and further from the diagonal of the density matrix.
- Evaluating

$$\langle \hat{O} \rangle = \frac{O_{ij}\rho_{ji}}{\rho_{kk}}$$

becomes more and more difficult.

Outline

Importance Sampling: Nick's Approach

Idea

Suppress psip population on density matrix elements $\rho_{\boldsymbol{ij}}$ with $d(\boldsymbol{i}, \boldsymbol{j})$ large.

- Multiply spawning rate from operators of excitation level d to operators of excitation level d' by $w_{d',d}$. Multiply reverse spawning rate by $w_{d,d'} = 1/w_{d',d}$.
- If $d' > d$, $w_{d',d} < 1$:
  - suppress spawning to operators further from the diagonal
  - enhance spawning to operators closer to the diagonal

Importance Sampling: Link to Standard Approach

- The psips now sample the modified density matrix

$$\tilde{\rho}_{ij} = w_{d,d-1} w_{d-1,d-2} \cdots w_{1,0} \rho_{ij} = \rho_{ij}^T \rho_{ij}$$

where $d = d(i, j) = d(j, i)$.

- If ρ satisfies

$$\frac{\partial \rho_{ij}}{\partial \beta} = L_{ij,kl} \rho_{kl}$$

then

$$\frac{\partial (\rho_{ij}^T \rho_{ij})}{d\beta} = \left(\rho_{ij}^T L_{ij,kl} \frac{1}{\rho_{kl}^T} \right) \rho_{kl}^T \rho_{kl}$$

Importance Sampling: Link to Standard Approach

- The psips now sample the modified density matrix

$$\tilde{\rho}_{ij} = w_{d,d-1} w_{d-1,d-2} \cdots w_{1,0} \rho_{ij} = \rho_{ij}^{T} \rho_{ij}$$

  where $d = d(i, j) = d(j, i)$.

- If ρ satisfies

$$\frac{\partial \rho_{ij}}{\partial \beta} = L_{ij,kl} \rho_{kl}$$

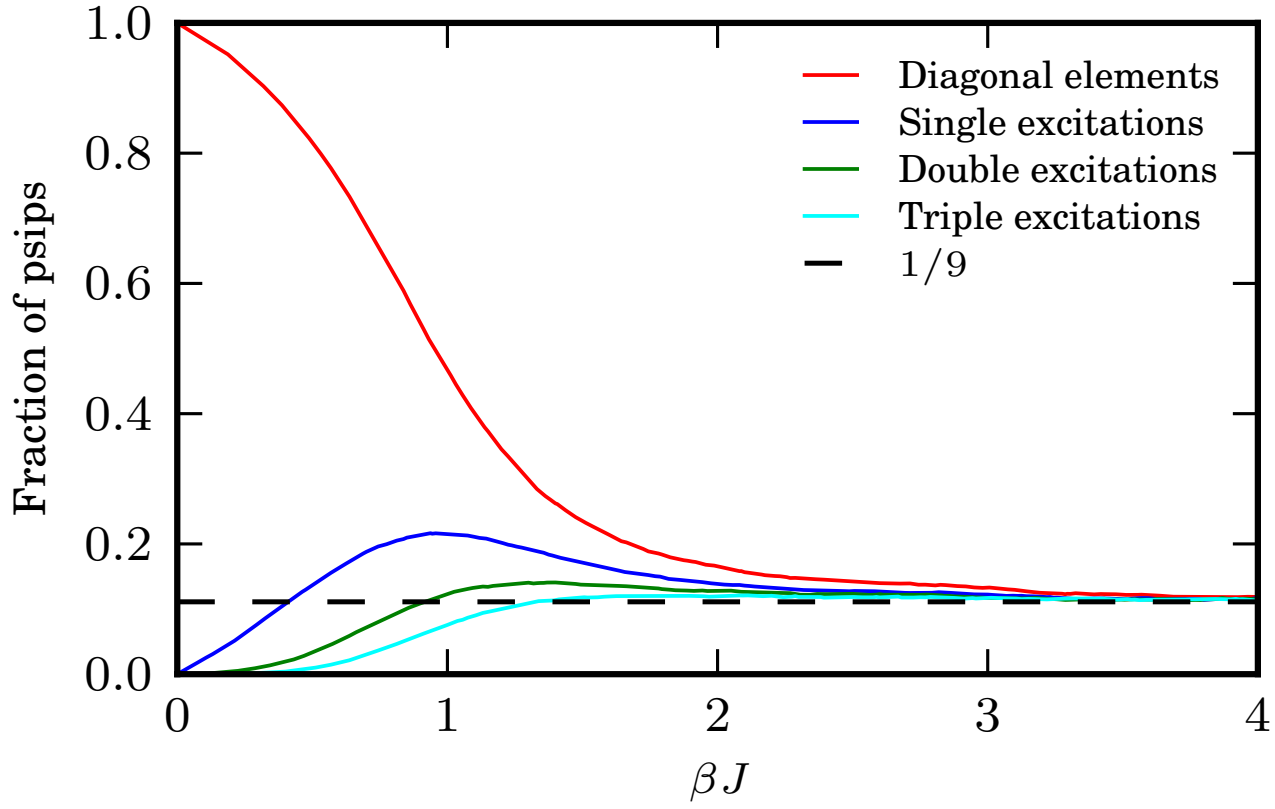
  then

$$\frac{d\tilde{\rho}_{ij}}{d\beta} = \left(\rho_{ij}^{T} L_{ij,kl} \frac{1}{\rho_{kl}^{T}} \right) \tilde{\rho}_{kl}$$

Outline

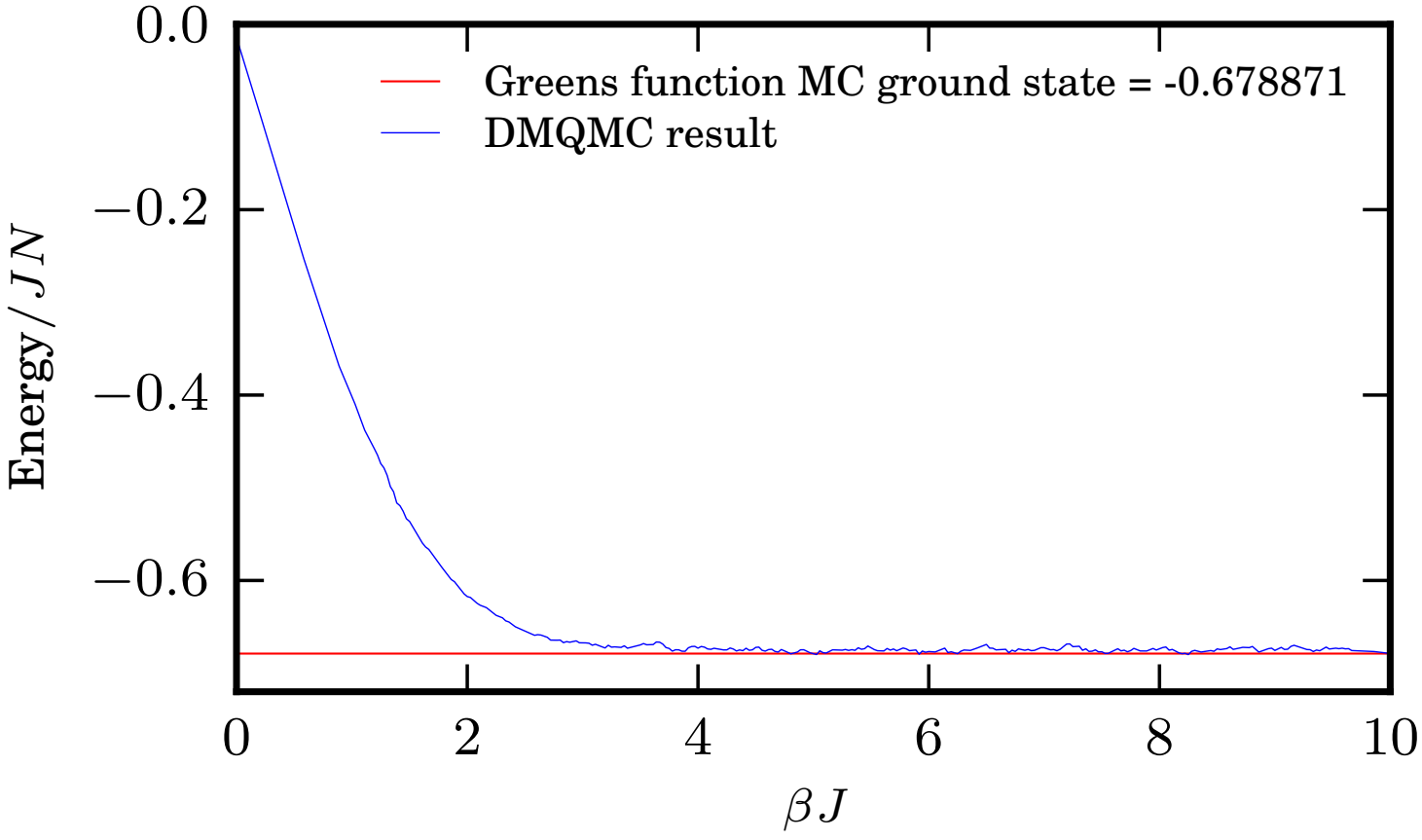
Importance-Sampled Results

- The weights are chosen to make the numbers of psips on each excitation level similar as $\beta \to \infty$.



- It is convenient to switch the weights on gradually as the simulation progresses (i.e., ρ^T is β-dependent).

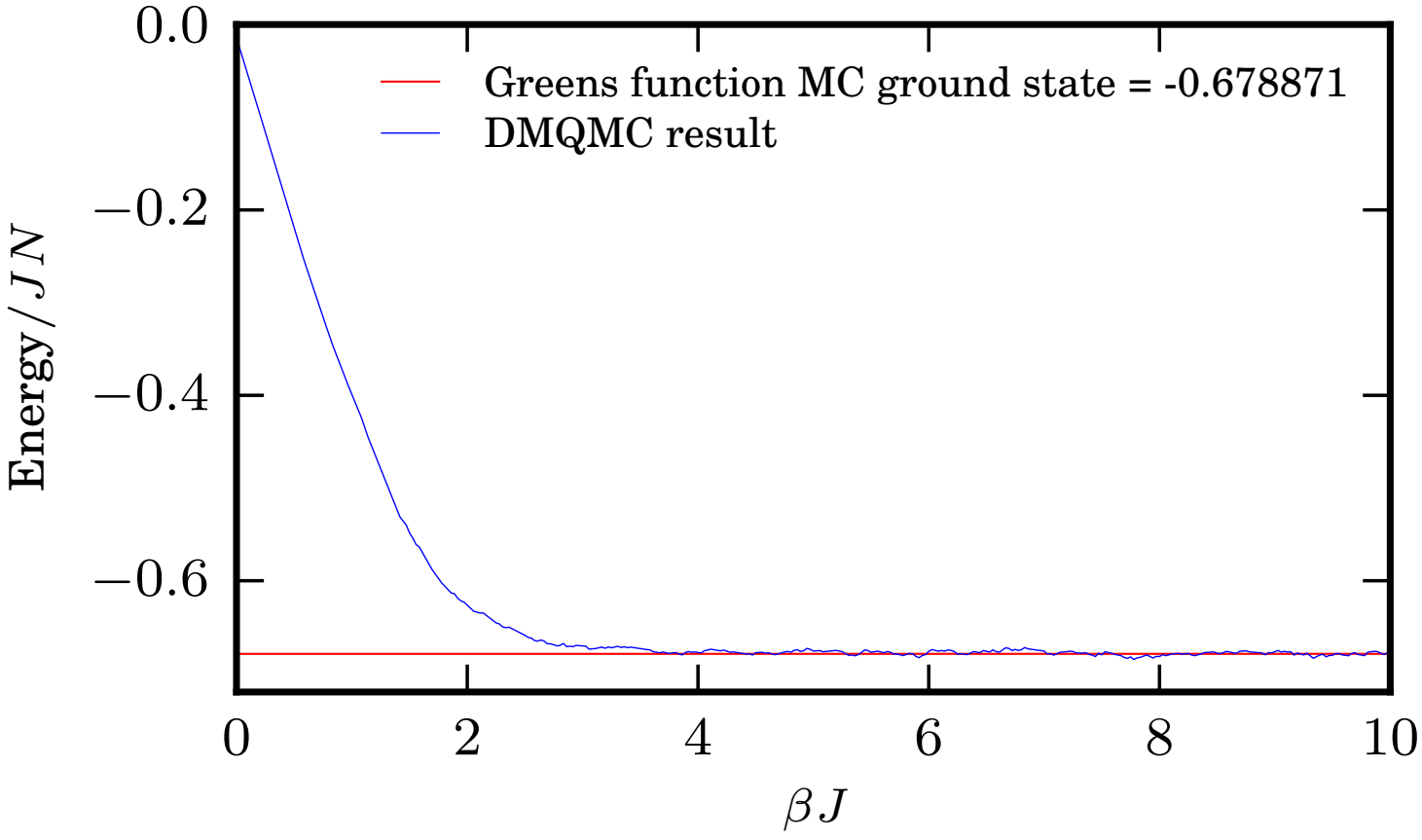
6×6 Heisenberg Model: Energy



6×6; 10^4 psips; 10^3 β-loops.

(note slight error)

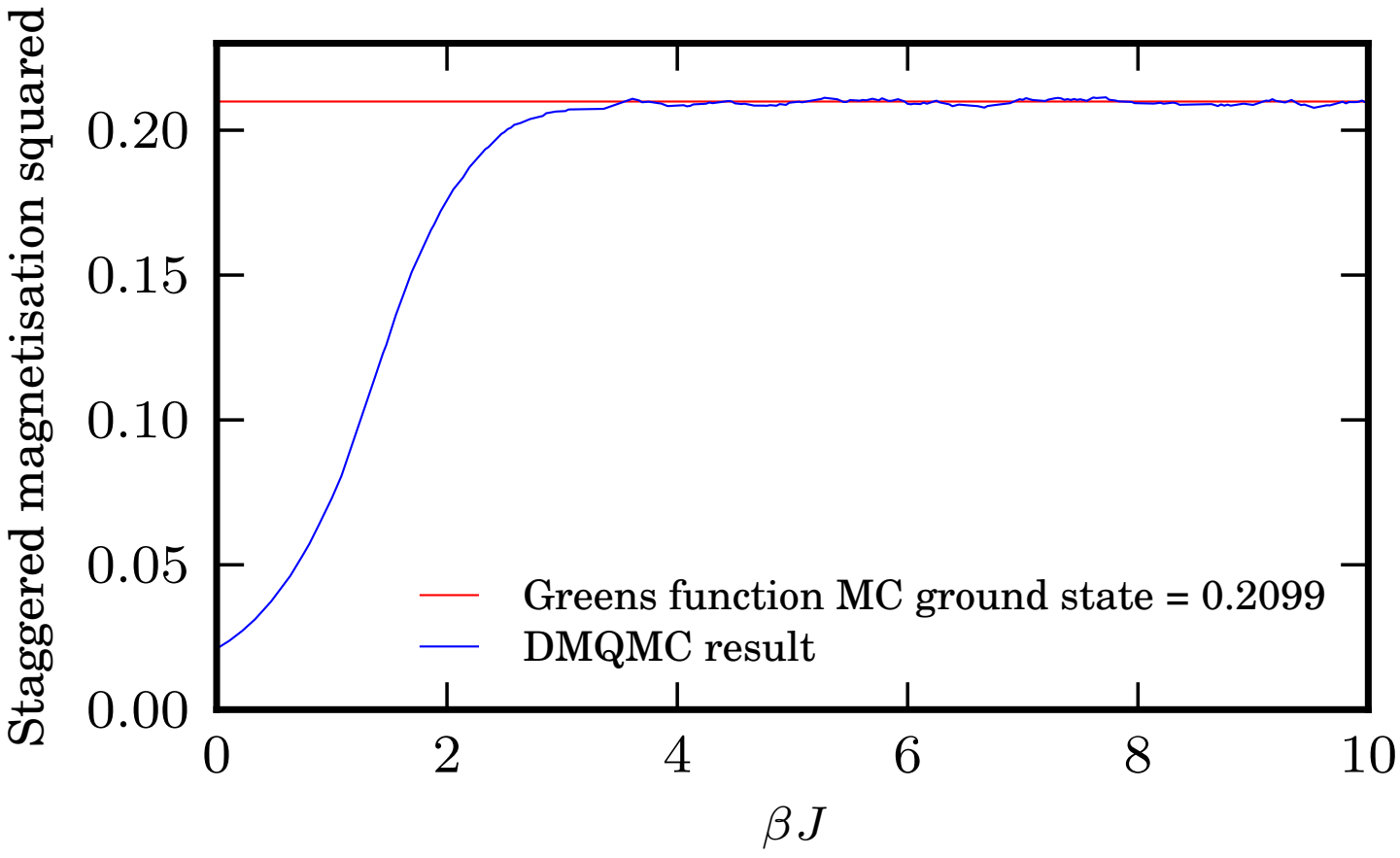
6 × 6 Heisenberg Model: Energy



6 × 6; 10^6 psips; 10 β-loops.

(error removed)

6×6 Heisenberg Model: Staggered Magnetization



6×6; 10^6 psips; 10 β-loops.

Larger Systems

Successfully studied an 8×8 Heisenberg model

- $M_s = 0$ Hilbert space dimension $> 10^{19}$
- Density matrix has $> 10^{38}$ elements

using the same psip population (10^6) and only 10–100 times more β-loops!

Summary

Achievement

- Importance-sampled DMQMC works surprisingly well.
- Can be applied to large sign-problem-free systems.
- Yields full thermodynamics at all temperatures at once.
- Yields full density matrix and hence arbitrary expectation values.

Summary

Achievement

- Importance-sampled DMQMC works surprisingly well.
- Can be applied to large sign-problem-free systems.
- Yields full thermodynamics at all temperatures at once.
- Yields full density matrix and hence arbitrary expectation values.

Not bad for two undergraduate students!

Outlook

Open Questions

- Substantial population control errors in large systems need investigating.
- Severity of sign problem? Analogue of initiator approach?

Where is it useful?

- The thermal properties of tiny molecules are not very interesting.
- When there is a sign problem, we may not be able to tackle large enough systems to study phase diagrams.
- Entanglement measures such as $\mathrm{Tr}(\rho_{\mathrm{red}} \ln \rho_{\mathrm{red}})$ and the concurrence depend on reduced density matrices. DMQMC seems able to calculate these better than other methods.