Unloading the dice

Minimising biases in Full Configuration Interaction Quantum Monte Carlo

William Vigor [1], James Spencer [2,3], Michael Bearpark [1], Alex Thom [1,4]

[1]Department of Chemistry, Imperial College London
[2]Department of Physics, Imperial College London
[3]Department of Materials, Imperial College London
[4]Department of Chemistry, University of Cambridge

Part I

Introduction: FCIQMC

Full Configuration Interaction

- Want to solve the eigenvalue problem $H|\Psi\rangle = E|\Psi\rangle$ where $|\Psi\rangle = \sum_I C_\mathbf{I}|D_\mathbf{I}\rangle$
- Traditionally Diagonalise matrix of $\langle D_\mathbf{I}|\hat{H}|D_\mathbf{J}\rangle$
- Unfortunately this turns out to be impossibly large $\binom{M}{N_e}$ by $\binom{M}{N_e}$.
- **Instead use Monte Carlo to sample this large space. FCIQMC:**

G. H. Booth, A. J. W. Thom, A. Alavi, J. Chem. Phys. **131**, 054106 (2009)

FCIQMC: Huge Systems to FCI accuracy

With an additional approximation people have studied:

- Molecular Systems with 10^{29} configurations

C. Daday *et. al.*, J. Chem. Theory Comput. **8**, 4441 (2012)

- A Uniform Electron Gas with a Hilbert Space of 10^{108} determinants.

J. J. Shepherd *et. al.*, Phys. Rev. B **85**, 081103 (2012)

- This would take about 10^{109} bytes of storage space.
- The world wide web is about 10^{21} bytes.

Full Configuration Interaction Quantum Monte Carlo: Projector

- Apply $e^{-(\hat{H}-S)\delta\tau}$ (stochastically) repeatedly to some initial vector $|I\rangle$.
- $\delta\tau$ is the time step, S is the shift (controls normalisation).
- As long as $\langle\Psi|I\rangle \neq 0$: the ground state emerges.
- Can use the approximation: $e^{-(\hat{H}-S)\delta\tau} \approx 1 - (\hat{H} - S)\delta\tau$
- Exact ground state still emerges if $\delta\tau < 2/(E_{max} - E_{FCI})$.

J. S. Spencer *et. al.*, J. Chem. Phys. **136**, 054110 (2012)
(This is just a stochastic powers method)

Full Configuration Interaction Quantum Monte Carlo: Sampling

- We use integer weights of determinants in the vector, although this is not the only choice.

F. R. Petruzielo *et. al.*, Phys. Rev. Lett. **109**, 230201 (2012)

- Positive (negative) coefficients are represented by a number of positive (negative) psips (Ψ particles).
- For this configuration:

$$|\Psi(\tau)\rangle = \quad +0.25 \quad |D_{\mathbf{0}}\rangle \quad +0.5 \quad |D_{\mathbf{1}}\rangle \quad -0.25 \quad |D_{\mathbf{2}}\rangle$$

J . B. Anderson J. Chem. Phys. **63**, 1499 (1975)

Full Configuration Interaction Quantum Monte Carlo: Sampling

- We use integer weights of determinants in the vector, although this is not the only choice.

F. R. Petruzielo *et. al.*, Phys. Rev. Lett. **109**, 230201 (2012)

- Positive (negative) coefficients are represented by a number of positive (negative) psips (Ψ particles).
- For this configuration:

$$|\Psi(\tau)\rangle = {\color{green}\bigcirc}\ |D_\mathbf{0}\rangle + \overset{\color{green}\bigcirc}{\color{green}\bigcirc}\ |D_\mathbf{1}\rangle + {\color{red}\bigcirc}\ |D_\mathbf{2}\rangle$$

J . B. Anderson J. Chem. Phys. **63**, 1499 (1975)

Full Configuration Interaction Quantum Monte Carlo

- Move τ forwards by $\delta\tau$:

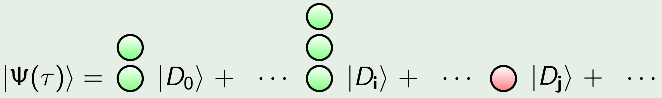
$$|\psi(\tau+\delta\tau)\rangle = \left(1-(\hat{H}-S)\delta\tau\right)|\psi(\tau)\rangle$$

$$|\psi(\tau+\delta\tau)\rangle = \left(1-\sum_{i\neq j}(\langle D_i|H|D_j\rangle)\delta\tau - \sum_i(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}\langle D_i|H|D_j\rangle\delta\tau - \sum_i(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

Spawning

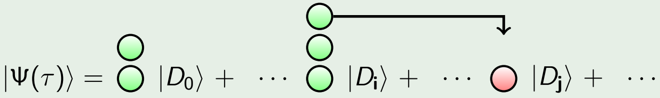


Full Configuration Interaction Quantum Monte Carlo

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}\langle D_i|H|D_j\rangle\delta\tau - \sum_i(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

Spawning

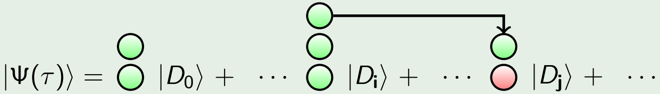


Full Configuration Interaction Quantum Monte Carlo

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i \neq j} \langle D_i | H | D_j \rangle \delta\tau - \sum_i (\langle D_i | H | D_i \rangle - S) \delta\tau \right) |\psi(\tau)\rangle$$

Spawning

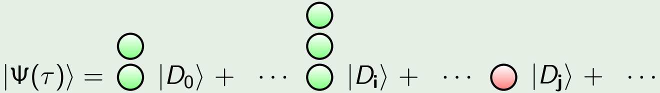


- with probability $\langle D_\mathbf{i} | H | D_\mathbf{j} \rangle \delta\tau$
- if $\langle D_\mathbf{i} | H | D_\mathbf{j} \rangle \delta\tau < 0$ psip has same sign as parent and vice-versa.

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}(\langle D_i|H|D_j\rangle)\delta\tau - \sum_{i}(\langle D_i|H|D_i\rangle - S)\delta\tau \right) |\psi(\tau)\rangle$$

Diagonal Death



Full Configuration Interaction Quantum Monte Carlo

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}(\langle D_i|H|D_j\rangle)\delta\tau - \sum_{i}(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

Diagonal Death

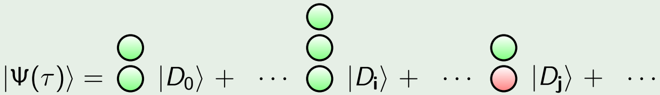
$$|\Psi(\tau)\rangle = |D_0\rangle + \cdots |D_\mathbf{i}\rangle + \cdots |D_\mathbf{j}\rangle + \cdots$$

- Death occurs with probability $(\langle D_\mathbf{i}|H|D_\mathbf{i}\rangle - S)\delta\tau$
- Cloning occurs (population becomes more negative or positive) if $(\langle D_\mathbf{i}|H|D_\mathbf{i}\rangle - S) > 0$

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}(\langle D_i|H|D_j\rangle)\delta\tau - \sum_{i}(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

Annihilation



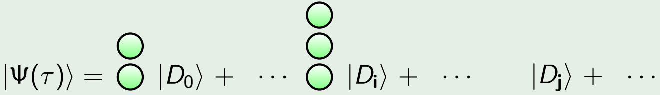
Full Configuration Interaction Quantum Monte Carlo

- Move τ forwards by $\delta\tau$:

$$|\psi(\tau+\delta\tau)\rangle = \left(1 - \sum_{i\neq j}(\langle D_i|H|D_j\rangle)\delta\tau - \sum_{i}(\langle D_i|H|D_i\rangle - S)\delta\tau\right)|\psi(\tau)\rangle$$

Annihilation

- psips with opposite signs on the same determinant annihilate.



Energy Estimates

Projected Energy

- Project onto a reference state:

$$E = \frac{\langle D_0|\hat{H}e^{-(\hat{H}-S)\tau}|D_0\rangle}{\langle D_0|e^{-(\hat{H}-S)\tau}|D_0\rangle} = \frac{\langle D_0|\hat{H}|\Psi_0\rangle}{\langle D_0|\Psi_0\rangle}$$

- Average over every step after the simulation has equilibrated.

Shift

- S is initially set to a constant until the total number of psips N reaches the desired level.
- After which S is updated every A steps accoring to:

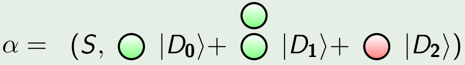
$$S(\tau + A\delta\tau) = S(\tau) - \frac{\gamma}{A\delta\tau}\log\frac{N(\tau + A\delta\tau)}{N(\tau)}$$

Part II

FCIQMC and Markov Chain Monte Carlo

- A Markov chain is a stochastic system in which each step only depends on the previous step.
- This means the system can be described by a matrix $\Gamma_{\alpha\beta}$ stochastic matrix: The probability that the system transitions from state α to state β in one 'Markov' step.
- Under some conditions Γ has an eigenvector γ with eigenvalue one.
- γ_α probability the system is in a state α after equilibration.
- FCIQMC seems to be described by an infinite number of states as the shift can take on any value.

An Example:



The Shift Again

- The shift is only function of the number of psips N on the last shift update step and the number on the step before the shift turned on N_s :

$$S(\tau + A\delta\tau) = S(\tau) - \frac{\gamma}{A\delta\tau}\log\frac{N(\tau + A\delta\tau)}{N(\tau)}$$

$$S(\tau + A\delta\tau) = S(\tau - A\delta\tau) - \frac{\gamma}{A\delta\tau}\log\frac{N(\tau)}{N(\tau - A\delta\tau)} - \frac{\gamma}{A\delta\tau}\log\frac{N(\tau + A\delta\tau)}{N(\tau)}$$

$$S(\tau) = S_0 - \frac{\gamma}{A\delta\tau}\log\frac{N(\tau)}{N_s}$$

- Thus we make one 'Markov' step every A steps of $\delta\tau$.
- $A = 1$ for mathematical simplicity

The FCIQMC Markov Chain: for two Determinants

We want $\Gamma_{\alpha,\beta}$. The population of psips on each determinant is changed by:

1. Death of psips on the determinant.
2. Spawning from the other determinant onto this determinant.

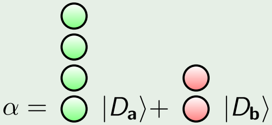
These events are Binomially distributed probability:

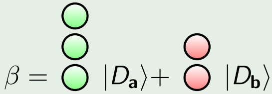
$$B(n, N, p) = \binom{N}{n} p^n (1-p)^{N-n} \tag{1}$$

Sum over these binomial coefficients such that the change could have happened.

The FCIQMC Markov Chain: for two Determinants

An Example





- If $\langle D_{\mathbf{b}}|H|D_{\mathbf{a}}\rangle < 0$ and $\langle D_{\mathbf{a}}|H|D_{\mathbf{a}}\rangle - S > 0$ then the probability that the population changes on **a**:

$$B(1,4,P_{d\mathbf{a}})B(0,2,P_{s\mathbf{a}}) + B(2,4,P_{d\mathbf{a}})B(1,2,P_{s\mathbf{a}}) \\ +B(3,4,P_{d\mathbf{a}})B(2,2,P_{s\mathbf{a}})$$

- After which we just do the same for **b** and product the two probabilities to compute $\Gamma_{\alpha\beta}$.

Conclusions: Markov Chain Monte Carlo and FCIQMC

- FCIQMC is a Markov chain.
- ξ, $\delta\tau$, N_s, A and the system specify the chain.
- The size of the stochastic matrix scales terribly with system size.
- We can diagonalise the Stochastic Matrix to compute the stationary distribution.

Part III

Population Control Bias in FCIQMC

Population control

- Let's scale γ when we change $\delta\tau$ and A, to fix ξ:

$$\xi = \frac{\gamma}{A\delta\tau} \tag{2}$$

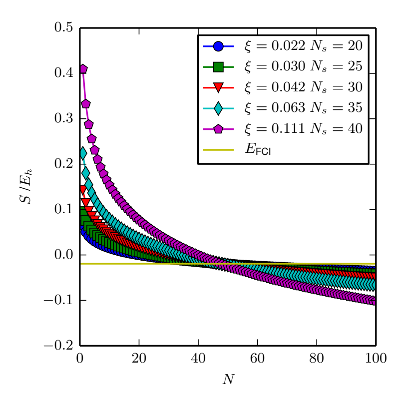
- If we assume that $S = E_{FCI}$ at $\langle N \rangle$

$$E_{FCI} = -\xi \log \frac{\langle N \rangle}{N_0}$$

$$\langle N \rangle = N_0 e^{-\frac{\xi}{E_{FCI}}}$$

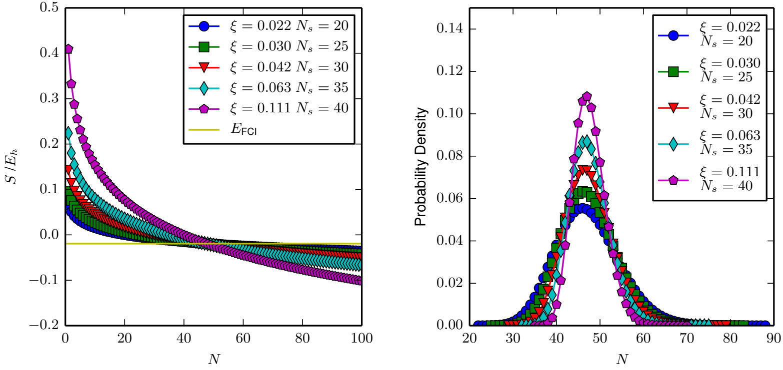
- Thus we can fix $\langle N \rangle$ if we know the correlation energy.

H2 in a STO-3g Basis set



Distribution of the Number of psips

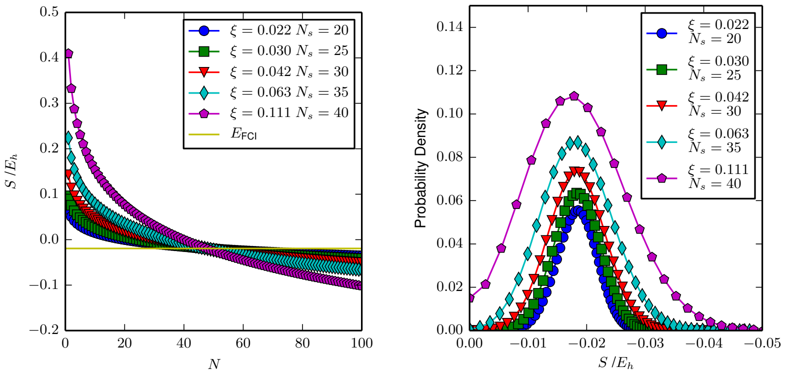
H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



- If ξ is big then the population is well controlled.

Distribution of the Shift

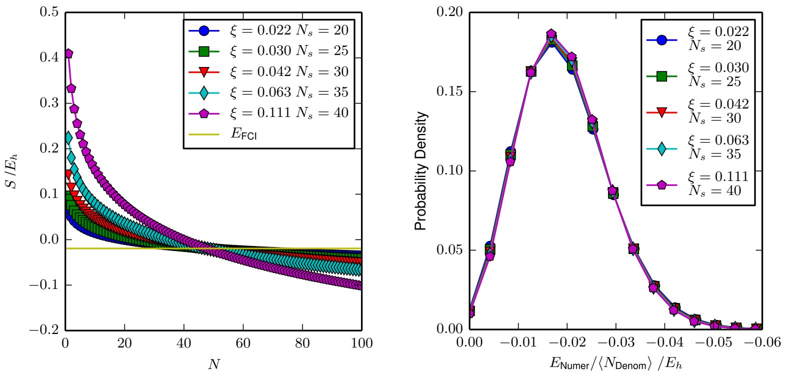
H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



- If ξ is big then shift has a large variance.

Distribution of the Projected Energy

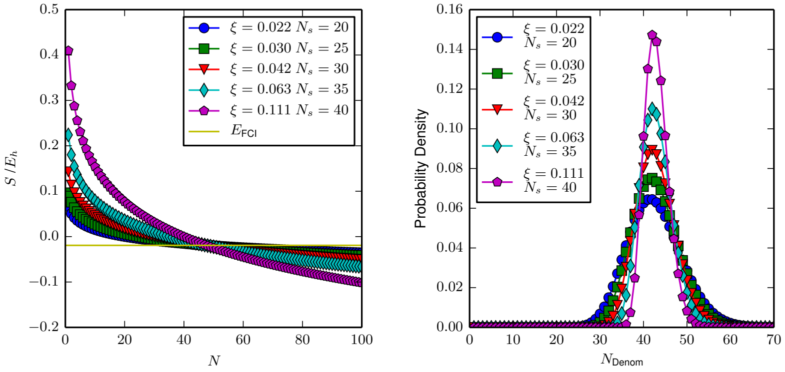
H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



- ξ has little effect on the numerator of the projected energy.

Distribution of the Projected Energy

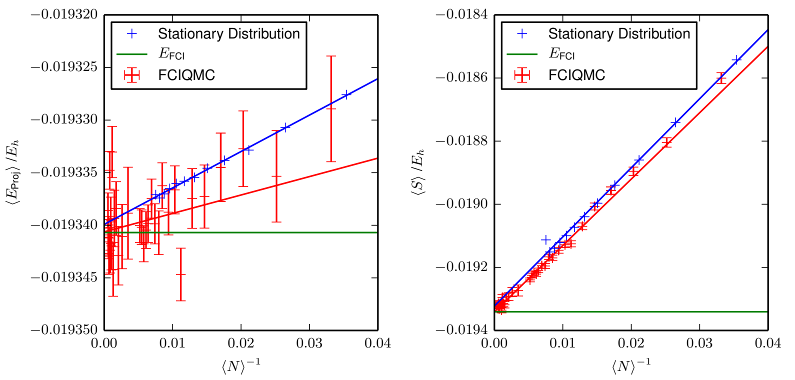
H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



- If ξ effects the denominator of the projected energy in the same way as the population.

Population Control Bias in H_2

H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



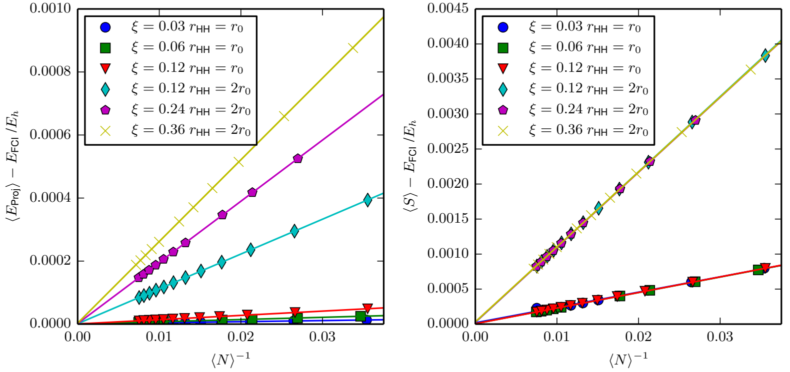
- The mean of both estimators of the energy is incorrect.

$$\mathcal{O}(1/\langle N \rangle)$$

N. Cerf *et. al.*, Phys. Rev. E **51**, 3679 (1995)

Population Control Bias in H_2

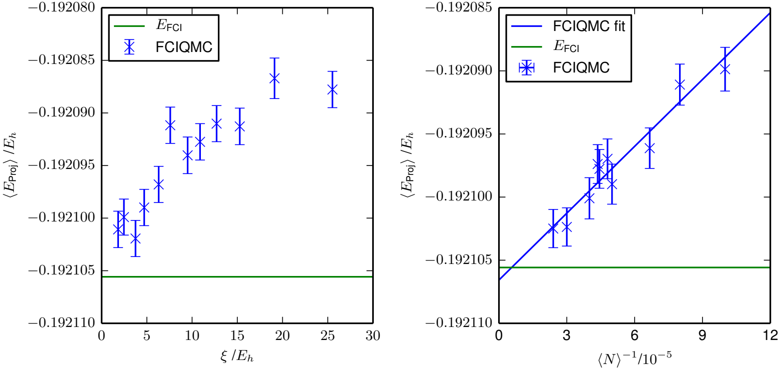
H2 in a STO-3g Basis set $r_0 = 0.7122$ Å



- If ξ is small then population control bias decreases.
- Strongly correlated systems seem to be more affected.

Population Control Bias in Ne cc-pVDZ

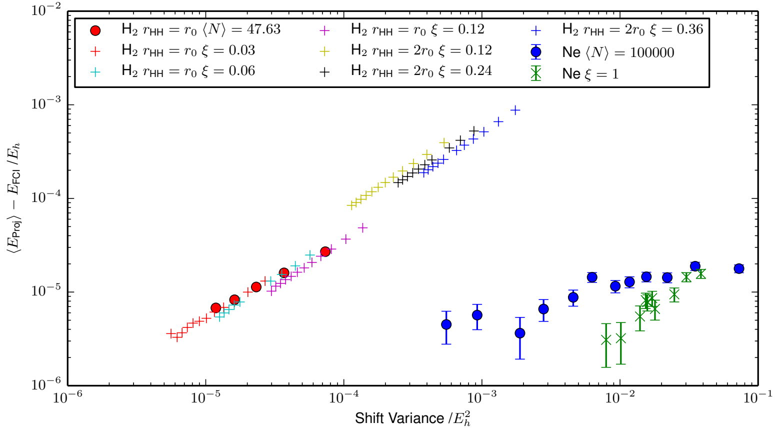
Ne cc-pVDZ



- We see similar thing for Ne cc-pVDZ.

Population Control Bias

Shift Variance



Minimise Population Control Bias

- Use as large as population as possible.
- Reduce ξ (don't let the population fall bellow the plateau).
- If you really care about accuracy extrapolate $1/\langle N \rangle \to \infty$ with a fixed ξ.

Acknowledgements

- **Thank you for listening**
- Alex Thom, Michael Bearpark, James Spencer
- EPSRC for a studentship
- FCIQMC calculations ran using HANDE:
- James Spencer, Alex Thom, Nick Blunt, Fionn Malone, James Shepherd, Matthew Foulkes, Thomas Rogers, Will Handley, Joseph Weston
- Imperial College High Performance Computing Service

See arXiv:1407.1753 for more details.