Introduction to Full CI Quantum Monte Carlo
(with applications to the Hubbard Model)

Ali Alavi
Max Planck Institute for Solid State Research
and
University of Cambridge
Overview

• Introduction
• FCIQMC
• Hubbard model
• i-FCIQMC
• Live demo
• Reduced Density Matrices
• 3-band Hubbard model
• Stochastic CASSCF
• Real Cuprates
• Excited states
Many-Electron Schrödinger equation

\[ H = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2 + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{i} \nu(r_i) \]

\[
H \Psi_0 = E_0 \Psi_0 \\
\Psi_0 = \Psi_0(x_1, \ldots, x_N) \quad x = (r, \sigma)
\]

Electrons are Fermions:

\[
\Psi_0(\ldots x_i, \ldots, x_j \ldots) = -\Psi_0(\ldots x_j, \ldots, x_i \ldots)
\]

Atomic units \( \hbar = m_e = |e| = 1 \quad E_h = 27.211 \text{ eV} \)
Ab initio strategies to get $E_0$

Quantum Chemical
Finite basis sets
Many-body approximations
Hartree-Fock (mean-field theory)
Many Body Perturbation theory,
Coupled cluster methods
↓
Full Configuration Interaction.
Systematically improvable.
Expensive

Density Functional Theory
Exchange Correlation functional $\Rightarrow$
*Uncontrolled approximation*, not systematically improvable.

*Widely used.*

Quantum Monte Carlo
Stochastic exploration of the configuration space
No basis sets
Fixed Node approximation (in diffusion QMC) due to
Fermion sign problem
$\Rightarrow$ *Uncontrolled error.*
Full configuration interaction

\[ H = \sum_i^N -\frac{1}{2} \nabla_i^2 + \sum_{i<j} \frac{1}{|r_i - r_j|} + \sum_i^N v(r_i) \]

\[ H\Psi_0 = E_0\Psi_0 \]

\[ \Psi_0 = \sum_i C_i \left| D_i \right> \text{ Include all (symmetry-allowed) determinants within basis} \]

Variationally minimise wrt \( C_i \) \[ \Rightarrow \sum_i \langle D_j | H | D_i \rangle C_i = E_0 C_j \]

**Ground-state eigenvalue problem in an exponentially large space**

\( \langle D_i | H | D_j \rangle \) can be positive or negative: this is a source of sign problem, but is NOT the Fermion sign problem!

Largest FCI calculation to date \( \sim 10^{10} \) determinants (\( N_2 \) molecule)


Hubbard model: \( 159 \times 10^9 \) determinants [Yamada, Imamura, Machida, on the Earth Simulator]
Slater determinant space: the Hilbert space for fermions

Suppose we have in hand a set of $2M$ spin-orbitals [e.g. RHF, UHF, pw]

$$\{u_1, u_2, \ldots, u_{2M}\}$$

$$|D_i\rangle = |D_{n_1,n_2\ldots,n_N}\rangle = \frac{1}{\sqrt{N!}} \det[u_{n_1}, u_{n_2}, \ldots, u_{n_N}]$$

Antisymmetric $N$-electron basis

$$N_{FCI} = \binom{M}{N_\alpha} \binom{M}{N_\beta}$$

$N_\alpha = 5, N_\beta = 5, M = 100$  

$\rightarrow N_{FCI} \approx 10^{16}$

HF determinant $D_0$  
(Fermi Vacuum)

2h,2p  
4h,4p
Hamiltonian matrix elements (Slater-Condon rules)

Since $H$ contains at most 2-body interactions:

$$\langle D_i | H | D_j \rangle = 0 \text{ if } D_i \text{ and } D_j \text{ differ by more than 2 spin-orbitals}$$

Maximum connectivity:

$$N(N-1)(2M-N)(2M-N-1)/4 \approx N^2M^2$$

Spin selection rule:

$$\langle D_i | H | D_j \rangle = 0 \text{ if } S_z[D_i] \neq S_z[D_j]$$

$$\langle D_i | H | D_i \rangle = \sum_{i<j} \left[ \langle ij | ij \rangle - \langle ij | ji \rangle \right]$$

$$\langle D_i | H | D_j \rangle = \sum_k \left[ \langle ik | ak \rangle - \langle ik | ka \rangle \right]$$

$$\langle D_i | H | D_j \rangle = \langle ij | ab \rangle - \langle ij | ba \rangle$$

Other symmetries may also exist:

- Translational invariance;
- Molecules: point group symmetry

for $D_i, D_j$ differing by one spin-orbital

for $D_i, D_j$ differing by two spin-orbitals
From “amplitudes” to “walkers”

Consider a population of $N_w$ “walkers” which inhabit Slater determinant space

$$\{i_1, i_2, \ldots, i_{N_w}\}$$

Each with an associated sign $s_i = +1$ or -1

$$\{s_1, s_2, \ldots, s_{N_w}\}$$

We will define the amplitude on $i$ to be the signed sum of walkers on $D_i$:

$$C_i \propto \sum_{j=1}^{N_w} s_j \delta_{i,i_j} = N_i$$
Pictorial example

\[ N_w = \sum_{i} |N_i| = 11 \]

Configuration space

Normalisation = \[ \frac{1}{\sqrt{\sum_{i} N_i^2}} = \frac{1}{\sqrt{19}} \]
A differential formulation for the CI coefficients

Let: \[ K_{ij} = H_{ij} - E_0^{(0)} \delta_{ij} \]

Consider the set of coupled first-order equations:

\[- \frac{dC_i}{dt} = \sum_j (K_{ij} - S \delta_{ij}) C_j \]

If \( \sum_j K_{ij} C_j = SC_i \Rightarrow \frac{dC_i}{dt} = 0 \Rightarrow \) The distribution is stationary and is an eigenstate of \( K \) (and hence \( H \))

If \( S = E_0 - E_0^{(0)} \) The stationary distribution is the exact ground-state

Any arbitrary initial distribution \( \{C_i\} \) will tend to the exact ground-state

However: this is not very useful, as we need the complete of \( \{C_j\} \) to complete the force calculation \( \Rightarrow \) the MEMORY BOTTLENECK OF FCI
Population Dynamics

We want to generate a population dynamics for our set of walkers so that the rate of change of walkers on a given determinant satisfies the imaginary-time Schrödinger equation:

\[
    K_{ij} = H_{ij} - E_{HF} \delta_{ij}
\]

\[
    - \frac{dC_i}{dt} = \sum_j (K_{ij} - S \delta_{ij})C_j = (K_{ii} - S)C_i + \sum_{j \neq i} K_{ij}C_j
\]

Reduce/enhance population on \(i\), proportional to \((K_{ii} - S)\) and \(C_i\)

Diagonal Death (or cloning) processes

- Reduce if \(K_{ij}C_j > 0\)
- Enhance if \(K_{ij}C_j < 0\)

Annihilation processes

Spawning processes (child of \(j\) on \(i\))

The sign of the child depends on the sign of \(K_{ij}\)
Overview of FCIQMC algorithm: a random Game of Life, death and annihilation

- Start with $N$ (positive) walkers on $D_0$, an initial value of $S$, and time-step $\tau$
- Spawning (birth) processes
- (Diagonal) death/cloning processes
- Annihilation step
- Adjust Shift (in constant $N$ mode)

$$S \rightarrow S - \frac{1}{A\tau} \ln \frac{N_{\text{current}}}{N_{\text{old}}}$$

Pictorial example

\[ N_w = \sum_i |N_i| = 11 \]

Normalisation = \[ \frac{1}{\sqrt{\sum_i N_i^2}} = \frac{1}{\sqrt{19}} \]
Spawning

spawning event

configuration space
Death event

configuration space

Death
Death
Annihilation

Hash algorithm: $O(N_w)$
Booth, Smart, Alavi, Mol. Phys., 112 (14), (2014), 1855-1869
Annihilation
The rules of the “Game of Life”
(derived from the underlying imaginary-time S.E.)

Probability of death

\[ p_d = \tau |H_{ii} - E_{HF} - S| \]

Probability to spawn new walker

\[ p_s = \tau \frac{|H_{ij}|}{p_{gen}(j|i)} \]

\[ \sum_j p_{gen}(j|i) = 1 \]

\[ p_{gen}(j|i) \sim (N^2M^2 + NM)^{-1} \]

If \( H_{ij} < 0 \), child has same sign as parent.
If \( H_{ij} > 0 \) child has opposite sign of parent.
The projected energy
(non-variational)

\[
E = \frac{\langle D_0| H |\Psi \rangle}{\langle D_0|\Psi \rangle} = \sum_j \frac{\langle D_0| H |D_j \rangle \langle D_j|\Psi \rangle}{\langle D_0|\Psi \rangle} = E_{HF} + \sum_{j \in \text{doubles}} \langle D_0| H |D_j \rangle \frac{C_j}{C_0}
\]

where \( \frac{C_j}{C_0} = \frac{N_j}{N_0} \)
$\text{Be}_2$ (cc-pVTZ). $N_{\text{FCI}} = 346,485$ determinants.
$\text{Be}_2$ (cc-pVTZ). $N_{\text{FCI}} = 346,485$ determinants
H₂O (all electron, cc-pVDZ, 452x10⁶ determinants)

Variable shift mode
Coherent exponential growth

"Annihilation plateau"
Incoherent exponential growth
**Comparison with existing FCI**

TABLE I. Results for systems with FCI comparisons. The geometries for the N$_2$ molecule were eqm: 2.068a$_0$, stretched: 4.2a$_0$, and C$_2$: 1.27273 Å. The geometry for the water molecule was taken from Ref. 35. The working space includes all point group symmetry of the molecule from $D_{2h}$ or the largest available subset thereof. All systems had core electrons frozen apart from H$_2$O. $N_{\text{FCI}}$ is the size of the FCI space in the $D_{2h}$ point group ($C_{2v}$ for H$_2$O. The digit in italics for $E_{\text{total}}$, represents the first uncertain digit. $N_c$ is the number of walkers required to achieve the plateau. $f_c = N_c / N_{\text{FCI}}$.

<table>
<thead>
<tr>
<th>System</th>
<th>$(N,M)$</th>
<th>$N_{\text{FCI}}/10^6$</th>
<th>$N_c/10^6$</th>
<th>$f_c$</th>
<th>$E_{\text{total}}$</th>
<th>$E_{\text{FCI}}$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne: aug-cc-pVDZ</td>
<td>(8,22)</td>
<td>6.69</td>
<td>0.21</td>
<td>0.031</td>
<td>$-128.70949$</td>
<td>$-128.709,476$</td>
<td>33</td>
</tr>
<tr>
<td>C$_2$: cc-pVDZ</td>
<td>(8,26)</td>
<td>27.9</td>
<td>15.0</td>
<td>0.538</td>
<td>$-75.7299$</td>
<td>$-75.729,853$</td>
<td>34</td>
</tr>
<tr>
<td>H$_2$O: cc-pVDZ</td>
<td>(10,24)</td>
<td>451</td>
<td>26</td>
<td>0.058</td>
<td>$-76.24186$</td>
<td>$-76.241,860$</td>
<td>35</td>
</tr>
<tr>
<td>N$_2$-eqm: cc-pVDZ</td>
<td>(10,26)</td>
<td>541</td>
<td>270</td>
<td>0.499</td>
<td>$-109.27649$</td>
<td>$-109.276,527$</td>
<td>33</td>
</tr>
<tr>
<td>N$_2$-stretched: cc-pVDZ</td>
<td>(10,26)</td>
<td>541</td>
<td>345</td>
<td>0.637</td>
<td>$-108.9669$</td>
<td>$-108.966,95$</td>
<td>36</td>
</tr>
</tbody>
</table>
New systems

Table II. Predicted FCI results. The geometries of the molecules were (in Å): CN (1.1941), HF (0.91622), CH$_4$ ($r_{CH}=1.087728$), CO (1.1448), H$_2$O ($r_{OH}=0.975512$, $\theta=110.565^\circ$) (Ref. 35), O$_2$ (1.2074), and NaH (1.885977). CN and O$_2$ orbitals were constructed from a restricted open-shell HF calculation with a spin multiplicity of two and three, respectively. CN, CH$_4$, CO, and O$_2$ had frozen core electrons. The number in brackets represents the error in the previous digit, obtained through a Flyvbjerg-Petersen blocking analysis (Ref. 37) of $E(\tau)$.

<table>
<thead>
<tr>
<th>System</th>
<th>$(N,M)$</th>
<th>$N_{FCI}/10^6$</th>
<th>$N_c/10^6$</th>
<th>$f_c$</th>
<th>$E_{total}$</th>
<th>$E_{CCSD(T)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be: cc-V5Z</td>
<td>(4,91)</td>
<td>2.11</td>
<td>0</td>
<td>0</td>
<td>-14.646 38(2)</td>
<td>-14.646 29</td>
</tr>
<tr>
<td>CN: cc-pVDZ</td>
<td>(9,26)</td>
<td>246</td>
<td>173</td>
<td>0.704</td>
<td>-92.493 8(3)</td>
<td>-92.491 64</td>
</tr>
<tr>
<td>HF: cc-pCVDZ</td>
<td>(10,23)</td>
<td>283</td>
<td>0.998</td>
<td>0.0035</td>
<td>-100.270 98(3)</td>
<td>-100.270 44</td>
</tr>
<tr>
<td>CH$_4$: cc-pVDZ</td>
<td>(8,33)</td>
<td>419</td>
<td>377</td>
<td>0.898</td>
<td>-40.387 52(1)</td>
<td>-40.389 74</td>
</tr>
<tr>
<td>CO: cc-pVDZ</td>
<td>(10,26)</td>
<td>1080</td>
<td>777</td>
<td>0.719</td>
<td>-113.056 44(4)</td>
<td>-113.054 97</td>
</tr>
<tr>
<td>H$_2$O: cc-pCVDZ</td>
<td>(10,28)</td>
<td>2410</td>
<td>47.4</td>
<td>0.0196</td>
<td>-76.280 91(3)</td>
<td>-76.280 28</td>
</tr>
<tr>
<td>O$_2$: cc-pVDZ</td>
<td>(12,26)</td>
<td>5409</td>
<td>2651</td>
<td>0.490</td>
<td>-149.987 5(2)</td>
<td>-149.985 62</td>
</tr>
<tr>
<td>NaH: cc-pCVDZ</td>
<td>(12,32)</td>
<td>205300</td>
<td>63.8</td>
<td>0.000 31</td>
<td>-162.609 0(1)</td>
<td>-162.609 01</td>
</tr>
</tbody>
</table>
Hubbard Model

\[ \hat{H} = -t \sum_{\langle p,q \rangle, \sigma} (c_{p \sigma}^\dagger c_{q \sigma} + \text{h.c.}) + U \sum_n n_{p \uparrow} n_{p \downarrow} \]

\[ |k\rangle = \frac{1}{\sqrt{\Omega}} \sum_p e^{-ik \cdot r_p} |p\rangle \]

\[ \hat{H} = \sum_{k, \sigma} \epsilon_k c_{k, \sigma}^\dagger c_{k, \sigma} + \frac{U}{\Omega} \sum_{k, q, Q} c_{k+Q}^\dagger c_{q}^\dagger c_{-Q, \downarrow} c_{k, \uparrow} c_{k, \downarrow} \]

3 \sqrt{2} \times 3 \sqrt{2}, 18 sites (L=3)

The diagrams illustrate the lattice and the basis vectors for the Brillouin zone:

- \( T_1 = (L, L) \)
- \( T_2 = (-L, L) \)

- \( b_1 = \frac{\pi}{L} (1, 1) \)
- \( b_2 = \frac{\pi}{L} (-1, 1) \)

\[ k_{n,m} = nb_1 + mb_2 \]

\[ \epsilon_k = -2t \left( \cos \frac{\pi(n+m)}{L} + \cos \frac{\pi(n-m)}{L} \right) \]
Hilbert space of K=0 sector at half-filling  

$N_{FCI} \approx 131 \times 10^8$
18 site-18 electron U=4 Hubbard model
Time evolution of the energy

![Graph showing the time evolution of the energy](image-url)
How to reduce $N_w$ while maintaining FCI accuracy?
Overview of initiator-FCIQMC

Start with $N$ (positive) walkers on $D_0$, an initial value of $S$, and time-step $\tau$

- Spawning (birth) processes
- Survival of the fittest [Cleland, Booth, Alavi, J Chem Phys, 132, 041103, (2010)]
- (Diagonal) death/cloning processes
- Annihilation step
- Adjust Shift (in constant $N$ mode)
Survival of the fittest and “initiators”

If $D$ is empty, child of $P$ spawned onto $D$ survives only if $P$ is an initiator ($N_P > n_{add}$)

The value of $n_{add}$ is not crucial, as long as it is sensibly chosen. We typically use $n_{add} = 2$ or 3.

Initiators can bring to life new determinants
Is the initiator method exact?

• In the limit of large walker number, all determinants acquire an occupation, and therefore all newly spawned walkers survive the test of “survival of the fittest”.

• Therefore the large walker number limit of “i-FCIQMC” is FCIQMC.

• Since the large walker limit of FCIQMC is FCI, we have that the large walker limit of i-FCIQMC is FCI.
Convergence of i-FCIQMC with walker number.

$N_2$ in cc-pVDZ

$(N_{\text{FCI}}=5.4 \times 10^8)$
The effect of varying $n_{add}$

CO in cc-pVQZ

$N_{\text{FCI}} = 4.7 \times 10^{14}$
Live Demo
Initiator method

(ninit=3)

18 site-18 electron U=4 Hubbard model
The $\text{Be}_2$ (VTZ) with i-FCIQMC: a simulation with 2,000 walkers
Energy, compared with subspace diagonalisations (c.f. Stochastic CI)
The instantaneous $i$-FCIQMC wavefunction is a very coarse-grained representation of the exact wavefunction (Be$_2$-VTZ, $N_{\text{FCI}} = 346,485$, $N_w = 2,000$).
But the time-averaged $i$-FCIQMC wavefunction with 2,000 walkers ($\text{Be}_2$, VTZ, $N_{\text{FCI}} = 346485$) is essentially perfect.
Electron affinities with i–FCIQMC: \( A + e^- \rightarrow A^- \)

comparison with Expt* (relativistically corrected)

\[ \text{aug-VDZ} = [4s3p2d] \sim 23, \quad \text{aug-VTZ} = [5s4p3d2f] \sim 46, \quad \text{aug-VQZ} = [6s5p4d3f2g] \sim 80, \]

\[ \text{aug-VXZ} \sim (1/3)(X+1)(X+3/2)(X+2)+(X+1)^2 \]


## First-row Diatomics

Size of spaces and required Nw’s to accumulate 50,000 walkers on the HF det.

<table>
<thead>
<tr>
<th>Diatomic</th>
<th>Basis</th>
<th>N, 2M</th>
<th>N_FCI</th>
<th>N_w</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2</td>
<td>VQZ</td>
<td>8, 216</td>
<td>6.1 x 10^{11}</td>
<td>2.6 x 10^{6}</td>
</tr>
<tr>
<td>CN</td>
<td>VQZ</td>
<td>9, 216</td>
<td>4.8 x 10^{13}</td>
<td>7.6 x 10^{7}</td>
</tr>
<tr>
<td>N2</td>
<td>VQZ</td>
<td>10, 216</td>
<td>2.4 x 10^{14}</td>
<td>3.0 x 10^{7}</td>
</tr>
<tr>
<td>CO</td>
<td>VQZ</td>
<td>10, 216</td>
<td>4.7 x 10^{14}</td>
<td>6.0 x 10^{7}</td>
</tr>
<tr>
<td>NO</td>
<td>VQZ</td>
<td>11, 216</td>
<td>1.5 x 10^{16}</td>
<td>1.1 x 10^{8}</td>
</tr>
<tr>
<td>O2</td>
<td>VQZ</td>
<td>12, 216</td>
<td>6.4 x 10^{16}</td>
<td>5.9 x 10^{7}</td>
</tr>
<tr>
<td>F2</td>
<td>VQZ</td>
<td>14, 216</td>
<td>1.3 x 10^{19}</td>
<td>5.0 x 10^{7}</td>
</tr>
</tbody>
</table>

eg. note that the 14-electron F_{2} is ~20% cheaper than the 10-electron CO, despite the fact that the space is 5 orders of magnitude larger!
## i-FCIQMC Diatomic dissociation energies in kcal mol\(^{-1}\)

<table>
<thead>
<tr>
<th>Basis</th>
<th>Be(_2)</th>
<th>C(_2)</th>
<th>CN</th>
<th>CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDZ</td>
<td>0.53(3)</td>
<td>129.95(8)</td>
<td>159.40(7)</td>
<td>241.49(6)</td>
</tr>
<tr>
<td>VTZ</td>
<td>1.78(6)</td>
<td>139.63(2)</td>
<td>171.71(6)</td>
<td>251.66(5)</td>
</tr>
<tr>
<td>VQZ</td>
<td>2.27(9)</td>
<td>143.44(5)</td>
<td>176.80(9)</td>
<td>255.92(9)</td>
</tr>
<tr>
<td>V(TQ)Z</td>
<td>2.67(10)</td>
<td>145.76(5)</td>
<td>179.8(1)</td>
<td>258.3(1)</td>
</tr>
<tr>
<td>VQZ+Δ(E^{\text{CCSD(T)}})</td>
<td>145.36(5)</td>
<td>179.71(9)</td>
<td>258.68(9)</td>
<td></td>
</tr>
<tr>
<td>(D_e^*) (Expt.)</td>
<td>2.658(6)</td>
<td>146.9(5)</td>
<td>180.4(2.4)</td>
<td>258.8(2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Basis</th>
<th>NO</th>
<th>N(_2)</th>
<th>O(_2)</th>
<th>F(_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDZ</td>
<td>132.57(5)</td>
<td>200.52(8)</td>
<td>105.17(6)</td>
<td>27.59(7)</td>
</tr>
<tr>
<td>VTZ</td>
<td>143.99(6)</td>
<td>216.86(9)</td>
<td>114.35(8)</td>
<td>35.5(1)</td>
</tr>
<tr>
<td>VQZ</td>
<td>148.9(1)</td>
<td>223.20(8)</td>
<td>117.5(1)</td>
<td>36.9(1)</td>
</tr>
<tr>
<td>V(TQ)Z</td>
<td>151.9(2)</td>
<td>227.3(1)</td>
<td>119.6(1)</td>
<td>38.4(2)</td>
</tr>
<tr>
<td>VQZ+Δ(E^{\text{CCSD(T)}})</td>
<td>152.0(2)</td>
<td>227.09(8)</td>
<td>120.1(1)</td>
<td>38.6(1)</td>
</tr>
<tr>
<td>(D_e^*) (Expt.)</td>
<td>152.63(4)</td>
<td>227.60(5)</td>
<td>120.42(5)</td>
<td>39.0(1)</td>
</tr>
</tbody>
</table>
Dissociation energies of some strongly correlated first-row molecules

(F12 correction added to VQZ)

### Errors (kcal mol\(^{-1}\)) in the De and comparison to other methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Basis</th>
<th>(\text{Be}_2^+)</th>
<th>(\text{C}_2)</th>
<th>(\text{CN})</th>
<th>(\text{N}_2)</th>
<th>(\text{CO})</th>
<th>(\text{NO})</th>
<th>(\text{O}_2)</th>
<th>(\text{F}_2)</th>
<th>MAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{i}-FCIQMC</td>
<td>VQZ</td>
<td>-0.4</td>
<td>-3.5</td>
<td>-3.6</td>
<td>-4.4</td>
<td>-2.9</td>
<td>-3.7</td>
<td>-2.9</td>
<td>-2.1</td>
<td>2.9</td>
</tr>
<tr>
<td>CCSD(T)(^a)</td>
<td>VQZ</td>
<td>-1.0</td>
<td>-4.0</td>
<td>-5.6</td>
<td>-5.0</td>
<td>-2.6</td>
<td>-4.5</td>
<td>-3.2</td>
<td>-2.4</td>
<td>3.5</td>
</tr>
<tr>
<td>CEEIS(^b)</td>
<td>VQZ</td>
<td>-3.4</td>
<td>-4.6</td>
<td>-2.5</td>
<td>-1.9</td>
<td>2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{i}-FCIQMC</td>
<td>V(TQ)Z</td>
<td>0.0</td>
<td>-1.1</td>
<td>-0.6</td>
<td>-0.3</td>
<td>-0.5</td>
<td>-0.7</td>
<td>-0.8</td>
<td>-0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>CCSD(T)(^a)</td>
<td>V(TQ)Z</td>
<td>-0.6</td>
<td>-1.5</td>
<td>-1.2</td>
<td>-0.6</td>
<td>-0.3</td>
<td>-1.0</td>
<td>-0.7</td>
<td>-0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>CEEIS(^b)</td>
<td>V(TQ)Z</td>
<td>-1.1</td>
<td>0.0</td>
<td>0.8</td>
<td>-0.4</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FN-DMC(^c)</td>
<td></td>
<td></td>
<td>-7.5</td>
<td>-3.2</td>
<td>-2.5</td>
<td>-7.0</td>
<td>-6.5</td>
<td>-5.7</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>\textit{i}-FCIQMC</td>
<td>VQZ+F12</td>
<td></td>
<td>-1.5</td>
<td>-0.7</td>
<td>-0.5</td>
<td>-0.1</td>
<td>-0.6</td>
<td>-0.3</td>
<td>-0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>CCSD(T)-F12</td>
<td>VQZ</td>
<td>-1.8</td>
<td>-1.0</td>
<td>-0.6</td>
<td>0.2</td>
<td>-0.8</td>
<td>-0.3</td>
<td>-0.4</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) D. Feller and J. Sordo, J Chem Phys 2000, 113, 485
\(^b\) L. Bytautas and K.Ruedenberg, J Chem Phys 2005, 122, 154110
Semi-Stochastic i-FCIQMC
Petruzielo, Holmes, Changlani, Nightingale, Umrigar, PRL, 109, 230201, (2012)

Select a small subset of determinants $\mathcal{D}$ (called the deterministic space)
Do part of the force update deterministically

$$
- \frac{dN_i}{d\tau} = (H_{ii} - S)N_i + \sum_{j \in \mathcal{D}} H_{ij}N_j + \sum_{j \in \mathcal{D}'} H_{ij}N_j
$$

Death step as per FCIQMC

Spawning step as per i-FCIQMC

Sum computed exactly for given $N_j$ in $\mathcal{D}$

We have parallel implementation of the SS method in the NECI code.

On the fly selection of $\mathcal{D}$ \quad $|\mathcal{D}|=10^6$

N.S. Blunt, S.D. Smart, J.A.F. Kersten, J.S. Spencer, G.H. Booth, and A. Alavi
The Journal of Chemical Physics 142, 184107 (2015)
Relative Efficiency of the Semi-Stochastic method

FIG. 4. The efficiency ($\epsilon_{E_0}$) of semi-stochastic simulations relative to an otherwise identical simulation without semi-stochastic, for the 14-electron homogeneous electron gas with 114 spin orbitals and $r_s = 1.0$ a.u., as the walker population is varied. It is found that the benefit of semi-stochastic tends to increase as the walker population increases, contrary to a simplistic intuition that there should be diminishing returns as stochastic error decreases due to the improved stochastic sampling.
Projected energy via a multi-determinant trial wavefunction

Select a subset of determinants $\mathcal{T}$ (need not be the same as $\mathcal{D}$)

$$\psi_T = \sum_{i \in \mathcal{T}} c_i |D_i\rangle$$

$c_i$ are a set of fixed coefficients obtained by diagonalising $H \in \mathcal{T}$

$$E = \frac{\langle \Psi | H | \Psi_T \rangle}{\langle \Psi | \Psi_T \rangle} = \sum_{i \in \mathcal{C} \mathcal{T}, j \in \mathcal{T}} C_i c_j H_{ij} \quad \sum_{i \in \mathcal{T}} C_i c_i
\quad h_i = \sum_{j \in \mathcal{T}} H_{ij} c_j, \quad i \in \mathcal{C} \mathcal{T}
= \frac{\sum_{i \in \mathcal{C} \mathcal{T}} C_i h_i}{\sum_{i \in \mathcal{T}} C_i c_i}
\quad \mathcal{C} \mathcal{T}$ is space of determinants connected to $\mathcal{T}$
FIG. 3 (color online). Relative efficiency of SQMC vs dimension $|D|$ of the deterministic space for the carbon dimer with a cc-pVTZ basis. Results are shown for trial wave functions of increasing size. The top two curves are for $D$ and $\mathcal{T}$ generated with two applications of our iterative scheme. The 165 and 1766 determinant wave functions with some quadruple excitations have much higher efficiency than the 4282 determinant wave function without any. For this system, $N \approx 10^9$. 
Parallel scaling
Time-step in a.u. for $N_2$ in cc-pVXZ basis sets

\[ \tau (N_2) \]

\[ M \sim X^3 \]

\[ \tau \sim (N^2M^2)^{-1} \]
Uniform Generation Scheme

\[ p_{gen}(A|I) = p_{gen}(B|I) = \ldots = \frac{1}{4} \]

\[ p_s(A|I) = \tau \frac{0.1}{\frac{1}{4}} = 0.4\tau \]

\[ p_s(B|I) = \tau \frac{0.001}{\frac{1}{4}} = 0.004\tau \]

The largest allowable \( \tau \) is set by \( p_s \leq 1 \)

\[ \rightarrow \tau = \frac{1}{0.4} = 2.5 \]

\[ \rightarrow p_s(A|I) = 1 \]

\[ p_s(B|I) = 0.01, etc \]

\[ p(\text{to spawn}) = (3/4) \times .01 + (1/4) \times 1 = 0.2575 \]

Rejection ratio is high

56
Hamiltonian-Weighted Excitation Generation

Construct an algorithm to select $j$ from $i$ so that

$$P_{gen}(j|i) \propto |H_{ij}|$$

In the ideal case (where proportionality is exact), the spawning rate is constant:

$$p_s = \tau \times \text{const}$$

$\tau$ can then be maximised with the condition $p_s \leq 1$ (and $p_d \leq 1$)

To do this exactly costs $O(N^2M^2)$
Weighted Generation Scheme

\[
\sum_A |H_{AI}| = 0.103
\]
\[
p_{gen}(A|I) = \frac{0.1}{0.103} = 0.971
\]
\[
p_{gen}(B|I) = \frac{0.001}{0.103} = 0.00971
\]
\[
p_s(A|I) = p_s(B|I) = 0.103\tau
\]

largest allowable \( \tau = 1/0.103 = 9.71 \)

\[
p_s(A|I) = p_s(B|I) = p_s(C|I) = p_s(D|I) = 1
\]
\[
p(to\ spawn) = 1
\]

Two advantages: allows larger timesteps, and minimises rejections
However, it is possible to devise a determinant selection algorithm based on a Cauchy-Schwarz decomposition of $H_{ij}$ which costs only $O(M)$ to compute!
In practice this allows $\tau$ to scale substantially better with system size, as well as lead to significant overall efficiency.
For opposite-spin excitations

\[ \sigma_i = \sigma_a \neq \sigma_j = \sigma_b \]

\[
\begin{array}{ccc}
D_i & & j \\
\hline
D_j & i & a \\
\hline 
& & b \\
\end{array}
\]

\[ \langle ij|ab \rangle \leq \sqrt{\langle ia|ia \rangle \langle jb|jb \rangle} \]

\[
p(ijab) = p(ab|ij)p(ij)
\]

\[
p(ab|ij) = p(a|i)p(b|j) \quad \text{[For same-spin, include exchange term]}
\]

Select hole \(a\) according to:

\[
p(a|i) \propto \sqrt{\langle ai|ai \rangle}
\]

\[
p(a|i) = \frac{\sqrt{\langle ai|ai \rangle}}{\sum_a \sqrt{\langle ai|ai \rangle}} \leftarrow O(M)
\]
tau for N2 with VXZ basis-sets

[Graph showing decreases in tau values for VDZ, VTZ, and VQZ basis sets for standard and newexcitgen methods.]

cost per walker per successful spawn per unit time

[Graph showing increases in cost per successful spawn for VDZ, VTZ, and VQZ basis sets for standard and newexcitgen methods.]
3-Band Hubbard Model of a Cuprate

10 (Cu) sites with 10 holes (half-filling)

\[
\hat{H} = \sum_{i,\sigma}(\varepsilon_d - \mu)\, d_{i,\sigma}^\dagger d_{i,\sigma} + \sum_{i,\nu,\sigma}(\varepsilon_{p\nu} - \mu)\, p_{i\nu,\sigma}^\dagger p_{i\nu,\sigma} \\
+ \sum_{i,j,\nu,\sigma}(t_{i,j,\nu}d_{i,\sigma}^\dagger p_{j\nu,\sigma} + h.c.) + \sum_{i,\nu,j,\kappa,\sigma\neq i,j}\, t'_{i\nu,j\kappa}p_{i\kappa,\sigma}^\dagger p_{j\nu,\sigma} \\
+ \sum_{i} U_d\, d_{i,\uparrow}^\dagger d_{i,\uparrow}d_{i,\downarrow}^\dagger d_{i,\downarrow}.
\]

(1)

Table 1. Parameters for a three-band model (in eV) calculated with a constrained first principles calculation for La$_2$CuO$_4$ done by Hybertsen et al. [12].

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>$t'$</th>
<th>$t'$'</th>
<th>$U_d$</th>
<th>$U_p$</th>
<th>$U_{pd}$</th>
<th>$U_{pp}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6</td>
<td>1.3</td>
<td>0.65</td>
<td>10.5</td>
<td>4</td>
<td>1.2</td>
<td>0</td>
</tr>
</tbody>
</table>
3-Band Hubbard model energies

\[ N_{FCI} = \left(\frac{30}{5}\right)^2 \approx 20 \times 10^9 \]

- RHF: \(-9.5206318541964\)
- UHF: \(-15.2905361816484\)
- Exact: \(-15.817 (5)\)

\[ N_w \approx 10^9 (\text{UHF basis}) \]

\[ N_w \approx 10^8 (\text{RHF basis}) \]
Which orbitals to use: RHF or UHF?

RHF

Energy of Determinant $E_i$ / a.u.

Coefficient $C_i$ in $\Psi = \sum_i C_i D_i$
Which orbitals to use: RHF or UHF?

UHF
Natural Orbitals: RNO or UNO?

RNO

Coefficient $C_i$ in $\Psi = \sum_i C_i D_i$

Energy of Determinant $E_i$ / a.u.
Natural Orbitals: RNO or UNO?

UNO

Coefficient $C_i$ in $\psi = \sum_i C_i D_i$

Energy of Determinant $E_i / \text{a.u.}$
Reduced density matrices

The 1- and 2-electron RDMs are defined in the orbital basis as:

\[ \gamma_{pq}^q = \langle \Psi | a_p^{\dagger} a_q | \Psi \rangle \]

\[ \Gamma_{pq}^{rs} = \langle \Psi | a_r^{\dagger} a_s^{\dagger} a_q a_p | \Psi \rangle \]

The 1-RDM can be obtained from the 2-RDM by tracing out an electron

\[ \gamma_{pq}^q = \frac{1}{N - 1} \sum_r \Gamma_{pr}^{qr} \]

Normalisation conditions:

\[ \sum_{pq} \Gamma_{pq}^{pq} = \binom{N}{2} \]

\[ \sum_p \gamma_{pq}^p = N \]
Properties such as $E$, $S^2$, electron density, forces, etc can be calculated via the 1 and 2-RDMs

$$\hat{O} = \sum_{ij} \hat{O}_{ij}$$

$$\langle \Psi | \hat{O} | \Psi \rangle = \Gamma_{pq}^{rs} o_{pq}^{rs}$$

**Energy**

$$E = h^q_p \gamma^p_q + \frac{1}{2} g_{pq}^{rs} \Gamma_{rs}^{pq}$$

**Nuclear gradients**

$$\frac{\partial E}{\partial \mathbf{R}} = \gamma^q_p \frac{\partial h^p_q}{\partial \mathbf{R}} + \frac{1}{2} \Gamma_{pq}^{rs} \frac{\partial g_{pq}^{rs}}{\partial \mathbf{R}}$$

**Spin**

$$\langle S^2 \rangle = \frac{3}{4} N + \frac{1}{4} \sum_{ij} \sum_{\sigma} \Gamma^{\sigma j \sigma}_{i \sigma j \sigma} - \frac{1}{2} \sum_{ij} \Gamma^{i \alpha j \beta}_{i \alpha j \beta} - \sum_{ij} \Gamma^{j \alpha i \beta}_{i \alpha j \beta}$$
Calculation of the RDMs

\[ \Gamma_{pq}^{rs} = \sum_i C_i^* C_j \text{ where } |j\rangle = a_r^\dagger a_s^\dagger a_q a_p |i\rangle \]

This is very expensive, because for each occupied determinant, \(i\), need to search over all its double excitations.

Instead, use the fact that in FCIQMC, we sample all double-excitations according to the Hamiltonian matrix elements.

Therefore, we can use the spawning step to stochastically sample the contributions to the 2-RDM!

\[ \Gamma_{pq}^{rs} = \sum_{i \in pq} \frac{C_i C_j}{p[j|i]} \times p[j|i] \quad \text{Spawning probability} \]
In practice, replace the product of the $C_i$ coefficients with the time-average of the products of the walker populations

$$\Gamma_{pq}^{rs} \propto \sum_{i \in pq} \frac{\langle N_i N_j \rangle_\tau}{p[j|i]} \times p[j|i]$$

Normalisation is fixed by:

$$\sum_{pq} \Gamma_{pq}^{pq} = \binom{N}{2}$$
When we do this, the result is not good!

Figure 6.1 – This data, corresponding to simulations on $C_2$ cc-pVTZ, is taken directly from Ref. [25], showing the very slow convergence of $E_{RDM}$ and significant error at large $N_w$, even once $E_P$ is well-converged.
Figure 7.1 – The diagonal elements of the 2-RDM are systematically overestimated in the uncorrected stochastic RDM method. This data corresponds to i-FCIQMC calculations run on $N_2$ cc-pVDZ, $r = 1.094\,\text{Å}$ with 8 frozen core electrons and $N_w = 25000$. The 2-RDM was accumulated stochastically according to the method described in Section 6.1 for $\sim 5 \times 10^6$ iterations. The reals calculation uses $\gamma = 4$, $N_{occ} = 1$. 
What is wrong?

The diagonal matrix elements have a bias

$$\Gamma_{pq} \propto \sum_{i \in pq} \langle N_i^2 \rangle_{\tau}$$

The instantaneous populations fluctuate about their exact value:

$$N_i = N_i^{ex} + \delta N_i$$

$$\langle \delta N_i \rangle_{\tau} = 0$$

But the time-average of the square shows a positive bias:

$$\langle N_i^2 \rangle_{\tau} = (N_i^{ex})^2 + \langle (\delta N_i)^2 \rangle_{\tau}$$
The solution: replica trick

Run two independent simulations in parallel and use the instantaneous populations on the two replicas to compute the RDM!

\[
\Gamma_{pq}^{pq} \propto \sum_{i \in pq} \langle N_i^{(1)} N_i^{(2)} \rangle_{\tau}
\]

Since the two populations are strictly uncorrelated, it is easy to show

\[
\langle \delta N_i^{(1)} \delta N_i^{(2)} \rangle_{\tau} = 0
\]

Therefore:

\[
\Gamma_{pq}^{pq} \propto \sum_{i \in pq} \langle N_i^{(1)} \rangle_{\tau} \langle N_i^{(2)} \rangle_{\tau}
\]
Figure 8.1 – The systematic sampling bias affecting the diagonal elements of the 2-RDM is removed when using the replica sampling method described in Section 8.1. This data corresponds to i-FCIQMC calculations run on N\textsubscript{2} cc-pVDZ, r = 1.094 Å with 8 frozen core electrons and N\textsubscript{w} = 25000. The 2-RDMs are well-converged with respect to simulation time, accumulated stochastically for \( \sim 5 \times 10^6 \) iterations. The reals calculations use \( \gamma = 4 \), \( N\text{occ} = 1 \). Section 7.1 details how these errors are calculated. Different bin widths are used for diagonal and off-diagonal elements to allow the features of each distribution to be clearly seen.

Overy, Booth, Blunt, Shepherd, Cleland, Alavi, JCP, 141, 244117 (2014)
Figure 8.17 – Comparing the uncorrected and replica-sampled RDMs through values of $E_{RDM}$ for $C_2$ cc-pVTZ. This refers directly back to Figure 6.1, though all values have been newly generated for this plot. Uncorrected RDMs were calculated with the integer i-FCIQMC algorithm (consistent with Ref. [25]), whilst replica-sampled RDMs use the real coefficients algorithm with $\chi = 4$, $N_{occ} = 1$, to represent the best quality RDM available with the techniques presented in this chapter.
Figure 9.1 – Energies and internuclear forces in the $H_2$ cc-pVTZ binding curve. The inset of the upper panel shows the i-FCIQMC gradients plotted as tangent lines on the
FIG. 1. Top: Potential energy profile for the N-N bond dissociation of N$_2$ relative to the energy of two isolated nitrogen atoms in the electronic ground state. Bottom: corresponding forces at one nitrogen atom computed using analytic gradients from i-FCIQMC reduced density matrices, compared to FCI with numerical differentiation. Results are identical within the accuracy of the numerical differentiation. The respective minimum energy ($E_{\text{min}} = -0.2685$ a.u.) and force ($F_{\text{min}} = 0.0$ a.u.) at an internuclear distance of 2.144 a.u. is indicated by the blue symbols. All results were obtained with a 6-31G basis set.

FIG. 2. Absolute forces acting on the oxygen and hydrogen atoms in a H$_2$O molecule computed using i-FCIQMC and FCI with a 6-31G and cc-pVTZ basis set (the sign corresponds to the z-component of the force vector). The data were acquired for symmetric displacements of the hydrogen atoms from the equilibrium geometry. The abscissa indicates the OH bond length of the respective molecular geometry.