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

$$H = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i<j}\frac{1}{r_{ij}} + \sum_i v(\mathbf{r}_i)$$

$$\begin{aligned} H\Psi_0 &= E_0\Psi_0 \\ \Psi_0 &= \Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N) \qquad\qquad \mathbf{x} = (\mathbf{r}, \sigma) \end{aligned}$$

Electrons are Fermions:

$$\Psi_0(....\mathbf{x}_i, ..., \mathbf{x}_j ...) = -\Psi_0(....\mathbf{x}_j, ..., \mathbf{x}_i ...)$$

Atomic units $\hbar = m_e = |e| = 1$ E_h = 27.211 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 ⇒

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
⇒*Uncontrolled error.*

Full configuration interaction

$$H = \sum_i^N -\frac{1}{2}\nabla_i^2 + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i^N v(\mathbf{r}_i)]$$

$$H\Psi_0 = E_0\Psi_0$$

$$\Psi_0 = \sum_{\mathbf{i}} C_{\mathbf{i}} |D_{\mathbf{i}}\rangle$$

Include all (symmetry-allowed) determinants within basis

Variationally minimise wrt C_i

$$\Rightarrow \sum_{\mathbf{i}} \langle D_{\mathbf{j}}|H|D_{\mathbf{i}}\rangle C_{\mathbf{i}} = E_0 C_{\mathbf{j}}$$

Ground-state eigenvalue problem in an exponentially large space

$\langle D_{\mathbf{i}}|H|D_{\mathbf{j}}\rangle$ can be positive or negative: this is a source of sign problem, but is NOT the Fermion sign problem!

Largest FCI calculation to date ~ 10^{10} determinants (N_2 molecule)
[E. Rossi, GL Bendazzoli, S. Evangelisti, D Meynau, Chem Phys Lett, 310, 530,(1999)]

Hubbard model: 159×10^9 determinants [Yamada, Imamure, 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]

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

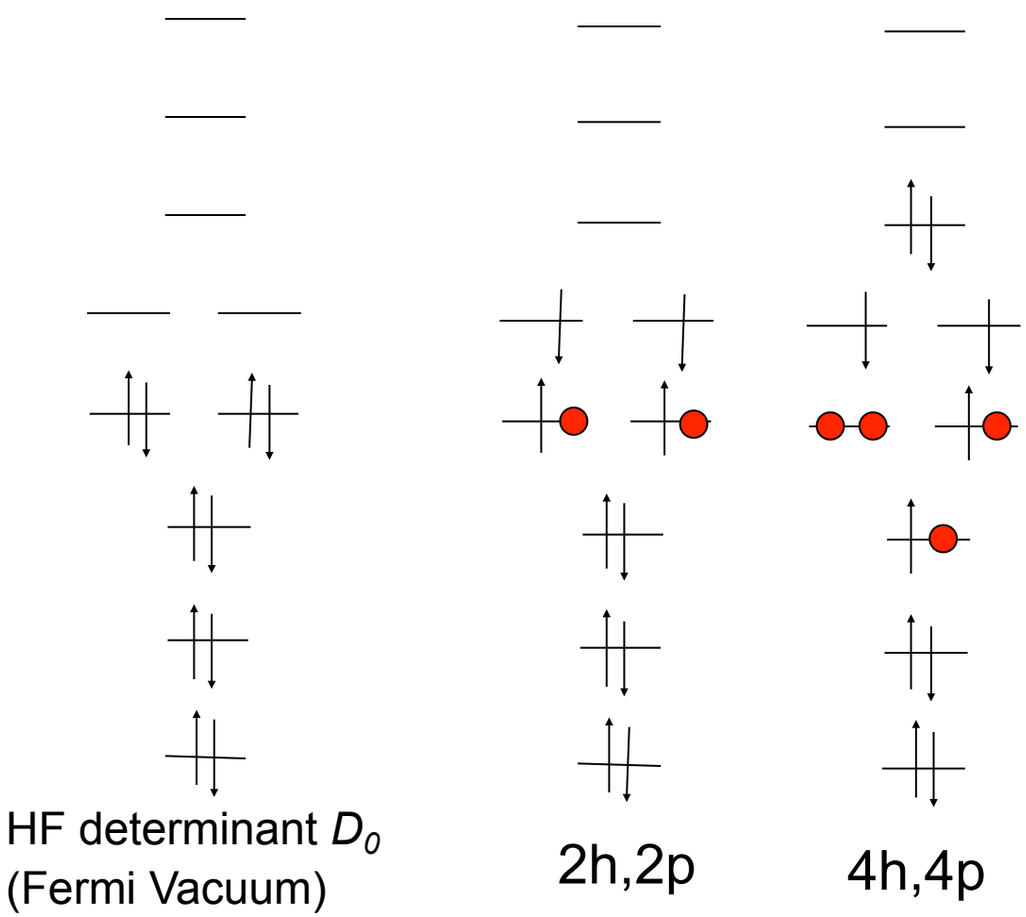
$$\left|D_{\mathbf{i}}\right\rangle = \left|D_{n_1 n_2 .. n_N}\right\rangle = \frac{1}{\sqrt{N!}} \det[u_{n_1} u_{n_2} ... u_{n_N}]$$

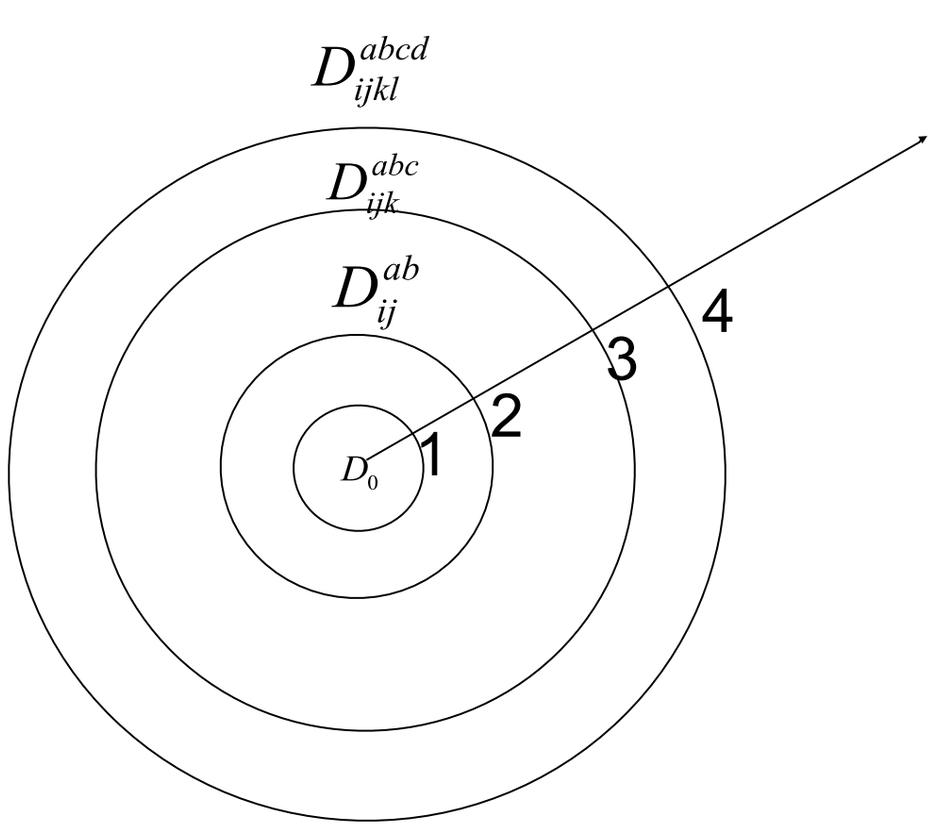
Antisymmetric *N*-electron basis

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

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

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

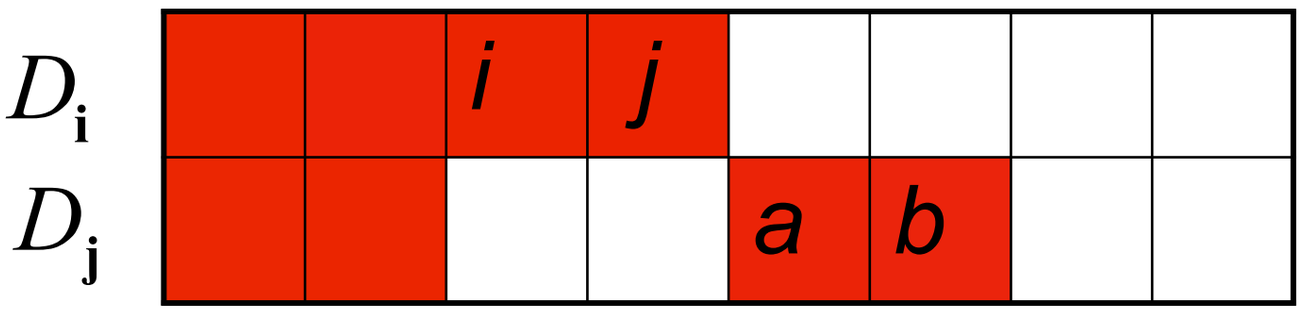




Hamiltonian matrix elements (Slater-Condon rules)

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

$\langle D_{\mathbf{i}} | H | D_{\mathbf{j}} \rangle = 0$ if $D_{\mathbf{i}}$ and $D_{\mathbf{j}}$ differ by more than 2 spin-orbitals



$$\langle D_{\mathbf{i}} | U | D_{\mathbf{j}} \rangle = \langle ij | r_{12}^{-1} | ab \rangle - \langle ij | r_{12}^{-1} | ba \rangle$$

Hamiltonian connects only single and double excitations:

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

Spin selection rule:

$\langle D_{\mathbf{i}} | H | D_{\mathbf{j}} \rangle = 0$ if $S_z[D_{\mathbf{i}}] \neq S_z[D_{\mathbf{j}}]$

Other symmetries may also exist
translational invariance;
Molecules:point group symmetry

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

$$\langle D_{\mathbf{i}} | H | D_{\mathbf{j}} \rangle = \sum_{k} \left[\langle ik | ak \rangle - \langle ik | ka \rangle \right]$$

for Di,Dj differing by one spin-orbital

$$\langle D_{\mathbf{i}} | H | D_{\mathbf{j}} \rangle = \langle ij | ab \rangle - \langle ij | ba \rangle$$

for Di,Dj differing by two spin-orbitals

From “amplitudes” to “walkers”

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

$$\{\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_{N_w}\}$$

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

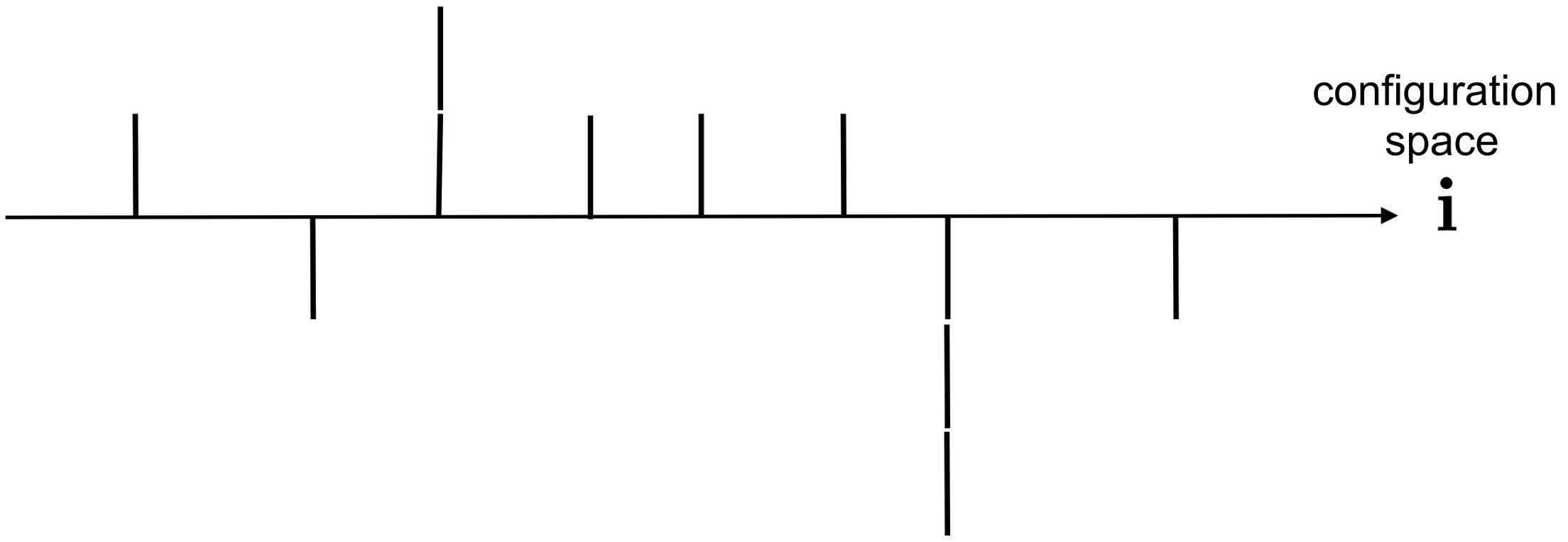
$$\{s_1, s_2, \ldots, s_{N_w}\}$$

We will *define* the amplitude on **i** to be the signed sum of walkers on $D_{\mathbf{i}}$:

$$C_{\mathbf{i}} \propto \sum_{j=1}^{N_w} s_j \delta_{\mathbf{i},\mathbf{i}_j} = N_{\mathbf{i}}$$

Pictorial example

$$N_w = \sum_{\mathbf{i}} |N_{\mathbf{i}}| = 11$$



$$\text{Normalisation} = \frac{1}{\sqrt{\sum_{\mathbf{i}} N_{\mathbf{i}}^2}} = \frac{1}{\sqrt{19}}$$

A differential formulation for the CI coefficents

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

$[E_0^{(0)} = \langle D_0 | H | D_0 \rangle = \text{HF energy.}]$

$K_{ii} \geq 0$

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 ⇒ 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:

Let $K_{ij} = H_{ij} - E_{HF}\delta_{ij}$

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

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

Diagonal Death (or cloning) processes

Annihilation processes

Reduce/enhance population on **i**, proportional to K_{ij} and C_j

Reduce if Kij Cj >0

Enhance if Kij Cj <0

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 τ

↓

Spawning (birth) processes

(Diagonal) death/cloning processes

Annihilation step

Pairwise removal of walkers of opposite sign on the same det

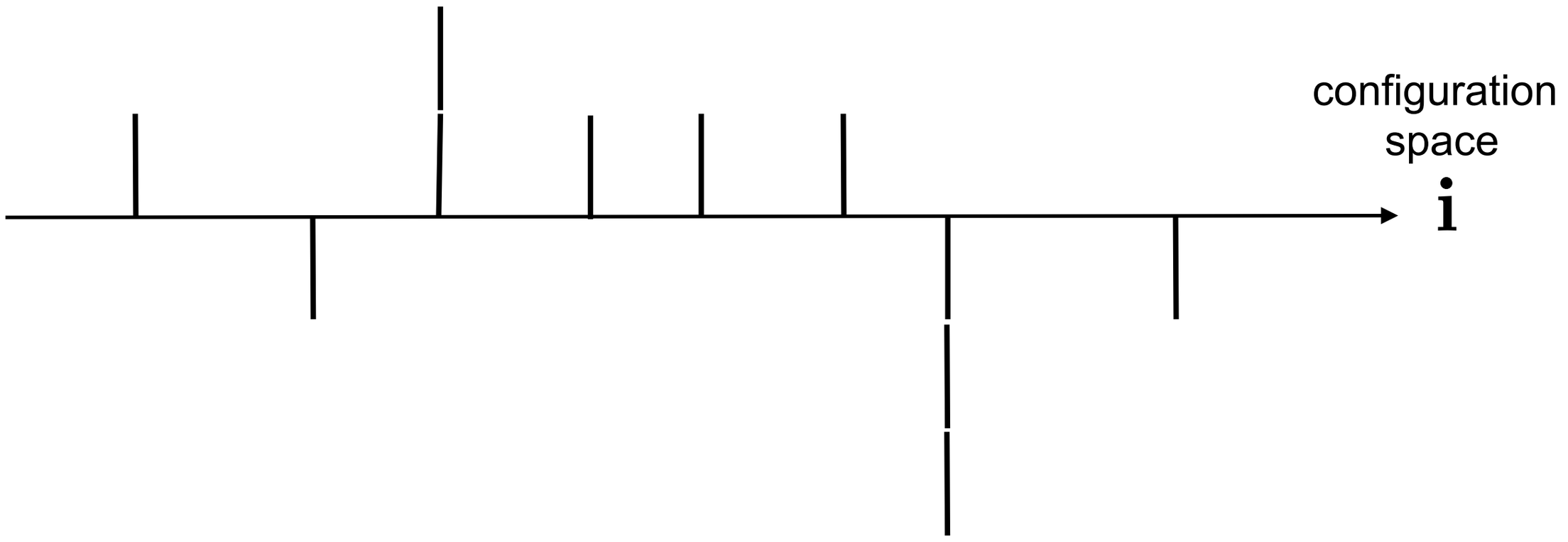
Adjust Shift (in constant N mode)

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

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

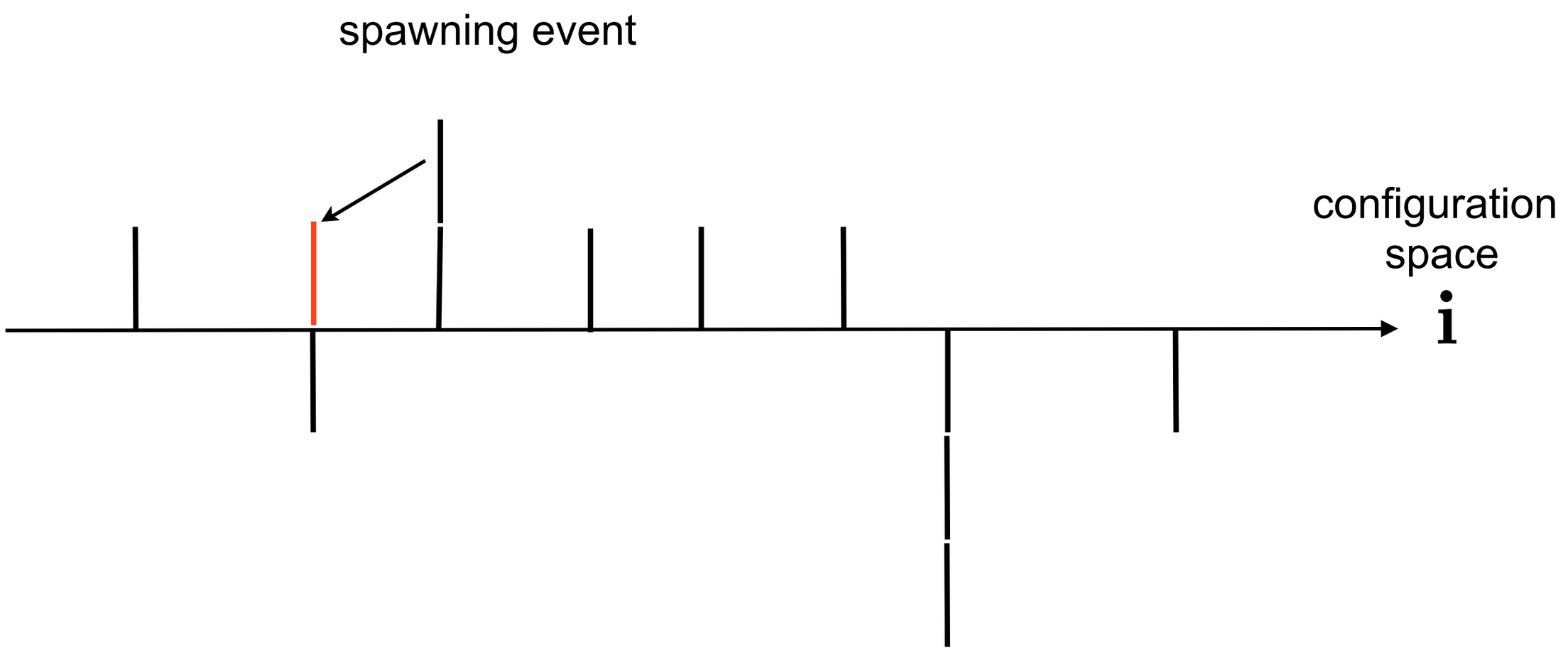
Pictorial example

$$N_w = \sum_{\mathbf{i}} |N_{\mathbf{i}}| = 11$$

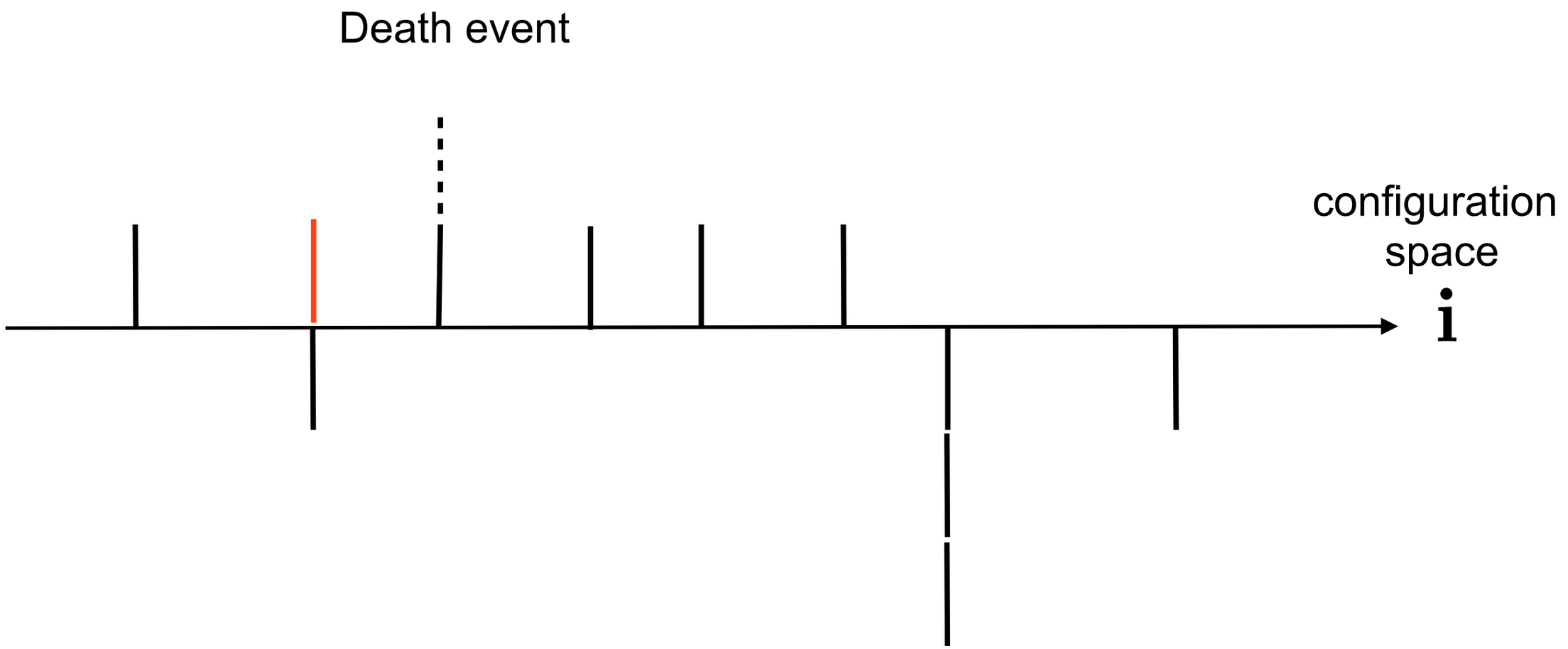


$$\text{Normalisation} = \frac{1}{\sqrt{\sum_{\mathbf{i}} N_{\mathbf{i}}^2}} = \frac{1}{\sqrt{19}}$$

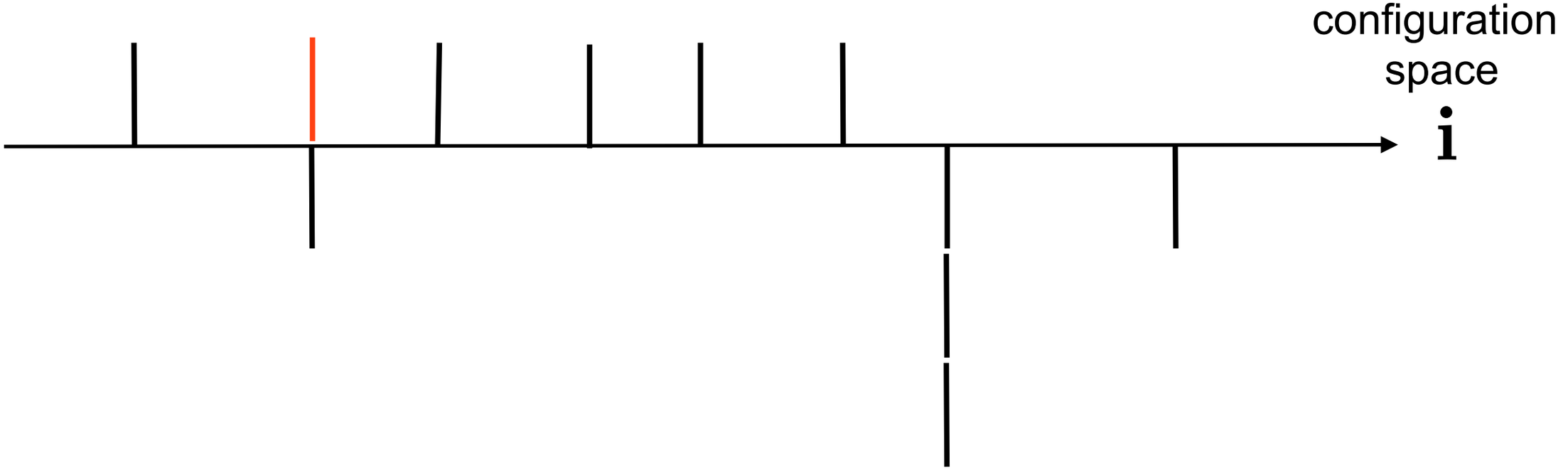
Spawning



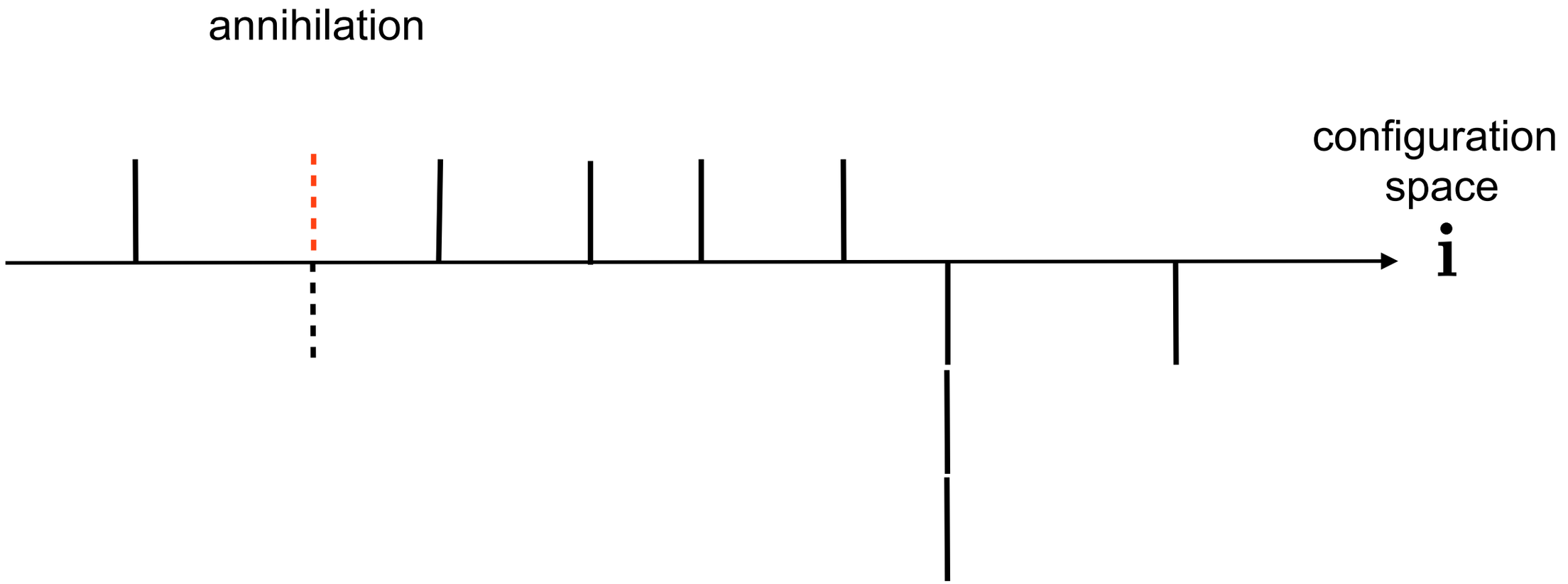
Death



Death



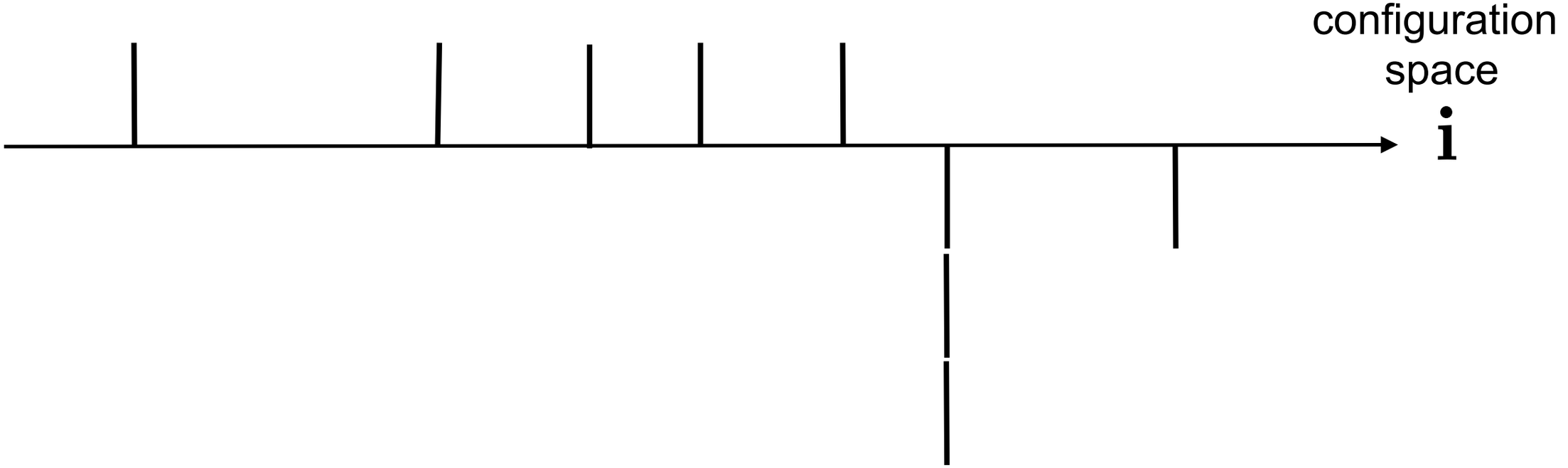
Annihilation



Hash algorithm: $O(N_w)$

Booth, Smart, Alavi, Mol. Phys., **112** (14), (2014),1855-1869

Annihilation

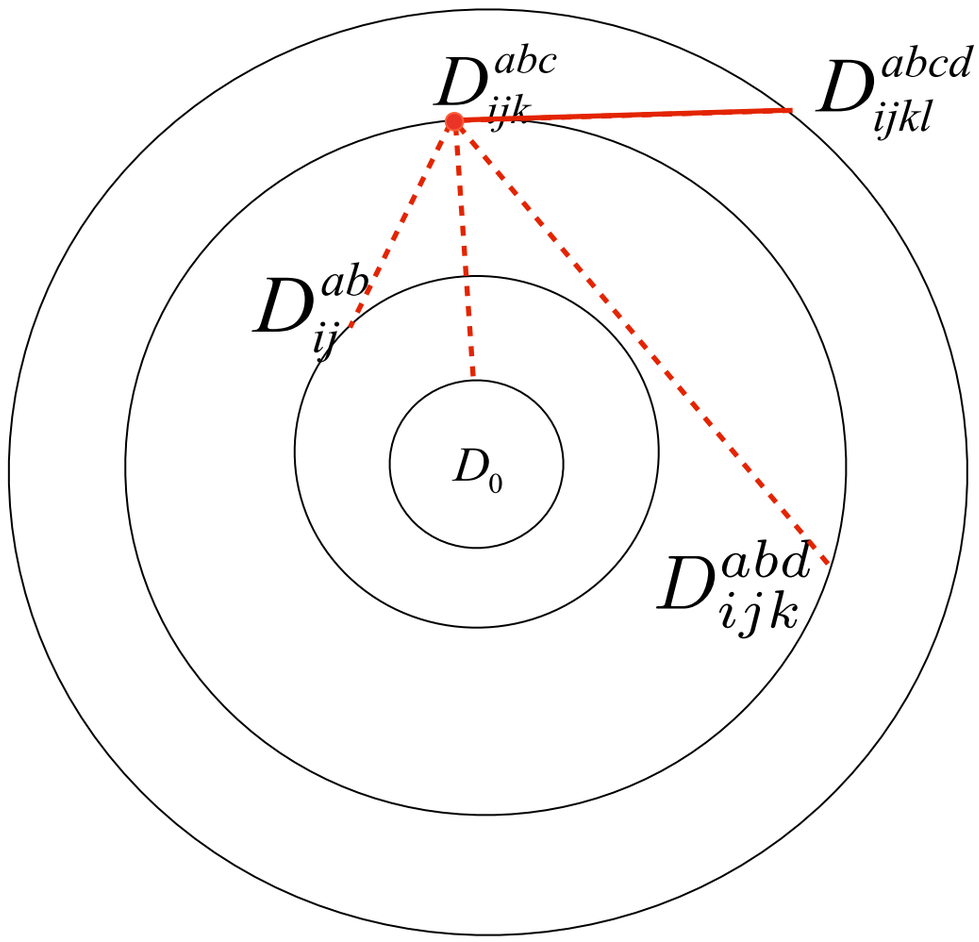


The rules of the “Game of Life”

(derived from the underlying imaginary-time S.E.)

Probability of death

$$p_d = \tau |H_{\mathbf{ii}} - E_{HF} - S|$$



Probability to spawn new walker

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

$$\sum_{\mathbf{j}} p_{gen}(\mathbf{j}|\mathbf{i}) = 1$$

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

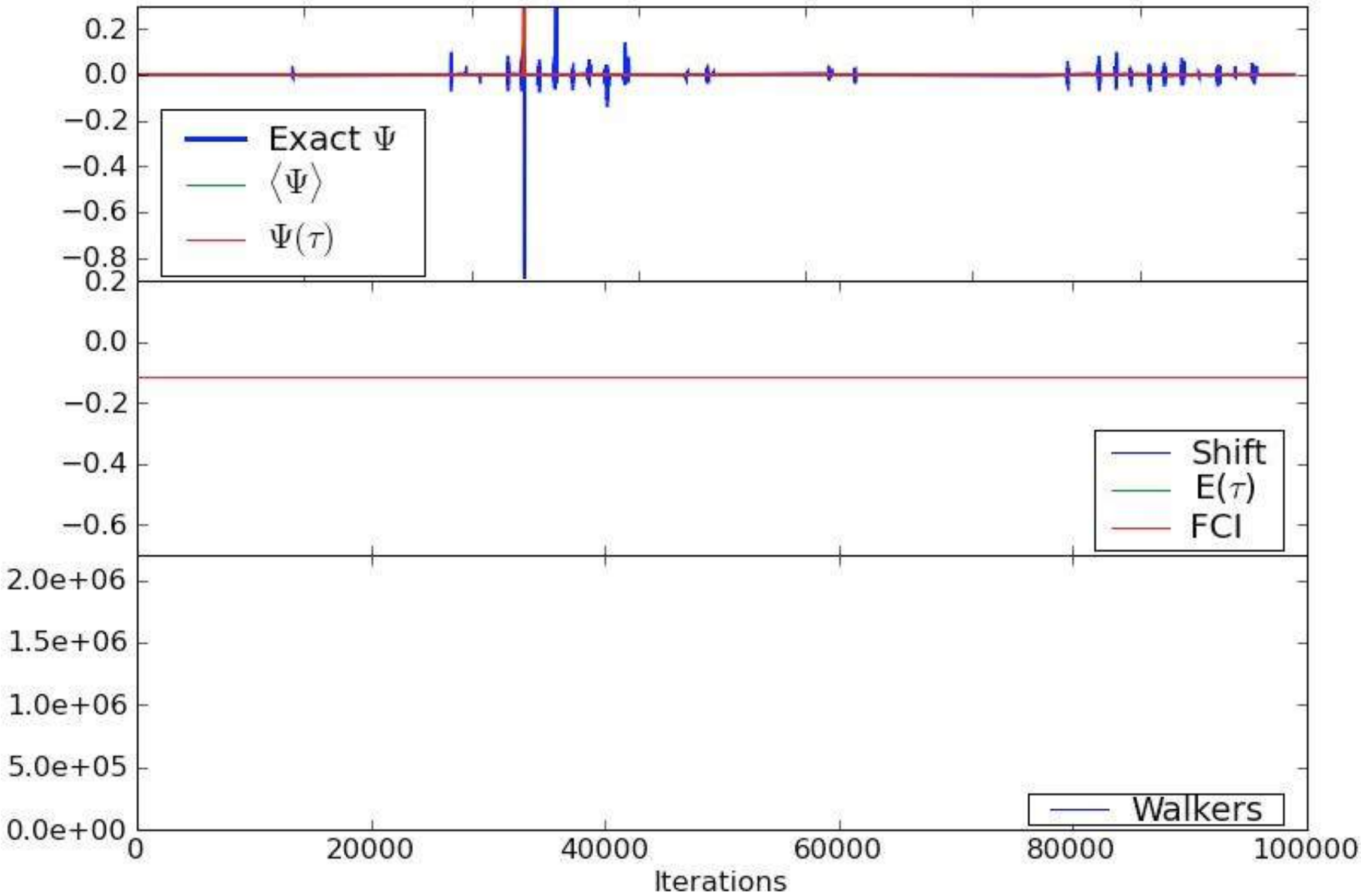
If Hij < 0 , child has same sign as parent.
If Hij > 0 child has opposite sign of parent

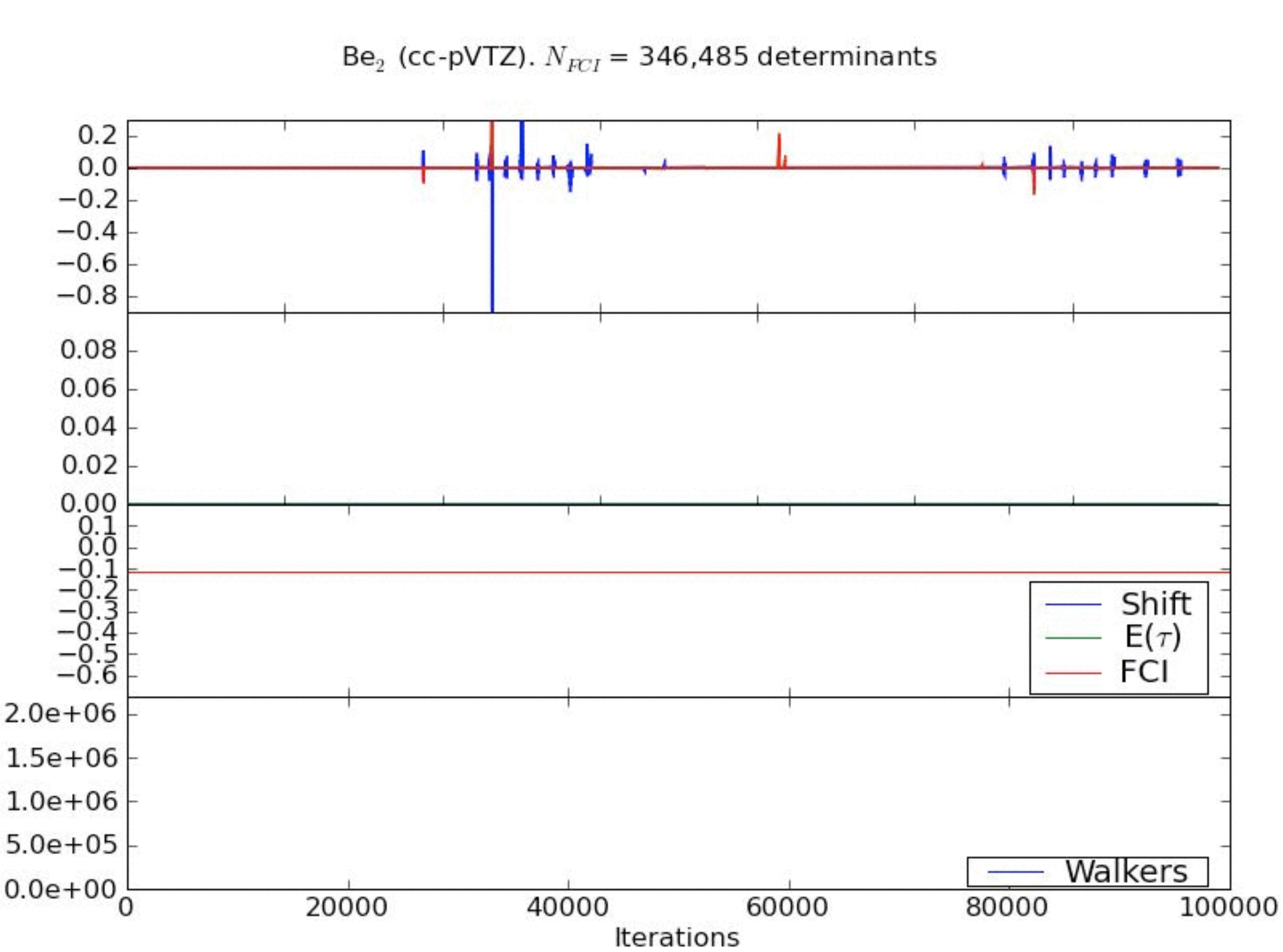
The projected energy

(non-variational)

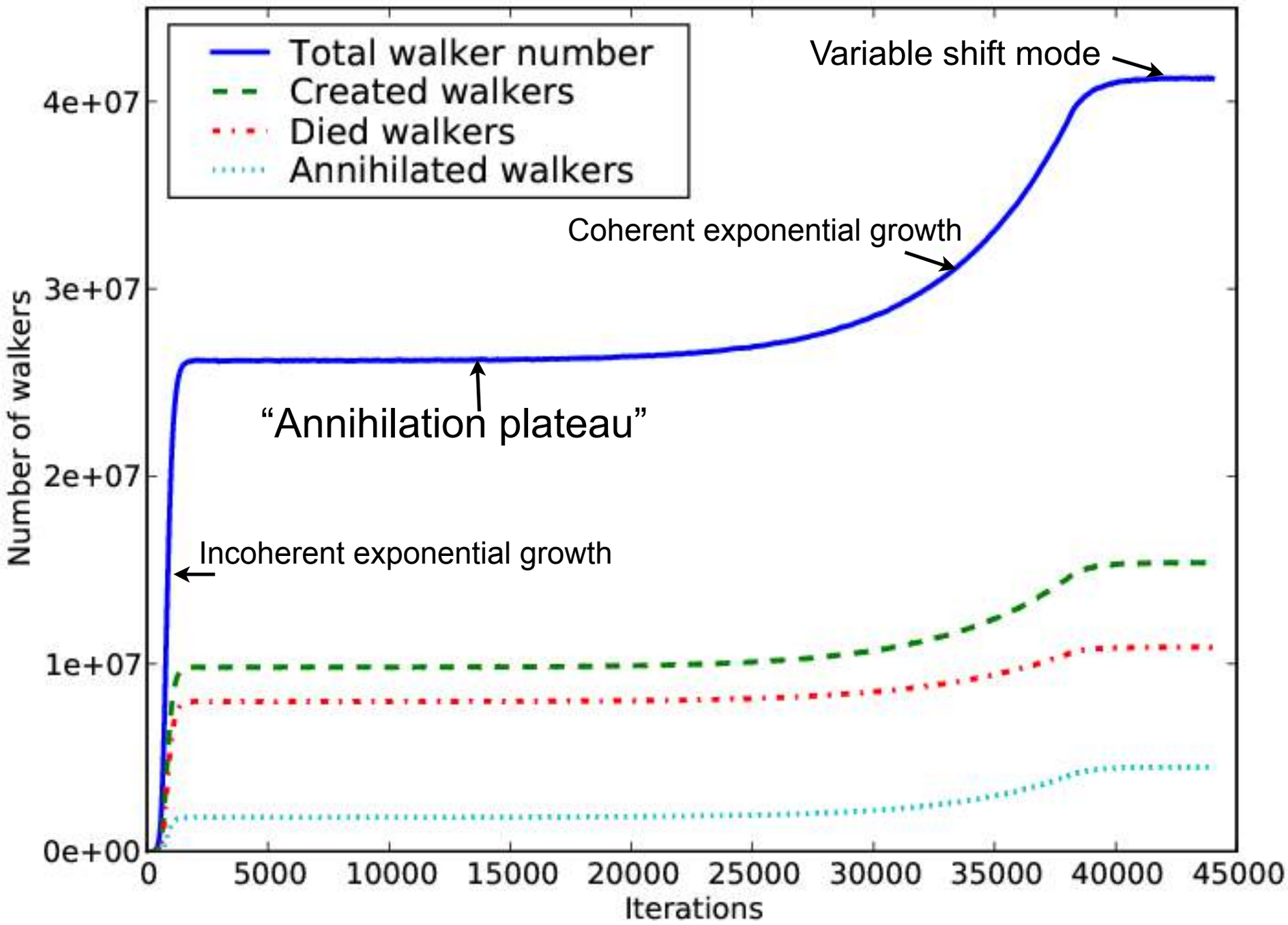
$$
\begin{aligned}
E &= \frac{\langle D_0|H|\Psi\rangle}{\langle D_0|\Psi\rangle} \\
&= \frac{\sum_{\mathbf{j}} \langle D_0|H|D_{\mathbf{j}}\rangle\langle D_{\mathbf{j}}|\Psi\rangle}{\langle D_0|\Psi\rangle} \\
&= E_{HF} + \sum_{\mathbf{j}\in\text{doubles}} \langle D_0|H|D_{\mathbf{j}}\rangle \frac{C_{\mathbf{j}}}{C_0}
\end{aligned}
$$

where $\dfrac{C_{\mathbf{j}}}{C_0} = \dfrac{N_{\mathbf{j}}}{N_0}$


Be$_2$ (cc-pVTZ). N_{FCI} = 346,485 determinants
0.2
0.0
−0.2
−0.4
−0.6
−0.8
Exact Ψ
⟨Ψ⟩
Ψ(τ)
0.2
0.0
−0.2
−0.4
−0.6
Shift
E(τ)
FCI
2.0e+06
1.5e+06
1.0e+06
5.0e+05
0.0e+00
Walkers
0
20000
40000
60000
80000
100000
Iterations


Be$_2$ (cc-pVTZ). N_{FCI} = 346,485 determinants
0.2
0.0
−0.2
−0.4
−0.6
−0.8
0.08
0.06
0.04
0.02
0.00
0.1
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
Shift
E(τ)
FCI
2.0e+06
1.5e+06
1.0e+06
5.0e+05
0.0e+00
Walkers
0
20000
40000
60000
80000
100000
Iterations

H2O (all electron, cc-pVDZ, 452×10^6 determinants)



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: 1.272 73 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_2O. N_{FCI} is the size of the FCI space in the D_{2h} point group (C_{2v} for H_2O. The digit in italics for E_{total}, represents the first uncertain digit. N_c is the number of walkers required to achieve the plateau. $f_c = N_c / N_{\mathrm{FCI}}$.

System	(N,M)	$N_{\mathrm{FCI}}/10^6$	$N_c/10^6$	f_c	E_{total}	E_{FCI}	Reference
Ne: aug-cc-pVDZ	(8,22)	6.69	0.21	0.031	−128.70949	−128.709,476	33
C_2: cc-pVDZ	(8,26)	27.9	15.0	0.538	−75.7299	−75.729,853	34
H_2O: cc-pVDZ	(10,24)	451	26	0.058	−76.24186	−76.241,860	35
N_2-eqm: cc-pVDZ	(10,26)	541	270	0.499	−109.27649	−109.276,527	33
N_2-stretched: cc-pVDZ	(10,26)	541	345	0.637	−108.9669	−108.966,95	36

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}=1.087 728), CO (1.1448), H_2O(r_{OH}=0.975 512, θ=110.565°) (Ref. 35), O_2 (1.2074), and NaH (1.885 977). 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 $E(\tau)$.

System	*(N,M)*	$N_{FCI}/10^6$	$N_c/10^6$	f_c	E_{total}	$E_{CCSD(T)}$
Be: cc-V5Z	(4,91)	2.11	0	0	−14.646 38(2)	−14.646 29
CN: cc-pVDZ	(9,26)	246	173	0.704	−92.493 8(3)	−92.491 64
HF: cc-pCVDZ	(10,23)	283	0.998	0.0035	−100.270 98(3)	−100.270 44
CH_4: cc-pVDZ	(8,33)	419	377	0.898	−40.387 52(1)	−40.389 74
CO: cc-pVDZ	(10,26)	1080	777	0.719	−113.056 44(4)	−113.054 97
H_2O: cc-pCVDZ	(10,28)	2410	47.4	0.0196	−76.280 91(3)	−76.280 28
O_2: cc-pVDZ	(12,26)	5409	2651	0.490	−149.987 5(2)	−149.985 62
NaH: cc-pCVDZ	(12,32)	205 300	63.8	0.000 31	−162.609 0(1)	−162.609 01

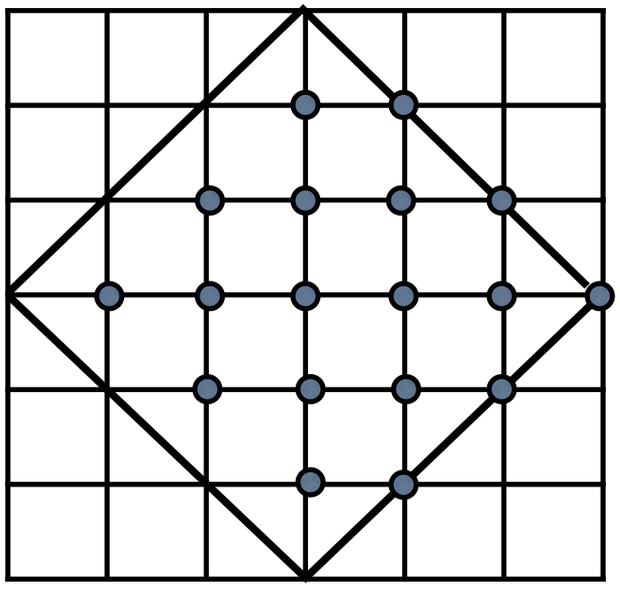
Hubbard Model

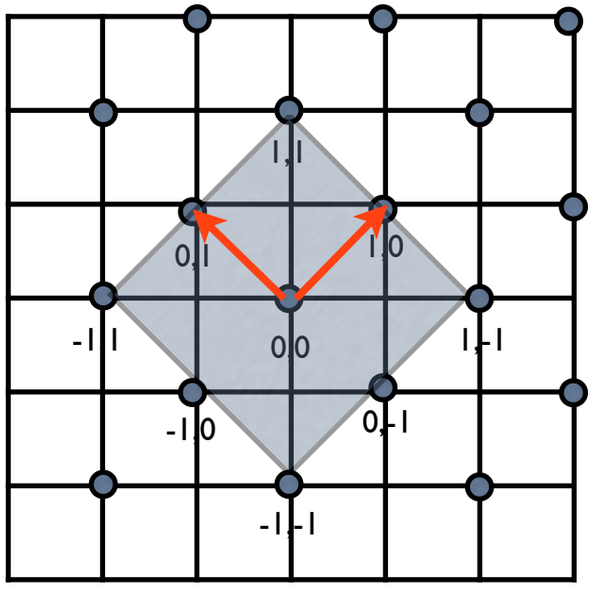
$$\hat{H} = -t \sum_{\langle p,q \rangle,\sigma} \left(c^{\dagger}_{p\sigma} c_{q\sigma} + \text{h.c.} \right) + U \sum_{p} n_{p\uparrow} n_{p\downarrow}$$

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{\Omega}} \sum_{p} e^{-i\mathbf{k}.\mathbf{r}_p} |p\rangle$$

$$\hat{H} = \sum_{\mathbf{k},\sigma} \epsilon_{\mathbf{k}} c^{\dagger}_{\mathbf{k},\sigma} c_{\mathbf{k},\sigma} + \frac{U}{\Omega} \sum_{\mathbf{k},\mathbf{q},\mathbf{Q}} c^{\dagger}_{\mathbf{k}+\mathbf{Q}\uparrow} c^{\dagger}_{\mathbf{q}-\mathbf{Q},\downarrow} c_{\mathbf{k}\uparrow} c_{\mathbf{q}\downarrow}$$

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



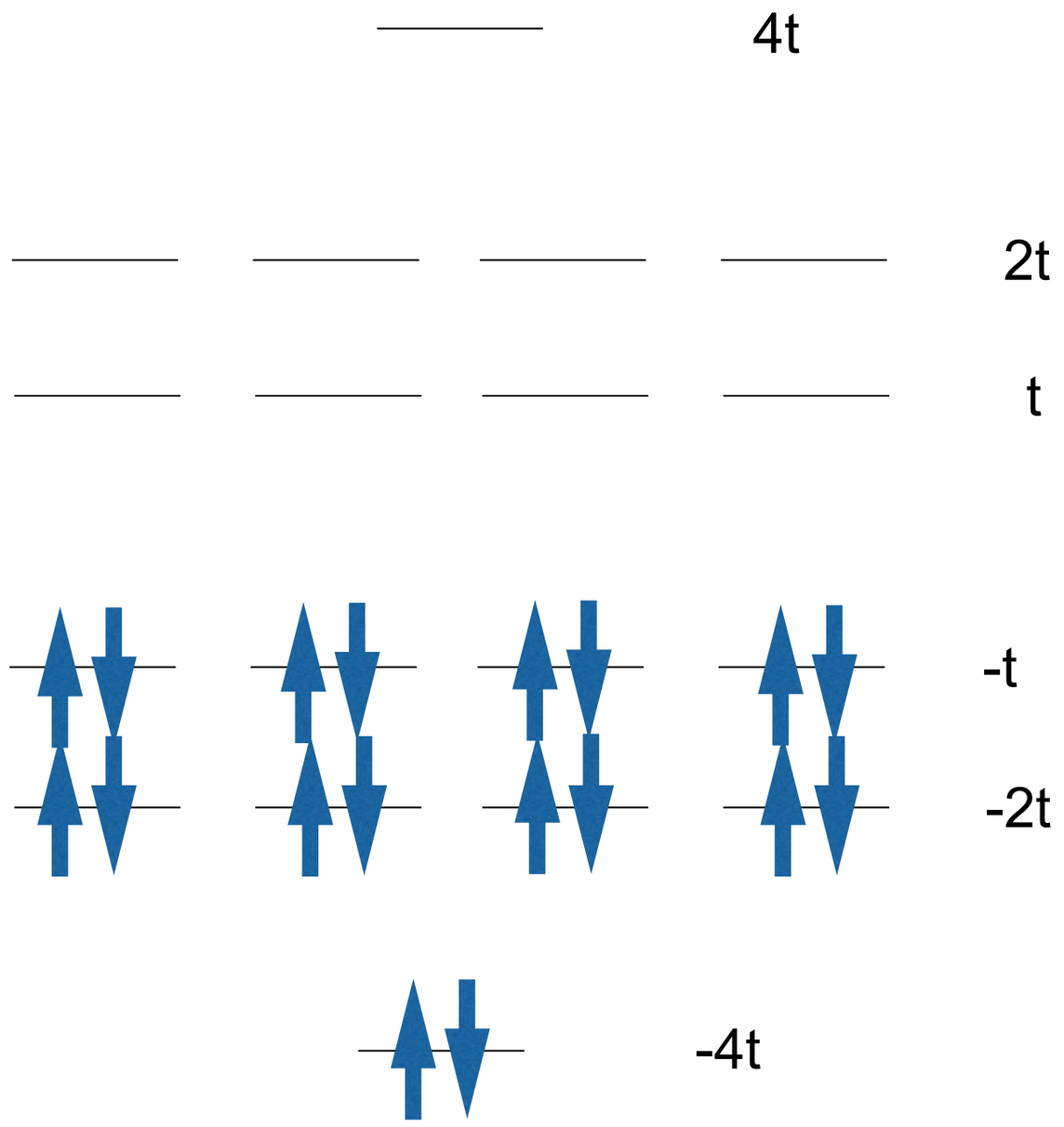


$$\mathbf{T_1} = (L, L)$$
$$\mathbf{T_2} = (-L, L)$$

$$\mathbf{b_1} = \frac{\pi}{L}(1, 1)$$
$$\mathbf{b_2} = \frac{\pi}{L}(-1, 1)$$
$$\mathbf{k}_{n,m} = n\mathbf{b_1} + m\mathbf{b_2}$$

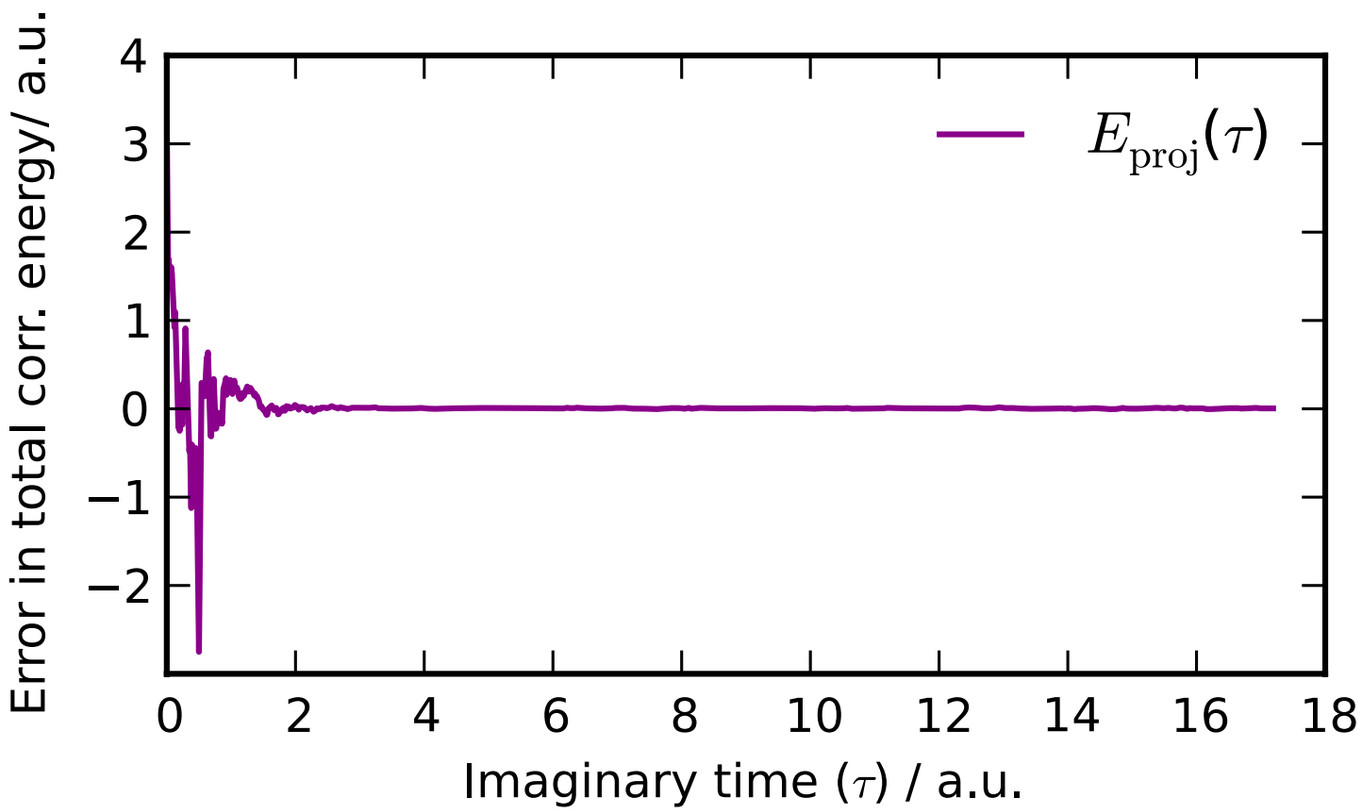
$$\epsilon_{\mathbf{k}} = -2t \left(\cos \frac{\pi(n+m)}{L} + \cos \frac{\pi(n-m)}{L} \right)$$

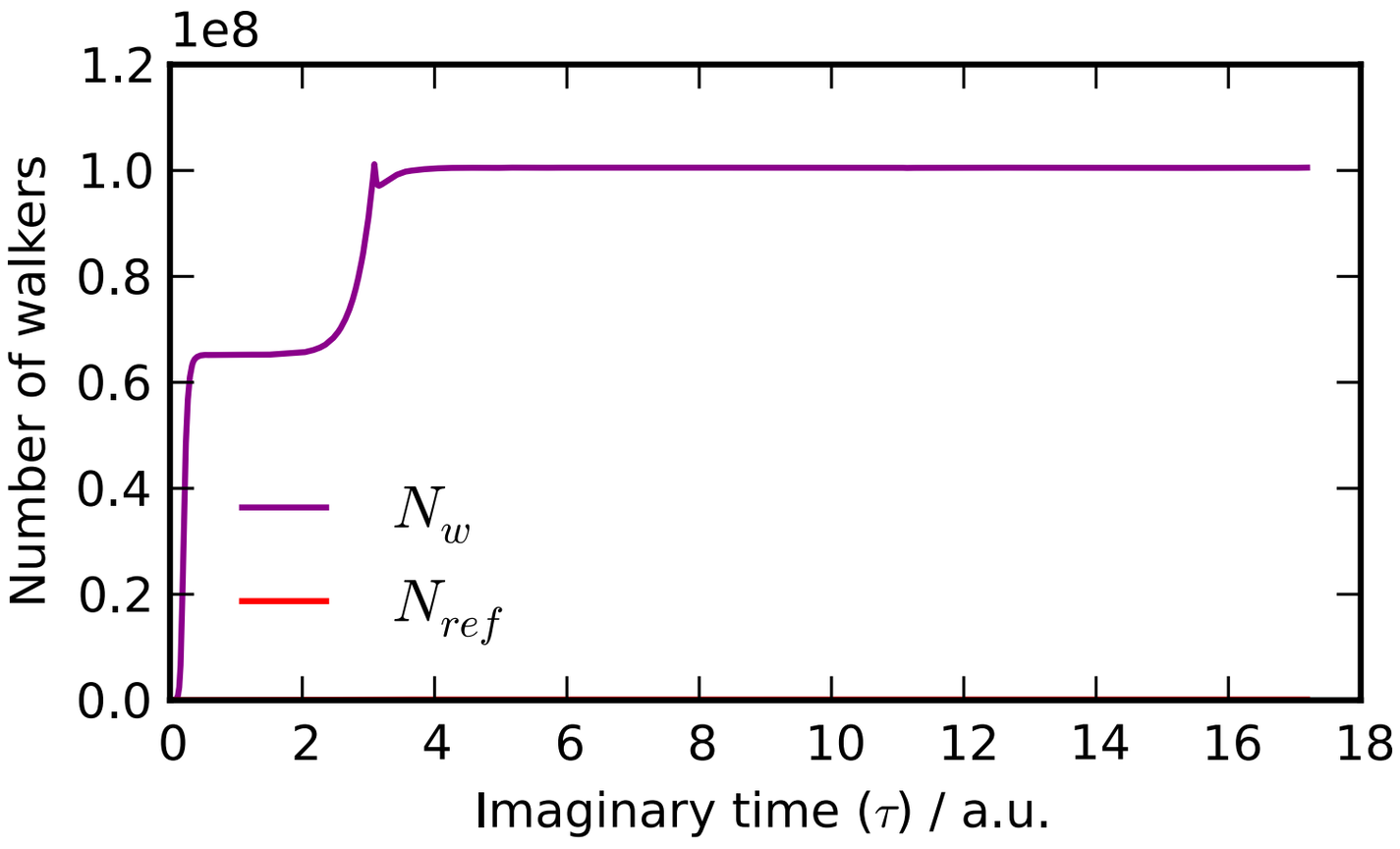
Hubbard model energy levels

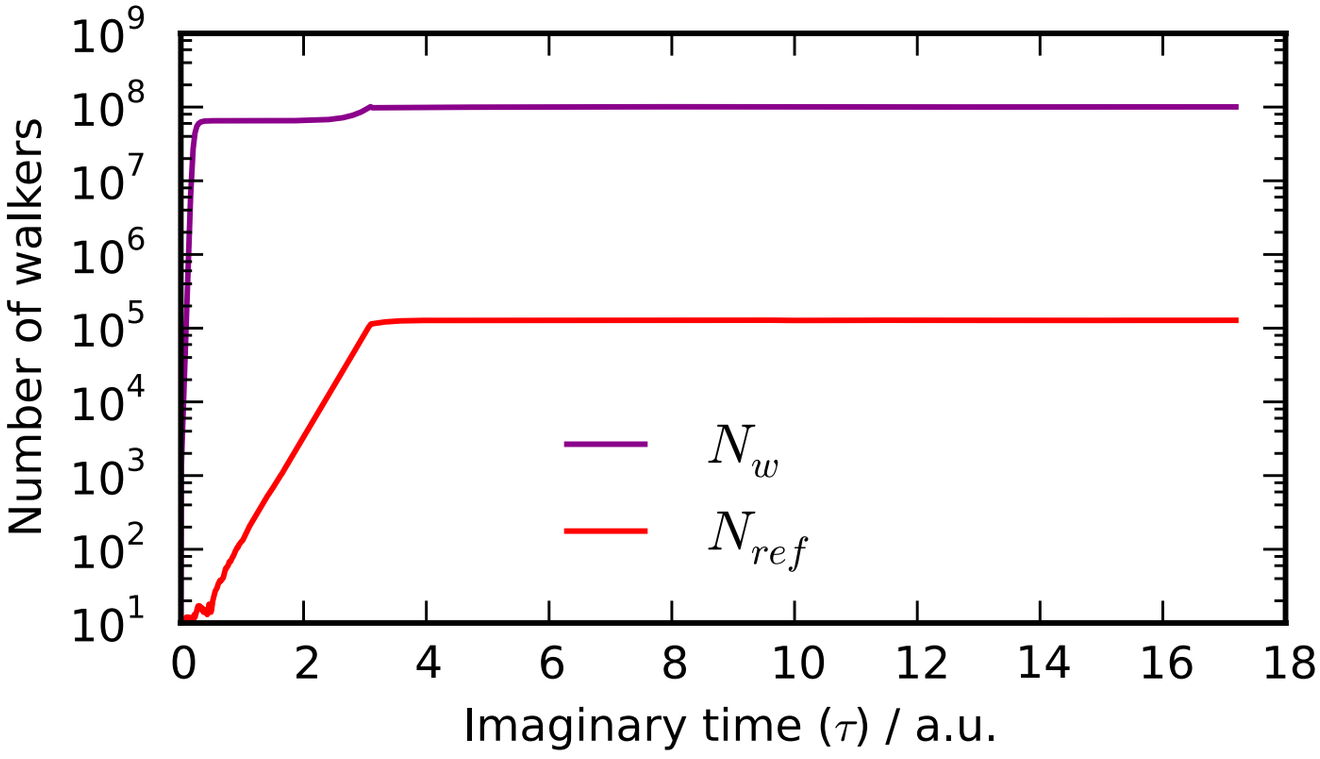


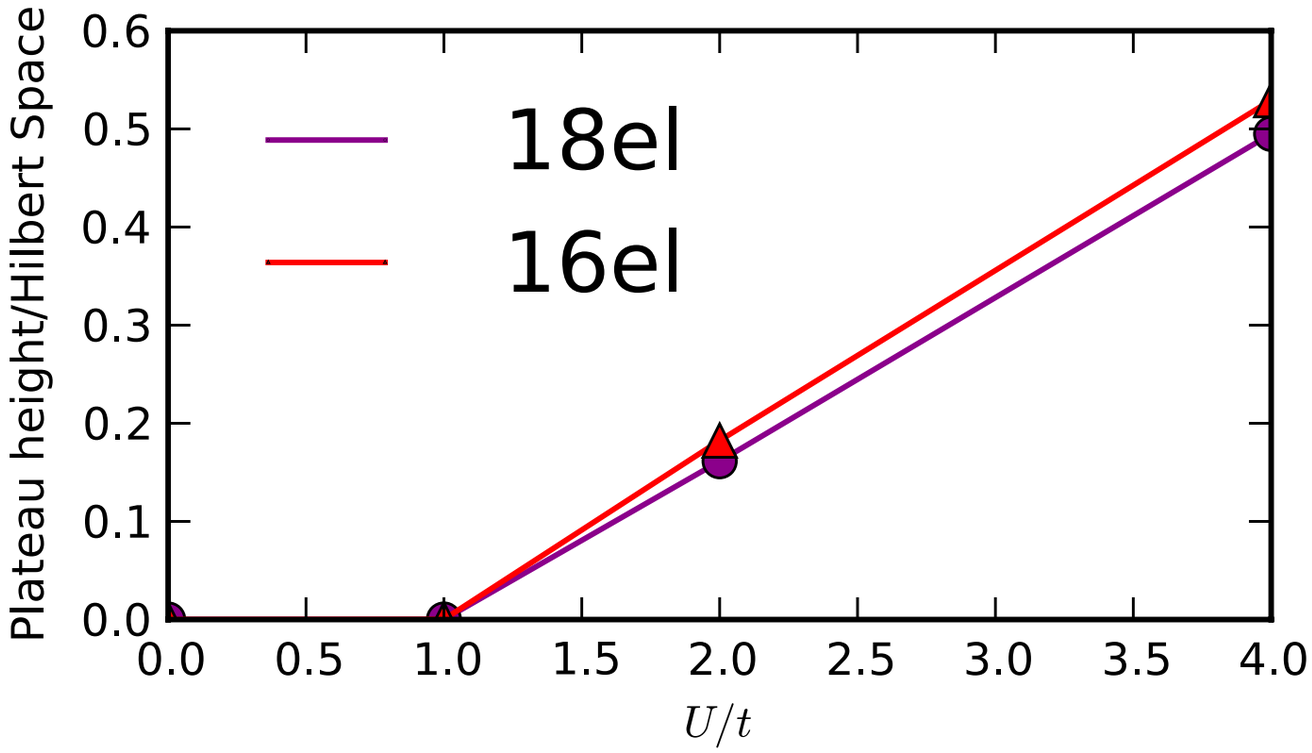
Hilbert space of K=0 sector at half-filling $N_{FCI} \approx 131 \times 10^8$

18 site-18 electron U=4 Hubbard model
Time evolution of the energy



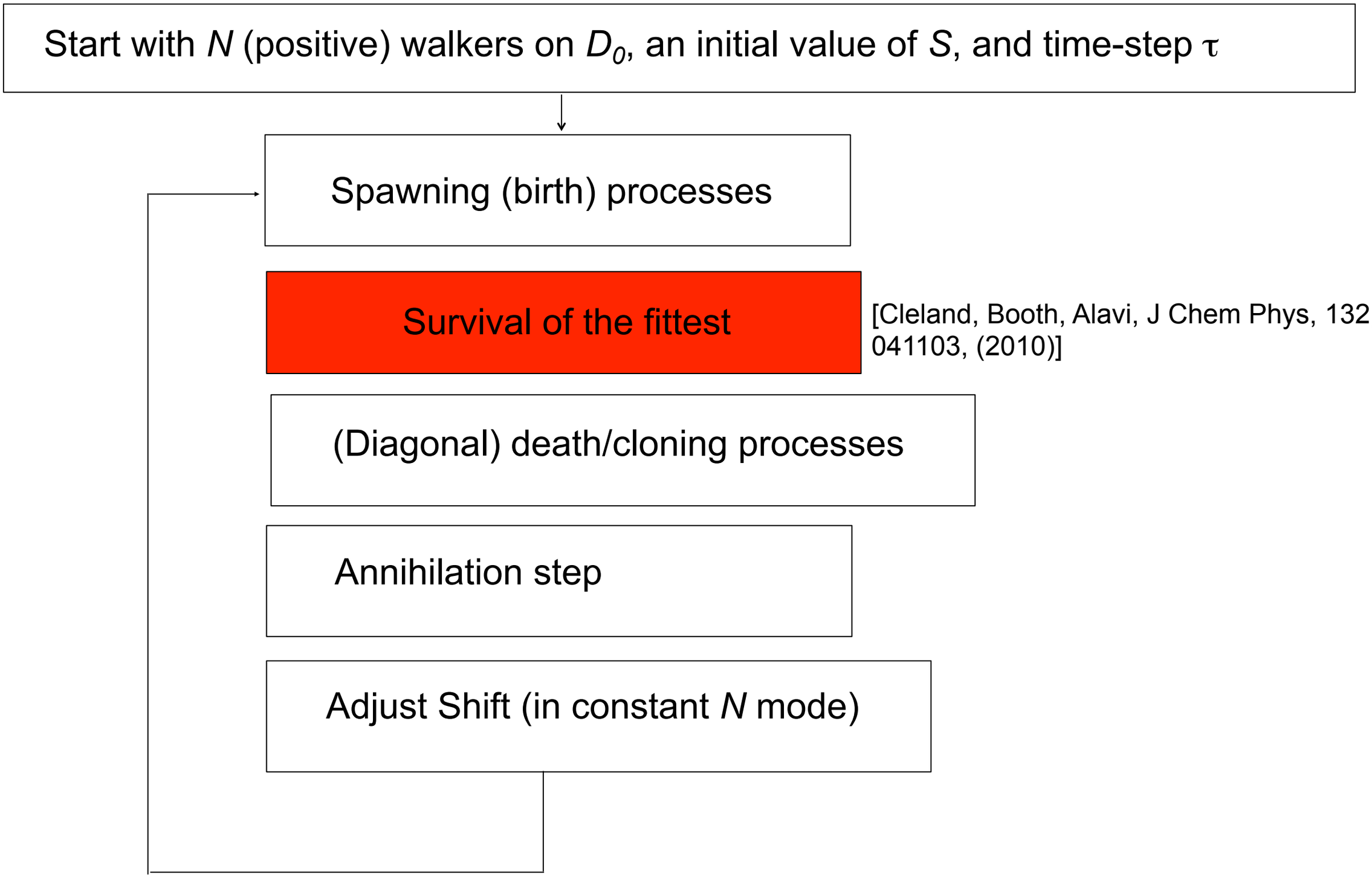

1e8
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Number of walkers
N_w
N_{ref}
0
2
4
6
8
10
12
14
16
18
Imaginary time (τ) / a.u.


Number of walkers
10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
0
2
4
6
8
10
12
14
16
18
Imaginary time (τ) / a.u.
N_w
N_{ref}


18el
16el
Plateau height/Hilbert Space
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
U/t

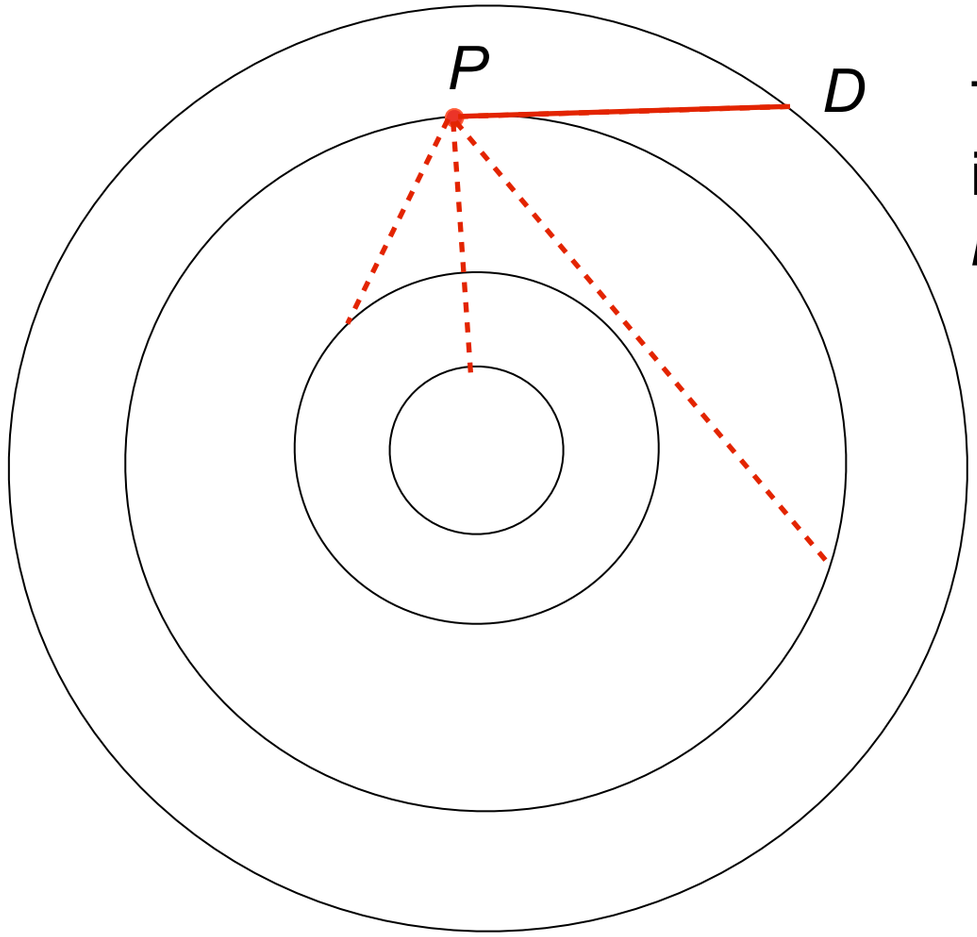
How to reduce N_w while maintaining FCI accuracy?

Overview of initiator-FCIQMC



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.

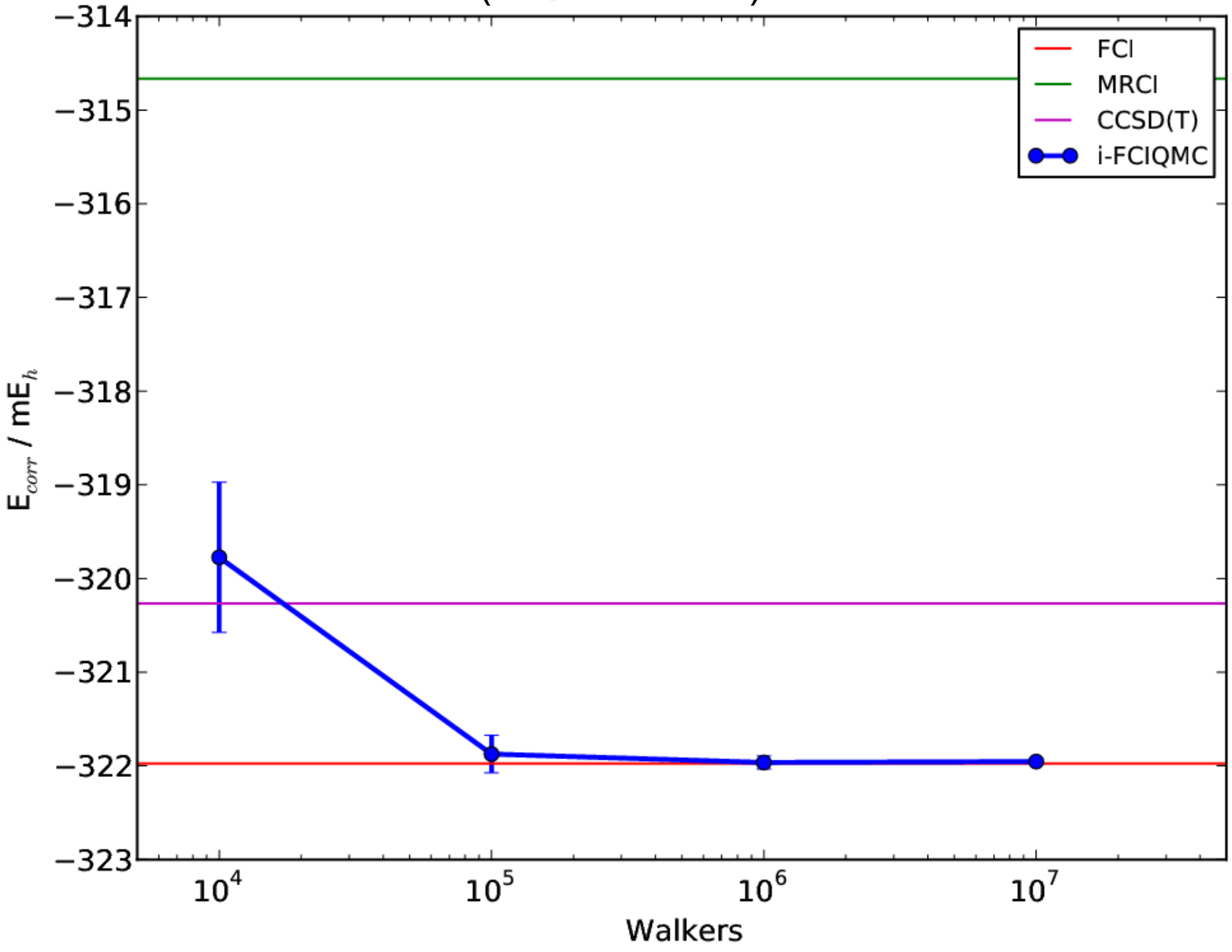
Initiators can bring to life new determinants

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

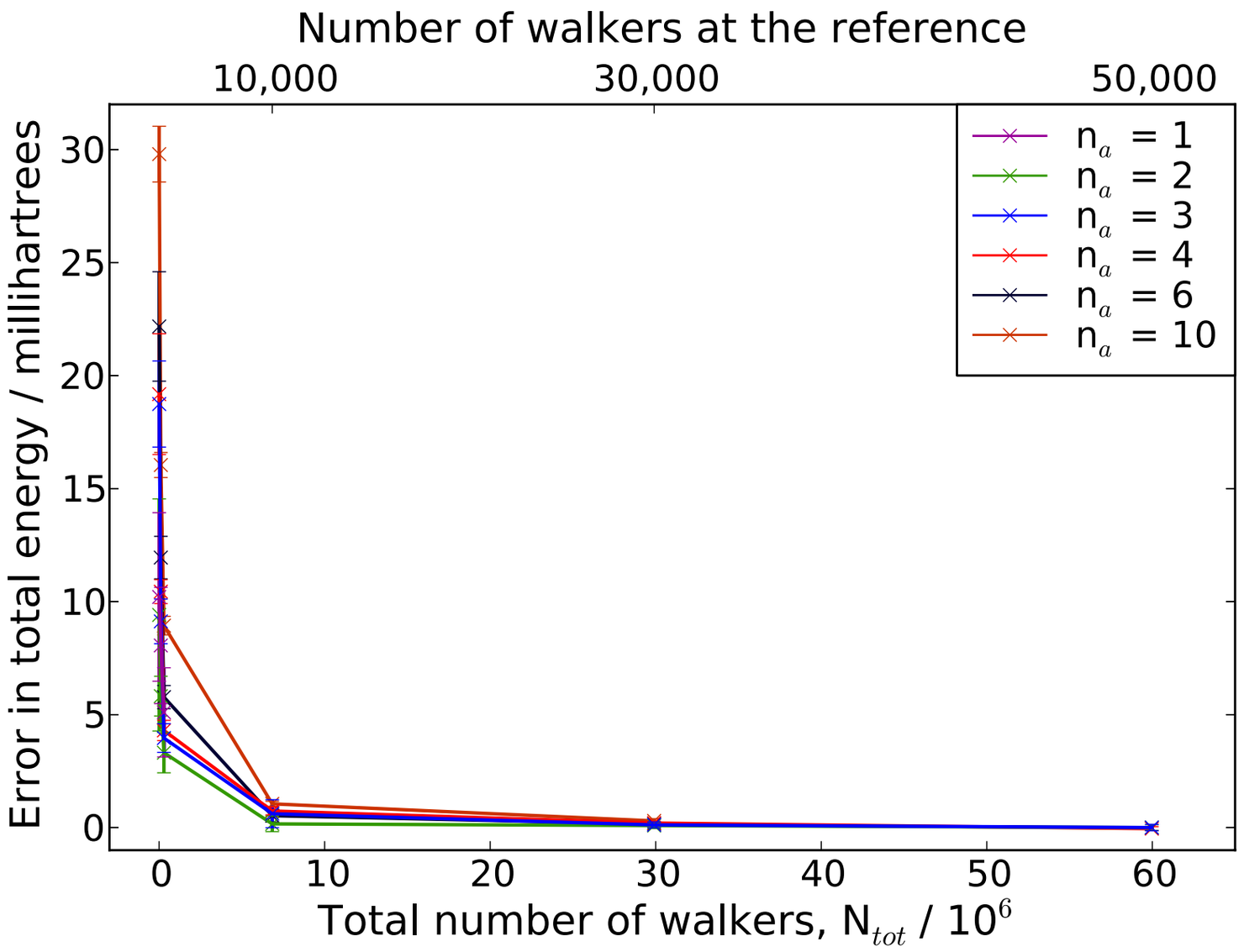
(N_{FCI}=5.4×10^8)



The effect of varying n_{add}

CO in cc-pVQZ

$N_{FCI}=4.7 \times 10^{14}$

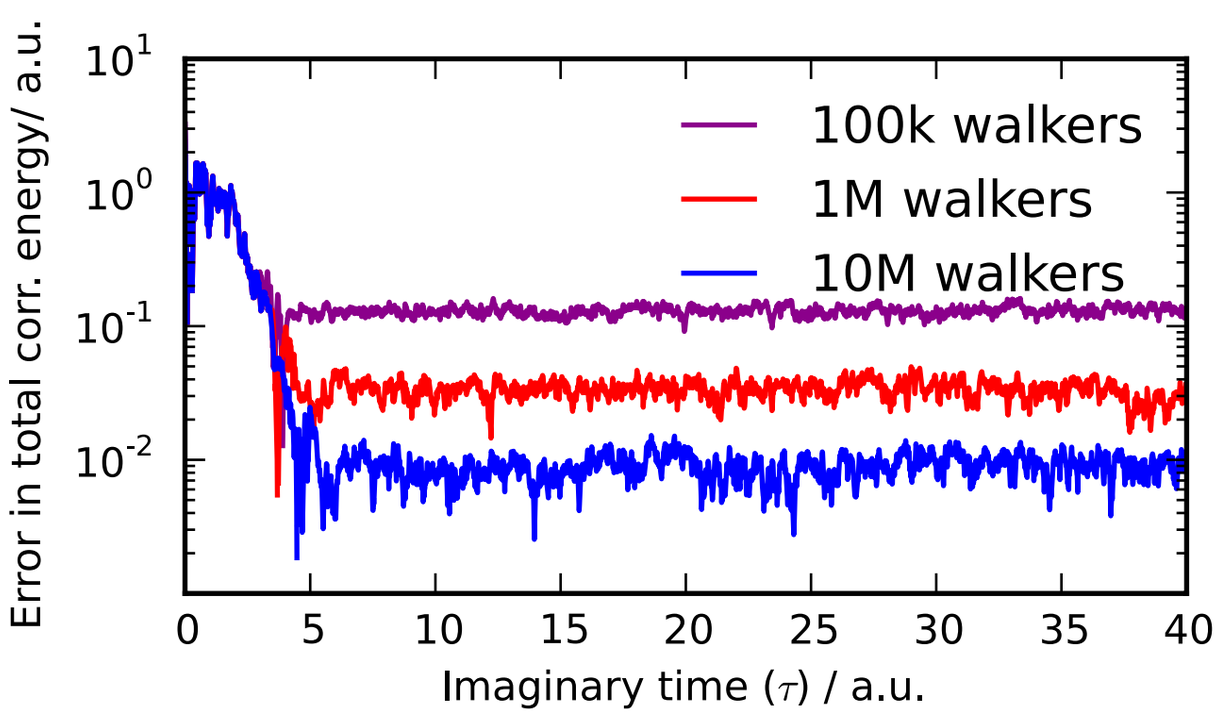


Live Demo

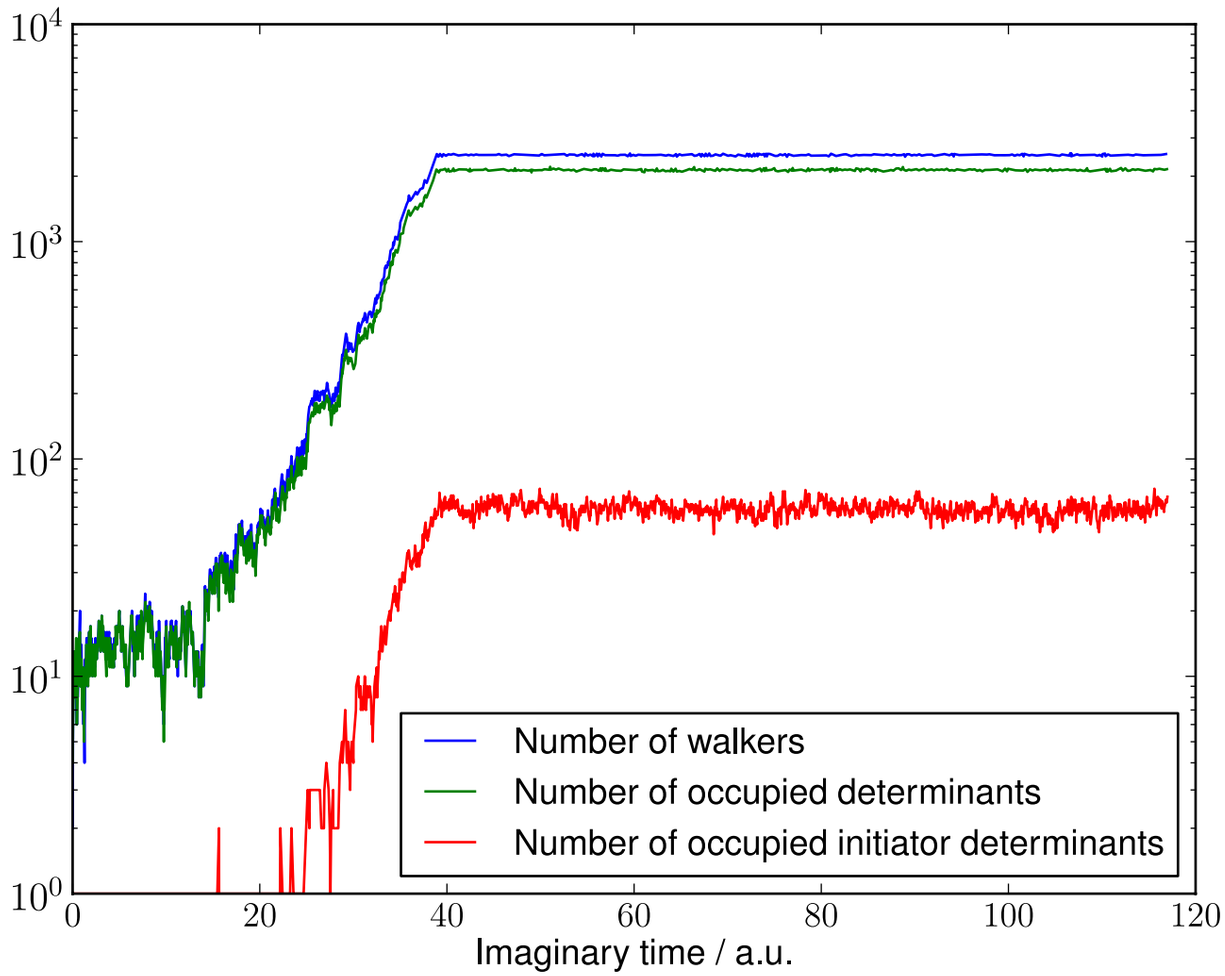
Initiator method

(ninit=3)

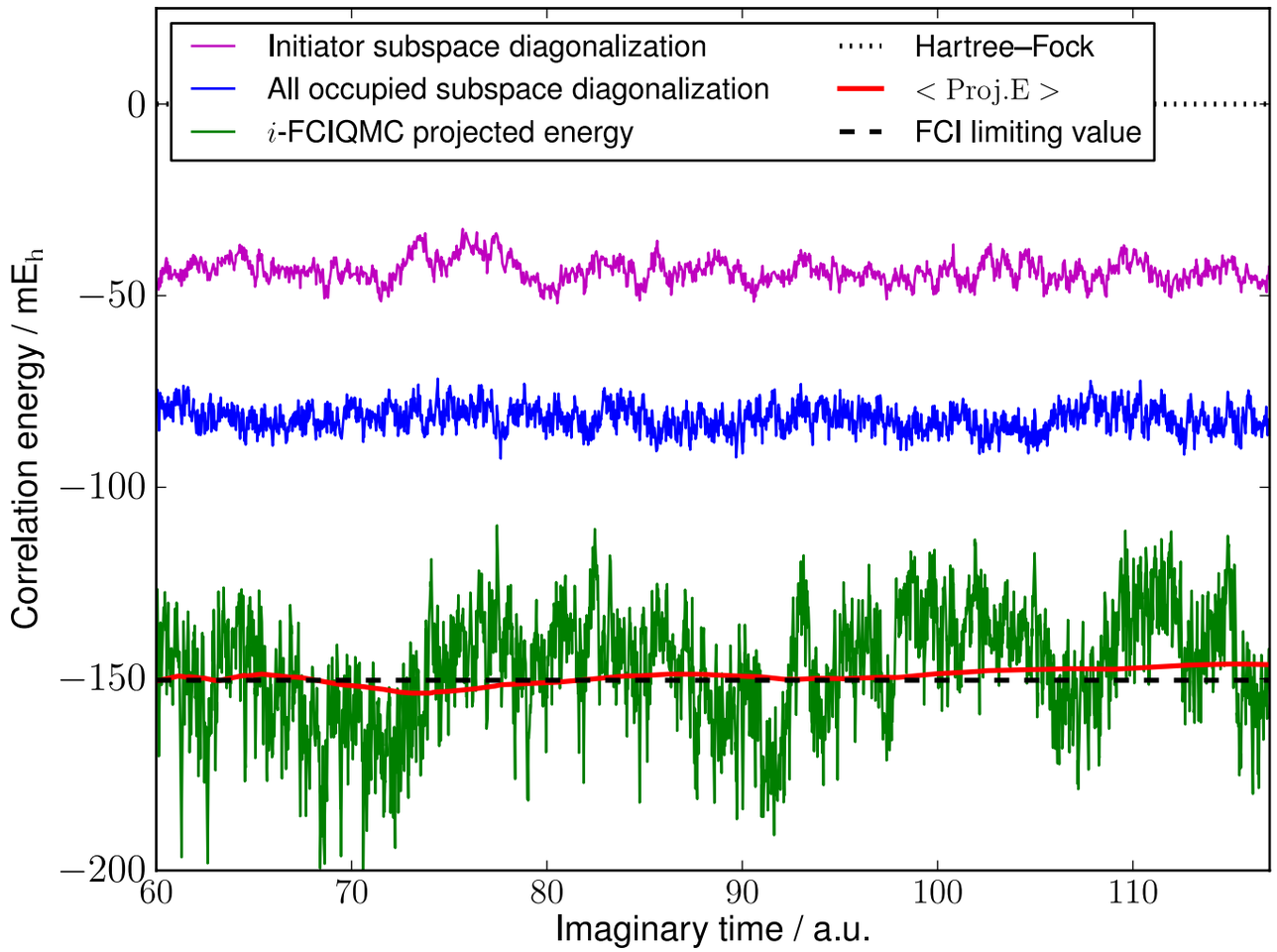
18 site-18 electron U=4 Hubbard model



The Be_2 (VTZ) with i-FCIQMC: a simulation with 2,000 walkers

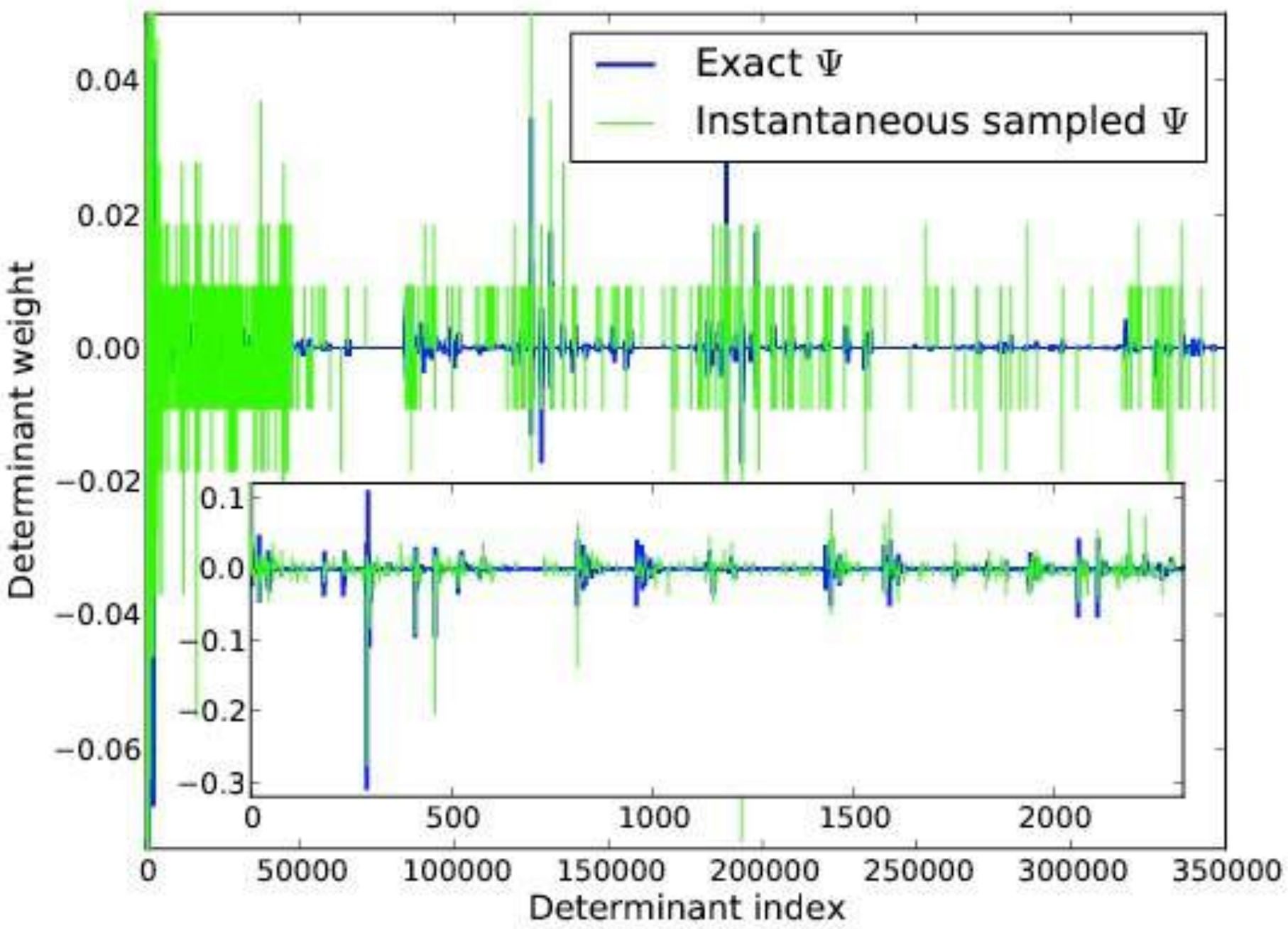


Energy, compared with subspace diagonalisations (c.f. Stochastic CI)

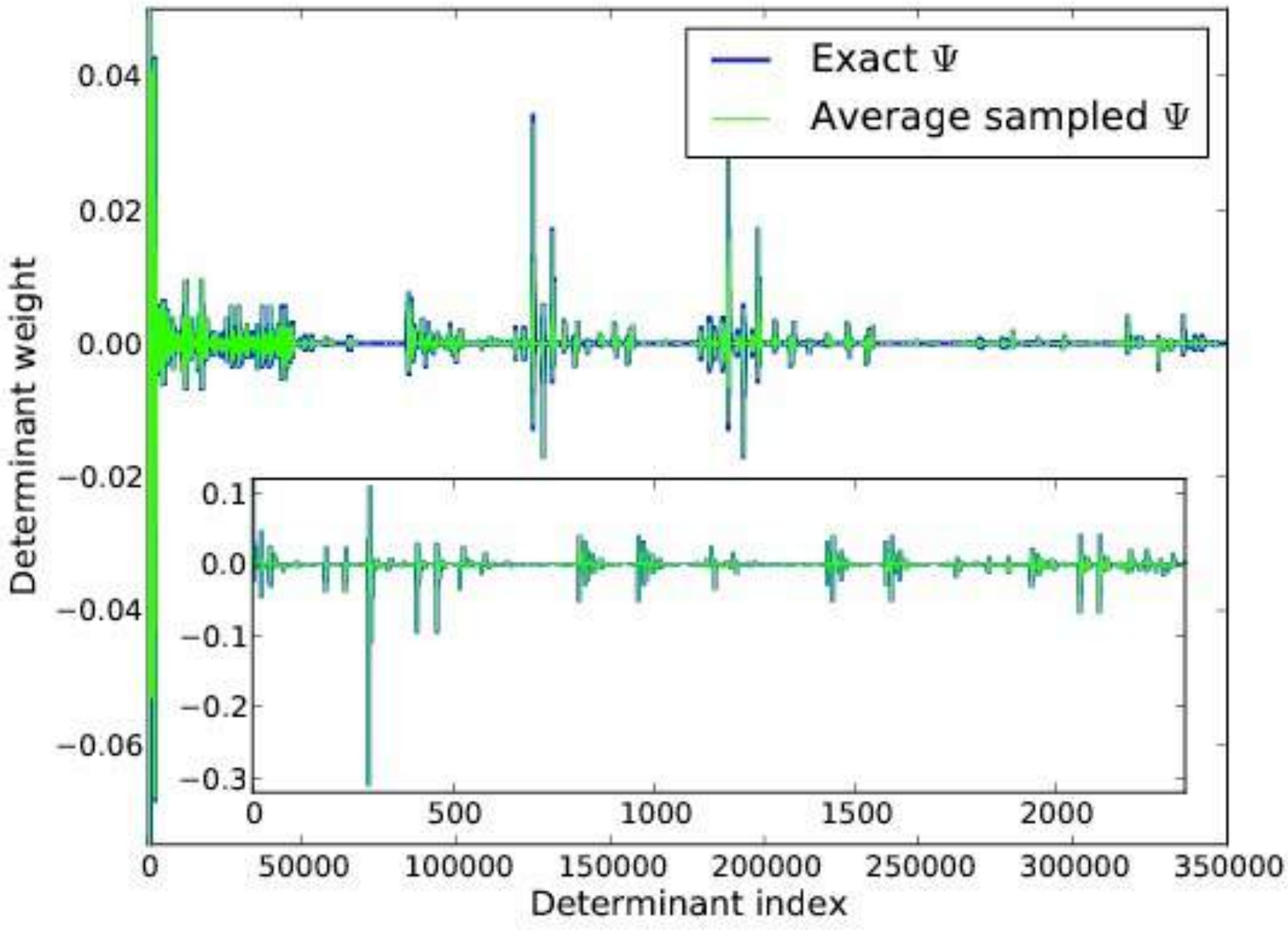


The instantaneous *i*-FCIQMC wavefunction is very coarse-grained representation of the exact wavefunction

(Be_2-VTZ, N_{FCI}= 346,485, N_w=2,000)



But the **time-averaged** *i*-FCIQMC wavefunction with 2,000 walkers (Be_2, VTZ, N_{FCI}= 346485) is essentially perfect

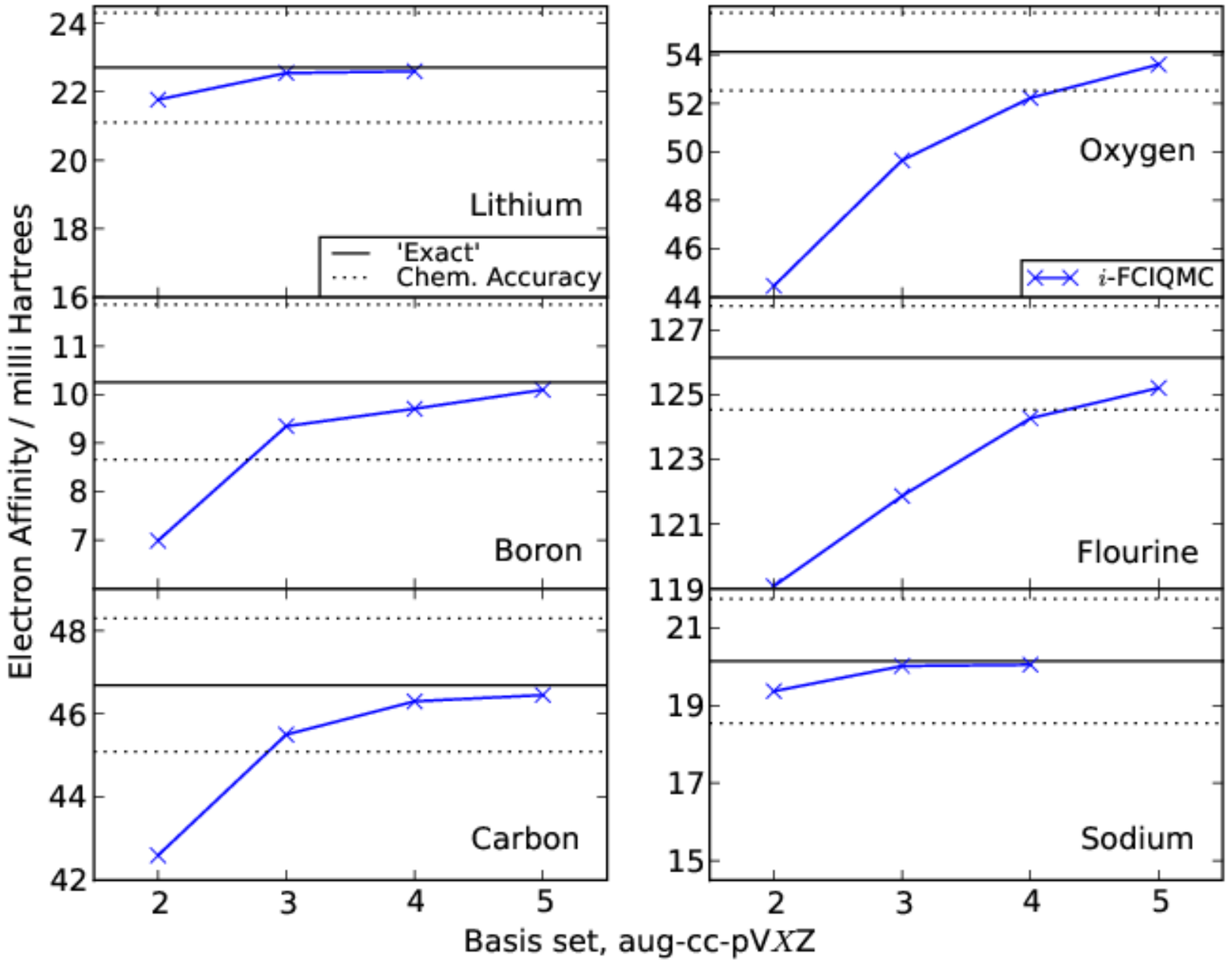


Electron affinities with i–FCIQMC: $A+e^- \rightarrow A^-$

comparison with Expt* (relativistically corrected)

aug-VDZ=[4s3p2d]~23, aug-VTZ=[5s4p3d2f]~46, aug-VQZ=[6s5p4d3f2g]~80,

aug-VXZ~(1/3)(X+1)(X+3/2)(X+2)+$(X+1)^2$ Cleland, Booth, Alavi,J Chem Phys 134 , 024112, (2011).



aug-V5Z F^-

$N = 10, M = 127$

$N_{\text{det}} = 4 \times 10^{15}$

$N_w = 8 \times 10^6$

$f_c = \frac{N_w}{N_{\text{det}}} = 4 \times 10^{-9}$

*T Koga, H Aoki, JM Garcia de la Vega, H Tatewaki, Theor. Chim Acta, 96, 248, (1997)

First-row Diatomics

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

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

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}

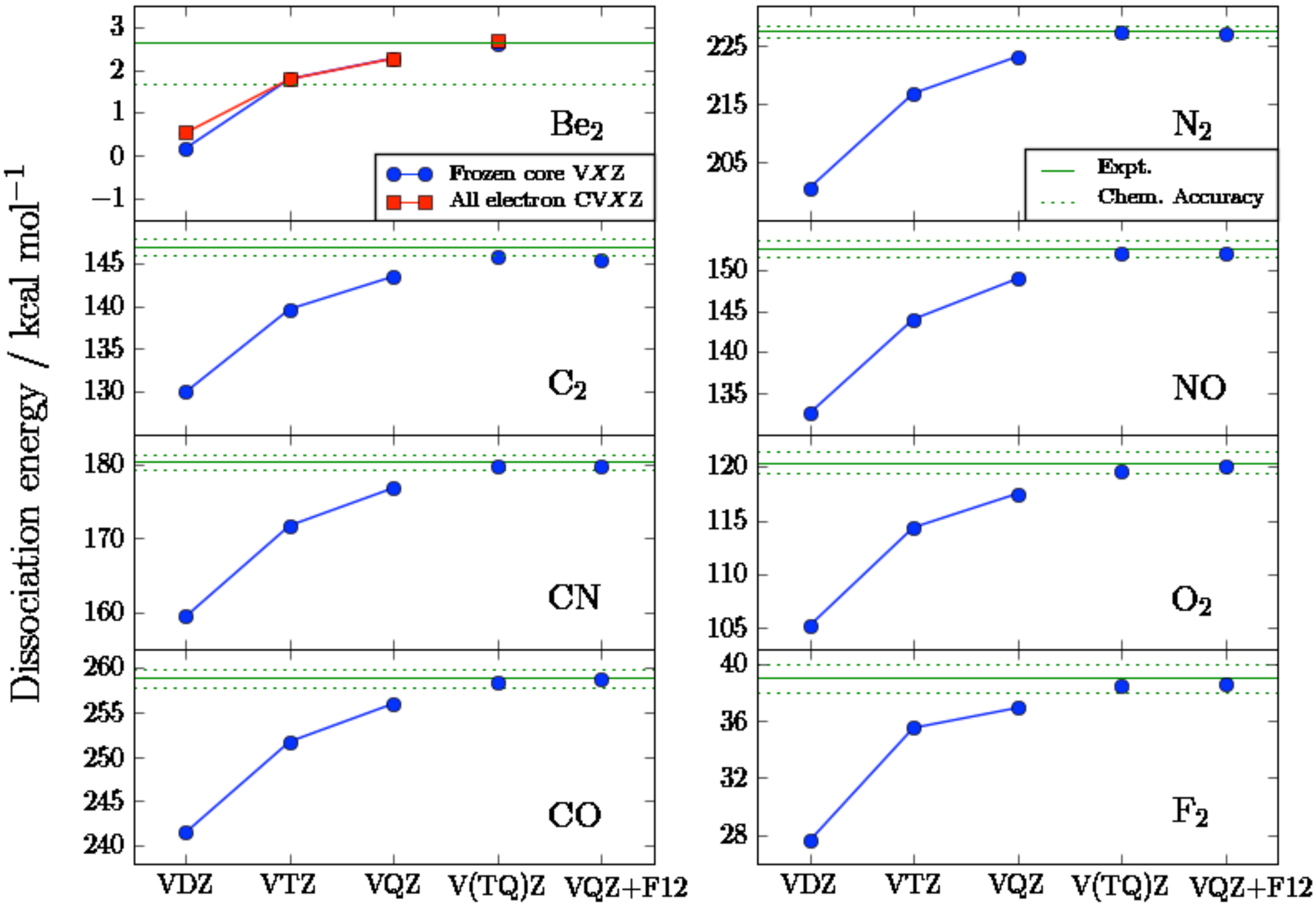
Basis	Be_2	C_2	CN	CO
VDZ	0.53(3)	129.95(8)	159.40(7)	241.49(6)
VTZ	1.78(6)	139.63(2)	171.71(6)	251.66(5)
VQZ	2.27(9)	143.44(5)	176.80(9)	255.92(9)
V(TQ)Z	2.67(10)	145.76(5)	179.8(1)	258.3(1)
VQZ+$\Delta E_{\mathrm{F12}}^{\mathrm{CCSD(T)}}$		145.36(5)	179.71(9)	258.68(9)
D_{e}^{*}(Expt.)	2.658(6)	146.9(5)	180.4(2.4)	258.8(2)

Basis	NO	N_2	O_2	F_2
VDZ	132.57(5)	200.52(8)	105.17(6)	27.59(7)
VTZ	143.99(6)	216.86(9)	114.35(8)	35.5(1)
VQZ	148.9(1)	223.20(8)	117.5(1)	36.9(1)
V(TQ)Z	151.9(2)	227.3(1)	119.6(1)	38.4(2)
VQZ+$\Delta E_{\mathrm{F12}}^{\mathrm{CCSD(T)}}$	152.0(2)	227.09(8)	120.1(1)	38.6(1)
D_{e}^{*}(Expt.)	152.63(4)	227.60(5)	120.42(5)	39.0(1)

Dissociation energies of some strongly correlated first-row molecules

(F12 correction added to VQZ)

Cleland, Booth, Overy, Alavi, J. Chem. Theory Comput. 2012, 8, 4138–4152



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

Method	Basis	$Be_2^{\dagger}$	C_2	CN	N_2	CO	NO	O_2	F_2	MAD
i-FCIQMC	VQZ	-0.4	-3.5	-3.6	-4.4	-2.9	-3.7	-2.9	-2.1	2.9
CCSD(T)[a]	VQZ	-1.0	-4.0	-5.6	-5.0	-2.6	-4.5	-3.2	-2.4	3.5
CEEIS[b]	VQZ		-3.4		-4.6			-2.5	-1.9	2.6
i-FCIQMC	V(TQ)Z	0.0	-1.1	-0.6	-0.3	-0.5	-0.7	-0.8	-0.6	0.6
CCSD(T)[a]	V(TQ)Z	-0.6	-1.5	-1.2	-0.6	-0.3	-1.0	-0.7	-0.9	0.9
CEEIS[b]	V(TQ)Z		-1.1		0.0			0.8	-0.4	0.5
FN-DMC[c]				-7.5	-3.2	-2.5	-7.0	-6.5	-5.7	5.4
i-FCIQMC	VQZ+F12		-1.5	-0.7	-0.5	-0.1	-0.6	-0.3	-0.4	0.4
CCSD(T)-F12	VQZ		-1.8	-1.0	-0.6	0.2	-0.8	-0.3	-0.4	0.5

(a) D. Feller and J. Sordo, J Chem Phys **2000**, 113, 485
(b) L. Bytautas and K.Ruedenberg, J Chem Phys **2005**, 122, 154110
(c) J.C. Grossman, J Chem Phys, **2002**, 117, 1434

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

$\sum'_{j\in\mathcal{D}}$: Sum computed exactly for given N_j in $\mathcal{D}$

$\sum'_{j\in\mathcal{D}'}$: Spawning step as per i-FCIQMC

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

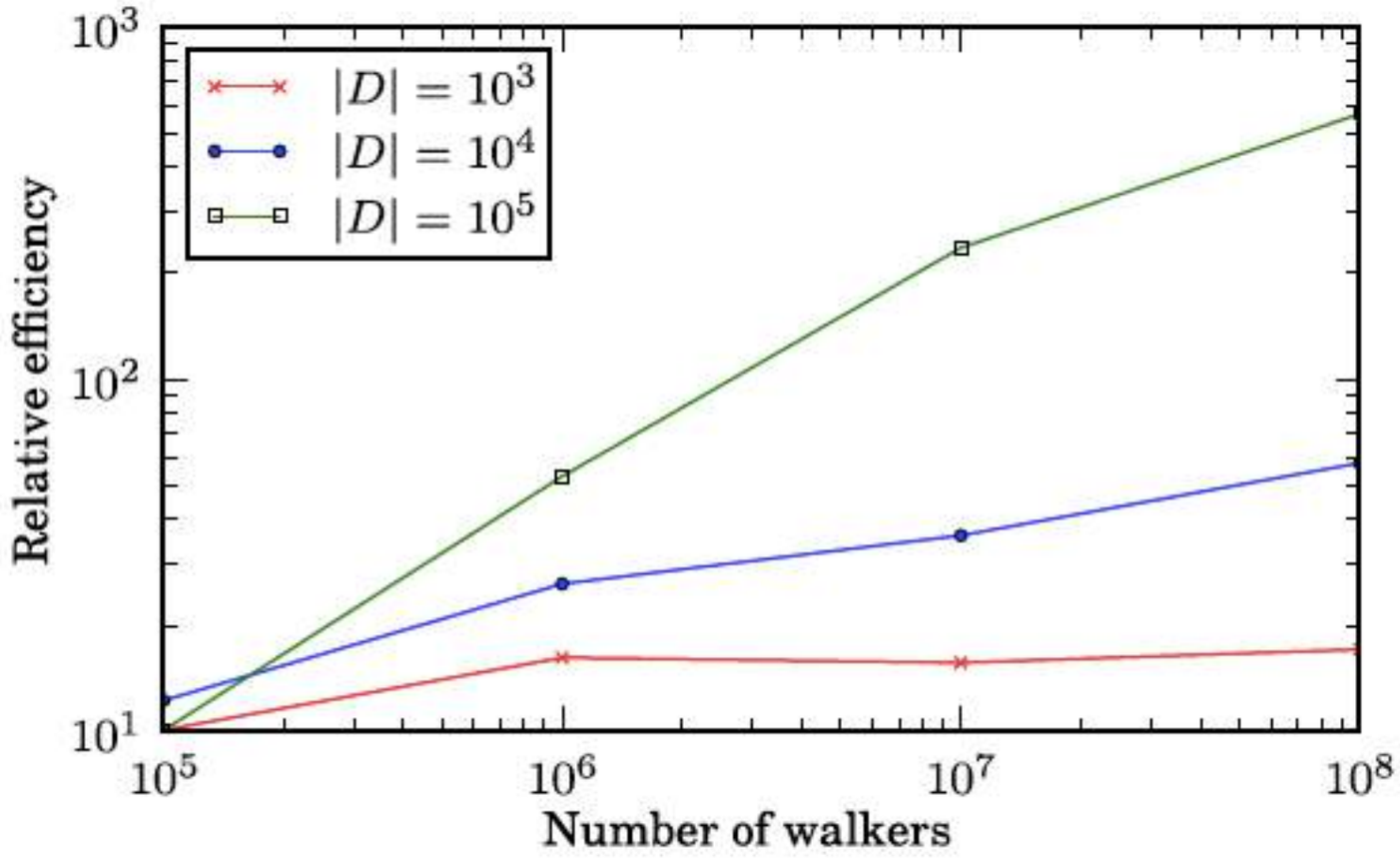


FIG. 4. The efficiency (ϵ_{E_0}) of semi-stochastic simulations relative to an otherwise identical simulation without semi-stochastic, for the 14-electron homogeneous electron gas with 114 spin orbitals and $r_s = 1.0$ a.u., as the walker population is varied. It is found that the benefit of semi-stochastic 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 $H \in \mathcal{T}$

$$\begin{aligned} E &= \frac{\langle \Psi | H | \Psi_T \rangle}{\langle \Psi | \Psi_T \rangle} \\ &= \frac{\sum_{i \in \mathcal{CT}, j \in \mathcal{T}} C_i c_j H_{ij}}{\sum_{i \in \mathcal{T}} C_i c_i} \\ &= \frac{\sum_{i \in \mathcal{CT}} C_i h_i}{\sum_{i \in \mathcal{T}} C_i c_i} \end{aligned}$$

$$h_i = \sum_{j \in \mathcal{T}} H_{ij} c_j, \quad i \in \mathcal{CT}$$

$\mathcal{CT}$ is space of determinants connected to $\mathcal{T}$

Petruzielo, Holmes, Changlani, Nightingale, Umrigar, PRL, **109**, 230201, (2012)

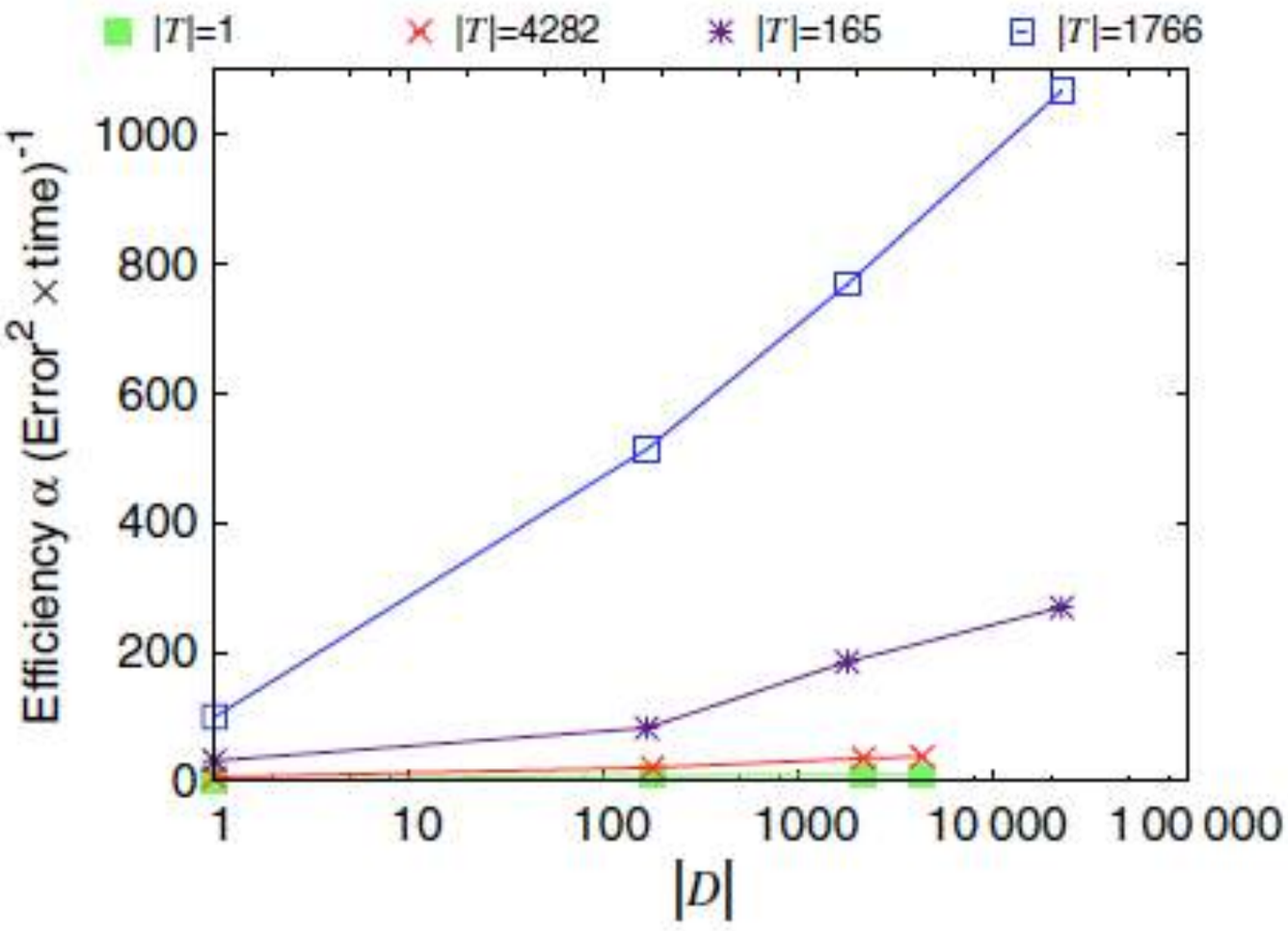
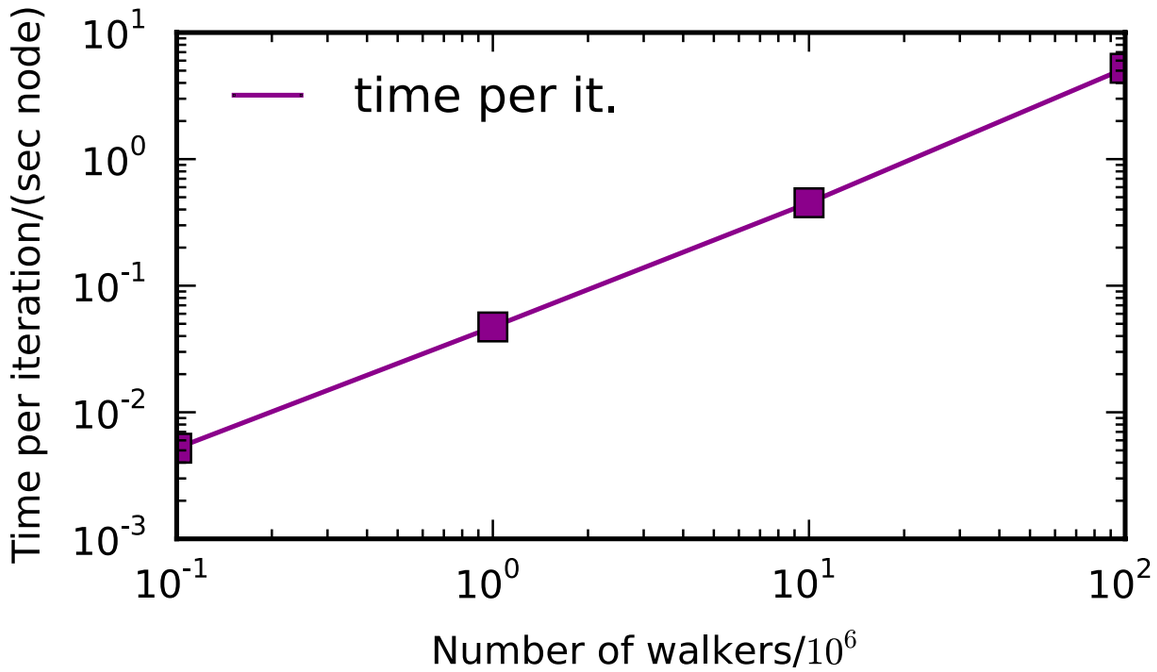
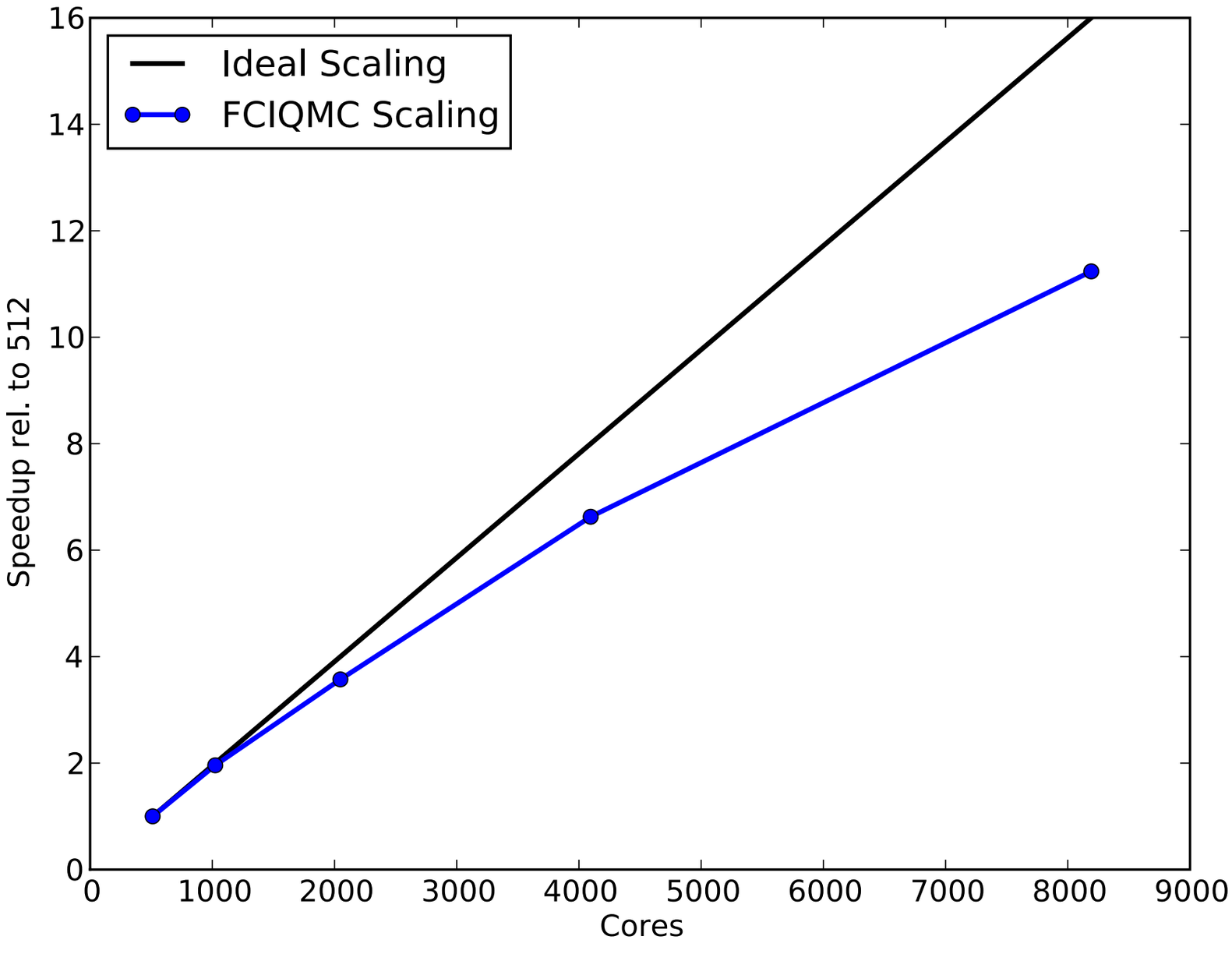


FIG. 3 (color online). Relative efficiency of SQMC vs dimension $|\mathcal{D}|$ of the deterministic space for the carbon dimer with a cc-pVTZ basis. Results are shown for trial wave functions of increasing size. The top two curves are for $\mathcal{D}$ and $\mathcal{T}$ generated with two applications 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, $N \approx 10^9$.

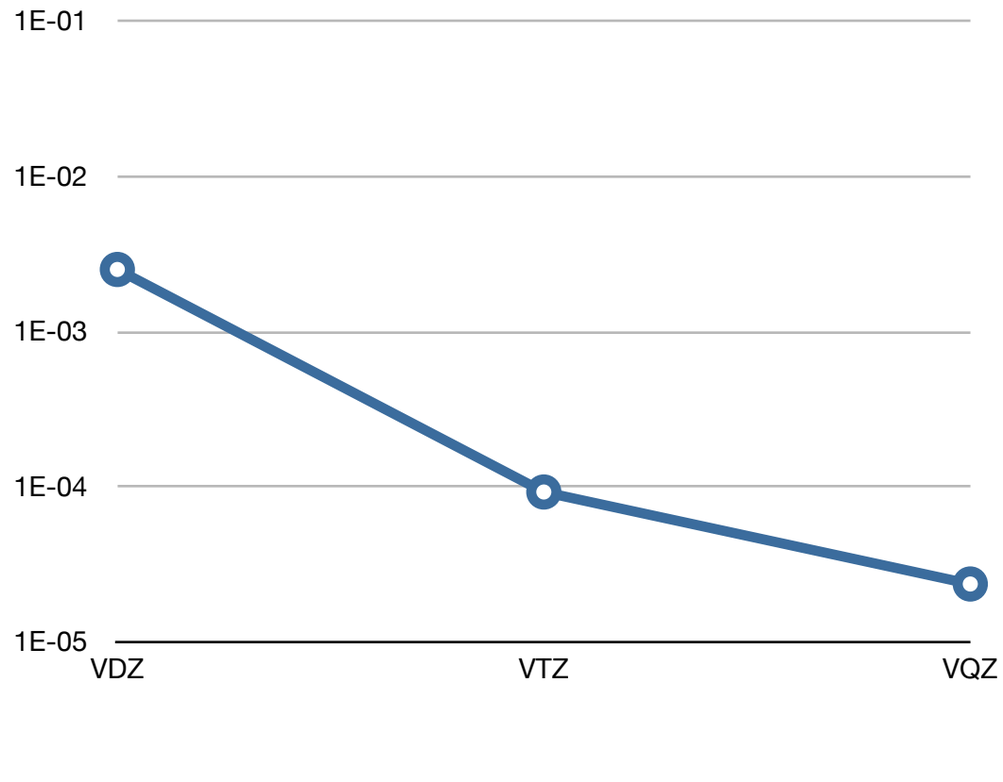

time per it.
Time per iteration/(sec node)
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-1}
10^{0}
10^{1}
10^{2}
Number of walkers/10^{6}

Parallel scaling



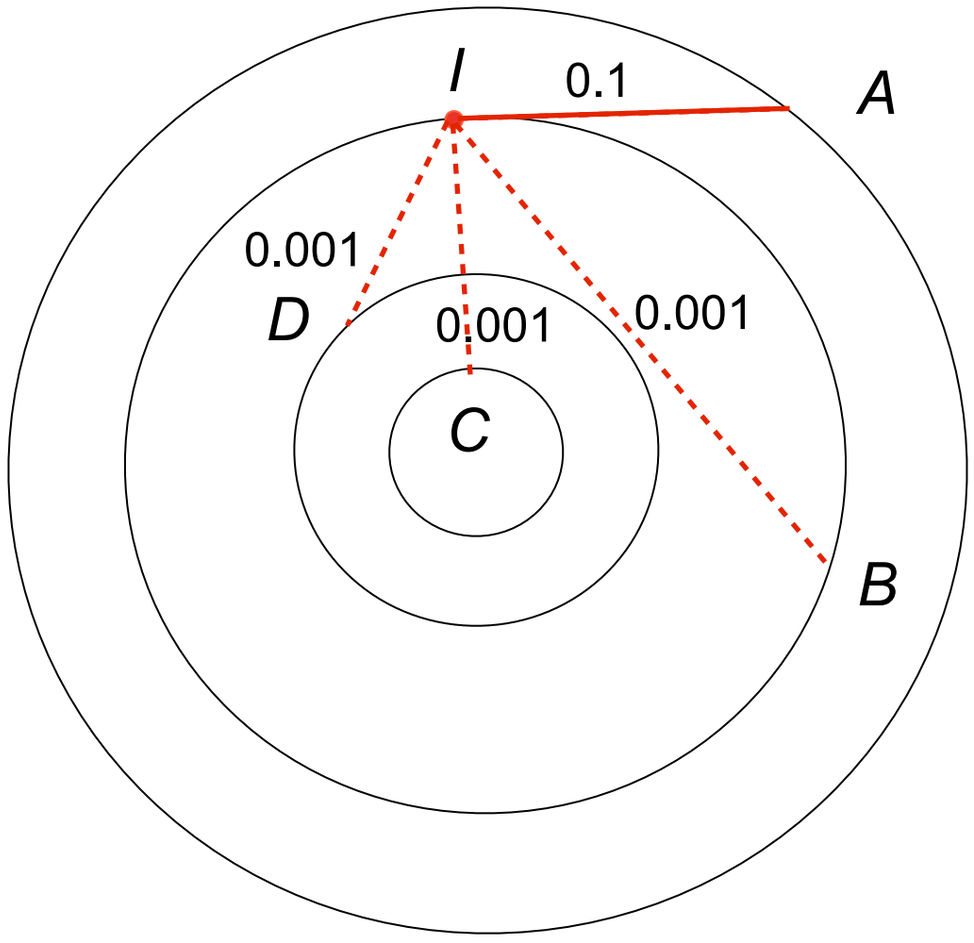
Time-step in a.u. for N_2 in cc-pV*X*Z basis sets

$$M \sim X^3$$



$$\tau \sim (N^2 M^2)^{-1}$$

Uniform Generation Scheme



$$p_{gen}(A|I) = p_{gen}(B|I) = .. = \frac{1}{4}$$

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

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

The largest allowable τ is set by $p_s \leq 1$

$$\rightarrow \tau = \frac{1}{0.4} = 2.5$$

$$\rightarrow p_s(A|I) = 1$$

$$p_s(B|I) = 0.01, etc$$

$$p(\text{to spawn}) = (3/4) * .01 + (1/4) * 1 = 0.2575$$

Rejection ratio is high

Hamiltonian-Weighted Excitation Generation

Construct an algorithm to select $\mathbf{j}$ from $\mathbf{i}$ so that

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

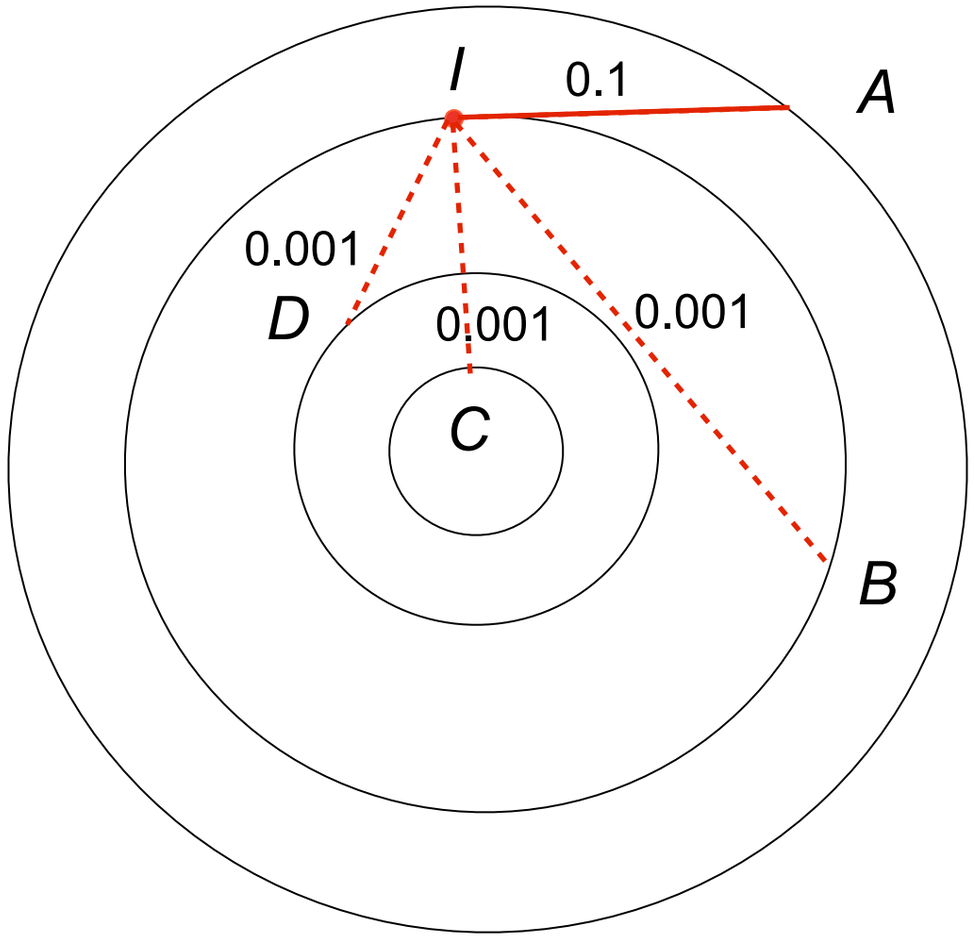
In the ideal case (where proportionality is exact), the spawning rate is constant:

$$p_s = \tau * const$$

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

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

Weighted Generation Scheme



$$\sum_A |H_{AI}| = 0.103$$

$$p_{gen}(A|I) = \frac{0.1}{0.103} = 0.971$$

$$p_{gen}(B|I) = \frac{0.001}{0.103} = 0.00971$$

$$p_s(A|I) = p_s(B|I) = 0.103\tau$$

$$\text{largest allowable } \tau = 1/0.103 = 9.71$$

$$p_s(A|I) = p_s(B|I) = p_s(C|I) = p_s(D|I) = 1$$

$$p(\text{to spawn}) = 1$$

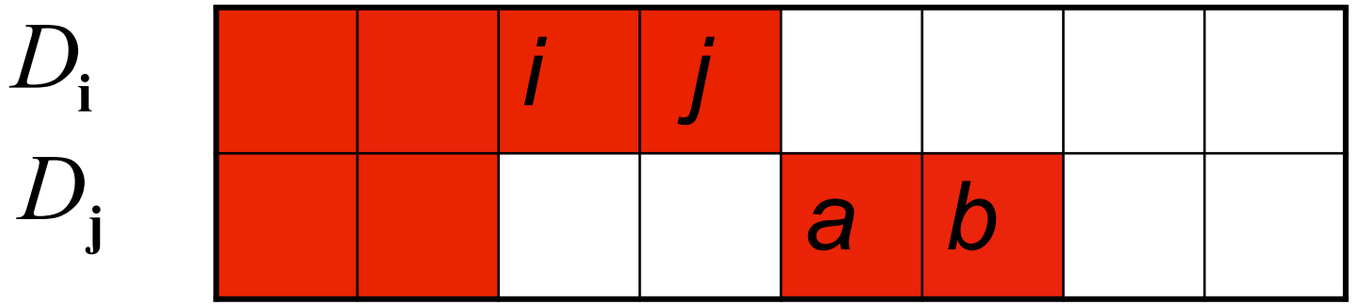
Two advantages: allows larger timesteps, and minimises rejections

However, it is possible to devise a determinant selection algorithm based on a **Cauchy-Schwarz** decomposition of $H_{\mathbf{ij}}$ which costs only $O(M)$ to compute!

In practice this allows τ to scale substantially better with system size, as well as lead to significant overall efficiency.

For opposite-spin excitations

$$\sigma_i = \sigma_a \neq \sigma_j = \sigma_b$$



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

$$p(ijab) = p(ab|ij)p(ij)$$

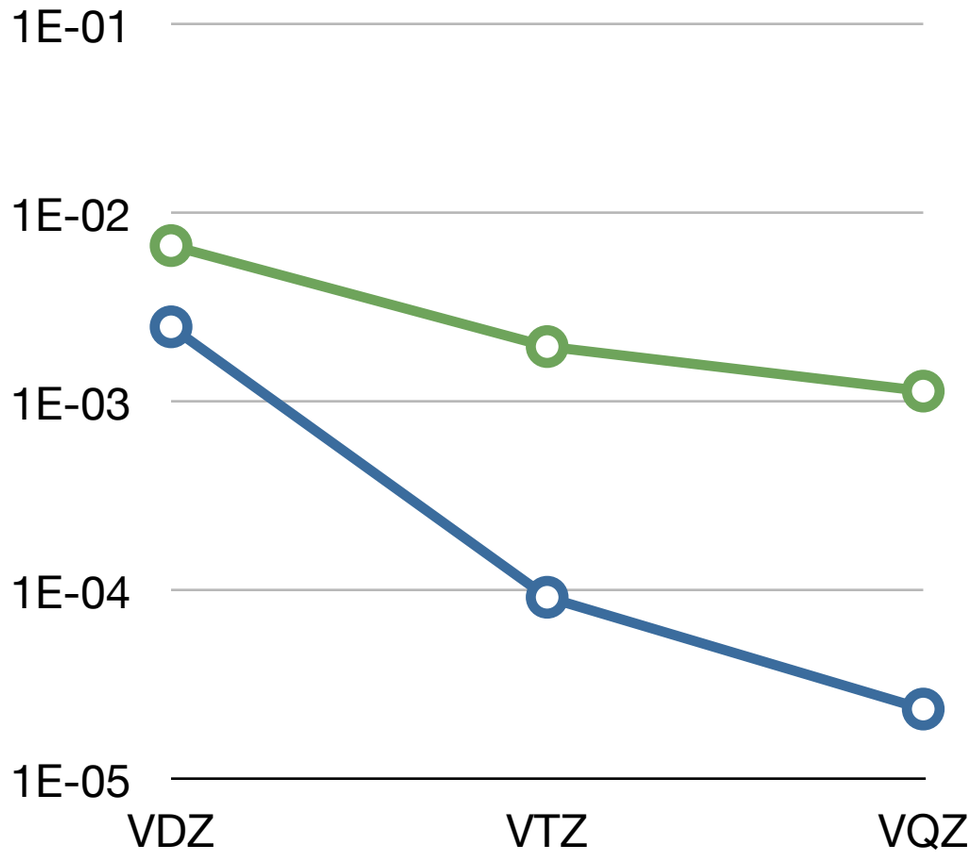
$$p(ab|ij) = p(a|i)p(b|j)$$ [For same-spin, include exchange term]

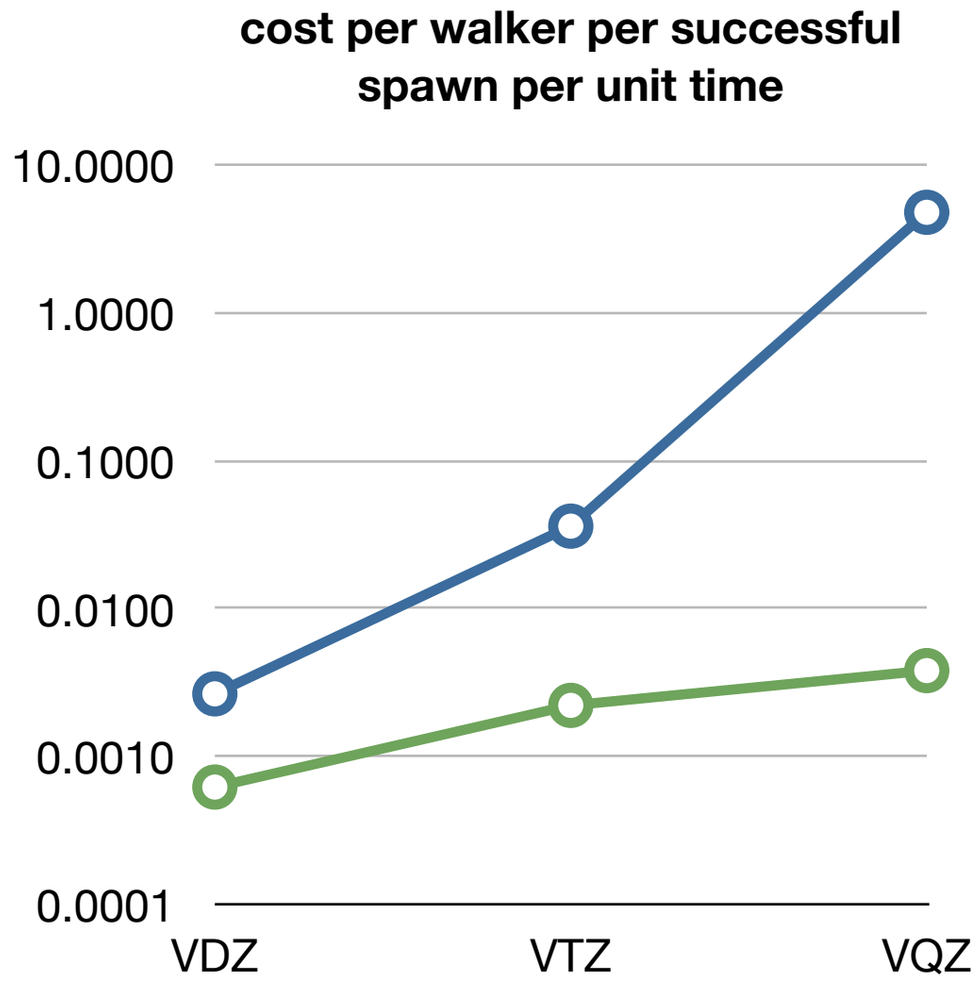
Select hole *a* according to:

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

$$p(a|i) = \frac{\sqrt{\langle ai|ai\rangle}}{\sum_a \sqrt{\langle ai|ai\rangle}} \leftarrow O(M)$$

tau for N2 with VXZ basis-sets

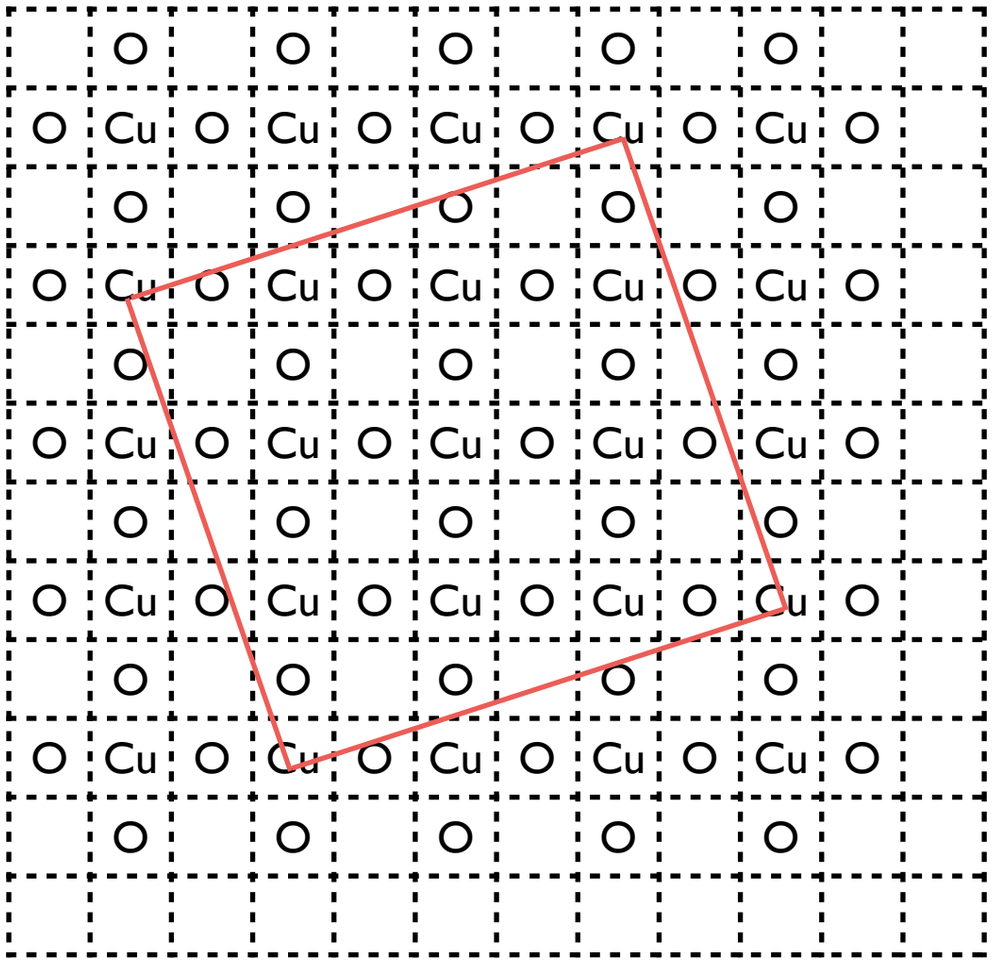


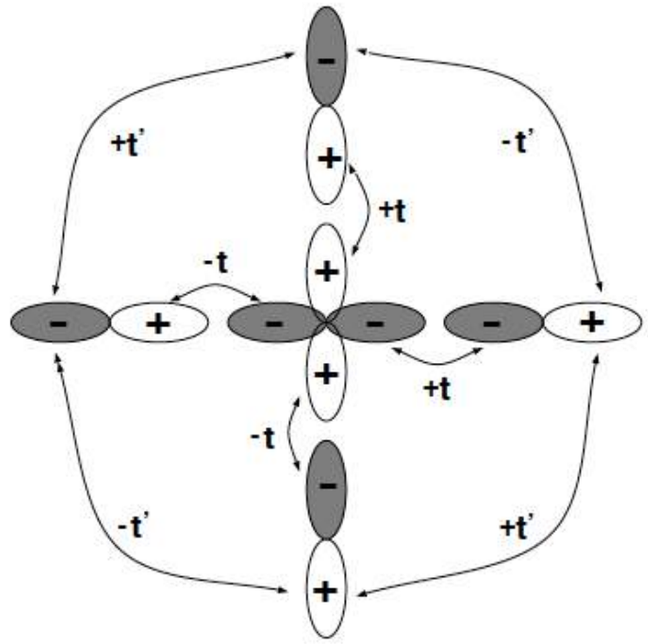


standard newexcitgen

3-Band Hubbard Model of a Cuprate

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





$$\hat{H} = \sum_{i,\sigma} (\varepsilon_d - \mu)\, d^\dagger_{i,\sigma} d_{i,\sigma} + \sum_{i,\nu,\sigma} (\varepsilon_{p_\nu} - \mu)\, p^\dagger_{i\nu,\sigma} p_{i\nu,\sigma}$$

$$+ \sum_{i,j,\nu,\sigma} (t_{i,j\nu}\, d^\dagger_{i,\sigma} p_{j\nu,\sigma} + h.c.) + \sum_{\substack{i,\nu,j,\kappa,\sigma \\ i\neq j,\nu\neq\kappa}} t'_{i\kappa,j\nu}\, p^\dagger_{i\kappa,\sigma} p_{j\nu,\sigma}$$

$$+ \sum_i U_d\, d^\dagger_{i,\uparrow} d_{i,\uparrow} d^\dagger_{i,\downarrow} d_{i,\downarrow}\,. \qquad (1)$$

Table 1. Parameters for a three-band model (in eV) calculated with a constrained first principles calculation for La_2CuO_4 done by Hybertsen et al. [12].

Δ	t	t'	U_d	U_p	U_{pd}	U_{pp}
3.6	1.3	0.65	10.5	4	1.2	0

3-Band Hubbard model energies

$$N_{FCI} = \binom{30}{5}^2 \approx 20 \times 10^9$$

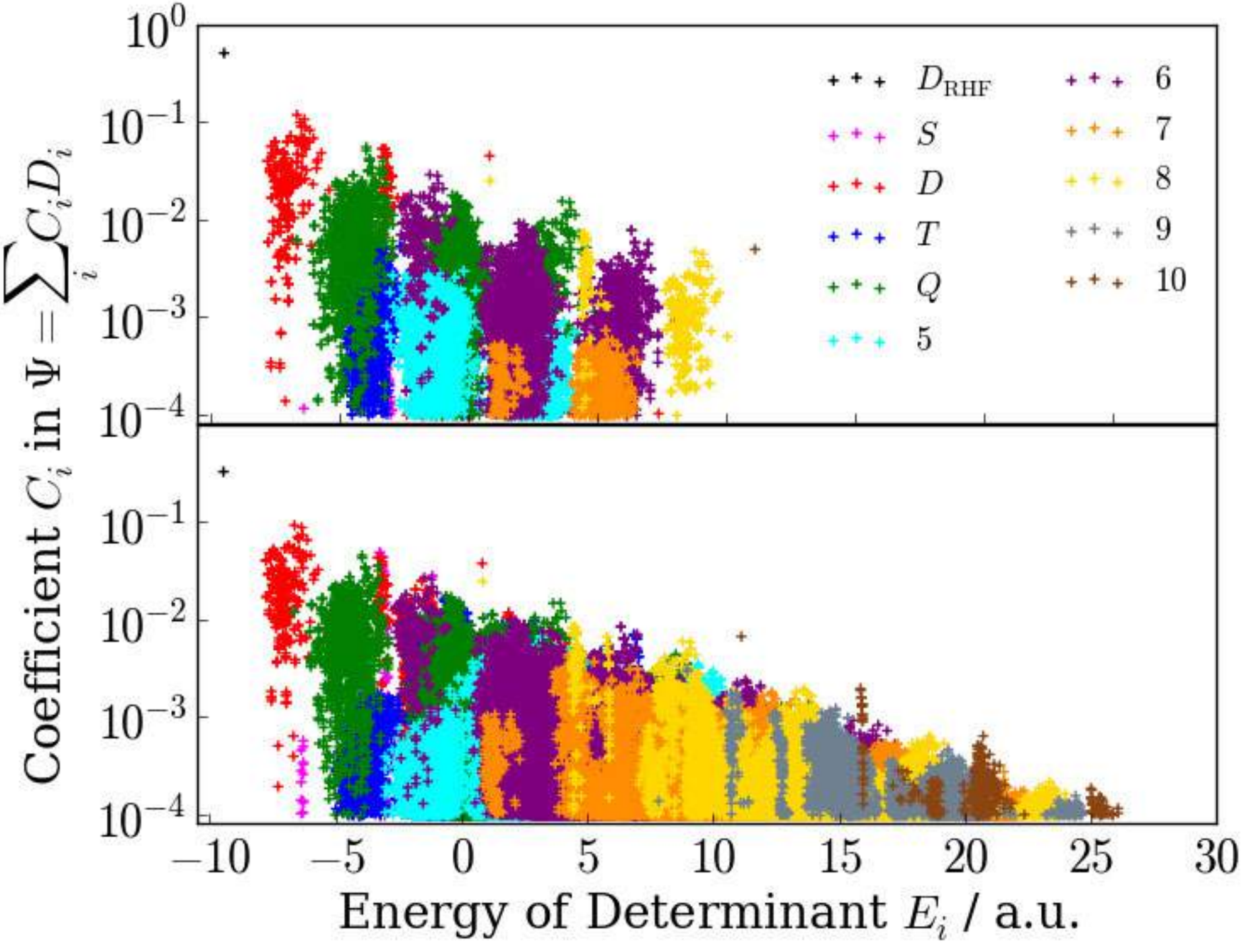
```
RHF:  -9.5206318541964
UHF: -15.2905361816484
Exact: -15.817 (5)
```

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

$$N_w \approx 10^8 (\text{RHF basis})$$

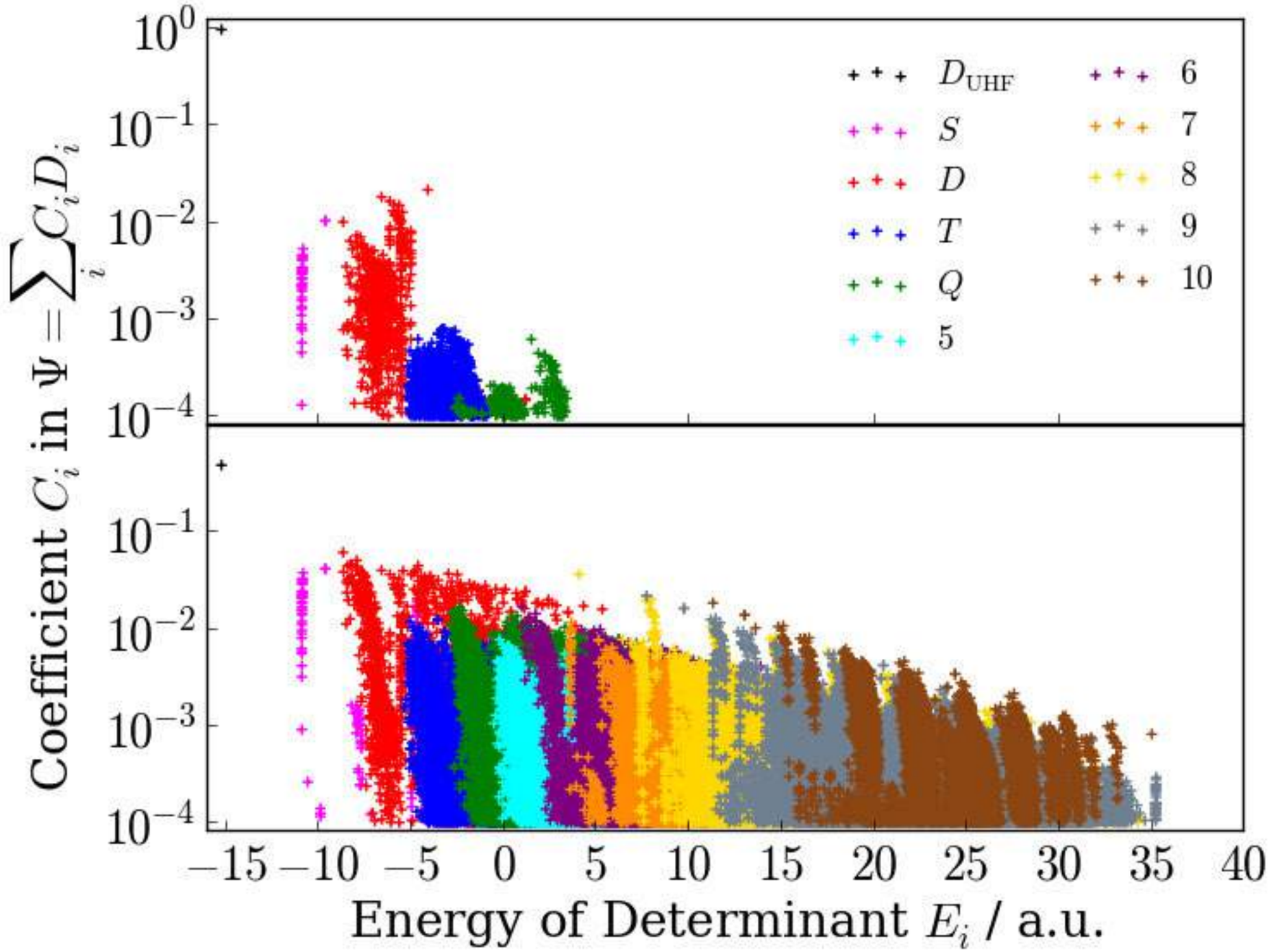
Which orbitals to use: RHF or UHF?

RHF



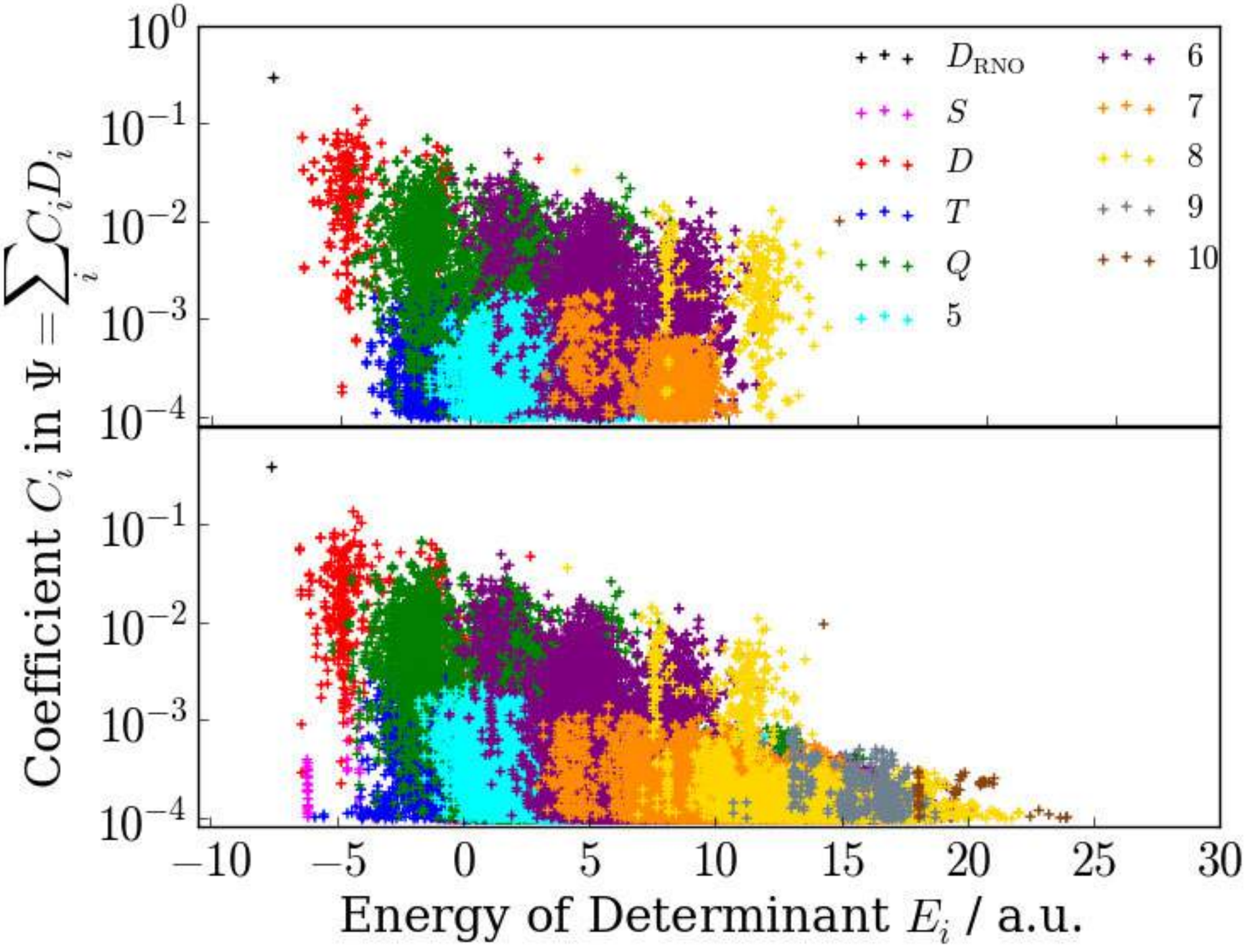
Which orbitals to use: RHF or UHF?

UHF



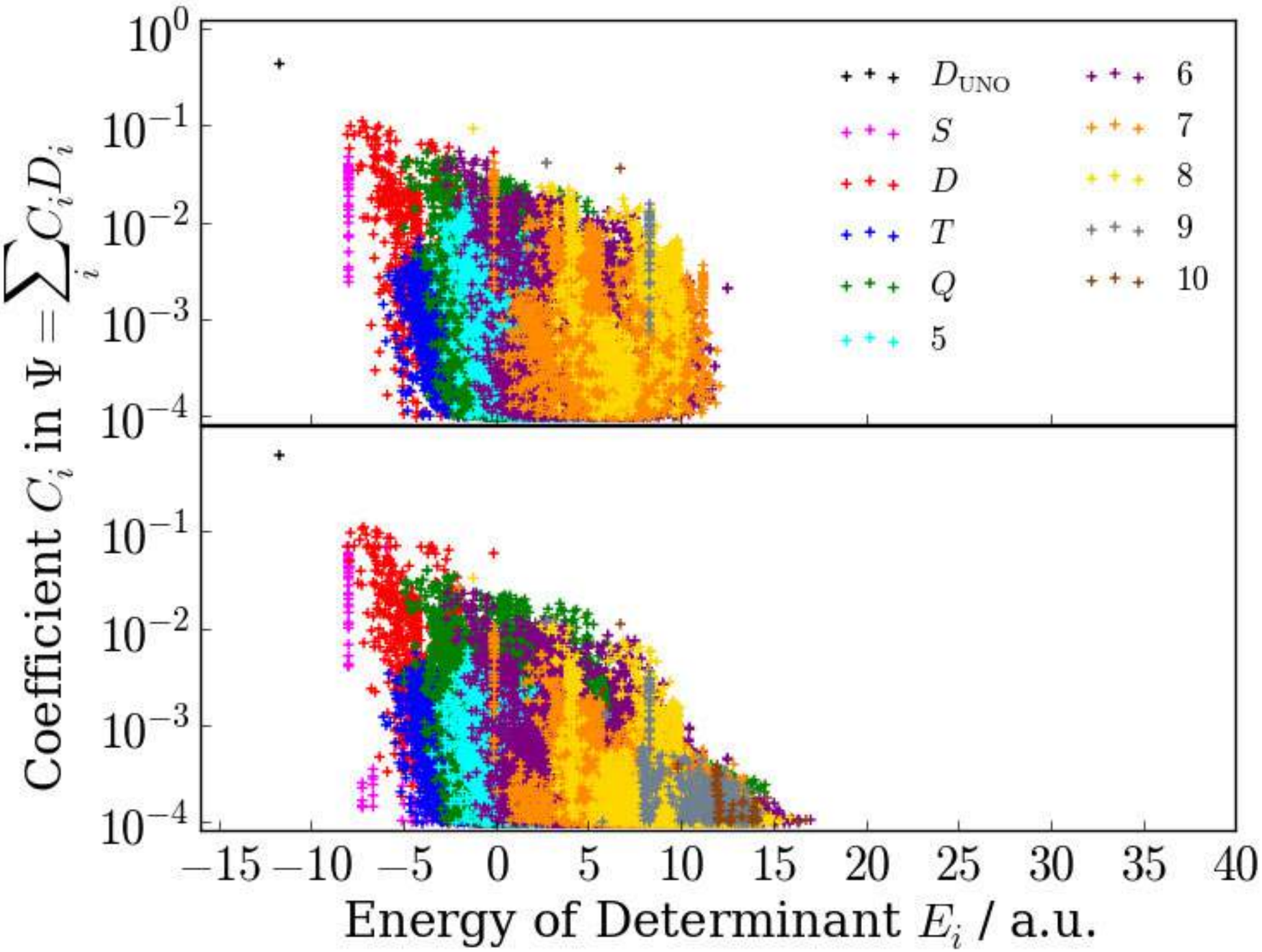
Natural Orbitals: RNO or UNO?

RNO



Natural Orbitals: RNO or UNO?

UNO



Reduced density matrices

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

$$\gamma_q^p = \langle \Psi | a_p^\dagger a_q | \Psi \rangle$$

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

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

$$\gamma_p^q = \frac{1}{N-1} \sum_r \Gamma_{pr}^{qr}$$

Normalisation conditions:

$$\sum_{pq} \Gamma_{pq}^{pq} = \binom{N}{2}$$

$$\sum_p \gamma_p^p = N$$

Properties such as *E*, S^2, electron density, forces, etc can be calculated via the 1 and 2-RDMs

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

$$\langle \Psi | \hat{O} | \Psi \rangle = \Gamma^{rs}_{pq} o^{pq}_{rs}$$

Energy

$$E = h^q_p \gamma^p_q + \frac{1}{2} g^{rs}_{pq} \Gamma^{pq}_{rs}$$

Nuclear gradients
Hellmann-Feynman+Pulay

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

Spin

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

Calculation of the RDMs

$$\Gamma^{rs}_{pq} = \sum_{\mathbf{i}} C^*_{\mathbf{i}} C_{\mathbf{j}} \text{ where } |\mathbf{j}\rangle = a^\dagger_r a^\dagger_s a_q a_p |\mathbf{i}\rangle$$

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!

$$\Gamma^{rs}_{pq} = \sum_{\mathbf{i} \in pq} \frac{C_{\mathbf{i}} C_{\mathbf{j}}}{p[\mathbf{j}|\mathbf{i}]} \times p[\mathbf{j}|\mathbf{i}]$$ ← Spawning probability

In practice, replace the product of the C_i coefficients with the time-average of the products of the walker populations

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

Normalisation is fixed by:

$$\sum_{pq} \Gamma_{pq}^{pq} = \binom{N}{2}$$

When we do this, the result is not good!

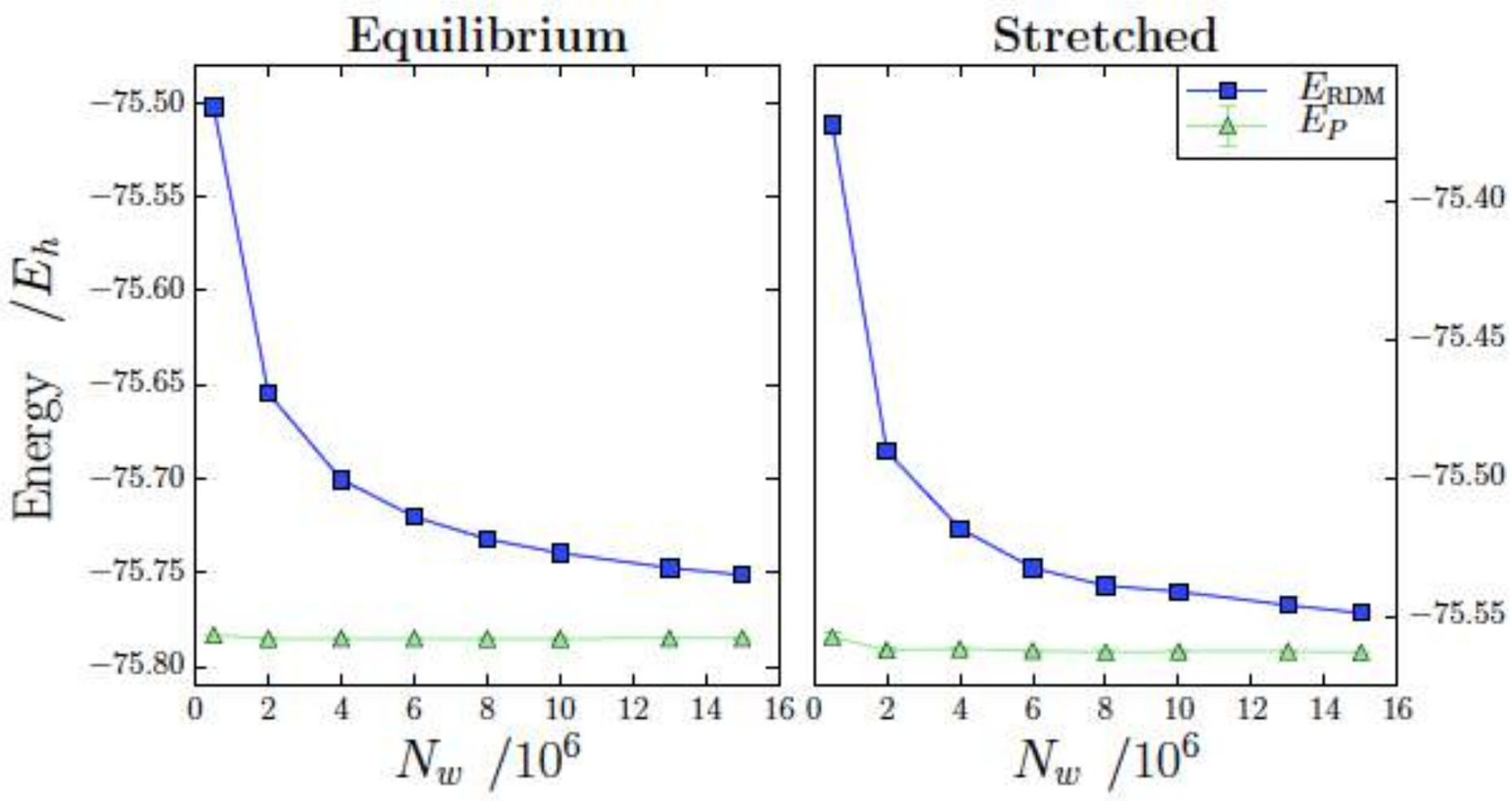


Figure 6.1 – This data, corresponding to simulations on C_2 cc-pVTZ, is taken directly from Ref. [25], showing the very slow convergence of E_{RDM} and significant error at large N_w, even once E_P is well-converged.

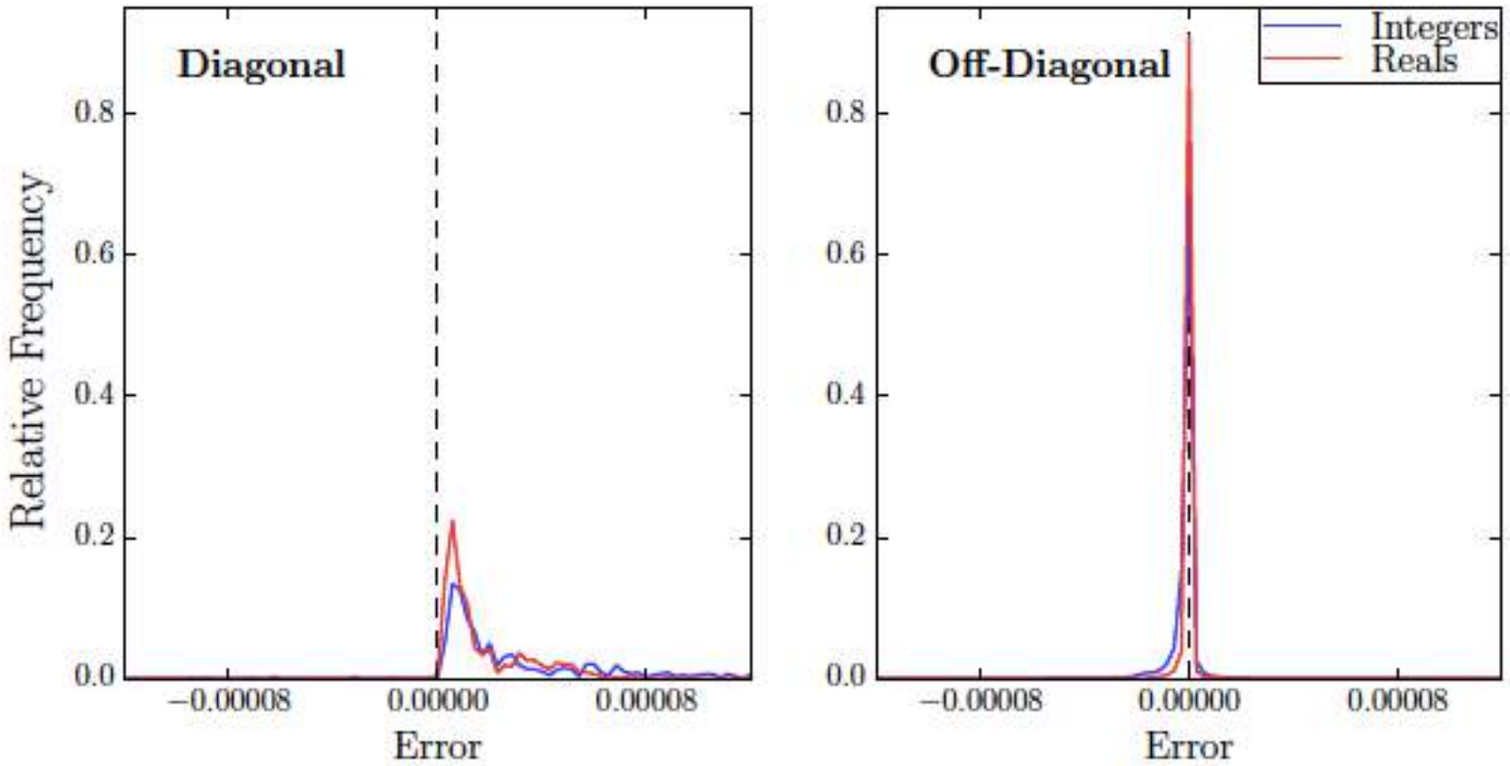


Figure 7.1 – The diagonal elements of the 2-RDM are systematically overestimated in the uncorrected stochastic RDM method. This data corresponds to i-FCIQMC calculations run on N_2 cc-pVDZ, $r = 1.094$ Å with 8 frozen core electrons and $N_w = 25000$. The 2-RDM was accumulated stochastically according to the method described in Section 6.1 for $\sim 5 \times 10^6$ iterations. The reals calculation uses $\gamma = 4$, $N_{\text{occ}} = 1$.

What is wrong?

The diagonal matrix elements have a bias

$$\Gamma^{pq}_{pq} \propto \sum_{\mathbf{i}\in pq} \langle N^2_{\mathbf{i}} \rangle_\tau$$

The instantaneous populations fluctuate about their exact value:

$$N_{\mathbf{i}} = N^{ex}_{\mathbf{i}} + \delta N_{\mathbf{i}}$$

$$\langle \delta N_{\mathbf{i}} \rangle_\tau = 0$$

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

$$\langle N^2_{\mathbf{i}} \rangle_\tau = (N^{ex}_{\mathbf{i}})^2 + \langle (\delta N_{\mathbf{i}})^2 \rangle_\tau$$

The solution: replica trick

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

$$\Gamma^{pq}_{pq} \propto \sum_{\mathbf{i}\in pq} \langle N^{(1)}_{\mathbf{i}} N^{(2)}_{\mathbf{i}} \rangle_\tau$$

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

$$\langle \delta N^{(1)}_{\mathbf{i}} \delta N^{(2)}_{\mathbf{i}} \rangle_\tau = 0$$

Therefore:

$$\Gamma^{pq}_{pq} \propto \sum_{\mathbf{i}\in pq} \langle N^{(1)}_{\mathbf{i}} \rangle_\tau \langle N^{(2)}_{\mathbf{i}} \rangle_\tau$$

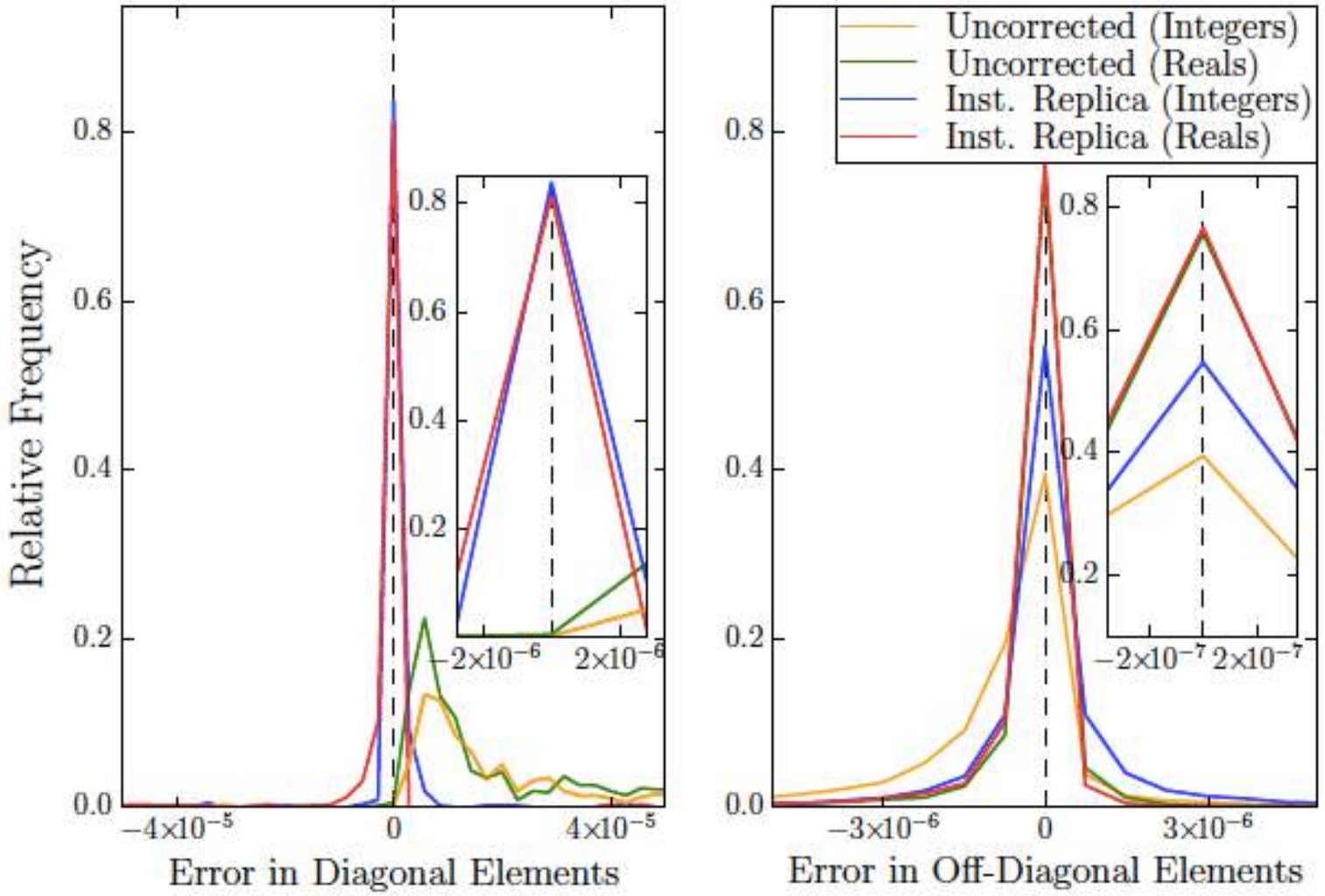


Figure 8.1 – The systematic sampling bias affecting the diagonal elements of the 2-RDM is removed when using the replica sampling method described in Section 8.1. This data corresponds to i-FCIQMC calculations run on N_2 cc-pVDZ, $r = 1.094\,\text{Å}$ with 8 frozen core electrons and $N_w = 25000$. The 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)

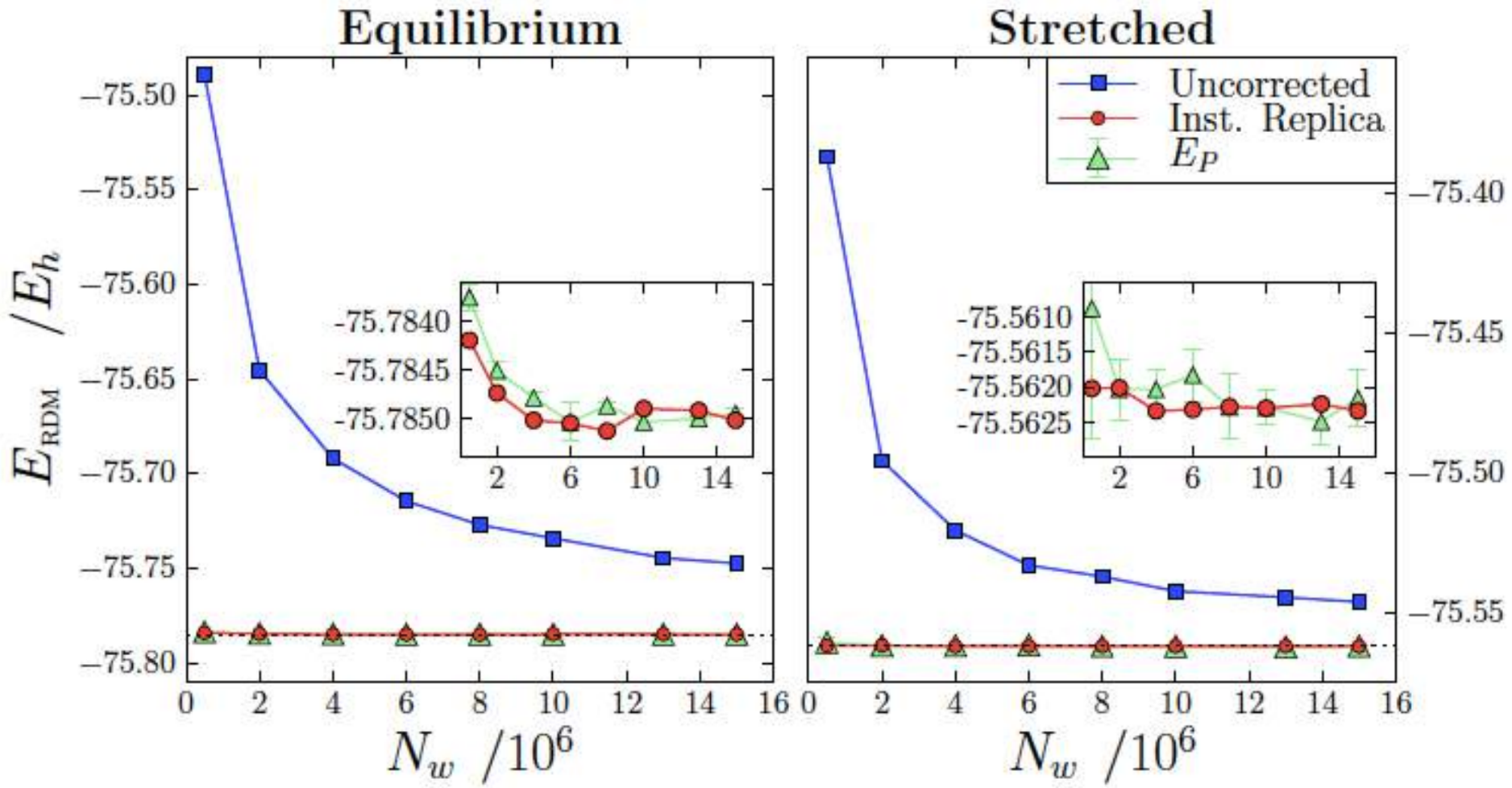


Figure 8.17 – Comparing the uncorrected and replica-sampled RDMs through values of E_{RDM} for C_2 cc-pVTZ. This refers directly back to Figure 6.1, though all values have been newly generated for this plot. Uncorrected RDMs were calculated with the integer i-FCIQMC algorithm (consistent with Ref. [25]), whilst replica-sampled RDMs use the real coefficients algorithm with $\chi = 4$, $N_{\mathrm{occ}} = 1$, to represent the best quality RDM available with the techniques presented in this chapter.

Nuclear gradients

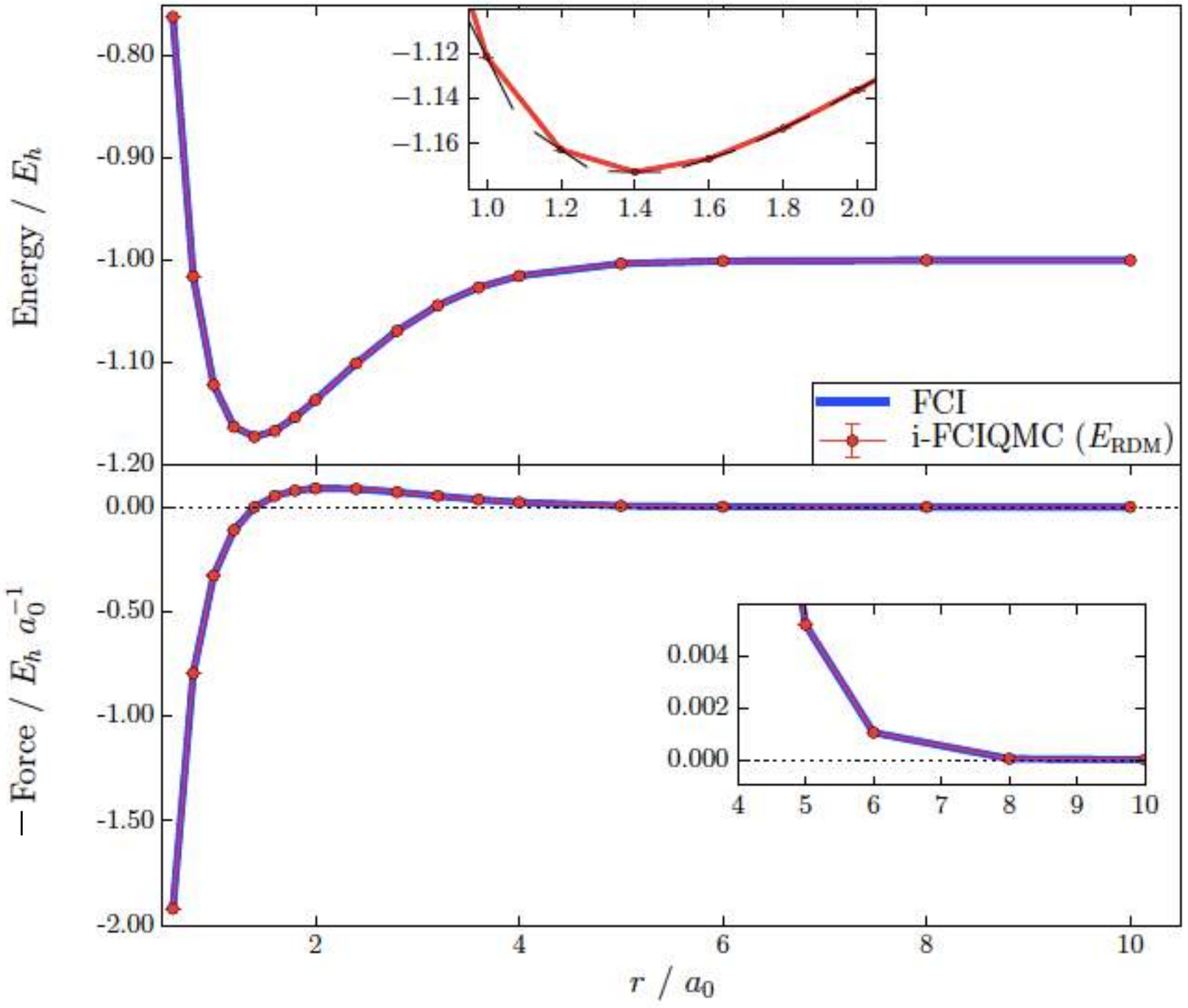


Figure 9.1 – Energies and internuclear forces in the H_2 cc-pVTZ binding curve. The inset of the upper panel shows the i-FCIQMC gradients plotted as tangent lines on the

N_2 in 6-31G

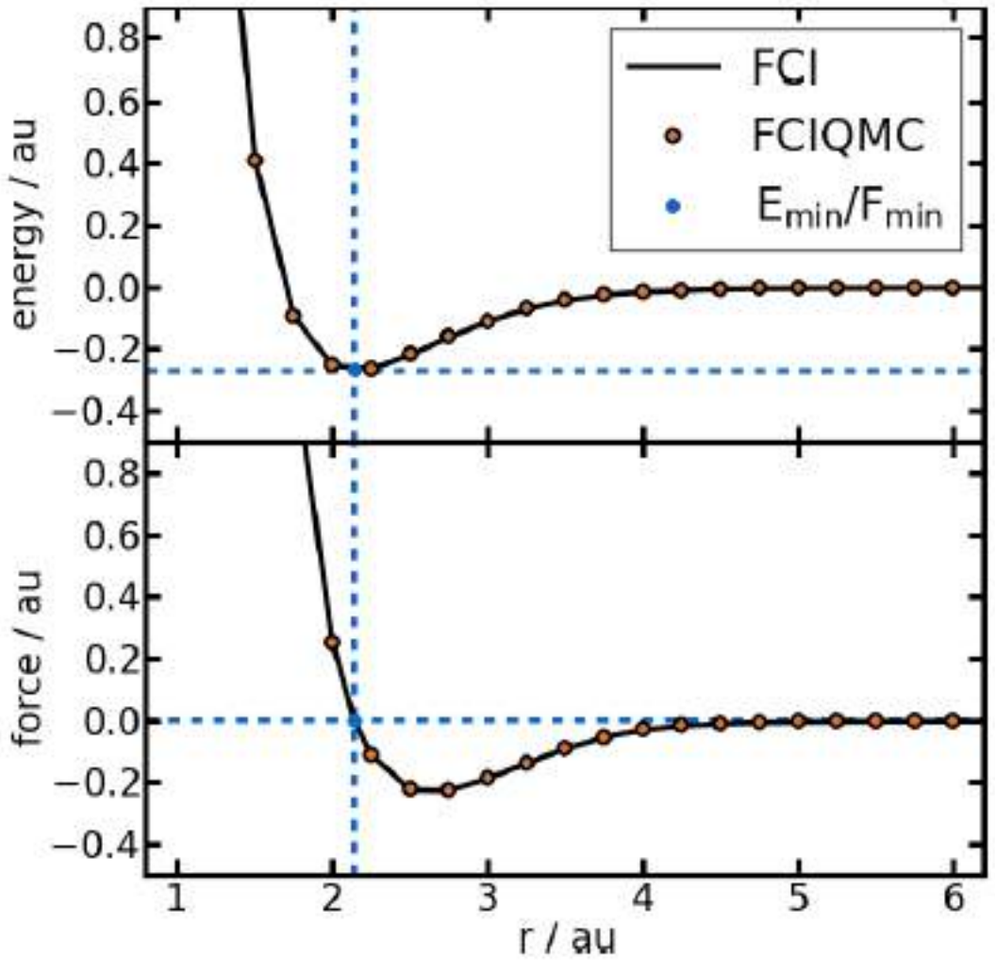


FIG. 1. Top: Potential energy profile for the N-N bond dissociation of N_2 relative to the energy of two isolated nitrogen atoms in the electronic ground state. Bottom: corresponding forces at one nitrogen atom computed using analytic gradients from *i*-FCIQMC reduced density matrices, compared to FCI with numerical differentiation. Results are identical within the accuracy of the numerical differentiation. The respective minimum energy (E_{min} = -0.2685 a.u.) and force (F_{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.

H_2O in 6-31G and VTZ

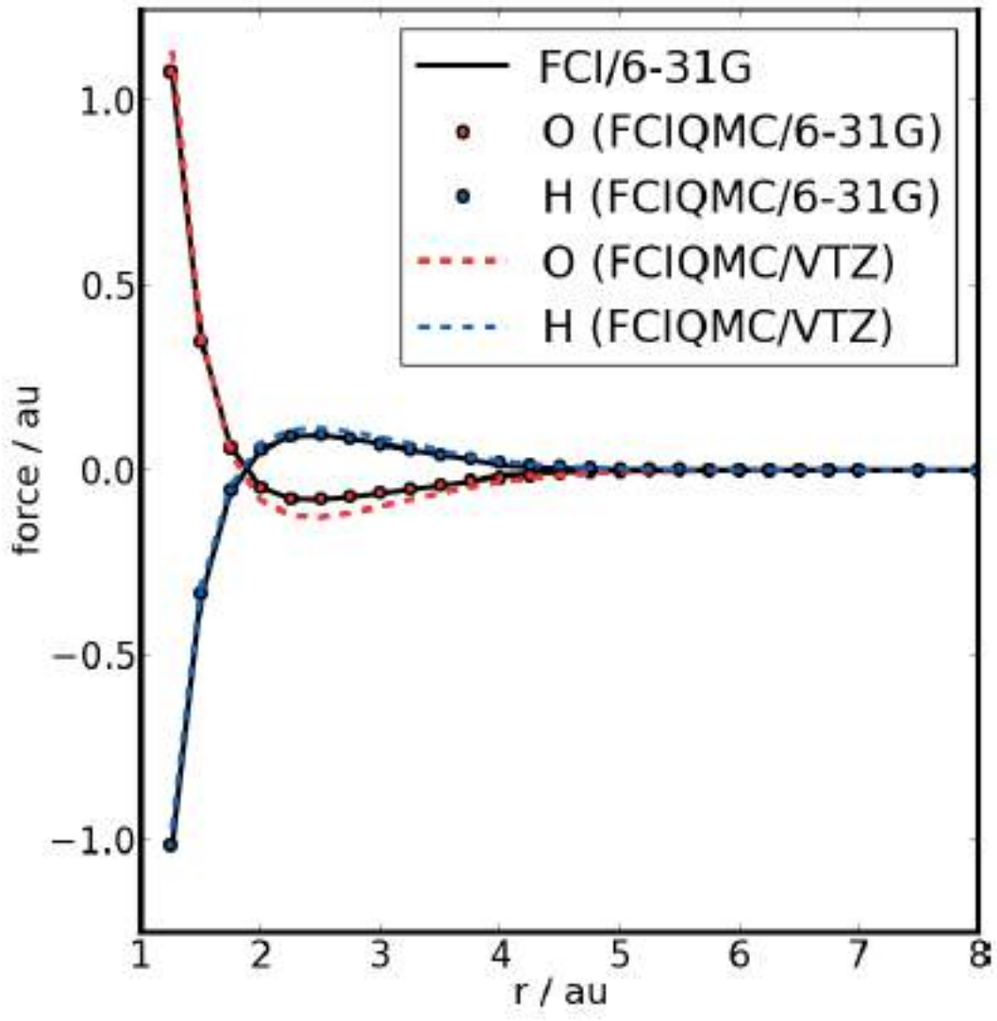


FIG. 2. Absolute forces acting on the oxygen and hydrogen atoms in a H_2O molecule computed using *i*-FCIQMC and FCI with a 6-31G and cc-pVTZ basis set (the sign corresponds to the z-component of the force vector). The data were acquired for symmetric displacements of the hydrogen atoms from the equilibrium geometry. The abscissa indicates the OH bond length of the respective molecular geometry.