# Essentially exact energies in enormous Hilbert spaces – the Semistochastic Heat-bath Configuration Interaction Method

Cyrus Umrigar

Physics Department, Cornell University, Ithaca.

Email: CyrusUmrigar@cornell.edu

INT, Seattle, 8 Aug 2018

# Solving the Many-Body Schrödinger Equation

Straightforward approach:

1. Expand the many-body wavefunction as a linear combination of (possibly nonorthogonal) basis states (determinants for Fermions).
2. Compute Hamiltonian and overlap matrices, $H$ and $S$ in this basis
3. Solve the generalized eigenvalue problem $Hc = ESc$

Problem:

The number of many-body states grows combinatorially in the number of single particle basis states and the number of particles, $\binom{N_{\mathrm{orb}}}{N_\uparrow} \times \binom{N_{\mathrm{orb}}}{N_\downarrow}$, e.g.

Molecules with 20 electrons in 200 orbitals: $\binom{200}{10}^2 = 5 \times 10^{32}$ (run on group cluster)

Molecules with 30 electrons in 300 orbitals: $\binom{300}{15}^2 = 6 \times 10^{49}$ (run on supercomputer)

HEG with 118 electrons in 30000 orbitals: $\binom{30000}{52}^2 = 5 \times 10^{329}$ (run on supercomputer)

(Partial) Solutions:

1. Quantum Monte Carlo methods (even $\infty$ Hilbert spaces).
2. Density Matrix Renormalization Group (DMRG) (White, Chan, ...).
3. SHCI: If only a small fraction (say $10^{14}$) of these states are important, then one can use smart methods for finding the most important of these (say $10^9$), and including rest of the $10^{14}$ using perturbation theory.

Cyrus J. Umrigar

# Semistochastic Heatbath Configuration Interaction Method

Although SHCI has a stochastic component, it is **not** a QMC method.
However, DMC/AFQMC convergence is improved by trial wavefunctions from SHCI.

Advantages of SHCI (relative to QMC methods):

1. No sign problem.
2. Samples are not correlated.

Gist of SHCI:

1. Useful for systems for which the imporant part of Hilbert space is much smaller than the entire space, e.g. all of quantum chemistry. Exponential in $N_{\mathrm{elec}}$ but not in $N_{\mathrm{orb}}$.
2. Treat part of it variationally, and the rest perturbatively.
3. Although it has a perturbative component, it yields essentially exact energies for interesting systems because the variational and perturbative space are systematically expanded until convergence is achieved.

History of SHCI:
A much improved version of Selected Configuration Interaction + Perturbation Theory (SCI+PT) methods.
First SCI+PT method is the CIPSI method of Malrieu et al (1973).

# References

0. Efficient Heat-Bath Sampling in Fock Space,
   A. A. Holmes, Hitesh J. Changlani and CJU, JCTC, **12**, 1561 (2016)

1. Heat-bath Configuration Interaction: An efficient selected CI algorithm inspired by heat-bath sampling,
   A. A. Holmes, Norm Tubman and CJU, JCTC, **12**, 3674 (2016)

2. Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory,
   Sharma, Holmes, Jeanmairet, Alavi and CJU, JCTC, **13**, 1595 (2017).

3. Excited states using semistochastic heat-bath configuration interaction,
   Holmes, CJU, and Sharma, J. Chem. Phys. **147**, 164111 (2017).

4. Excited States of Methylene, Polyenes, and Ozone from Heat-Bath Configuration Interaction Alan D. Chien, Adam A. Holmes, Matthew Otten, C. J. Umrigar, Sandeep Sharma and Paul M. Zimmerman, J. Phys. Chem. A **122**, 2714 (2018).

Extensions from Sandeep Sharma's group in Colorado:

1. Cheap and Near Exact CASSCF in Large Active Spaces, James Smith, Bastien Mussard, Adam Holmes, and Sandeep Sharma, JCTC, **13**, 5468 (2017)

2. One-Step Treatment of SpinOrbit Coupling and Electron Correlation in Large Active Spaces, Bastien Mussard and Sandeep Sharma, JCTC (2018)

# The people who did most of the work



Adam Holmes
Developed the method

Matthew Otten
Efficient Parallelization
Application to solids
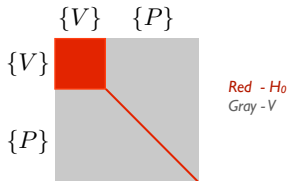
Junhao Li
Super efficient algorithm, HEG

Sandeep Sharma
Semistochastic PT
New directions:
Orbital optimization
Spin-orbit

# Selected CI + PT methods

$$H = H_0 + V$$

$$H_0 = \sum_{i,j} H_{ij}|D_i\rangle\langle D_j| + \sum_a H_{aa}|D_a\rangle\langle D_a|.$$

$$V = \sum_{a,i} H_{ai}|D_a\rangle\langle D_i| + \sum_{a\neq b} H_{ab}|D_a\rangle\langle D_b|$$



Red - $H_0$
Gray - V

$$|\Psi_1\rangle = \frac{1}{E_0 - H_0}V|\Psi_0\rangle = \sum_a \frac{\sum_i H_{ai}c_i}{E_0 - E_a}|D_a\rangle$$

$$\Delta E_2 = \langle\Psi_0|V|\Psi_1\rangle = \sum_a \frac{\left(\sum_i H_{ai}c_i\right)\left(\sum_i H_{ai}c_i\right)}{E_0 - E_a}$$

$$= \sum_a \frac{\left(\sum_i H_{ai}c_i\right)^2}{E_0 - E_a},$$

Note: $E_0 + \Delta E_2$ is identical to the mixed estimator in QMC if the projected wavefunction is replaced by $\Psi_0 + \Psi_1$.

Cyrus J. Umrigar

# Selected CI + PT methods (e.g. CIPSI)

1. Iteratively generate a variational wavefunction as a sum of Slater determinants:

    1. Generate connections to states currently in the wavefunction and keep those for which the perturbative estimate of the absolute coefficient in the wavefunction expansion is sufficiently large

    $$\left| c_a^{(1)} \right| = \left| \frac{\sum_{i \in \mathcal{V}} H_{ai} c_i}{E_0 - H_{aa}} \right|$$

    2. Diagonalize in this expanded space.

2. Compute lowest-order (2nd-order) perturbative correction to the variational energy,

    $$E^{(2)} = \sum_{a \in \mathcal{C}(\mathcal{V})} \frac{\left( \sum_{i \in \mathcal{V}} H_{ai} c_i \right)^2}{E_0 - H_{aa}}$$

**Configuration Interaction Perturbing a Multideterminant Wavefunction Selected Iteratively (CIPSI), B. Huron, J.-P. Malrieu and P Rancurel, J. Chem. Phys. 58, 5745, (1973)**

Cyrus J. Umrigar

# Selected CI + PT methods (e.g. CIPSI)

1. Iteratively generate a variational wavefunction as a sum of Slater determinants:

    1. Generate connections to states currently in the wavefunction and keep those for which the perturbative estimate of the absolute coefficient in the wavefunction expansion is sufficiently large

$$\left| c_a^{(1)} \right| = \left| \frac{\sum_{i \in \mathcal{V}} H_{ai} c_i}{E_0 - H_{aa}} \right|$$

    2. Diagonalize in this expanded space.

2. Compute lowest-order (2nd-order) perturbative correction to the variational energy,

$$E^{(2)} = \sum_{a \in \mathcal{C}(\mathcal{V})} \frac{\left( \sum_{i \in \mathcal{V}} H_{ai} c_i \right)^2}{E_0 - H_{aa}}$$

**At both steps all the connections to the wavefunction are generated. Expensive!**

**Configuration Interaction Perturbing a Multideterminant Wavefunction Selected Iteratively (CIPSI), B. Huron, J.-P. Malrieu and P Rancurel, J. Chem. Phys. 58, 5745, (1973)**

# How expensive?

For a system with 20 electrons in a basis of 120 orbitals (240 spin orbitals) and $10^8$ variational determinants, the number of connections to the variational wavefunction is:

$$10^8 \times \frac{20^2}{2} \times \frac{100^2}{2} \; = \; 1 \times 10^{14}$$

This is expensive in terms of computer time, and it exceeds the available computer memory.

# Semistochastic Heat-bath Configuration Interaction (SHCI) method

Three main improvements relative to CIPSI:

1. At both the variational and the perturbative stages of the calculation find the important connected determinants without ever even looking at the vast majority of the connected determinants.
   Heat-bath Configuration Interaction, Holmes, Tubman and CJU, JCTC (2016)

2. Remove the memory bottleneck in the perturbative step by doing Semistochastic PT.
   Semistochastic Heat-Bath Configuration Interaction Method: Selected Configuration Interaction with Semistochastic Perturbation Theory, Sharma, Holmes, Jeanmairet, Alavi and CJU, JCTC (2017).

3. a) Fast algorithm for constructing Hamiltonian using "helper" arrays,
   b) At various points it is necessary to find a determinant in a list. Can be done by creating sorted list ($\mathcal{O}(N \log(N))$) and binary searching ($\mathcal{O}(\log(N))$) or by hashing ($\mathcal{O}(1)$). We mostly hash.
   Li, Otten, Holmes, Sharma and CJU, unpublished

We elaborate on points 1 and 2 next (3 is too detailed to present in a talk).

Cyrus J. Umrigar

# Structure of quantum chemical Hamiltonian

Since electronic Hamiltonian has only 1- and 2-body interactions:

$$\hat{H} = \sum_{pr} f_{rp} a_r^\dagger a_p + \frac{1}{2} \sum_{pqrs} g_{rspq} a_r^\dagger a_s^\dagger a_q a_p + h_{\text{nuc}}$$

only determinants that differ by 1 or 2 excitations have nonzero $H_{ij}$:

| Type | # connected to a det | # distinct values in full $H$ |
|---|---|---|
| **Single:** $\|H(r \leftarrow p)\|$ $= \left\| f_{rp} + \sum_q^{\mathcal{O}} (g_{rqpq} - g_{qrpq}) \right\|$ | | |
| **Double:** $\|H(rs \leftarrow pq)\|$ $= \|g_{rspq} - g_{srpq}\|$ | | |

# Structure of quantum chemical Hamiltonian

Since electronic Hamiltonian has only 1- and 2-body interactions:

$$\hat{H} = \sum_{pr} f_{rp} a_r^\dagger a_p + \frac{1}{2} \sum_{pqrs} g_{rspq} a_r^\dagger a_s^\dagger a_q a_p + h_{\text{nuc}}$$

only determinants that differ by 1 or 2 excitations have nonzero $H_{ij}$:

| Type | # connected to a det | # distinct values in full $H$ |
|---|---|---|
| **Single:** $\lvert H(r \leftarrow p) \rvert$ $= \left\lvert f_{rp} + \sum_q^{\mathcal{O}} (g_{rqpq} - g_{qrpq}) \right\rvert$ | $\mathcal{O}\left(N N_{\text{orb}}\right)$ | |
| **Double:** $\lvert H(rs \leftarrow pq) \rvert$ $= \lvert g_{rspq} - g_{srpq} \rvert$ | | |

# Structure of quantum chemical Hamiltonian

Since electronic Hamiltonian has only 1- and 2-body interactions:

$$\hat{H} = \sum_{pr} f_{rp} a_r^\dagger a_p + \frac{1}{2} \sum_{pqrs} g_{rspq} a_r^\dagger a_s^\dagger a_q a_p + h_{\text{nuc}}$$

only determinants that differ by 1 or 2 excitations have nonzero $H_{ij}$:

| Type | # connected to a det | # distinct values in full $H$ |
|:---:|:---:|:---:|
| **Single:** $\lvert H(r \leftarrow p) \rvert$ $= \left\lvert f_{rp} + \sum_q^{\mathcal{O}} (g_{rqpq} - g_{qrpq}) \right\rvert$ | $\mathcal{O}\left(N N_{\text{orb}}\right)$ | |
| **Double:** $\lvert H(rs \leftarrow pq) \rvert$ $= \lvert g_{rspq} - g_{srpq} \rvert$ | $\mathcal{O}\left(N^2 N_{\text{orb}}^2\right)$ | |

# Structure of quantum chemical Hamiltonian

Since electronic Hamiltonian has only 1- and 2-body interactions:

$$\hat{H} = \sum_{pr} f_{rp} a_r^\dagger a_p + \frac{1}{2} \sum_{pqrs} g_{rspq} a_r^\dagger a_s^\dagger a_q a_p + h_{\text{nuc}}$$

only determinants that differ by 1 or 2 excitations have nonzero $H_{ij}$:

| Type | # connected to a det | # distinct values in full $H$ |
|:---:|:---:|:---:|
| **Single:** $\lvert H(r \leftarrow p) \rvert$ $= \left\lvert f_{rp} + \sum_q^{\mathcal{O}} (g_{rqpq} - g_{qrpq}) \right\rvert$ | $\mathcal{O}\left(N N_{\text{orb}}\right)$ | combinatorially many! |
| **Double:** $\lvert H(rs \leftarrow pq) \rvert$ $= \lvert g_{rspq} - g_{srpq} \rvert$ | $\mathcal{O}\left(N^2 N_{\text{orb}}^2\right)$ | |

# Structure of quantum chemical Hamiltonian

Since electronic Hamiltonian has only 1- and 2-body interactions:

$$\hat{H} = \sum_{pr} f_{rp} a_r^\dagger a_p + \frac{1}{2} \sum_{pqrs} g_{rspq} a_r^\dagger a_s^\dagger a_q a_p + h_{\text{nuc}}$$

only determinants that differ by 1 or 2 excitations have nonzero $H_{ij}$:

| Type | # connected to a det | # distinct values in full $H$ |
|---|---|---|
| **Single:** $\|H(r \leftarrow p)\|$ $= \left\| f_{rp} + \sum_q^{\mathcal{O}} (g_{rqpq} - g_{qrpq}) \right\|$ | $\mathcal{O}\left(N N_{\text{orb}}\right)$ | combinatorially many! |
| **Double:** $\|H(rs \leftarrow pq)\|$ $= \|g_{rspq} - g_{srpq}\|$ | $\mathcal{O}\left(N^2 N_{\text{orb}}^2\right)$ | $\mathcal{O}\left(N_{\text{orb}}^4\right)$ |

$\Rightarrow$ Most connections are double, but they can be speeded up since the number of distinct values is manageable!

Cyrus J. Umrigar

# Difference between HCI and CIPSI

Current variational wavefunction:

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

Select additional determinants using an importance measure:

$$\text{Select determinant } a \text{ if} = \begin{cases} \left| c_a^{(1)} \right| = \left| \frac{\sum_i H_{ai} c_i}{H_{aa} - E_0} \right| > \epsilon_1, & \text{CIPSI} \quad (1973) \\[2em] \max_i \left( |H_{ai} c_i| \right) > \epsilon_1, & \text{HCI} \quad (2016) \end{cases}$$

**1) Avoid calculating the sum in the numerator.**
**2) Avoid calculating $H_{aa}$ in the denominator.**
**3) Touch only those states that will be included in the wavefunction by presorting 2-body integrals for every pair of orbitals.**

The denominator is never small because all the states for which $H_{aa} - E_0$ is small have already been included in the variational wavefunction.

Cyrus J. Umrigar

# HCI versus CIPSI determinant selection



Comparison of CIPSI and HCI energies for (24e, 30o) $Cr_2$
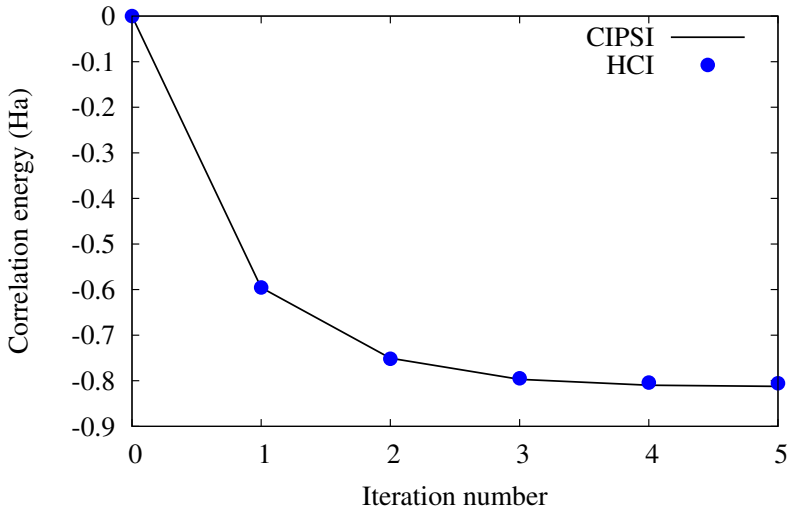
At each iteration, the same # of determinants were used for CIPSI and HCI

Cyrus J. Umrigar

# HCI can be used to preselect the determinants that contribute the most in PT

1. Use HCI criterion,

$$\max_i \left( |H_{ai} c_i| \right) > \epsilon_1$$

to preselect determinants.

2. Use approximate $2^{nd}$-order perturbation theory expression for the energy,

$$\frac{(\sum_j H_{ai} c_i)^2}{H_{aa} - E_0} > \epsilon_1',$$

to select determinants out of the preselected ones.

Efficient, since most of the non-contributing determinants are eliminated at step 1, and only a few tens of percent at step 2.

Cyrus J. Umrigar

# HCI, Stage I: Generating a variational wavefunction

Iterative process: Start with a selected space consisting of only one determinant (say HF). Each iteration,

1. Add to the selected space the determinants $a$ for which $|H_{ai}c_i| > \epsilon_1$ for at least one determinant $i$

2. Diagonalize in the selected space

# HCI, Stage II: Perturbative correction

Compute the perturbative energy correction:

$$\Delta E^{(2)} = \sum_a^{\mathcal{C}} \frac{\left(\sum_i^{\mathcal{V}} H_{ai}c_i\right)^2}{E_0 - H_{aa}} \approx \sum_a^{\mathcal{C}} \frac{\left(\sum_i^{\mathcal{V}(\epsilon_2)} H_{ai}c_i\right)^2}{E_0 - H_{aa}}$$

where $\sum_i^{\mathcal{V}(\epsilon_2)}$ denotes that only terms with $|H_{ai}c_i| > \epsilon_2$ are included.

$\Rightarrow$ Doesn't waste **any** time on the terms that are left out of the sum!

Cyrus J. Umrigar

# Perturbative correction for large systems/bases

Perturbative correction:

$$\Delta E^{(2)} \approx -\sum_a^{\mathcal{C}} \frac{\left(\sum_i^{\mathcal{V}(\epsilon_2)} H_{ai} c_i\right)^2}{H_{aa} - E_0}$$

where $\sum_a^{\mathcal{C}(\epsilon_2)}$ denotes that only terms with $|H_{ai} c_i| > \epsilon_2$ are included in the sum.

Because of square, cannot accumulate contributions.

Straightforward way:

Have to accumulate the contributions from each $a$ separately before squaring.

<span style="color:red">This uses a lot of memory:</span>

$N_V = 10^8$

$N_{\mathrm{conn}} = N_{\mathrm{elec}}^2 N_{\mathrm{orb}}^2 = 20^2 \times 200^2 = 1.6 \times 10^7$ (overestimate, since duplicates exist)

$N_{\mathrm{conn}}^{\mathrm{uniq}} = 10^4$

$N_V \times N_{\mathrm{conn}}^{\mathrm{uniq}} = 10^{12}$ words = 8TB

<span style="color:blue">Instead we have developed a semistochastic algorithm that is memory- and time-efficient, and is easily parallelizable.</span>

# Stochastic perturbative correction

Use $N_s$ samples, each consisting of $N_d$ determinants $\{D_i\}$ sampled from $\mathcal{V}$ with probability

$$p_i = \frac{|c_i|}{\sum_i |c_i|}.$$

Each sample contains $N_d^{\mathrm{diff}} < N_d$ distinct determinants $D_i$ with $w_i$ repeats, such that

$$\sum_i^{N_d^{\mathrm{diff}}} w_i = N_d$$

$w_i$, is distributed according to the multinomial distribution, the mean and second moment of which are

$$\langle w_i \rangle = p_i N_d$$
$$\langle w_i w_j \rangle = \langle w_i \rangle \langle w_j \rangle + \mathrm{cov}(w_i, w_j) \ = \ p_i p_j N_d (N_d - 1).$$

where $\langle \cdot \rangle$ denotes expectation values evaluated for a sample of $N_d$ determinants.

Cyrus J. Umrigar

# Stochastic perturbative correction

The unbiased estimate of the second-order perturbation energy can be calculated:

$$\Delta E_2 = \sum_a \frac{1}{E_0 - E_a} \left[ \sum_{ij}^{\mathcal{V}} H_{ai} H_{aj} c_i c_j \right]$$

$$= \sum_a \frac{1}{E_0 - E_a} \left[ \sum_{i \neq j}^{\mathcal{V}} H_{ai} H_{aj} c_i c_j + \sum_i^{\mathcal{V}} H_{ai}^2 c_i^2 \right]$$

$$= \left\langle \sum_a \frac{1}{E_0 - E_a} \left[ \sum_{i \neq j}^{N_d^{\text{diff}}} \frac{w_i w_j c_i c_j H_{ai} H_{aj}}{\langle w_i w_j \rangle} + \sum_i^{N_d^{\text{diff}}} \frac{w_i c_i^2 H_{ai}^2}{\langle w_i \rangle} \right] \right\rangle$$

$$= \left\langle \sum_a \frac{1}{E_0 - E_a} \left[ \sum_{i \neq j}^{N_d^{\text{diff}}} \frac{w_i w_j c_i c_j H_{ai} H_{aj}}{p_i p_j N_d (N_d - 1)} + \sum_i^{N_d^{\text{diff}}} \frac{w_i c_i^2 H_{ai}^2}{p_i N_d} \right] \right\rangle$$

$$= \frac{1}{N_d (N_d - 1)} \left\langle \sum_a \frac{1}{E_0 - E_a} \left[ \left( \sum_i^{N_d^{\text{diff}}} \frac{w_i c_i H_{ai}}{p_i} \right)^2 + \sum_i^{N_d^{\text{diff}}} \left( \frac{w_i (N_d - 1)}{p_i} - \frac{w_i^2}{p_i^2} \right) c_i^2 H_{ai}^2 \right] \right\rangle$$

Stochastic, but, unlike QMC, no sign problem and uncorrelated samples!

Cyrus J. Umrigar

# Semistochastic perturbative correction

$$\Delta E_2(\epsilon_2) = \left\langle \Delta E_2^{\text{stoch}}(\epsilon_2) \right\rangle + \left[ \Delta E_2^{\text{det}}(\epsilon_2^{\text{big}}) - \left\langle \Delta E_2^{\text{stoch}}(\epsilon_2^{\text{big}}) \right\rangle \right]$$

1. The term in [ ] has zero expectation value, but cancels most of the fluctuations in the first term at no additional cost.

2. Use small $\epsilon_2$, say $10^{-8}$ to have negligible bias. $\epsilon_2^{\text{big}}$ is as small as memory allows. Hence get super-linear speedup on parallel computer! No statistical error when $\epsilon_2^{\text{big}} \to \epsilon_2$.

3. Larger memory also allows larger statistical samples, and this also contributes to a super-linear speedup since it is more efficient to have a small number of large samples than a large number of small samples.

Cyrus J. Umrigar

SHCI Energy of $Cr_2$ with x2c Ham, cc-pVDZ-DK basis

$E_{var}$ +
$E_{tot}$ *

E (Ha)

$E_{var}$-$E_{tot}$
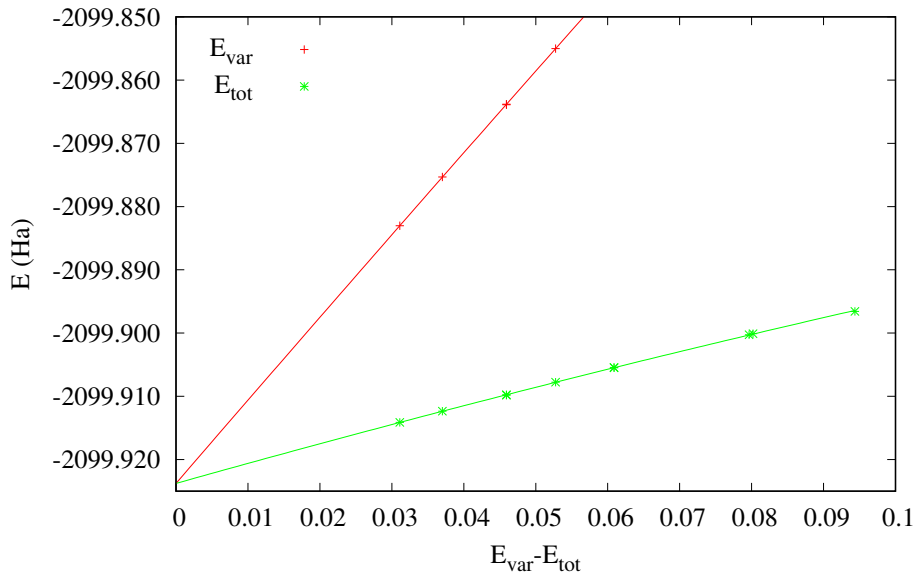
For large systems quadratic (nearly linear) extrap. gives accurate energy.
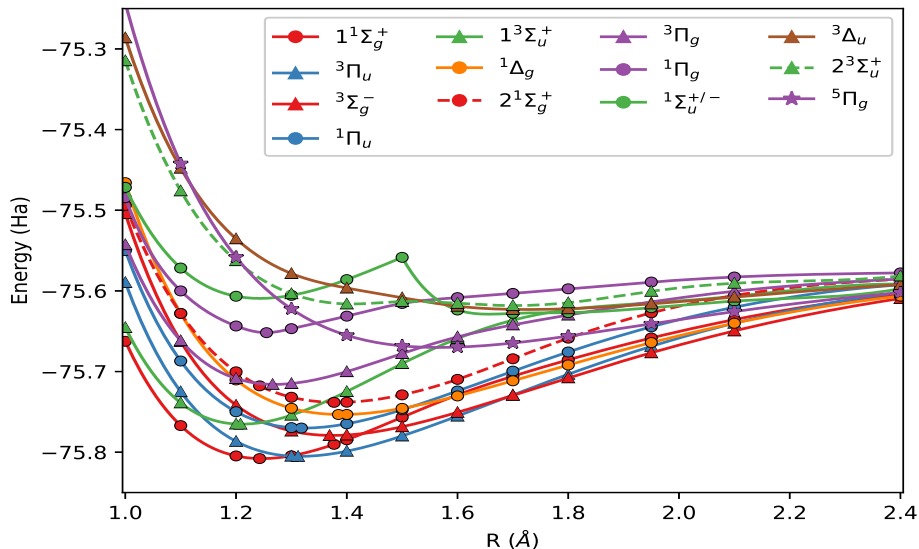Cyrus J. Umrigar

# Excited states

Excited states are straightforward: Apply the HCI selection criterion

$$\max_i \left( |H_{ai} c_i| \right) > \epsilon_1$$

to each of the states and take the union of the selected states.

# Lowest 14 excited states of cc-pV5Z C$_2$

## HCI Excited States of the cc-pV5Z Carbon dimer



Cyrus J. Umrigar

# $Cr_2$ calculations

1. $1^{st}$ paper: 24-electron and 48-electron (all-electron) calculations in
   Aldrichs VDZ basis, at $r_e$.
   $(^{30}C_{12})^2 = 7.5 \times 10^{15}$,     $(^{42}C_{24})^2 = 1.2 \times 10^{23}$ dets.
   Time: Few minutes on a single core!

2. $2^{nd}$ paper: 12-electron (Ar core) calculations in cc-pVDZ-DK,
   cc-pVTZ-DK and cc-pVQZ-DK bases, at $r_e$.
   $\binom{68}{6}^2 = 1.2 \times 10^{16}$,   $\binom{118}{6}^2 = 1.1 \times 10^{19}$,   $\binom{190}{6}^2 = 3.6 \times 10^{21}$ dets.
   Time: 20 minutes on 1 node to 2 hours on 4 nodes.

3. current work: 28-electron (Ne core) calculations in cc-pVDZ-DK to
   cc-pwCVQZ bases, for full PE curve.
   $\binom{76}{14}^2 = 5 \times 10^{29}$ dets.: 1 day on 50 nodes with 20 cores, 48MB/node.
   $\binom{288}{14}^2 = 6 \times 10^{49}$ dets.: 2 days on 8 nodes with 40 cores, 3TB/node.
   The latter has 1B variational dets, but is not fully converged.

Cyrus J. Umrigar

# Cr$_2$ at r = 1.5 Å in Ahlrichs VDZ basis

Hilbert space: $\left(^{42}C_{24}\right)^2 = 1.2 \times 10^{23}$.

Heat-bath CI, with $\epsilon_1 = 1$ mHa and $\epsilon_2 = 10$ $\mu$Ha, yields:

Table: Energies (Ha)

| Method | (24e, 30o) Mg core | (48e, 42o) all electron |
|---|---|---|
| HCI($\epsilon_1 = 1$, $\epsilon_2 = 0.01$ mHa) [1] | $-2086.421\,30$ | $-2086.444\,04$ |
| DMRG[2] | $-2086.420\,95(3)$ | $-2086.444\,78(32)$ |

These HCI calculations took 2 and 8 minutes, respectively, on a single core

$\Rightarrow$ All-electron calculations are not much more challenging for HCI than frozen core (this basis has core-valence correlations but not core-core)!

[1] A. A. Holmes, Norm Tubman and C. J. Umrigar, *JCTC*, **12**, 3674 (2016)

[2] Olivares-Amaya, Hu, Nakatani, Sharma, Yang and Chan, *J. Chem. Phys.* **142**, 034102 (2015)

Cyrus J. Umrigar

# Distribution of excitation orders, $Cr_2$ 12 electrons
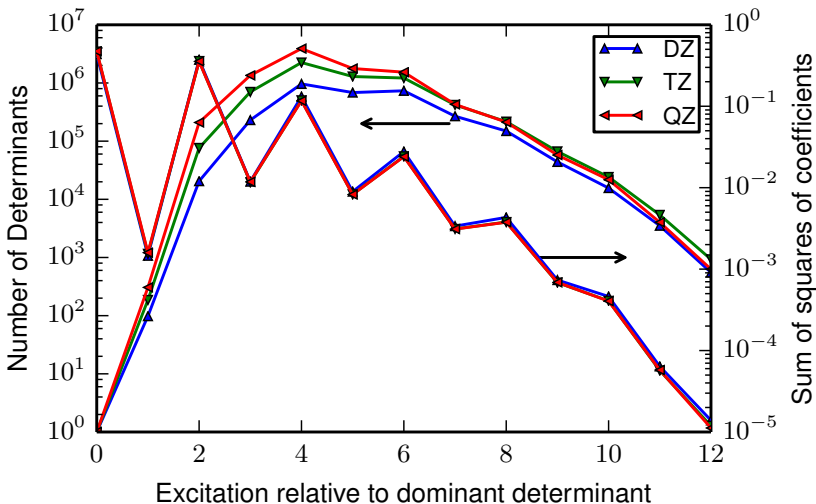


So CISDTQ would be nowhere near good enough.
When we correlate 28 electrons (frozen Ne core) we have upto 16-fold excitations!

Cyrus J. Umrigar

# Cr$_2$ Ar core
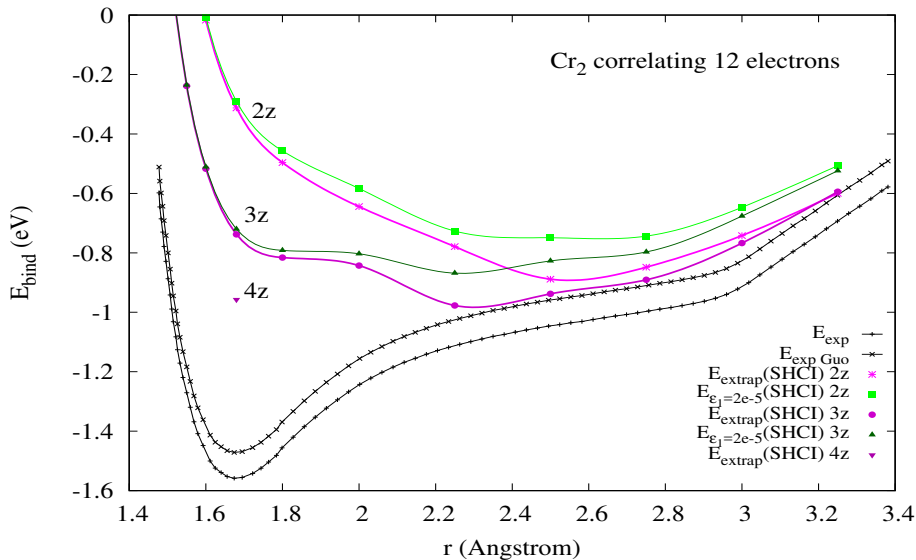


Short-distance behaviour bad because of lack of semicore correlation.

Cyrus J. Umrigar

# Cr$_2$ Ne core
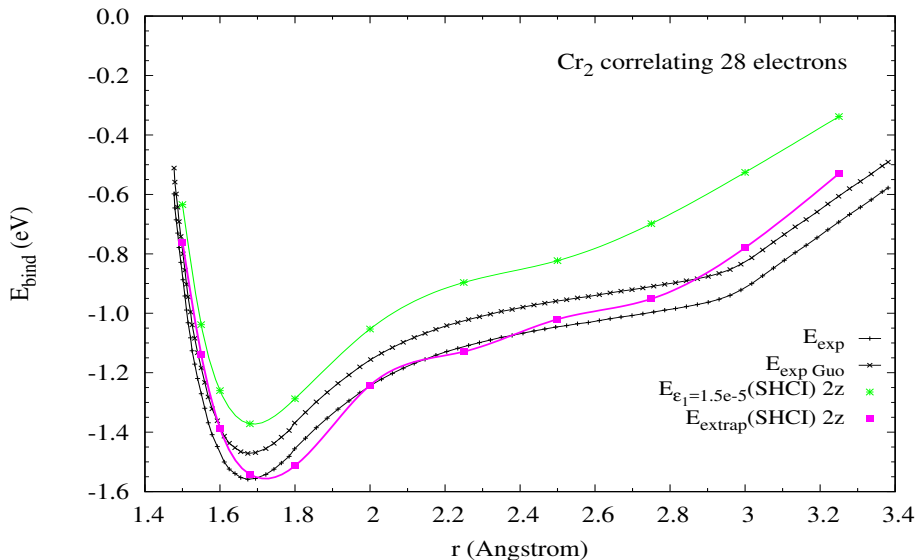


Good agreement, though probably fortuitous. One more $\epsilon_1$ value needed to get better convergence.
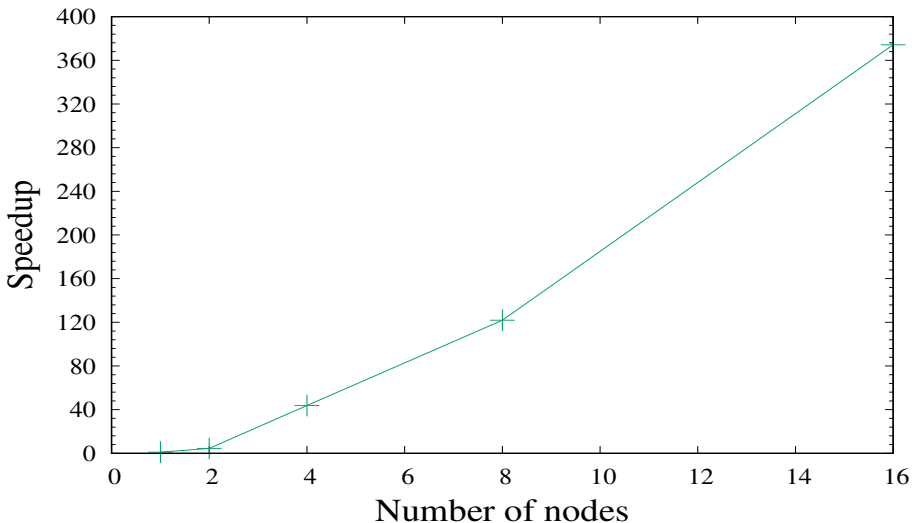
Cyrus J. Umrigar

# Parallel Efficiency

Speedup of variational part of SHCI for $C_2$ using cc-pV5Z basis
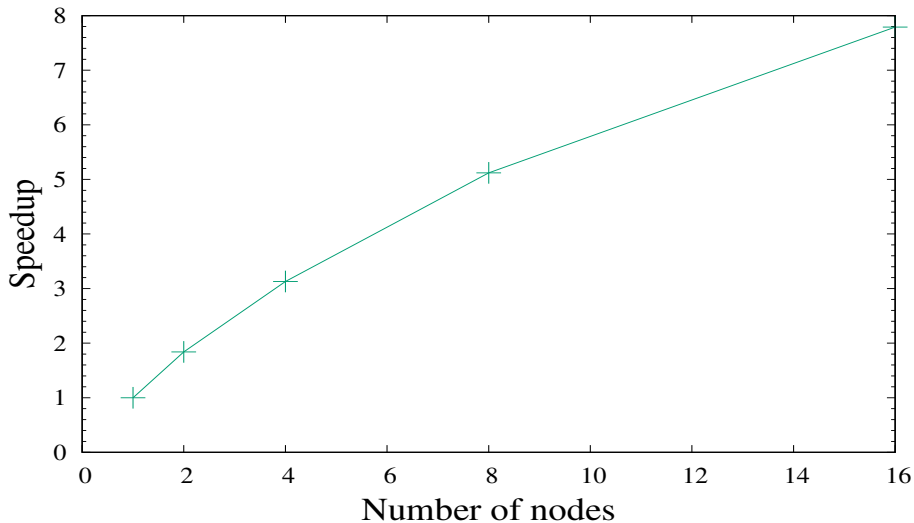


50% parallel efficiency for the variational step.

Cyrus J. Umrigar

# Parallel Efficiency



Speedup of perturbative part of SHCI for $C_2$ using cc-pV5Z basis

Highly superlinear parallel efficiency for the perturbative step!!

# Choice of Orbitals

Natural orbitals give faster convergence than HF orbitals.

Optimized orbitals give yet faster convergence.

Cheap and Near Exact CASSCF in Large Active Spaces,
James Smith, Bastien Mussard, Adam Holmes, and Sandeep Sharma, JCTC, **13**, 5468 (2017)

Split-localized optimized orbitals are yet more efficient for molecules with several atoms.

# Two Semistochastic Algorithms in our group
# S-FCIQMC and SHCI

1. FCIQMC $\rightarrow$ S-FCIQMC (Stochastic $\rightarrow$ Semistochastic) (2012)

2. CIPSI $\rightarrow$ SHCI   (Deterministic $\rightarrow$ Semistochastic) (2017)

Advantages of SHCI relative to S-FCIQMC:

1. No sign problem

2. No autocorrelation time (which can be very large for large basis sets in S-FCIQMC)

Consequently SHCI calculations are typically orders of magnitude faster than S-FCIQMC.

Possible disadvantages of SHCI relative to S-FCIQMC:

1. Memory constraint may be more severe.
2. Parallel efficiency of the variational part of the calculation may be a bit less than for S-FCIQMC.

Cyrus J. Umrigar

# Comparison to DMRG

SHCI more efficient when there are a lot of electrons per atom.

DMRG more efficient for quasi 1-D systems and for systems with a lot of natural orbitals with fillings close to 1 (as opposed to 0 or 2).

# SHCI wavefunctions in DMC

Diffusion Monte Carlo (DMC) works directly at the infinite basis set limit.
However, fixed-node (or fixed-phase) approximation is necessary.

Fixed-node error can be made negligible by systematically improving the trial wavefunctions and calculations with as many as $2 \times 10^5$ determinants with fully optimized orbitals and determinantal coefficients are feasible.

Usual procedure for picking determinants is to do a CASSCF or a RASSCF calculation and pick the determinants with the largest coefficients.

This has 2 problems:

1. Some higher-order excitations may be important and some excitations outside the CAS space may be important.
2. Determinants that are important in the presence of a Jastrow are not necessarily the ones that are important in the absence of a Jastrow.

HCI immediately eliminates the first problem.

The $2^{nd}$ problem can be eliminated by using an HCI-like procedure within QMC to select important determinants.

Cyrus J. Umrigar

# SHCI wavefunctions in AFQMC and DMC

The usual procedure in AFQMC and DMC is to use trial wavefns from CAS or RAS calculations (reoptimized in the presence of a Jastrow in the case of DMC). For small expansions these have the most important determinants, but not for larger expansions.

Energy of $Cu^+$, Hao Shi et al,, unpublished



Cyrus J. Umrigar
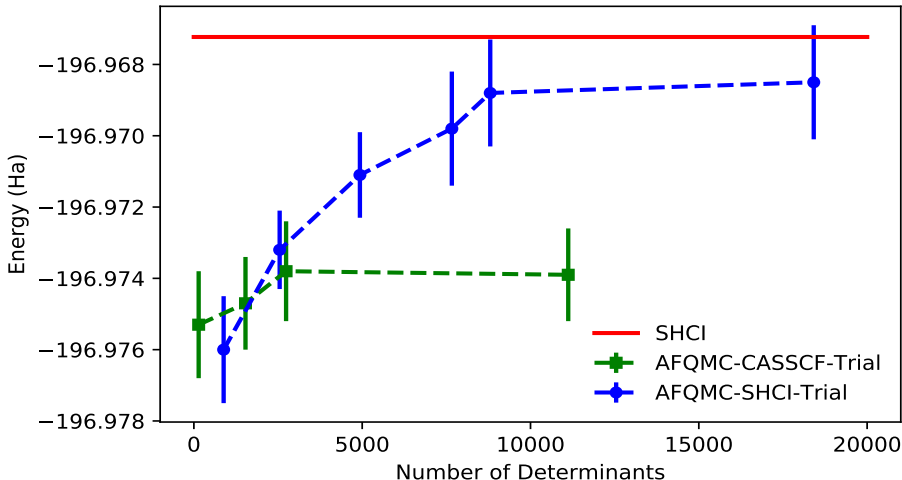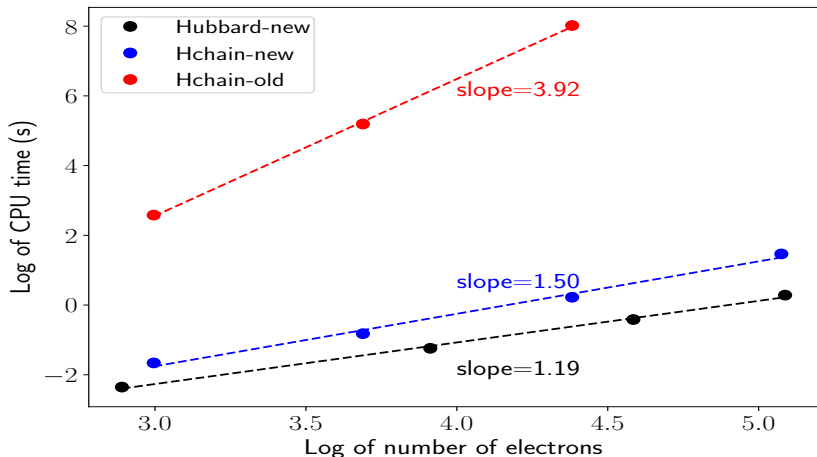
# SHCI wavefunctions in AFQMC and DMC

## Energy of Butadiene
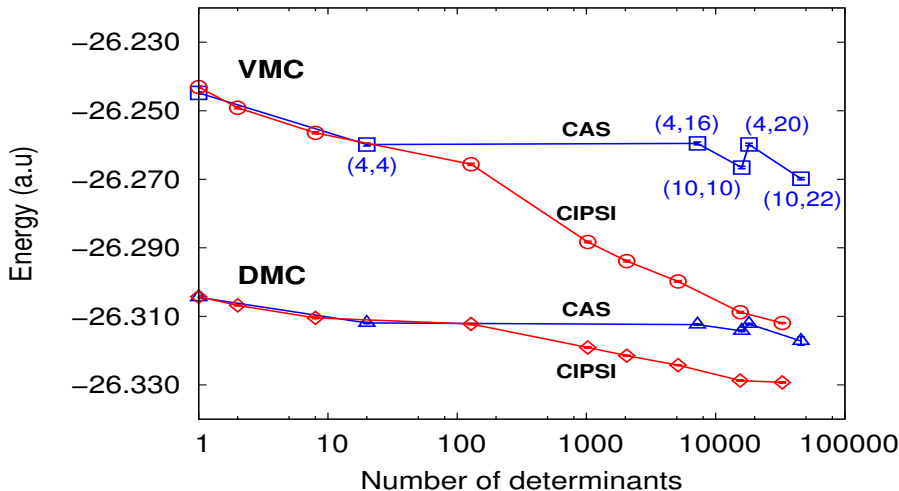
Could do yet better by selecting the determinants in the presence of a Jastrow.

Cyrus J. Umrigar

# Scaling of orbital-space VMC (Sharma)

HCI greatly improves the scaling of orbital-space VMC, using localized orbitals.