QMC with large CI trial wave functions

A. Scemama, E. Giner, T. Applencourt, Y. Garniron, P.-F. Loos, M. Caffarel
18/07/2017

Lab. Chimie et Physique Quantiques, IRSAMC, UPS/CNRS, Toulouse, France
Wave function methods
Chlorine atom in cc-pVDZ

- 9 ↑ and 8 ↓ electrons in 19 MOs
- Full-CI: 92378 ↑ determinants and 75582 ↓ determinants
- 6.98 \(10^9\) possible Slater determinants
- The size of the Full-CI space is huge, but the space is empty: \(\sim 10^6\) coefficients \(|c_k| > 10^{-12}\)

\[\implies\text{Selected CI algorithm to select iteratively the most important Slater determinants}\]
Brief history of Selected CI

1969 Bender / Davidson et al: one-shot selection with perturbative criterion
1969 Whitten / Ackmeyer: Independently proposed the same criterion, but with iterative selection
1973 Malrieu et al: Add a PT2 calculation to the selected determinants (CIPSI)

Groups using/developing CIPSI independently since then

- Angeli, Cimiraglia, Persico: Italy
- Barone: Italy
- Illas, Rubio, Ricart: Spain
- Malrieu, Daudey, Spiegelman: France
Iterative selected CI has been re-invented many times.
Iterative selected CI has been re-invented many times.

CIPSI

MRDCI (< 1980)
Iterative selected CI has been re-invented many times.

- CIPSI
- MRDCI (< 1980)
- Adaptative CI (Evangelista, 2014)
Iterative selected CI has been re-invented many times.

- CIPSI
- MRDCI (< 1980)
- Adaptative CI (Evangelista, 2014)
- Accelerated CI (Neese, 2015)
Iterative selected CI has been re-invented many times.

- CIPSI
- MRDCI (< 1980)
- Adaptative CI (Evangelista, 2014)
- Accelerated CI (Neese, 2015)
- Heat-Bath CI (Holmes, 2016)
Brief history of Selected CI

Iterative selected CI has been re-invented many times.

- CIPSI
- MRDCI (< 1980)
- Adaptative CI (Evangelista, 2014)
- Accelerated CI (Neese, 2015)
- Heat-Bath CI (Holmes, 2016)
- Iterative CI (Liu, 2016)
Iterative selected CI has been re-invented many times.

- CIPSI
- MRDCI (< 1980)
- Adaptative CI (Evangelista, 2014)
- Accelerated CI (Neese, 2015)
- Heat-Bath CI (Holmes, 2016)
- Iterative CI (Liu, 2016)

All differ by the selection criterion and the implementation. But they refer to the same natural idea.
Selected CI algorithm

1. Define a **reference** wave function:

\[ |\Psi\rangle = \sum_{i \in \mathcal{D}} c_i |i\rangle \quad E_{\text{var}} = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \]
Selected CI algorithm

1. Define a reference wave function:

\[ |\psi\rangle = \sum_{i \in D} c_i |i\rangle \quad \text{E}_{\text{var}} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \]

2. Generate external determinants \( \alpha \):

\[ A = \left\{ (\forall i \in D) \left( \forall \hat{T} \in \mathcal{T}_1 \cup \mathcal{T}_2 \right) : |\alpha\rangle = \hat{T}|i\rangle \right\} \]
Selected CI algorithm

1. Define a reference wave function:

\[ |\psi\rangle = \sum_{i \in D} c_i |i\rangle \quad \text{and} \quad E_{\text{var}} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \]

2. Generate external determinants \( \alpha \):

\[ A = \left\{ (\forall i \in D) \left( \forall \hat{T} \in T_1 \cup T_2 \right) : |\alpha\rangle = \hat{T}|i\rangle \right\} \]

3. Second order perturbative contribution of each \( |\alpha\rangle \) :

\[ \Delta E_\alpha = \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\text{var}} \text{ } - \text{ } \langle \alpha | \hat{H} | \alpha \rangle} \]
Selected CI algorithm

1. Define a reference wave function:

\[ |\psi\rangle = \sum_{i \in \mathcal{D}} c_i |i\rangle \quad \text{and} \quad E_{\text{var}} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \]

2. Generate external determinants \( \alpha \):

\[ \mathcal{A} = \left\{ (\forall i \in \mathcal{D}) \left( \forall \hat{T} \in \mathcal{T}_1 \cup \mathcal{T}_2 \right) : |\alpha\rangle = \hat{T}|i\rangle \right\} \]

3. Second order perturbative contribution of each \( |\alpha\rangle \) :

\[ \Delta E_{\alpha} = \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\text{var}} - \langle \alpha | \hat{H} | \alpha \rangle} \]

4. Select the \( |\alpha\rangle \)'s with the largest \( \Delta E_{\alpha} \) and add them into \( \mathcal{D} \)
Selected CI algorithm

1. Define a reference wave function:

\[ |\psi\rangle = \sum_{i \in \mathcal{D}} c_i |i\rangle \quad E_{\text{var}} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \]

2. Generate external determinants \( \alpha \):

\[ \mathcal{A} = \left\{ (\forall i \in \mathcal{D}) \left( \forall \hat{T} \in \mathcal{T}_1 \cup \mathcal{T}_2 \right) : |\alpha\rangle = \hat{T}|i\rangle \right\} \]

3. Second order perturbative contribution of each \( |\alpha\rangle \):

\[ \Delta E_\alpha = \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\text{var}} - \langle \alpha | \hat{H} | \alpha \rangle} \]

4. Select the \( |\alpha\rangle \)'s with the largest \( \Delta E_\alpha \) and add them into \( \mathcal{D} \)

5. Diagonalize \( \hat{H} \) in \( \mathcal{D} \) \( \Rightarrow \) update \( |\psi\rangle \) and \( E_{\text{var}} \)
Selected CI algorithm

1. Define a **reference** wave function:

\[ |\psi\rangle = \sum_{i \in \mathcal{D}} c_i |i\rangle \quad \quad E_{\text{var}} = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} \]

2. Generate **external determinants** \( \alpha \):

\[ \mathcal{A} = \left\{ (\forall i \in \mathcal{D}) \left( \forall \hat{T} \in \mathcal{T}_1 \cup \mathcal{T}_2 \right) : |\alpha\rangle = \hat{T} |i\rangle \right\} \]

3. Second order perturbative contribution of each \( |\alpha\rangle \):

\[ \Delta E_\alpha = \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\text{var}} - \langle \alpha | \hat{H} | \alpha \rangle} \]

4. Select the \( |\alpha\rangle \)'s with the largest \( \Delta E_\alpha \) and add them into \( \mathcal{D} \)

5. Diagonalize \( \hat{H} \) in \( \mathcal{D} \) \( \implies \) update \( |\psi\rangle \) and \( E_{\text{var}} \)

6. Iterate
Remarks

- When all $|\alpha\rangle$’s are selected the Full-CI is obtained.
- CIPSI is more an algorithm than a method.
- CIPSI can be seen as the deterministic counterpart of FCIQMC.
- Any WF method can be realized with the CIPSI algorithm.
- Rules on the generation of $|\alpha\rangle$’s define the wave function method (CISD, CAS, MRCI, . . .).
- Rules on the selection of $|\alpha\rangle$’s define the person you need to cite for the trademark.
- With the CIPSI selection, at any time $E_{PT2} = \sum_\alpha \Delta E_\alpha$ estimates the distance to the solution.
Making CIPSI efficient

1. Initiator-like approximation: Generate only $\alpha$ determinants from the most important determinants (typically 99% of the norm) (Evangelisti, Chem. Phys., 1983). Typically $\sim 2000$ generators for $10^7$ determinants.

2. Each $\alpha$ determinant is connected to only a subset of determinants of the wave function $\implies \Delta E_\alpha$ is done in constant time (Cimiraglia, JCP, 1985).

3. Evaluation of $H_{\alpha\alpha}$ needs 10 flops in the worst case: doesn’t scale with the number of electrons (Cimiraglia, JCP, 1985).

Our contributions:

1. Efficient implementation of Slater Condon’s rules (ArXiv 2013).

2. Knowing that $H_{ij}$ is zero can be done in 5.3 CPU cycles, $\sim 50 \times$ faster than a random memory access.

3. Sorting: $O(N \log N)$ for mathematicians but $O(N)$ for computer scientists (radix sort algorithm).
Making CIPSI efficient

1. Initiator-like approximation: Generate only \( \alpha \) determinants from the most important determinants (typically 99% of the norm) (Evangelisti, Chem. Phys., 1983). Typically \( \sim 2000 \) generators for \( 10^7 \) determinants.

2. Each \( \alpha \) determinant is connected to only a subset of determinants of the wave function \( \implies \Delta E_\alpha \) is done in constant time (Cimiraglia, JCP, 1985).

3. Evaluation of \( H_{\alpha\alpha} \) needs 10 flops in the worst case: doesn’t scale with the number of electrons (Cimiraglia, JCP, 1985).

---

Our contributions:

- Efficient implementation of Slater Condon’s rules (ArXiv 2013)
- Knowing that \( H_{ij} \) is zero can be done in 5.3 CPU cycles, \( \sim 50 \times \) faster than a random memory access.
- Sorting: \( O(N \log N) \) for mathematicians but \( O(N) \) for computer scientists (radix sort algorithm).
1. Initiator-like approximation: Generate only $\alpha$ determinants from the most important determinants (typically 99% of the norm) (Evangelisti, Chem. Phys., 1983). Typically $\sim 2000$ generators for $10^7$ determinants

2. Each $\alpha$ determinant is connected to only a subset of determinants of the wave function $\Rightarrow \Delta E_\alpha$ is done in constant time (Cimiraglia, JCP, 1985)

3. Evaluation of $H_{\alpha\alpha}$ needs 10 flops in the worst case: doesn’t scale with the number of electrons (Cimiraglia, JCP, 1985)

Our contributions:

1. Efficient implementation of Slater Condon’s rules (ArXiv 2013)
2. Knowing that $H_{ij}$ is zero can be done in 5.3 CPU cycles, $\sim 50\times$ faster than a random memory access
3. Sorting: $O(N \log(N))$ for mathematicians but $O(N)$ for computer scientists (radix sort algorithm)
Making CIPSI efficient

- Selection and PT2 scale *linearly* with the number of variational determinants
- Efficient hybrid stochastic PT2\(^1\) which converges to the exact (zero error) value in *finite* time. Error convergence rate is \(O(1/t^{3.5})\).
- Selection is embarrassingly parallel (client/server task parallelism)
- Practical bottleneck is today the Davidson diagonalization (work in progress)

What we can do today

Big picture: we can address the same problems as other selected CI methods, including FCI-QMC.
These are not heroic runs but large runs:

- \( \text{Cr}_2 \) cc-pVQZ, 20 million dets + PT2, (JCP 2017)
- Butadiene, 60 million dets + PT2 (work in progress)
- \( \text{FeO}_4 \), def2-TZVPP, 50 million dets + PT2 (work in progress)
- \( \text{Cu}_2\text{O}_2^{2+} \), 6-31G benchmark (work in progress)
- We could not do more than 60 million because of a stupid 32-bit integer limit in a library call (will be fixed soon)
CIPSI vs HCl

Energy: \[ \sum_{\alpha \notin D} \frac{\langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle}{E_{\text{var}} - \langle \alpha | \hat{H} | \alpha \rangle} \]

Variance: \[ \sum_{\alpha \notin D} \langle \psi | \hat{H} | \alpha \rangle \langle \alpha | \hat{H} | \psi \rangle \]

Copper, augmented ANO-TZ

Difference:
CIPSI: Energy minimization  HCl: Variance minimization
QMC
Jastrow factors?

- **Full-CI**: post-Hartree-Fock method

\[ J = 1 \]

To see what these WFs give with a Jastrow, see Anouar Benali’s talk on Friday.
Jastrow factors?

- Full-CI: post-Hartree-Fock method
- DMC: post-Full-CI method

\(^2\)To see what these WFs give with a Jastrow, see Anouar Benali’s talk on Friday.
Jastrow factors?

- Full-CI: post-Hartree-Fock method
- DMC: post-Full-CI method

What else?

\[ J^2 \]

To see what these WFs give with a Jastrow, see Anouar Benali’s talk on Friday.
Jastrow factors?

- Full-CI: post-Hartree-Fock method
- DMC: post-Full-CI method

What else?

So in this talk, I have used my favorite Jastrow factor:\(^2\): \[ J = 1 \]

---

\(^2\)To see what these WFs give with a Jastrow, see Anouar Benali’s talk on Friday.
Efficient scheme for CI wave functions in QMC

Ψ can be expressed in a bi-linear form \(^3\):

\[
\Psi(R) = \sum_{k}^{N_{\text{det}}} c_k D_k = \sum_{i}^{N_{\text{det} \uparrow}} \sum_{j}^{N_{\text{det} \downarrow}} C_{ij} D_{i \uparrow}(R_{\uparrow}) D_{j \downarrow}(R_{\downarrow})
\]  

- \(D_{\uparrow}(R_{\uparrow})\): vector of \(N_{\text{det} \uparrow}\) elements
- \(D_{\downarrow}(R_{\downarrow})\): vector of \(N_{\text{det} \downarrow}\) elements
- \(C\): \(N_{\text{det} \uparrow} \times N_{\text{det} \downarrow}\) matrix. The matrix contains \(N_{\text{det}}\) elements.

\(C\) is constant in a QMC run \(\Longrightarrow\) preprocessing.

\(^3\)QMC with very large multideterminant wavefunctions

Efficient scheme for CI wave functions in QMC

At every MC step, we need to evaluate: (↑ electrons and ↓ electrons)

\[ \Psi = (D_{↑}^\dagger CD_\downarrow) \] (2)

\[ \nabla_i \Psi = \nabla_i D_{↑}^\dagger (CD_\downarrow) \text{ or } (D_{↑}^\dagger C) \nabla_i D_\downarrow \] (3)

\[ \Delta_i \Psi = \Delta_i D_{↑}^\dagger (CD_\downarrow) \text{ or } (D_{↑}^\dagger C) \Delta_i D_\downarrow \] (4)

\[ V_{\text{non-loc pseudo}}^{\Psi} = V_{\text{non-loc pseudo}}^{\Psi} D_{↑}^\dagger (CD_\downarrow) \text{ or } (D_{↑}^\dagger C). V_{\text{non-loc pseudo}}^{\Psi} D_\downarrow \] (5)

- Only \( \Psi \) requires \( \mathcal{O}(N_{\text{det}}) \) operations (tiny prefactor)
- Others are \( \mathcal{O}(N_{\text{elec}}^{↑} \times N_{\text{det}}^{↑}) \Rightarrow \) expensive work is \( \mathcal{O}(\sqrt{N_{\text{det}}}) \)

\(^{3}\)QMC with very large multideterminant wavefunctions

Efficient scheme for CI wave functions in QMC

At every MC step, we need to evaluate: (↑ electrons and ↓ electrons)

\[
\Psi = (D_{\uparrow}^\dagger (C) D_{\downarrow}) \tag{2}
\]

\[
\nabla_i \Psi = \nabla_i D_{\uparrow}^\dagger (C D_{\downarrow}) \text{ or } (D_{\uparrow}^\dagger C) \cdot \nabla_i D_{\downarrow} \tag{3}
\]

\[
\Delta_i \Psi = \Delta_i D_{\uparrow}^\dagger (C D_{\downarrow}) \text{ or } (D_{\uparrow}^\dagger C) \cdot \Delta_i D_{\downarrow} \tag{4}
\]

\[
V_{\text{non−loc pseudo}} \Psi = V_{\text{non−loc pseudo}} D_{\uparrow}^\dagger (C D_{\downarrow}) \text{ or } (D_{\uparrow}^\dagger C) \cdot V_{\text{non−loc pseudo}} D_{\downarrow} \tag{5}
\]

- Only \( \Psi \) requires \( O(N_{\text{det}}) \) operations (tiny prefactor)
- Others are \( O(N_{\text{elec}} \times N_{\text{det}}) \) \( \Rightarrow \) expensive work is \( O(\sqrt{N_{\text{det}}}) \)
- We can now use \( \sim \) 2 000 000 determinants in DMC.

\(^3\)QMC with very large multideterminant wavefunctions
Pseudo-potentials

Back 30 years ago$^4$:

\textbf{C. ECP–QMC}

For QMC we render the nonlocal ECP operator in a local form. This cannot in general be done exactly, but to a very good approximation this can be accomplished simply by allowing $U_{\text{ECP}}$ to act on $\Psi_{\text{val}}$ as implied in Eq. (6b). This leads to an additional term in the local energy, namely

$$U_{\text{ECP, Local}} = \sum_A \sum_i^{N_{\text{val}}} \left( U_{l_{\text{max}} + 1}^{A} (r_{iA}) + \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^{l} Y_{lm} (\Omega_{iA}) \right)$$

$$\times U_{l}^{A} (r_{iA}) \langle Y_{lm} (\Omega_{iA}) | \Psi_{\text{val}} \rangle / \Psi_{\text{val}} \right), \quad (10)$$

where $\Psi_{\text{val}}$ becomes the QMC valence importance function.

$$\langle Y_{lm} (\Omega_{iA}) | \Psi_{\text{val}} \rangle / \Psi_{\text{val}} = \sum_j D_{ji}^{-1} \langle Y_{lm} (\Omega_{iA}) | \phi_j (i) \rangle, \quad (11)$$

$^4$Valence quantum Monte Carlo with abinitio effective core potentials
Pseudo-potentials

- $\langle Y_{lm}(\Omega_iA)|\Phi_j(i)\rangle$ can be computed analytically
- For efficiency: pre-computed on a grid
- No more quadrature points to compute
- No more need for “T-moves or not T-moves”
- Cost: Same as Laplacian ($\sim 15 - 20\%$ of a MC step)
- Of course, also applicable to single-determinant for DFT trial wave functions
Some results
1. FCI/DMC — $\text{H}_2\text{O}$

All-electron quasi-Full-CI trial wave functions:

TABLE I. Number of determinants and corresponding variational energies for CIPSI expansions used in DMC for each cc-pCVnZ ($n = 2$ to $5$) basis set. Last column: Deviations of the variational energy to the best FCI estimates of Almora-Díaz.\textsuperscript{27} Energies in atomic units.

<table>
<thead>
<tr>
<th>Basis set</th>
<th>FCI size</th>
<th># dets used in DMC</th>
<th>$E^\text{var}_0$</th>
<th>FCI, Almora-Díaz\textsuperscript{27}</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc-pCVDZ</td>
<td>$\sim 10^{10}$</td>
<td>172 256</td>
<td>$-76.282\ 136$</td>
<td>$-76.282\ 865$</td>
<td>0.0007</td>
</tr>
<tr>
<td>cc-pCVTZ</td>
<td>$\sim 2 \cdot 10^{14}$</td>
<td>640 426</td>
<td>$-76.388\ 287$</td>
<td>$-76.390\ 158$</td>
<td>0.0018</td>
</tr>
<tr>
<td>cc-pCVQZ</td>
<td>$\sim 2 \cdot 10^{17}$</td>
<td>666 927</td>
<td>$-76.419\ 324$</td>
<td>$-76.421\ 148$</td>
<td>0.0018</td>
</tr>
<tr>
<td>cc-pCV5Z</td>
<td>$\sim 7 \cdot 10^{19}$</td>
<td>1 423 377</td>
<td>$-76.428\ 550$</td>
<td>$-76.431\ 105$</td>
<td>0.0025</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Basis set[Ndets]</th>
<th>$T_{CPU}(Ndets)/T_{CPU}(1det)$</th>
<th>$E_0^{\text{DMC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc-pCVDZ[172 256]</td>
<td>$\sim 101$</td>
<td>$-76.415\ 71(20)$</td>
</tr>
<tr>
<td>cc-pCVTZ[640 426]</td>
<td>$\sim 185$</td>
<td>$-76.431\ 82(19)$</td>
</tr>
<tr>
<td>cc-pCVQZ[666 927]</td>
<td>$\sim 128$</td>
<td>$-76.436\ 22(14)$</td>
</tr>
<tr>
<td>cc-pCV5Z[1 423 377]</td>
<td>$\sim 235$</td>
<td>$-76.437\ 44(18)$</td>
</tr>
</tbody>
</table>

1. FCI/DMC — H$_2$O

**TABLE III.** Comparison of nonrelativistic ground-state total energies of water obtained with the most accurate theoretical methods. Energies in a.u.

<table>
<thead>
<tr>
<th>Method</th>
<th>Energy (a.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clark <em>et al.</em>,$^{20}$ DMC (upper bound)</td>
<td>$-76.4368(4)$</td>
</tr>
<tr>
<td>This work, DMC (upper bound)</td>
<td>$-76.43744(18)$</td>
</tr>
<tr>
<td>Almora-Diaz,$^{27}$ CISDTQnSx (upper bound)</td>
<td>$-76.4343$</td>
</tr>
<tr>
<td>Helgaker <em>et al.</em>,$^{29}$ R12-CCSD(T)</td>
<td>$-76.4392$</td>
</tr>
<tr>
<td>Muller and Kutzelnigg,$^{30}$ R12-CCSD(T)</td>
<td>$-76.4373$</td>
</tr>
<tr>
<td>Almora-Diaz,$^{27}$ FCI + CBS</td>
<td>$-76.4386(9)$</td>
</tr>
<tr>
<td>Halkier <em>et al.</em>,$^{31}$ CCSD(T) + CBS</td>
<td>$-76.4386$</td>
</tr>
<tr>
<td>Bytautas and Ruedenberg,$^{32}$ FCI + CBS</td>
<td>$-76.4390(4)$</td>
</tr>
<tr>
<td>This work, DMC + CBS</td>
<td>$-76.43894(12)$</td>
</tr>
<tr>
<td>Experimentally derived estimate$^{25}$</td>
<td>$-76.4389$</td>
</tr>
</tbody>
</table>

1. FCI/DMC — \( \text{H}_2\text{O} \)

\[
\begin{array}{c}
\text{E(FCI) } = -76.43860(90) \text{ Almora-Diaz, 2014} \\
\text{E(QMC) } = -76.43894(12) \text{ this work, 2016} \\
\text{E(exp) } = -76.4389 \text{ expt. derived value}
\end{array}
\]

## 2. Pseudopotentials — \( \text{C}_2 \)

<table>
<thead>
<tr>
<th></th>
<th>( \text{C}_2 ), cc-pVTZ and VTZ-BFD</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{C} ) (a.u.)</td>
<td>( \text{C}_2 ) (a.u.)</td>
<td>AE (kcal/mol)</td>
</tr>
<tr>
<td>Hartree-Fock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all-e</td>
<td>-37.6867</td>
<td>-75.4015</td>
<td>17.6</td>
</tr>
<tr>
<td>pseudo-</td>
<td>-5.3290</td>
<td>-10.6880</td>
<td>18.8</td>
</tr>
<tr>
<td>CIPSI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all-e</td>
<td>-37.7810</td>
<td>-75.7852</td>
<td>140.1</td>
</tr>
<tr>
<td>pseudo-</td>
<td>-5.4280</td>
<td>-11.0800</td>
<td>140.6</td>
</tr>
<tr>
<td>DMC-HF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all-e</td>
<td>-37.8293(1)</td>
<td>-75.8597(3)</td>
<td>126.3(2)</td>
</tr>
<tr>
<td>pseudo-</td>
<td>-5.4167(1)</td>
<td>-11.0362(3)</td>
<td>127.2(2)</td>
</tr>
<tr>
<td>DMC-CIPSI, ( \epsilon = 10^{-6} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all-e</td>
<td>-37.8431(2)</td>
<td>-75.9166(2)</td>
<td>144.6(2)</td>
</tr>
<tr>
<td>pseudo-</td>
<td>-5.4334(1)</td>
<td>-11.0969(3)</td>
<td>144.3(2)</td>
</tr>
<tr>
<td>Estimated exact AE</td>
<td></td>
<td></td>
<td>147( \pm 2 )</td>
</tr>
</tbody>
</table>
2. Pseudopotentials — $F_2$

![Graph showing E (a.u.) vs R (Angstroms) for $F_2$, quasi-FCI / DMC with lines for All-electron cc-pVDZ and Pseudopotential VDZ-BFD.](attachment:image.png)
2. Pseudopotentials — $F_2$

![Graph showing the_comparison of F2, quasi-FCI / DMC, All-electron cc-pVTZ, and Pseudopotential VTZ-BFD](image)

- E (a.u.)
- R (Angstroms)
If the method used to generate $\Psi_T$ is not size-consistent, the DMC will lose the property of additivity of the energies.

This effect is responsible for some of the *Localization Error*. 
2. Pseudopotentials — Discussion

If the method used to generate $\Psi_T$ is not size-consistent, the DMC will lose the property of additivity of the energies.

This effect is responsible for some of the Localization Error. A scheme for very large systems (low-quality Jastrow + Kohn-Sham determinant)

- Kohn-Sham determinant
- Use a Jastrow to reduce the fluctuations
- Localize the pseudopotentials only on the determinantal component
- The FN-DMC energy will no longer depend on the Jastrow, similarly to all-electron calculations
- Additivity of energies is conserved in $H$ whatever the Jastrow
- No quadrature points to compute
3. Dissociation of FeS

Recent work of Mood and Lüchow\textsuperscript{5}:

- **Experiment**: $^5\Delta$ ground state
- **DFT**: mostly $^5\Delta$, some $^5\Sigma$, Post-HF: $^5\Sigma$
- **Single determinant DMC** gives inaccurate results
- **Full optimization** (Jastrow, CI, MOs) + FN-DMC confirms the $^5\Delta$ ground state
- **What do we get with CIPSI?**
- **What do we get with CIPSI+DMC?**

\textsuperscript{5}Full WF optimization with QMC and its effect on the dissociation energy of FeS, K. H. Mood, A. Lüchow, ArXiv[physics.chem-ph]: 1702.06535v3
3. Dissociation of FeS

FeS, CIPSI(8M dets)+PT2 / VTZ-ANO-BFD

E (a.u.) vs. r (Angstroms)
3. Dissociation of FeS

FeS, FN-DMC / CIPSI(8M dets) / VTZ-ANO-BFD

$\Delta$ 500 dets CASSCF
$\Sigma$ 500 dets CASSCF
$\Delta$ 18 000 dets FCI
$\Sigma$ 18 000 dets FCI
$\Delta$ 235 000 dets FCI
$\Sigma$ 235 000 dets FCI
$\Sigma$ 1 000 000 dets FCI
$\Delta$ 1 000 000 dets FCI
3. Dissociation of FeS

<table>
<thead>
<tr>
<th></th>
<th>Mood/Lüchow</th>
<th>This work</th>
<th>Exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energies</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^5\Sigma$</td>
<td>-134.0571(4)</td>
<td>-134.0638(4)</td>
<td>-134.0696(6)</td>
</tr>
<tr>
<td>$^5\Delta$</td>
<td>-134.0579(4)</td>
<td>-134.0642(4)</td>
<td>-134.0714(6)</td>
</tr>
<tr>
<td>Fe</td>
<td>-123.8126(4)</td>
<td>-123.8321(4)</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>-10.1314(1)</td>
<td>-10.1325(3)</td>
<td></td>
</tr>
<tr>
<td>$D_0$ (eV)</td>
<td>3.159(15)</td>
<td>2.770(17)</td>
<td>2.965(21)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.25(15)</td>
</tr>
<tr>
<td>Frequencies</td>
<td></td>
<td>18 000</td>
<td>235 000 dets</td>
</tr>
<tr>
<td>$^5\Sigma$</td>
<td>518(7)</td>
<td>558(25)</td>
<td>535(38)</td>
</tr>
<tr>
<td>$^5\Delta$</td>
<td>499(11)</td>
<td>542(37)</td>
<td>544(31)</td>
</tr>
<tr>
<td>Equilibrium</td>
<td></td>
<td>18 000</td>
<td>235 000 dets</td>
</tr>
<tr>
<td>$^5\Sigma$</td>
<td>2.00(1)</td>
<td>1.9882(3)</td>
<td>1.9933(4)</td>
</tr>
<tr>
<td>$^5\Delta$</td>
<td>2.031(7)</td>
<td>2.0124(2)</td>
<td>1.9909(7)</td>
</tr>
</tbody>
</table>

Frequencies are given in dets.
Summary

1. quasi-Full-CI wave functions:
   - DMC can be used as a “post-Full-CI” method
   - Black-box: no need to define a CAS
   - Results are reproducible and well defined
   - FCI is orbital-invariant and size-consistent
   - Fixed-node error is systematically improvable (Water)
   - Sensitivity to the basis set: atomization energies are always underestimated because atoms are described better than molecules, but less severely than in FCI
2. Pseudopotentials:

- No numerical integration needed
- No need for T-moves
- Size-consistent wave functions: localization error behaves well (F\(_2\), FeS)
3. What can we do to reduce the dependence on the basis set?

- **CIPSI+Jastrow** (E. Giner et al). See Anouar’s talk on Friday.
- Another old idea: basis functions at the middle of bonds
- CI with range-separated DFT (E. Giner). LDA can play the role of the Jastrow. Preliminary results on F\textsubscript{2} increase the atomization energy by 2 kcal/mol:

<table>
<thead>
<tr>
<th>R(Å)</th>
<th>CIPSI</th>
<th>LDA/CIPSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>-199.437(1)</td>
<td>-199.437(1)</td>
</tr>
<tr>
<td>1.4</td>
<td>-199.4867(6)</td>
<td>-199.4899(6)</td>
</tr>
</tbody>
</table>

- Combining Selected CI with \( f_{12} \) (P. F. Loos): very promising results
- These schemes are applicable to any post-HF method
Perspectives

4. What can we do for larger systems?

Quasi-Full-CI wave functions are too large. We can apply the selected CI algorithm to:

- **CAS+SD**: not size-consistent but very good nodes (careful with pseudos)
- **MR-CCSD(T) / MR-(SC)2**: PhD. thesis of Y. Garniron
- **DD-CI**: Excited states, magnetic systems
- **B_k** method method (Davidson et al 1981): Dress the Hamiltonian with the PT2 \( \overset{\Rightarrow}{\rightarrow} \) diagonalize in the presence of the \( |\alpha\rangle \)
- **JM-Heff-PT2**: Effective Hamiltonian with a size-consistent MR-PT2 scheme (Giner, JCP 2017)
- **Any other post-HF method**

\(^6\) Alternative definition of excitation amplitudes in state-specific MRCC, Y. Garniron, E. Giner, J.-P. Malrieu, A. Scemama *J. Chem. Phys.*, **146**:15,
All calculations performed with our codes (open source, GitHub):

- Quantum Package (Selected CI)
- QMC=Chem

Acknowledgments:

- Ken Jordan, Anouar Benali: invitation to the workshop
- Arne Lüchow: discussions on FeS
- J.-P. Malrieu: collaborations on wave function methods
- GENCI + CALMIP: Computational resources