Recent progress in explicitly correlated electronic structure methods

Fred Manby
Centre for Computational Chemistry, School of Chemistry
University of Bristol, Bristol BS8 1TS, UK

09–11/09/04 — LCC2004 Torino
Outline

• Introduction

• Local R12 methods

• Approximations in R12 methods
Slow convergence in orbital methods

• Correlation consistent basis sets cc-pV\(_{n}Z\), \(n = D,T,Q,5,6\ldots\)

• Basis set errors in correlation energies \(\varepsilon = \mathcal{O}(n^{-3})\)

• Size of sets is \(m = \mathcal{O}(n^3)\)

• \textit{Ab initio} methods cost \(t = \mathcal{O}(m^4)\)

• Hence \(\varepsilon = \mathcal{O}(t^{-1/4})\)

• Asymptotically 10 000 times harder for one order of magnitude

Explicit correlation
Hylleraas CI

- Hylleraas had tried CI with hydrogenic functions for He
- Frustrated by *slow convergence*
- Introduced new coordinate system

\[
s = r_1 + r_2 \quad t = r_1 - r_2 \quad u = r_{12}
\]

- Hylleraas found even small expansions very accurate: \( e^{-\alpha s}(1 + \lambda u) \)
Hylleraas and orbital CI compared

Log errors in helium energies vs. number of configurations

\[ E_{\text{exact}} = -2.903724377034 \ldots \]
MP2 theory: pairs and Hylleraas functional

- MP2 correlation energy separates into pair contributions (Sinanoglu)

$$E_2 = \sum_{i \leq j} \epsilon_{ij}^{(2)}$$

- Hylleraas functional for one pair

$$H[u_{ij}] = \langle u_{ij} | \hat{f}_1 + \hat{f}_2 - \epsilon_i - \epsilon_j | u_{ij} \rangle + 2 \langle u_{ij} | r_{12}^{-1} | ij \rangle \geq \epsilon_{ij}^{(2)}$$

- Separation requires strong orthogonality

$$\int d\vec{r}_1 \psi_k(\vec{r}_1) u_{ij}(\vec{r}_1, \vec{r}_2) = 0 \quad \forall \ k \in \text{occ}$$

Explicit correlation
An explicitly correlated MP2 theory

- Explicitly correlated pair function
  \[ |u_{ij}\rangle = T_{ab}^{ij}|ab\rangle + t_{kl}^{ij}\hat{Q}_{12}r_{12}|kl\rangle \]

- Minimize Hylleraas functional to obtain correlation energy

- Strong orthogonality operator \( \hat{Q}_{12} \), for example
  \[ \hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) \quad \text{where} \quad \hat{O}_1 = |i\rangle \langle i| \]

- 3- and 4-electron integrals arise

- Avoided by the resolution of the identity
Resolution of the identity (RI)

\[
\langle ijk | r_{12} r_{23}^{-1} | klm \rangle \approx \langle ijk | r_{12} \hat{P}_2' r_{12}^{-1} | klm \rangle \\
= \langle ij | r_{12} | kp' \rangle \langle p' k | r_{12}^{-1} | lm \rangle
\]

where

\[
\hat{P}_2' = |p'\rangle \langle p'|
\]
MP2-R12 theory

- Use ansatz

\[ |u_{ij}⟩ = T_{ab}^{ij} |ab⟩ + t_{kl}^{ij} \hat{Q}_{12}r_{12} |kl⟩ \]

- Make some approximations to simplify expressions

- Use RIs to remove all remaining many-particle integrals \( \Rightarrow \)

\[ (ip|jq) \quad (ip|r_{12}|jq) \quad (ip|[\hat{t}_1, r_{12}]|jq) \]
MP2-R12 theory

Approximations:

- **GBC:** \( \hat{f}|i\rangle = \epsilon_i|i\rangle \) then \([\hat{f}_1, \hat{Q}_{12}] = 0\)

- **EBC:** \( \hat{f}|p\rangle = \epsilon_p|p\rangle \) (couplings vanish)

\[
\hat{Q}_{12}(\hat{f}_1 + \hat{f}_2)\hat{Q}_{12}r_{12}|kl\rangle = \hat{Q}_{12}(\hat{f}_1 + \hat{f}_2)r_{12}|kl\rangle
= \hat{Q}_{12}[\hat{f}_1 + \hat{f}_2, r_{12}]|kl\rangle + \hat{Q}_{12}r_{12}(\hat{f}_1 + \hat{f}_2)|kl\rangle
= \hat{Q}_{12}[\hat{f}_1 + \hat{f}_2, r_{12}]|kl\rangle + (\epsilon_k + \epsilon_l)\hat{Q}_{12}r_{12}|kl\rangle
\]

\[
[\hat{f}_1, r_{12}] = [\hat{t}_1 - \hat{K}_1, r_{12}] \approx [\hat{t}_1, r_{12}]
\]

- Resolution of identity for remaining many-electron integrals

Explicit correlation
Density fitting

• All \textit{ab initio} methods require ERIs

\[ \langle pr| r_{12}^{-1}| qs \rangle = (pq|r s) = \int d\vec{r}_1 \int d\vec{r}_2 \frac{\psi^*_p(1)\psi_q(1)\psi^*_r(2)\psi_s(2)}{r_{12}} \]

• \(|pq\rangle\) and \(|rs\rangle\) are densities

• Can describe them in a basis set

\[ |pq\rangle \approx |\tilde{pq}\rangle = D^A_{pq}|A\rangle \]

• The error \(|pq\rangle - |\tilde{pq}\rangle\) should be made small

• Then compute approximate integral

\[ (pq|r s) \approx (\tilde{pq}|\tilde{rs}) = D^A_{pq}(A|B)D^B_{rs} \]
The history of density fitting

1959 Boys and Shavitt, to avoid 3-centre 2-electron Slater integrals (for H₃)

1972 Baerends et al., to compute Coulomb energy in Xα (ie DFT)

1973 Whitten — importance of minimizing Coulomb energy of error |pq − ˜pq|

1979 Dunlap at al. — realized that Coulomb norm led to errors quadratic in |pq − ˜pq|

later Ahlrichs, Feyereisen, others
Density fitting continued

• The best criterion

\[ \Delta_{pq} = (pq \mid \tilde{p}q \mid pq \mid \tilde{p}q) \]

• Find \( D_{pq}^A \) by minimizing \( \Delta_{pq} \)

\[ D_{pq}^A = [J^{-1}]_{AB} J_{pq}^B = \bar{J}_{pq}^A \]

where

\[ J_{AB} = \int d\vec{r}_1 \int d\vec{r}_2 \frac{A(1)B(2)}{r_{12}} \quad J_{pq}^A = \int d\vec{r}_1 \int d\vec{r}_2 \frac{A(1)\psi_p^*(2)\psi_q(2)}{r_{12}} \]

• We approximate \((pq \mid rs) \approx (\tilde{p}q \mid \tilde{r}s) = \bar{J}_{pq}^A J_{rs}^A\)
Density fitting in MP2-R12

\[
(ip| r_{12}| jq) = \bar{J}_{ip}^A R_{jq}^A + R_{ip}^A \bar{J}_{jq}^A - \bar{J}_{ip}^A R_{AB} \bar{J}_{jq}^B
\]

\[
(ip| [\hat{t}_1, r_{12}]| jq) = \bar{Y}_{ip}^A R_{jq}^A + X_{ip}^A \bar{J}_{jq}^A - \bar{Y}_{ip}^A R_{AB} \bar{J}_{jq}^B
\]

where

\[
R_{AB} = (A|r_{12}|B)
\]

\[
R_{ip}^A = (A|r_{12}|ip)
\]

\[
X_{ip}^A = (A|[\hat{t}_1, r_{12}]|ip)
\]

\[
Y_{ip}^A = (A|[\hat{t}_1, r_{-12}]|ip)
\]

Great improvement in speed, essentially no loss of accuracy
**Benchmark reactions for DF-LMP2-R12**

<table>
<thead>
<tr>
<th></th>
<th>Reaction</th>
<th></th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CO + SO$_3$ → CO$_2$ + SO$_2$</td>
<td>2</td>
<td>C$_2$H$_2$ + H$_2$O → CH$_3$CHO</td>
</tr>
<tr>
<td>3</td>
<td>CO + Cl$_2$ → COCl$_2$</td>
<td>4</td>
<td>furan + H$_2$S → thiophene + H$_2$O</td>
</tr>
<tr>
<td>5</td>
<td>CO + CH$_3$OH → HCOOCH$_3$</td>
<td>6</td>
<td>CO + NH$_3$ → HCONH$_2$</td>
</tr>
<tr>
<td>7</td>
<td>CO + H$_2$O → CO$_2$ + H$_2$</td>
<td>8</td>
<td>CS$_2$ + 2 H$_2$O → CO$_2$ + 2 H$_2$S</td>
</tr>
<tr>
<td>9</td>
<td>H$_2$CCO + HCHO → C$_2$H$_4$O + CO</td>
<td>10</td>
<td>H$_2$O$_2$ + H$_2$ → 2 H$_2$O</td>
</tr>
<tr>
<td>11</td>
<td>C$_2$H$_2$ + H$_2$ → C$_2$H$_4$</td>
<td>12</td>
<td>C$_2$H$_4$ + H$_2$ → C$_2$H$_6$</td>
</tr>
<tr>
<td>13</td>
<td>HCHO + H$_2$ → CH$_3$OH</td>
<td>14</td>
<td>C$_2$H$_6$ + H$_2$ → 2 CH$_4$</td>
</tr>
<tr>
<td>15</td>
<td>CO + H$_2$ → HCHO</td>
<td>16</td>
<td>CH$_4$ + 4 H$_2$O$_2$ → CO$_2$ + 6H$_2$O</td>
</tr>
<tr>
<td>17</td>
<td>NH$_3$ + 4 H$_2$O$_2$ → HNO$_3$ + 5 H$_2$O</td>
<td>18</td>
<td>CO + H$_2$O$_2$ → CO$_2$ + H$_2$O</td>
</tr>
<tr>
<td>19</td>
<td>SO$_2$ + H$_2$O$_2$ → SO$_3$ + H$_2$O</td>
<td>20</td>
<td>HNCO + NH$_3$ → NH$_2$CONH$_2$</td>
</tr>
</tbody>
</table>

Geometries optimized at MP2/AVTZ level, cc-pVQZ fitting basis

Explicit correlation
Comparison of MP2-R12 and DF-MP2-R12

R12-corrections in kcal/mol (orbital basis=AVQZ)

MP2-R12 values from W. Klopper

Explicit correlation
Density fitted local MP2-R12 (DF-LMP2-R12) with H-J Werner

In a local orbital basis, the following (optional) approximations are possible:

- Compute the $r_{12}$ correction only for strong pairs
- Use a local AO basis for the RI-approximation:
  \[
  \sum_{pq} R_{mp'}^{kl} K_{mp'}^{ij} \approx \sum_{\mu'\nu'\in[ij]_{RI}} R_{m\mu'}^{kl} [S^{-1}]_{\mu'\nu'} K_{m\nu'}^{ij}
  \]
- Use a local fitting basis:
  \[
  K_{m\nu'}^{ij} \approx \sum_{A\in[i]_{fit}} \bar{J}_{im}^{A} J_{j\nu'}^{A}
  \]

If all three approximations are possible, linear scaling can be achieved.

Explicit correlation
Pair correlation energies as function of orbital distance

(Gly)$_2$ / cc-pVTZ

$E_{ext}$

$E_{r12}$

$E_{tot}$

0.17438 0.005948 0.00044851

strong weak distant

Explicit correlation
Convergence of $R_{12}$-correction with AO domain size

$(\text{Gly})_4$ cc-pVTZ (uncontracted)

Explicit correlation
Comparison of canonical and local MP2-R12 approximations

R12-corrections in kcal/mol

-4
-3
-2
-1
0
1
2
3 MP2-R12 correction / kcal mol

Orbital basis=AVQZ
RI-basis=AV5Z (uncontracted)

Explicit correlation
Errors of CCSD(T) reaction energies relative to experiment

Aug-cc-pVQZ

Explicit correlation
Basis set effect on MP2-R12 correlation energies

Explicit correlation
Conclusions about DF-LMP2-R12

• Very efficient MP2-R12 implementation
  ◦ Largest calculation so far: 92 atoms, 2500 basis functions

• Treating the strong pairs only accounts for $>95\%$ of the R12 correction

• Stable results only with rather large RI basis sets

• MP2-R12 gives surprisingly bad results
What is going wrong in MP2-R12?

- Resolution of the identity
- Other approximations (GBC, EBC, $\hat{K}, r_{12} \approx 0$)
- Form of first order wavefunction

Explicit correlation
Evaluation of 3-electron integrals

- Exact
- Resolution of the identity (RI) in AO basis

- RI in auxiliary basis: Klopper and Samson (JCP 2002)
- Density fitting (DF): Manby and May (to be published)
- Combination of RI and DF: Ten-no and Manby (JCP 2003)
- Quadrature: Boys and Handy, Ten-no (JCP 2004)
What is going wrong in MP2-R12?

- Resolution of the identity
- Other approximations (GBC, EBC, $[\hat{K}, r_{12}] \approx 0$)
- Form of first order wavefunction
Effect of approximations in R12

- EBC can be avoided
- \([\hat{K}, r_{12}] \approx 0\) can be avoided
- Only GBC is left
# Effect of GBC in R12

Valence correlation energies and errors in millihartree

<table>
<thead>
<tr>
<th>basis</th>
<th>‘exact’</th>
<th>Δ(EBC)</th>
<th>Δ(GBC)</th>
<th>Δ(EBC + GBC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne</td>
<td>AVDZ</td>
<td>−289.3</td>
<td>−2.2</td>
<td>−0.3</td>
</tr>
<tr>
<td></td>
<td>AVTZ</td>
<td>−307.6</td>
<td>−1.3</td>
<td>−0.2</td>
</tr>
<tr>
<td></td>
<td>AVQZ</td>
<td>−314.4</td>
<td>−1.4</td>
<td>−0.1</td>
</tr>
<tr>
<td>HF</td>
<td>AVDZ</td>
<td>−292.3</td>
<td>−1.2</td>
<td>−0.9</td>
</tr>
<tr>
<td></td>
<td>AVTZ</td>
<td>−309.5</td>
<td>−0.8</td>
<td>−0.2</td>
</tr>
<tr>
<td></td>
<td>AVQZ</td>
<td>−316.1</td>
<td>−0.5</td>
<td>+0.1</td>
</tr>
</tbody>
</table>

Thanks to Ed Valeev for these data

Explicit correlation
What is going wrong in MP2-R12?

- Resolution of the identity
- Other approximations (GBC, EBC, $[\hat{K}, r_{12}] \approx 0$)
- Form of first order wavefunction
Testing a minimal Gaussian geminal approach

- Frozen Gaussian geminal program (MP2-F12) done (JCP 2004)

\[ |u_{ij} \rangle = T_{ab}^{ij} |ab\rangle + t_{kl}^{ij} \hat{Q}_{12} f_{12} |kl\rangle \]

where

\[ f_{12} = \sum c_\mu e^{-\mu r_{12}^2} \]

- Aim to vary geminal, but can use single Gaussian as an experiment

- Following plots show energies of ansatz

\[ |u_{ij} \rangle = T_{ab}^{ij} |ab\rangle + t_{kl}^{ij} \hat{Q}_{12} e^{-\mu r_{12}^2} |kl\rangle \]
Ne atom; cc-pVTZ

Explicit correlation
Water; cc-pVTZ

Explicit correlation
Conclusions

• Many practical schemes for 3-electron integrals

• Ansätze possible without EBC and with exchange (Klopper+Samson)

• Ansätze possible without the GBC approximation (Valeev)

• Remaining error only arises from wavefunction ansatz

• Geminals with R12 technology will combine speed and accuracy
Acknowledgements

• People

  ○ Hans-Joachim Werner — DF-LMP2-R12
  ○ Andrew May — PhD student, wrote MP2-F12
  ○ Edward Valeev — GBC-free method
  ○ Seichiiro Ten-no, Robert Polly — other explicit correlation discussions

• Funding

  ○ The Royal Society
  ○ Engineering and Physical Sciences Research Council