Comparison of QMC and ab-initio methods for 8 constitutional isomers of $\text{C}_4\text{H}_6$

Martin Krupička

Department of Structure and Function of Saccharides
Slovak Academy of Sciences
Bratislava, Slovakia
Goals

- Hands-on exercise
- Find the systematic approach for solving problems of organic chemistry area
- Create cookbook for “black box” QMC calculations
- Namely the wave-function optimization

- Constitutional isomers of $\text{C}_4\text{H}_6$ contain diverse chemical functionalities
Methods

- 8 small, closed shell organic molecules
- Heat of formations from CCCBDB
- Experimental or calculated ZPVE
- Compare with DFT, CCSD(T), HF and PM6
QM part

- Geometry optimization as part of G3 procedure: MP2(Full)/6-31G(d)
- HF wf: HF/6-311+G(2df,2pd) -> QMC
- B3LYP/6-311+G(2df,2pd)
- CCSD(T)/cc-pVTZ
## Experimental data

<table>
<thead>
<tr>
<th></th>
<th>Hfg 298K</th>
<th>hfg 0K</th>
<th>ZPE exp</th>
<th>ZPE Calc</th>
<th>Correction H, calc HF/6-31G(d)</th>
<th>Eelec from HF 0K calc ZPE</th>
<th>Eelec from HF 298K corr H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,3-Butadiene</td>
<td>26.45</td>
<td>30.16</td>
<td>51.46</td>
<td>51.25</td>
<td>60.83</td>
<td>-21.09</td>
<td>-34.38</td>
</tr>
<tr>
<td>1-Butyne</td>
<td>40.02</td>
<td>43.34</td>
<td>42.92</td>
<td>51.02</td>
<td>60.71</td>
<td>-7.68</td>
<td>-20.70</td>
</tr>
<tr>
<td>Bicyclo[1.1.0]butane</td>
<td>52.60</td>
<td>--</td>
<td>52.42</td>
<td>52.12</td>
<td>61.22</td>
<td></td>
<td>-8.62</td>
</tr>
<tr>
<td>2-Butyne</td>
<td>35.17</td>
<td>38.52</td>
<td>--</td>
<td>50.74</td>
<td>60.88</td>
<td>-12.22</td>
<td>-25.71</td>
</tr>
<tr>
<td>1,2-Butadiene</td>
<td>39.32</td>
<td>42.59</td>
<td>--</td>
<td>50.56</td>
<td>60.28</td>
<td>-7.98</td>
<td>-20.96</td>
</tr>
<tr>
<td>Cyclobutene</td>
<td>37.96</td>
<td>42.10</td>
<td>52.89</td>
<td>52.20</td>
<td>61.38</td>
<td>-10.10</td>
<td>-23.42</td>
</tr>
<tr>
<td>1-Methylocyclopene</td>
<td>59.01</td>
<td>62.48</td>
<td>51.51</td>
<td>50.89</td>
<td>60.49</td>
<td>11.58</td>
<td>-1.48</td>
</tr>
<tr>
<td>Methyleneocyclopropane</td>
<td>48.57</td>
<td>52.40</td>
<td>51.56</td>
<td>51.36</td>
<td>60.68</td>
<td>1.04</td>
<td>-12.11</td>
</tr>
</tbody>
</table>

Correlation of ZPVE: 0.529
Correlation of $E_{elec}$: 0.999
Without 1-butyne: 0.934

- HoF at 298K for all species
- HoF at 0K for but bicyclobutane
- Exp. ZPVE for 6 species, one of them strange

Source: CCCBDB.nist.gov
QMC idea - KISS

- single determinant Slater-Jastrow WF
- universal (automated) and fast procedure for WF optimization
- reliable DMC calculation
Jastrow factor construction

\[ J(\{r_i\}, \{r_I\}) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} u(r_{ij}) + \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^{N} \chi_I(r_{iI}) + \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} f_I(r_{iI}, r_{jI}, r_{ij}) \]

- cutoff
- dependent upon atom type or specific for each atom
- expansion order
Slater-Jastrow wavefunction choice

Small: 6-31G(d,p)
Medium: 6-311+G(2df,2pd)
Large: aug-cc-pVQZ

\[ N_u = N_\chi = 4,6,8 \]
\[ N_f = 0,3 \]

Medium (triple-zeta) basis set
\[ N_u = N_\chi = 6 \]
\[ N_f = 2 \]
Atom type dependency of Jastrow factor

Questionable results with Jastrow factor terms separate for each atom
Moreover, not feasible for larger molecules
Cutoff

u term – expansion order 6, cutoff 5-6

chi term – expansion order 6, cutoff 5-6

Optimized cutoffs for test molecules
DMC timestep

Timestep error acceptable below 0.001

Did extrapolate_tau procedure with 3 parameters
Wavefunction optimization

1. Generated WF in G03, at HF/6-311+G(2df,2pd) level with SCF=Tight
2. Jastrow factor: truncation order 3, spin dependency uu=dd/=ud, u term: expansion order = 6, cutoff 5.0
   chi term: expansion order = 6, cutoff 5.0 for both C’s and H’s
   f term: expansion order $N_{ee} = N_{eN} = 2$ for both C’s and H’s
3. Variance optimization for linear parameters in Jastrow factor, 3 cycles, 300.000 steps, 100.000 configs, decorr. period 10 (T~1600s)
4. Variance optimization, same as previous (T~6000s) – no significant improvement
5. Variance optimization for linear parameters in Jastrow factor, 5 cycles, 1.000.000 steps, 500.000 configs, decorr. period 10 (T~10.000s) – no significant improvement
6. Parameter choice – choose lowest energy, when variance lowest within confidence interval, or other way round (almost always worked)

All calculation on 2 nodes, 8 CPU each, GbE interconnect
DMC

1. DMC, 3000 equil. steps, 50,000 steps with 10,000 workers, longer equilibration for shorter timesteps (T~140,000s)
2. Extrapolate to zero timestep
3. Extrapolation probably unnecessary, when using sufficiently low timestep (<0.001)
Problems

• 2-butyne always led to population explosion in DMC
  – Jastrow factor terms shows proper behavior
• Large basis set leads to population explosion in DMC, shorter timestep helps
Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Intercept Value</th>
<th>Intercept Standard Error</th>
<th>Slope Value</th>
<th>Slope Standard Error</th>
<th>Adj. R-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>-2.50001E-10</td>
<td>0.84786</td>
<td>1</td>
<td>--</td>
<td>0.94463</td>
</tr>
<tr>
<td>B3LYP</td>
<td>-6.24997E-10</td>
<td>0.9663</td>
<td>1</td>
<td>--</td>
<td>0.92687</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>7.49999E-10</td>
<td>0.41407</td>
<td>1</td>
<td>--</td>
<td>0.98733</td>
</tr>
<tr>
<td>G3</td>
<td>-1.375E-9</td>
<td>0.58226</td>
<td>1</td>
<td>--</td>
<td>0.97463</td>
</tr>
<tr>
<td>DMC</td>
<td>-1.28571E-9</td>
<td>0.65174</td>
<td>1</td>
<td>--</td>
<td>0.9741</td>
</tr>
</tbody>
</table>
Timings

Jastrow factor optimization (linjas)

Jastrow factor optimization

DMC run
Conclusion

- Optimized wavefunction possible to obtain by “black-box” process
- DMC electronic energies comparable with correlated ab-initio methods
- DMC feasible for routine calculations
Cookbook

- Sufficient basis set, even the small seemed to perform well,
- Polarized double zeta or triple zeta, not higher. Too large basis sets causes trouble in DMC
- Reasonable Jastrow factor ($N_u = N_{\chi} = 6$, f term significantly slows the calculation; when needed reasonable $N_f = 2$)
- Optimal cutoff lengths similar in different molecules, energy not strongly dependent on cutoff, so maybe just few steps of varmin_linjas sufficient
- DMC timestep sufficiently small, 0.0005 performed well
- DMC equilibration quite lengthy for such small timestep
Acknowledgement

• (R)evcat - MRTN-CT-2006-035866

• Science and Technology Assistance Agency under contract APVV-0607-07
• Centers of Excellence program of the Slovak Academy of Sciences (COMCHEM, Contract no. II/1/2007)

• Department colleagues

• Mike D. Towler

• You for your kind attention and peaceful discussion