Comparison of QMC and ab-initio methods for 8 constitutional isomers of C_4H_6

Martin Krupička

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

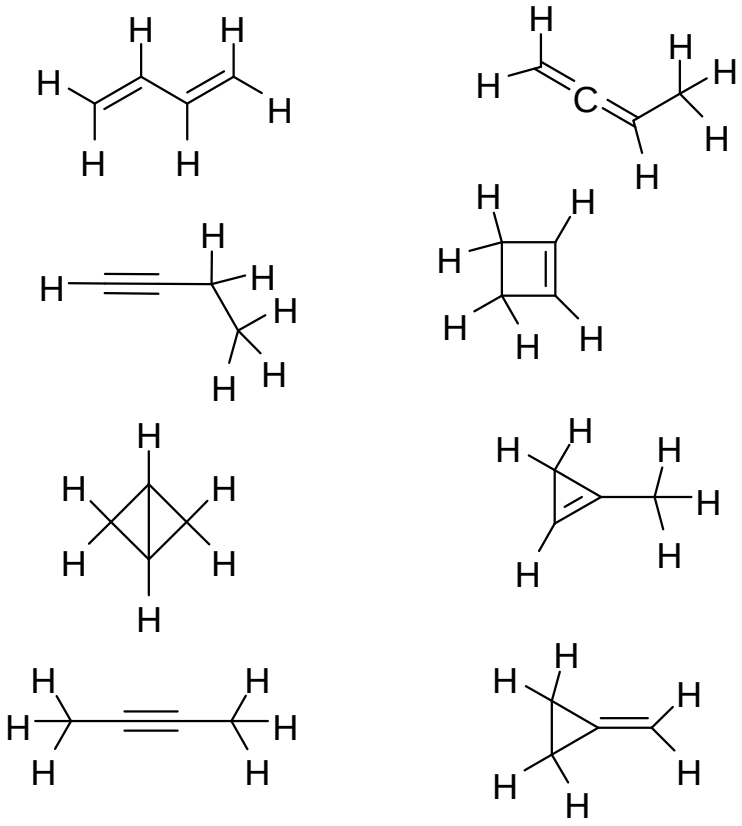
Goals

- 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 C_4H_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

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

Correlation of ZPVE: 0.529

Without 1-butyne: 0.934

Correlation of E_{elec}: 0.999

- 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

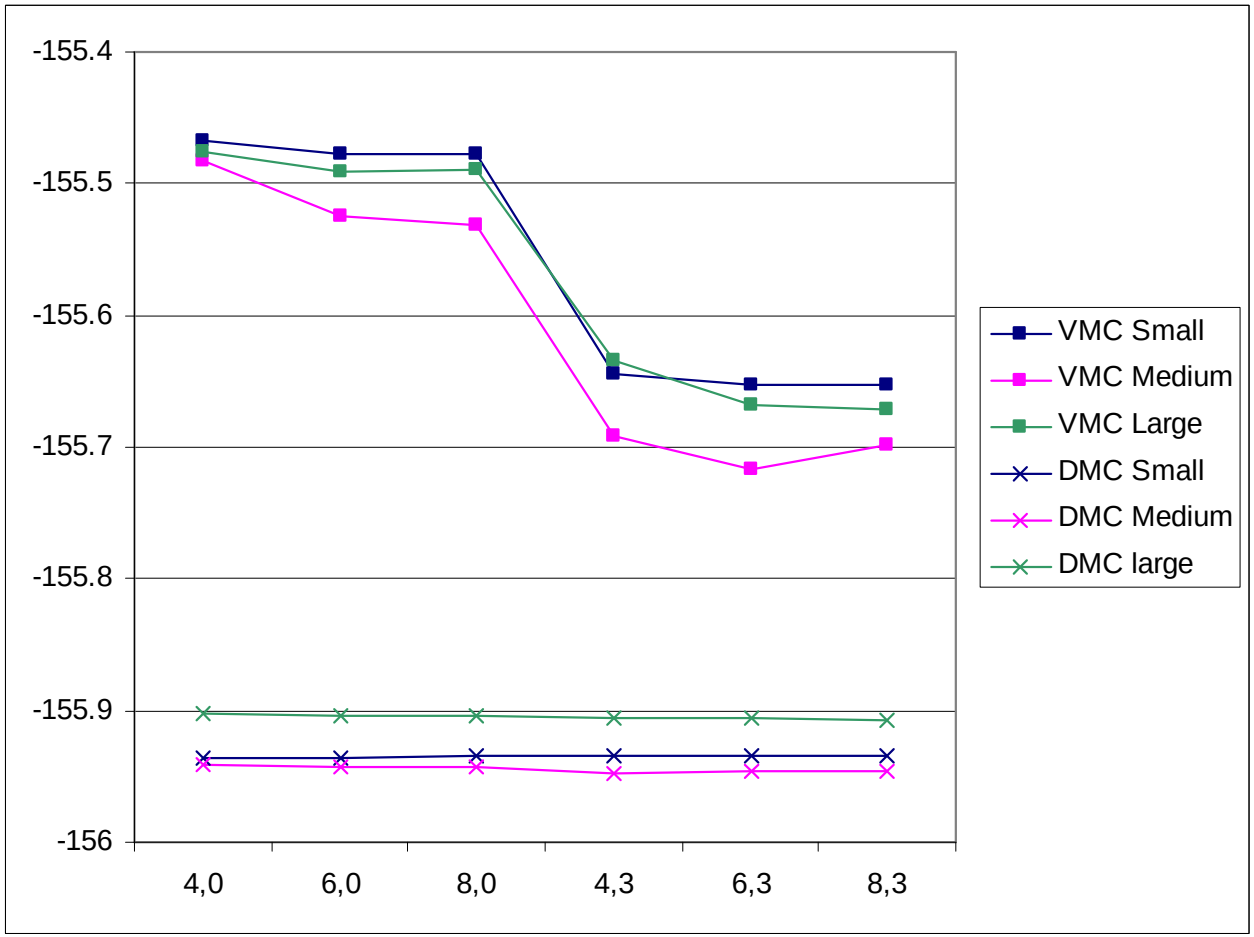
- single determinant Slater-Jastrow WF
- universal (automated) and fast procedure for WF optimization
- reliable DMC calculation

Jastrow factor construction

$$J(\{\mathbf{r}_i\},\{\mathbf{r}_I\}) = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} u(r_{ij}) + \sum_{I=1}^{N_{\text{ions}}}\sum_{i=1}^{N}\chi_I(r_{iI}) + \sum_{I=1}^{N_{\text{ions}}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} f_I(r_{iI}, r_{jI}, r_{ij})$$

- 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

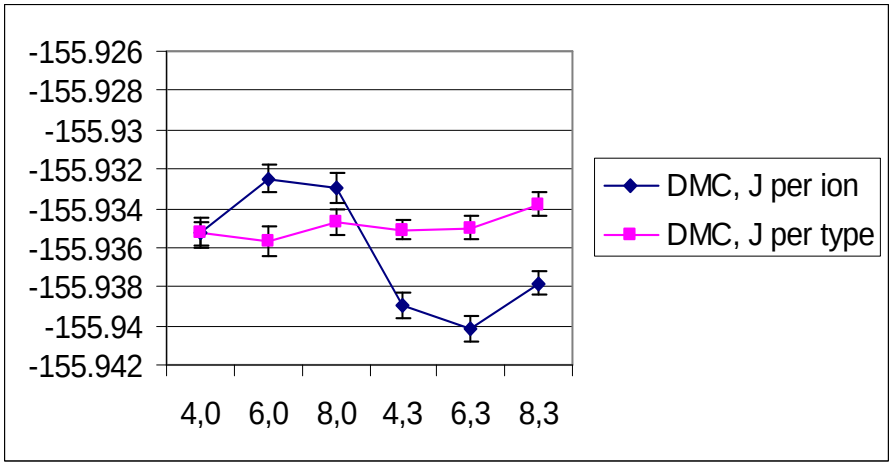
$N_u = N_\chi = 4,6,8$

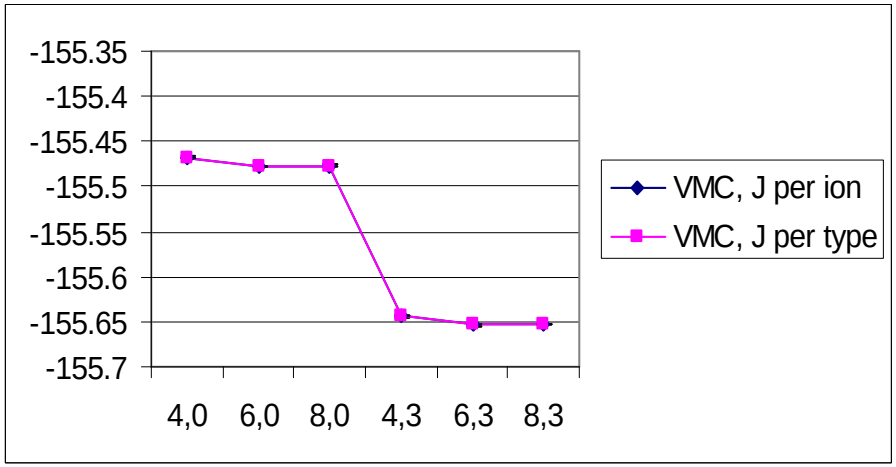
$N_f = 0,3$

Medium (triple-zeta) basis set
$N_u = N_\chi = 6$

$N_f = 2$

Atom type dependency of Jastrow factor



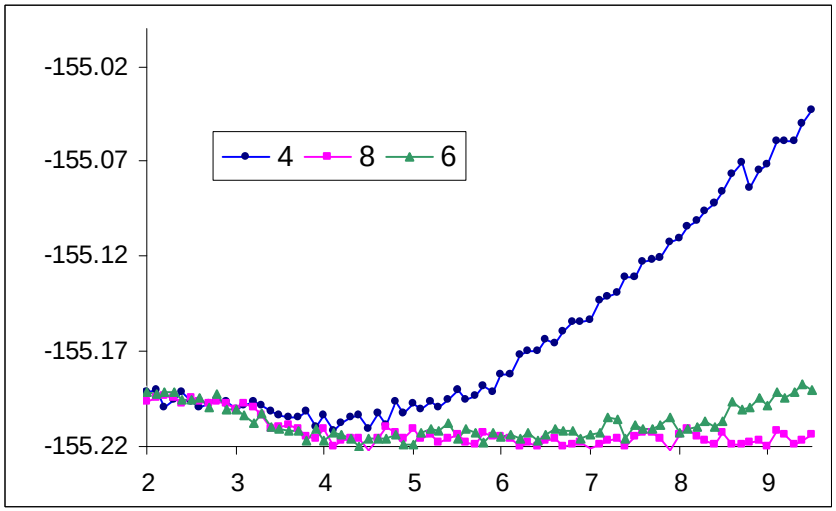


Questionable results with Jastrow factor terms separate for each atom

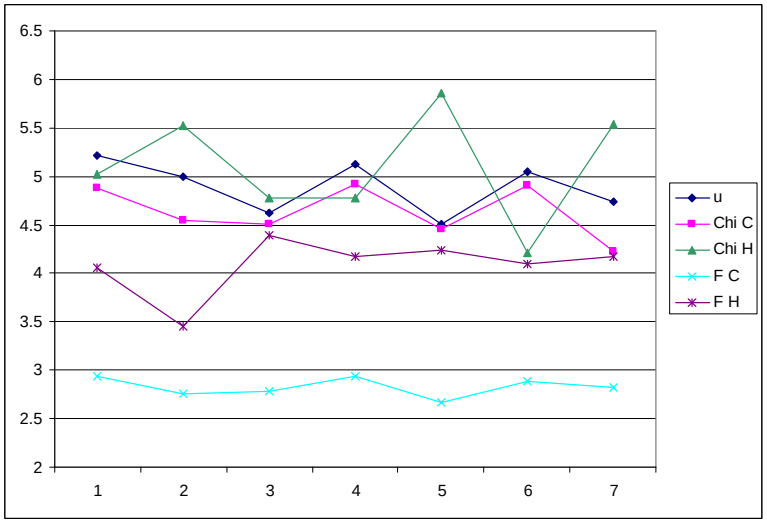
Moreover, not feasible for larger molecules

Cutoff

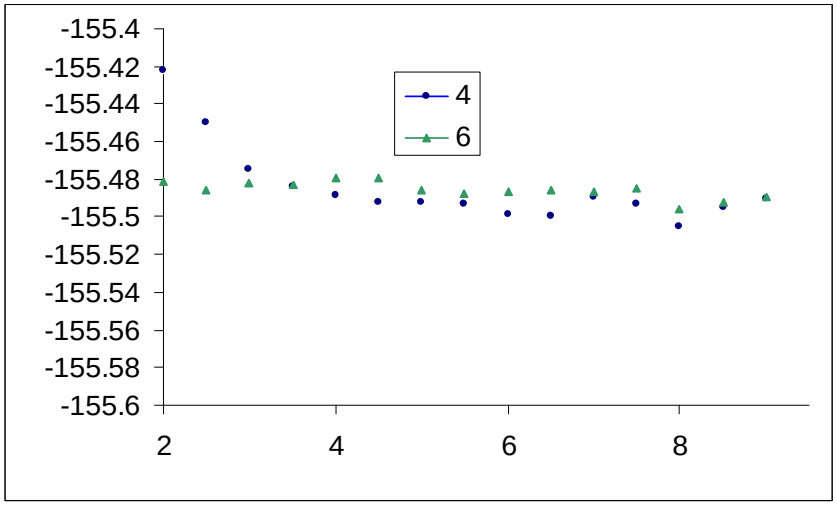
u term – expansion order 6, cutoff 5-6



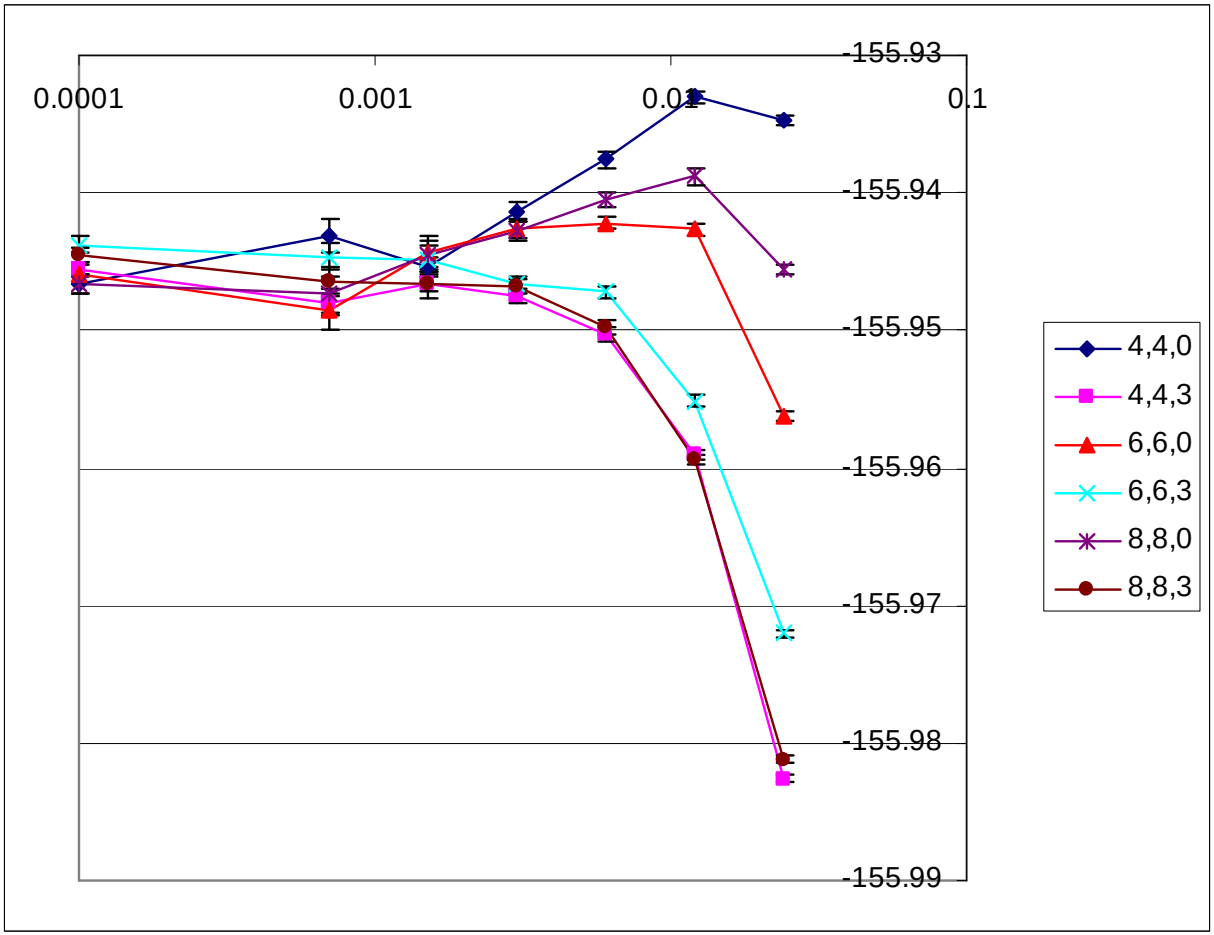
Optimized cutoffs for test molecules



chi term – expansion order 6, cutoff 5-6



DMC timestep



Timestep error acceptable below 0.001

Did extrapolate_tau procedure with 3 parameters

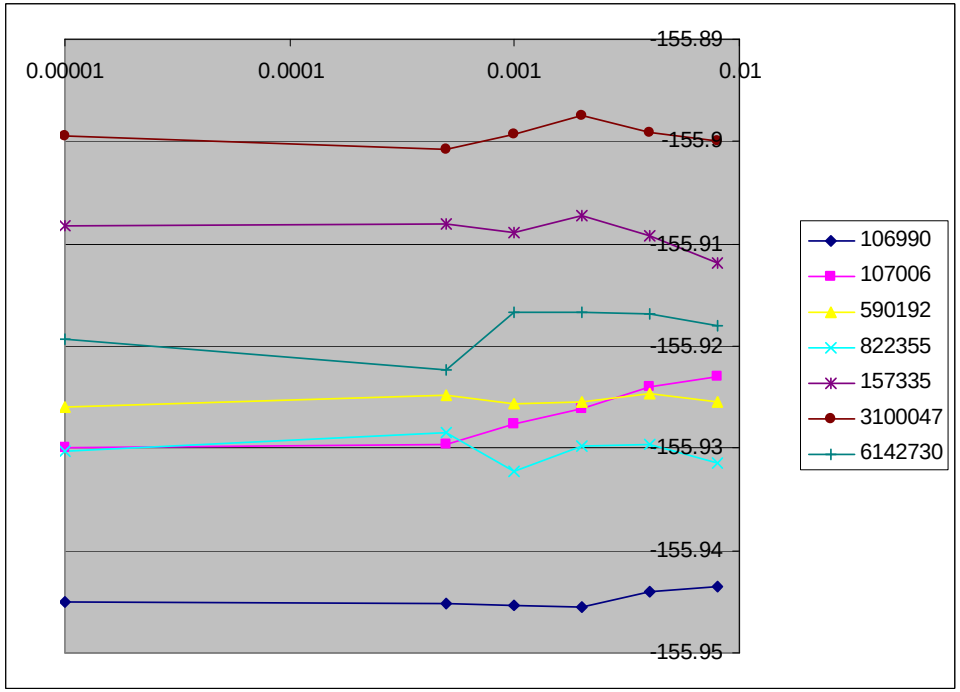
Wavefunction optimization

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

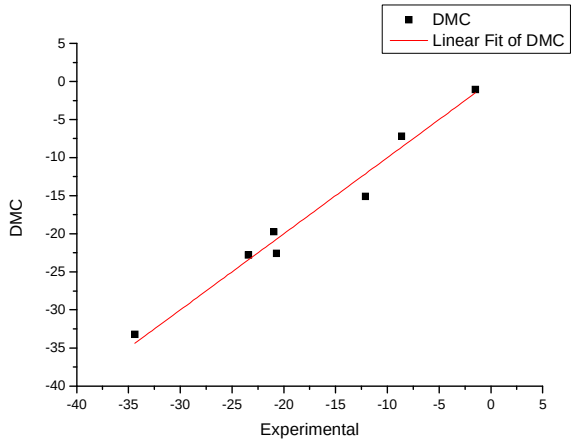
1. DMC, 3000 equil. steps, 50.000 steps with 10.000 workers, longer equilibration for shorter timesteps (T~140.000s)
2. Extrapolate to zero timestep
3. Extrapolation probably unnecessary, when using sufficiently low timestep (<0.001)

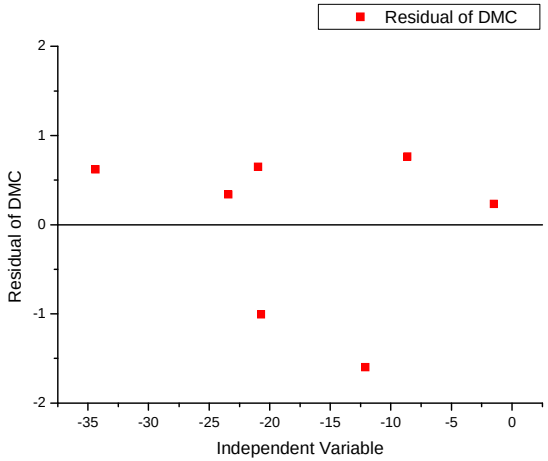


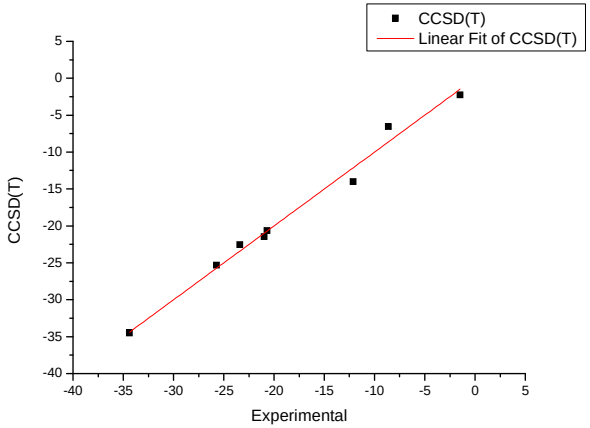
Problems

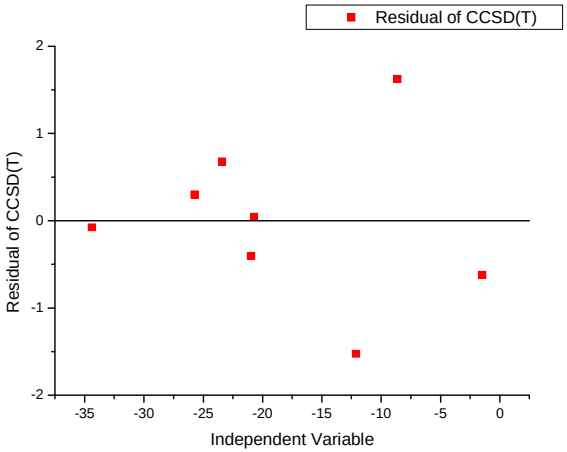
- 2-butyne always led to population explosion in DMC
  - Jastrow factor terms shows proper behavior
- Large basis set leads to population explosion in DMC, shorter timestep helps

Results





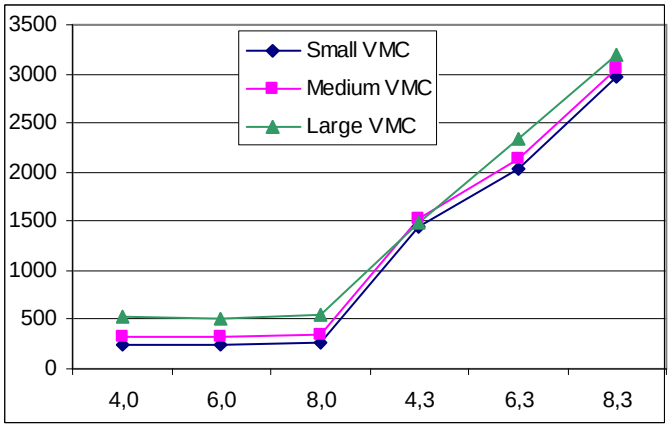




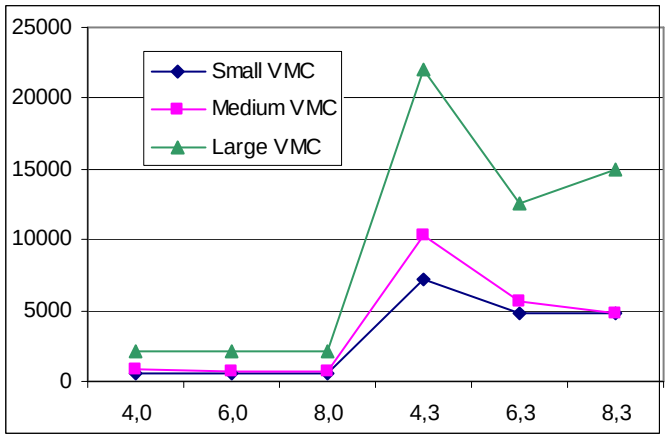
	Intercept Value	Intercept Standard Error	Slope Value	Slope Standard Error	Statistics Adj. R-Square
HF	-2.50001E-10	0.84786	1	--	0.94463
B3LYP	-6.24997E-10	0.9663	1	--	0.92687
CCSD(T)	7.49999E-10	0.41407	1	--	0.98733
G3	-1.375E-9	0.58226	1	--	0.97463
DMC	-1.28571E-9	0.65174	1	--	0.9741

Timings

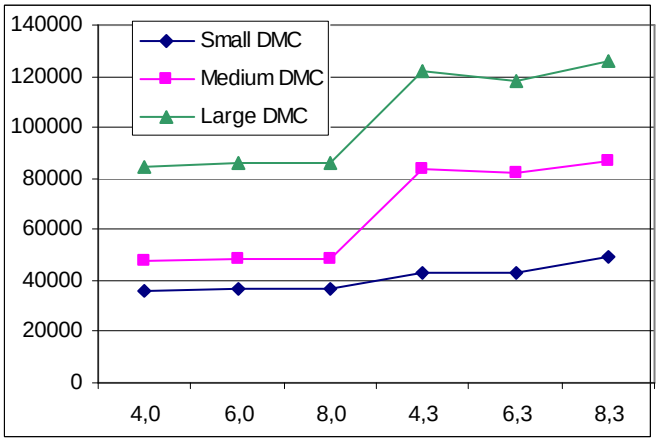
Jastrow factor optimization (linjas)



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 lenghts 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