

'Mindless' QMC Benchmarking

Martin Korth

Czech Academy of Science Prague / University of Münster

QMC in the Apuan Alps V — 30/07/2009



Outline

Quantum Monte Carlo @ Home

- Volunteer Computing for Quantum Chemistry!

Benchmarking in Quantum Chemistry

... and the new 'Mindless' Benchmarking approach

'Mindless' DFT Benchmarking

... and how to learn chemistry with random numbers

'Mindless' QMC Benchmarking

- FNDMC for Quantum Chemistry?

Part I

Quantum Monte Carlo @ Home



QMC@HOME: A distributed QMC supercomputer

Quantum Monte Carlo calculations ...

- ▶ an advantage of QMC that is becoming increasingly important with high-density (multi-core, multi-socket) and distributed (cluster, grid) computing: **massively parallel calculations**
- ▶ QMC is even suited for a very special 'flavor' of distributed computing: **Volunteer Computing** (VC)

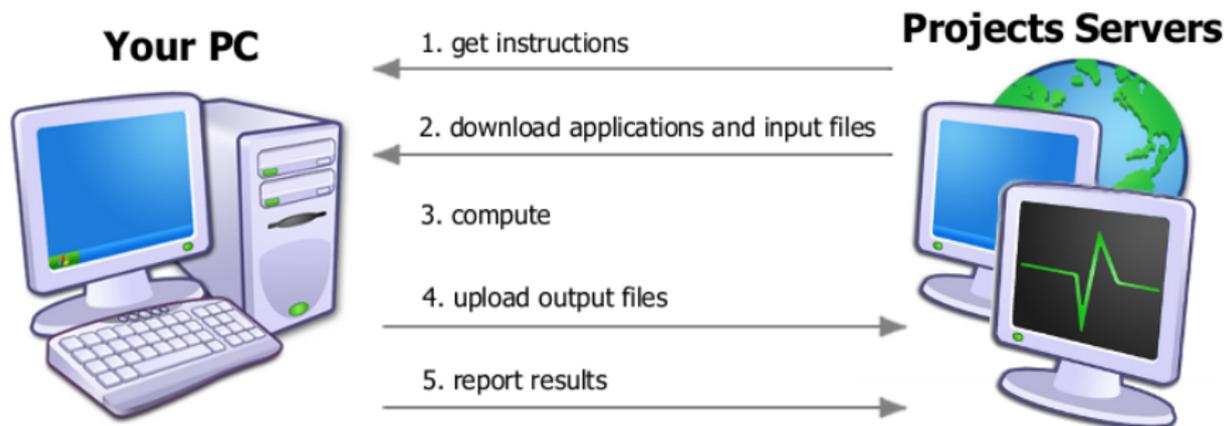
... via Volunteer Computing

- ▶ majority of the world's computing power no longer concentrated in supercomputer centers, instead distributed in hundreds of millions of personal computers
- ▶ **VC invites the public to donate computing power to science**

QMC@HOME: How does it work?

Volunteer Computing part

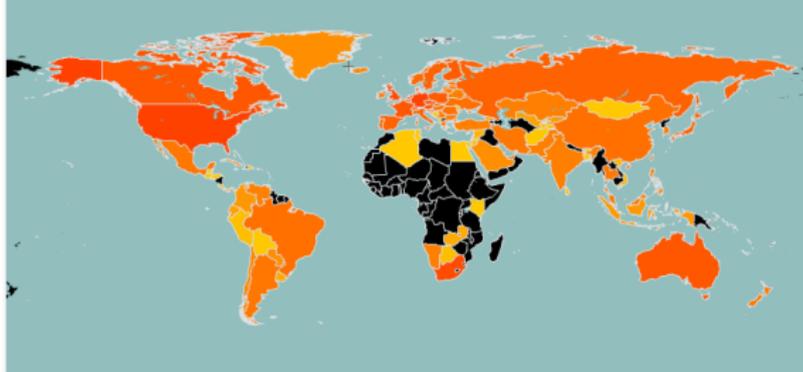
- ▶ BOINC – a software platform for Volunteer Computing



QMC part

- ▶ QAHmolqc – based on the QMC code **Amolqc** by **Arne Lüchow**

Project Statistics



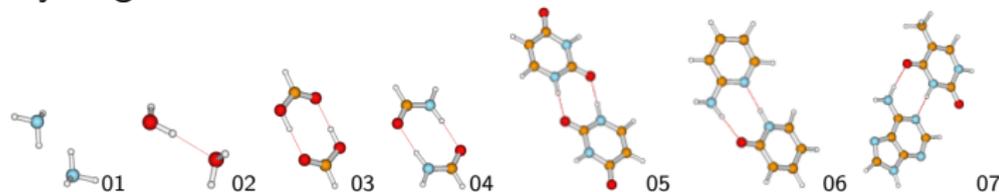
- ▶ over 63,000 registered users and over 145,000 registered hosts
 - ▶ over 11,000 highly active compute nodes
 - ▶ over **22 TeraFLOPS** average computing power
 - ▶ equivalent to rank 258 on the international top500.org supercomputer list (rank 23 on the German list)
 - you need over 3000 Xeon cores to get there!
- **A supercomputer for the price of a mid-size server system!**



In need for a few hundred processors?
Contact me for help setting up your
BOINC project!

S22 benchmark set Jurečka *et al.*, Phys. Chem. Chem. Phys., 2006, 8, 1985.

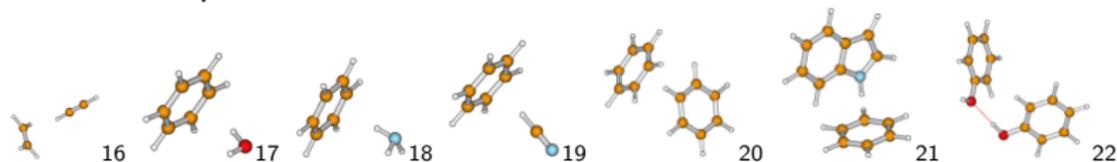
Hydrogen-bond dominated:



Dispersion dominated:



Mixed complexes:



For details see: MK/Lüchow/Grimme, J. Phys. Chem. A, **2008**, 112, 2104.

Comparison with QMC results from other groups: energetic ordering of small water clusters

Water hexamer calculations via QMC@HOME

- ▶ HF or B3LYP, BQZ(-g)/ECP, time step 0.005, 100 walkers

Water hexamer calculations from Santra et al. (JCP 2008, 129, 194111)

- ▶ B3-LYP/VTZ(-f)/ECP, time step 0.0125, 800 walkers

Results:

Isomer	HF/BQZ	B3-LYP/BQZ	Santra et al.
PRISM	0.00(23)	0.00(15)	0.00(15)
CAGE	0.13(26)	0.32(12)	0.33(14)
BOOK	0.59(22)	0.58(12)	0.57(15)
CYCLIC	1.63(28)	1.63(13)	1.54(15)

QMC@HOME: Further activities

QMC GPU project in cooperation with NVIDIA

- ▶ enable QAHmolqc for FNDMC GPU computing via QMC@HOME
- ▶ in our case: big gains even without any speedup

QMC desktop grid project with Fujitsu Technology Services

- ▶ 'burning' tests for consumer computers consume a lot of energy
- ▶ we are now able to use this energy for scientific computing

Other MEST methods

- ▶ from 'Quantum Monte Carlo @ home' to 'Quantum Mechanical Computations @ home'
- ▶ Robert Hünenbein (Grimme Group, Münster): DFT and MP2 with Orca on QMC@HOME (data parallel, e.g. for parameter sweeps)

Part II

Benchmarking in Quantum Chemistry

Why benchmarking?

The goals

- ▶ evaluate ('new' or improved) computational methods to ...
- ▶ ... judge the applicability of a method for specific cases
- ▶ ... find starting points for further improvements
- ▶ ... sell 'your' method(s)

The problems

- ▶ the availability of reference data:
the 'reference' problem (not considered here)
- ▶ the composition of the test sets:
the 'selection' problem

What about the existing benchmark sets?

'First generation'

- ▶ experimental references explicitly required → very stable molecules, a lot of redundant information
- ▶ atomization energies → the worst case scenario of bond breaking, quite far away from everyday quantum chemistry (PBE!)
- ▶ e.g. G1, G2, G3, ...

'Second generation'

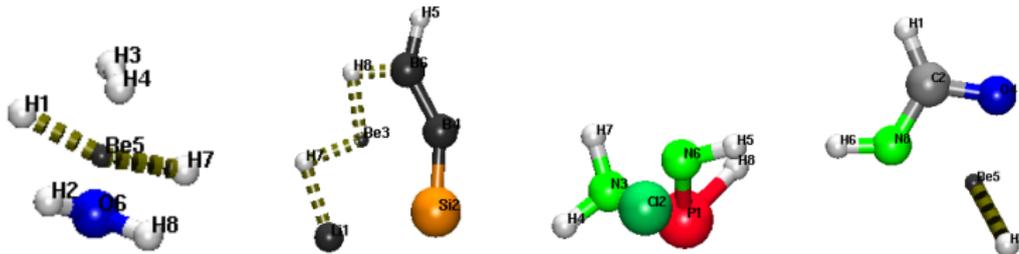
- ▶ model systems with theoretical references → limited to a very narrow structural space by chemical intuition (vs. 'chemical universe' of possible structures)
 - ▶ relative (reaction) energies → more meaningful to judge performance for 'chemical' questions
 - ▶ e.g. S22, IDCH7, ISO34, ...
- **existing benchmark sets are strongly biased**

The new approach: Diversity Oriented Benchmarking

Construct unbiased benchmark sets for thermochemistry with randomly generated molecules:

- ▶ put 8 randomly chosen atoms on the corners of a cube
- ▶ optimize this 'Artificial Molecule' (AM) with PBE-D/TZVP
- ▶ sort out e.g. multi-reference cases (via T1/D1 diagnostic)
- ▶ decompose AMs systematically into small molecules, e.g.
$$2 \text{ AM}(\text{NLiBHCHHB}) + 8 \text{ H}_2 \rightarrow 4 \text{ BH}_3 + 2 \text{ CH}_4 + 2 \text{ LiH} + 1 \text{ N}_2$$

The opening of the narrow structural space of chemical intuition produces demanding test cases in an unforeseeable manner!



The 'mindless' details

- ▶ the basic idea: generate '**Artificial Molecules**' (AMs)
- ▶ not a molecule in the classical sense, instead a randomly chosen minima on the energetic hypersurface of a random conglomerate of atoms
- ▶ BUT we have to make sure that the AMs are of use for QC:
 - ▶ choose general conditions (i.e. the constraints to randomness):
How many AMs? How many atoms per AM? What elements with what probability? What spatial arrangement? What complexity of electronic structure? What reaction scheme?
 - ▶ generate random geometries
 - ▶ check for wanted complexity
 - ▶ generate reference data
- ▶ vast amount of completely different benchmark sets can be systematically generated, characteristically depending on the countless possible answers to the above named questions

Two example benchmark sets

- ▶ both with 300 initial AMs, each with exactly 8 atoms
- ▶ main group elements up to chlorine, excluding Nobel gases
- ▶ different elemental occurrence: '9-3-1' set and 'organic' set
- ▶ PBE-D/TZVP optimization (w. large iter. limits) starting from corners of a cube with an edge length of 2 a.u.
- ▶ all systems uncharged, roughly 25 percent doublet open-shell, HOMO/LUMO-gap larger than 0.5eV and T1/D1-diagnostic (from CCSD(T)/cc-pVDZ scan) smaller than 0.02/0.10
- ▶ systematic decomposition into hydrides and diatomic molecules, only H_2 as additional reactant, whole number stoichiometry

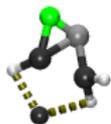
83 '931' and 84 'ORG' AMs with very diverse and unusual structures (despite being small, single-reference, main-group)

Element distributions (in percent)

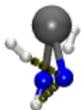
Element	MB08-931			MB08-ORG		
	Initial	Final	Ratio ^a	Initial	Final	Ratio ^a
H	69.2	64.8	0.94	61.8	58.2	0.94
Li	3.3	3.5	1.061	1.9	1.0	0.53
Be	3.3	3.6	1.091	0.5	0.0	0.00
B	3.3	4.7	1.42	1.0	1.8	1.80
C	3.3	3.9	1.18	15.5	17.6	1.14
N	3.3	3.8	1.15	3.9	4.3	1.10
O	3.3	2.9	0.88	3.9	4.5	1.15
F	3.3	3.2	0.97	1.0	1.3	1.30
Na	1.1	1.1	1.00	1.0	0.3	0.30
Mg	1.1	1.2	1.09	1.0	1.2	1.20
Al	1.1	2.0	1.82	1.0	1.5	1.50
Si	1.1	1.4	1.27	1.9	2.4	1.26
P	1.1	1.2	1.09	1.9	2.7	1.42
S	1.1	1.7	1.55	1.9	1.9	1.00
Cl	1.1	1.4	1.27	1.9	1.3	0.68

^a Ratio of final to initial distribution.

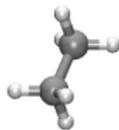
Examples from the combined MB08-165 set



a) AM(NLiBHCHHB)



b) AM(HHHHOOHNa)



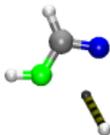
i) AM(HCHHCCHH)



j) AM(HHHHCCHH)



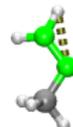
c) AM(HHCCSFBeH)



d) AM(HCHOBeHNN)



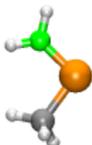
k) AM(HCHHHHHO)



l) AM(HCNHHHNN)



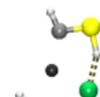
e) AM(BHNaHHHHH)



f) AM(HCSiHHHNNH)



m) AM(HHHCHAICH)



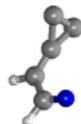
n) AM(LiHHCHCIHS)



g) AM(LiSiBeHBHBH)



h) AM(HHOHHBeHH)



o) AM(CHHCCCCO)



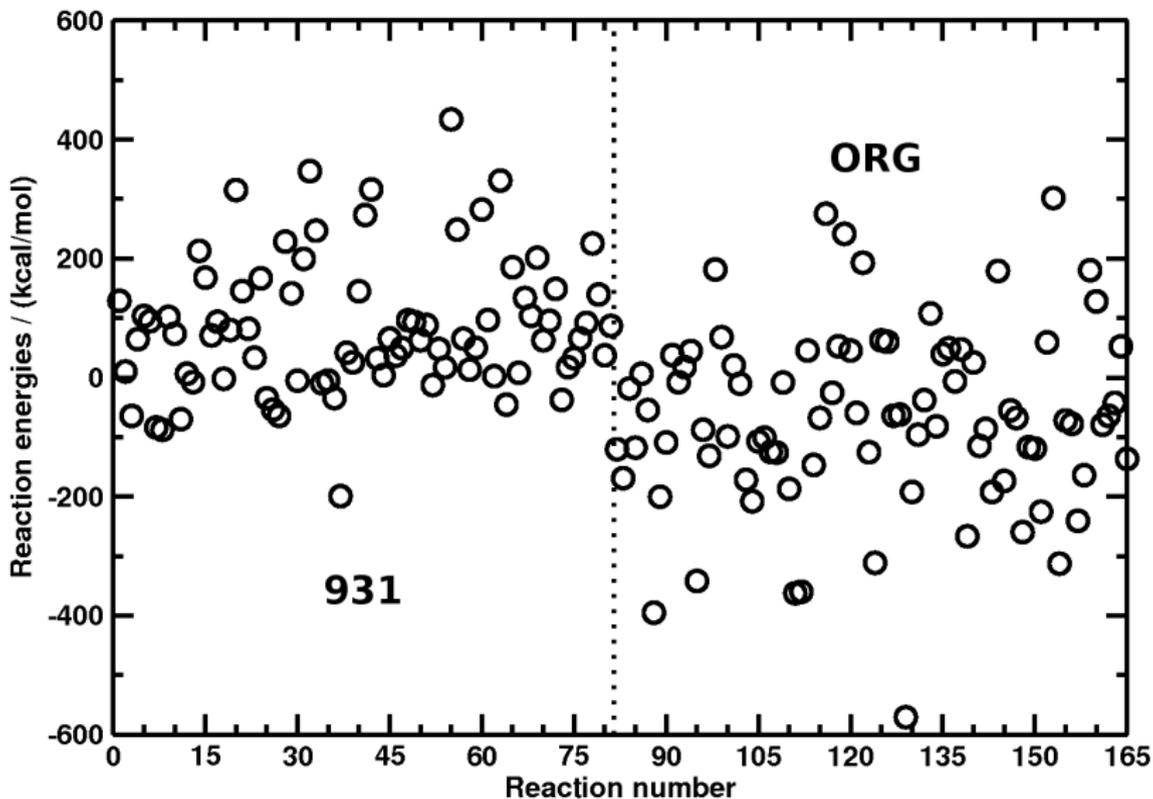
p) AM(HCHPHCHH)

More 'mindless' details

Reference values

- ▶ RCCSD(T)/CBS (extrapolated from cc-pVTZ and cc-pVQZ)
- ▶ core correlation effects estimated from difference
AE-RCCSD(T)/cc-pCVTZ and RCCSD(T)/cc-pVTZ
- ▶ core effects small (MAD 1 kcal/mol, MAX 5 kcal/mol) compared
to average reaction energy (117 kcal/mol)
- ▶ final references correspond to all-electron (AE) complete basis set
limit (CBS) CCSD(T) data
- ▶ commonly assumed to reach (sub-) 'chemical accuracy'
(1 kcal/mol) for single-reference cases

Distribution of reaction energies (est. AE-CCSD(T)/CBS)



Advantages of the MB approach

In general

- ▶ basic principle completely general, albeit every generated test set is necessarily biased
- ▶ transparency of the selection process: unperceived biases are turned into known constraints (as restrictions have to be explicitly specified)
- ▶ related work: 'random' minima search (e.g. Saunders 2004, Schleyer 2006, Needs 2006+)

The MB08-165 set

- ▶ reaction energies – chemistry!
- ▶ beyond standard fit-sets – only 'robust' methods survive (those with 'extrapolative power')!

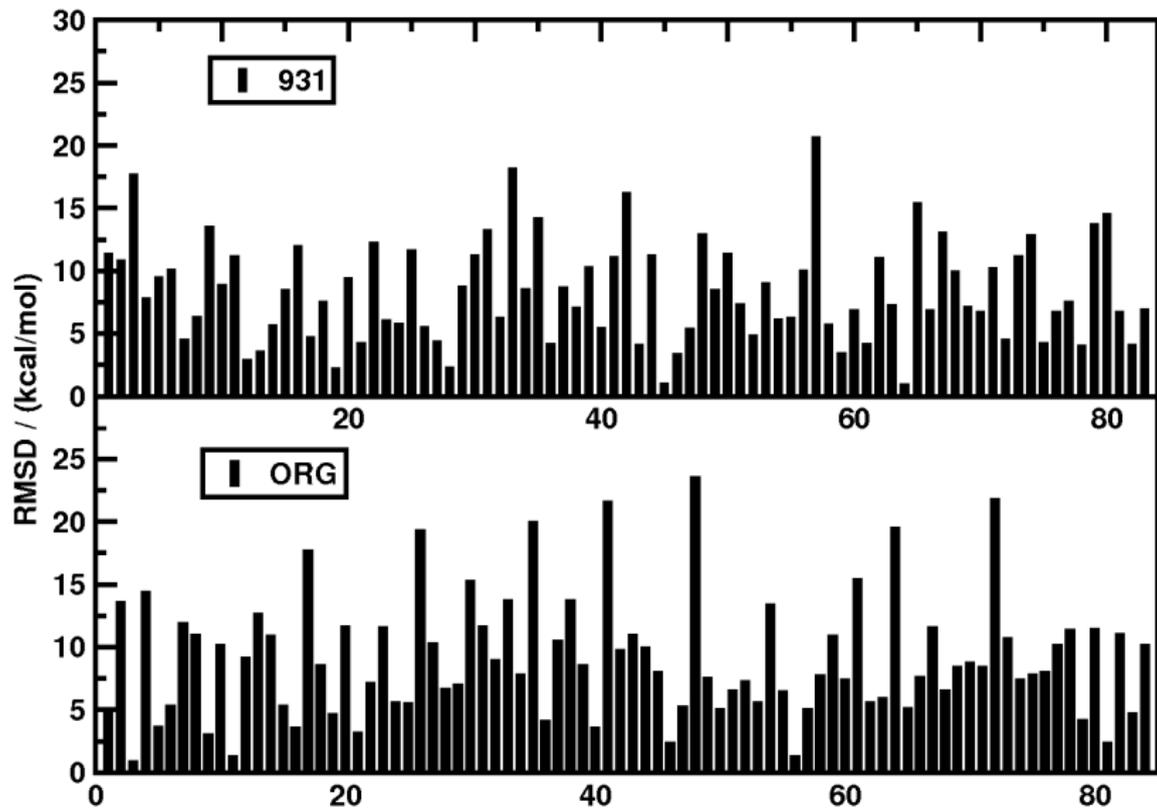
Part III

'Mindless' DFT Benchmarking

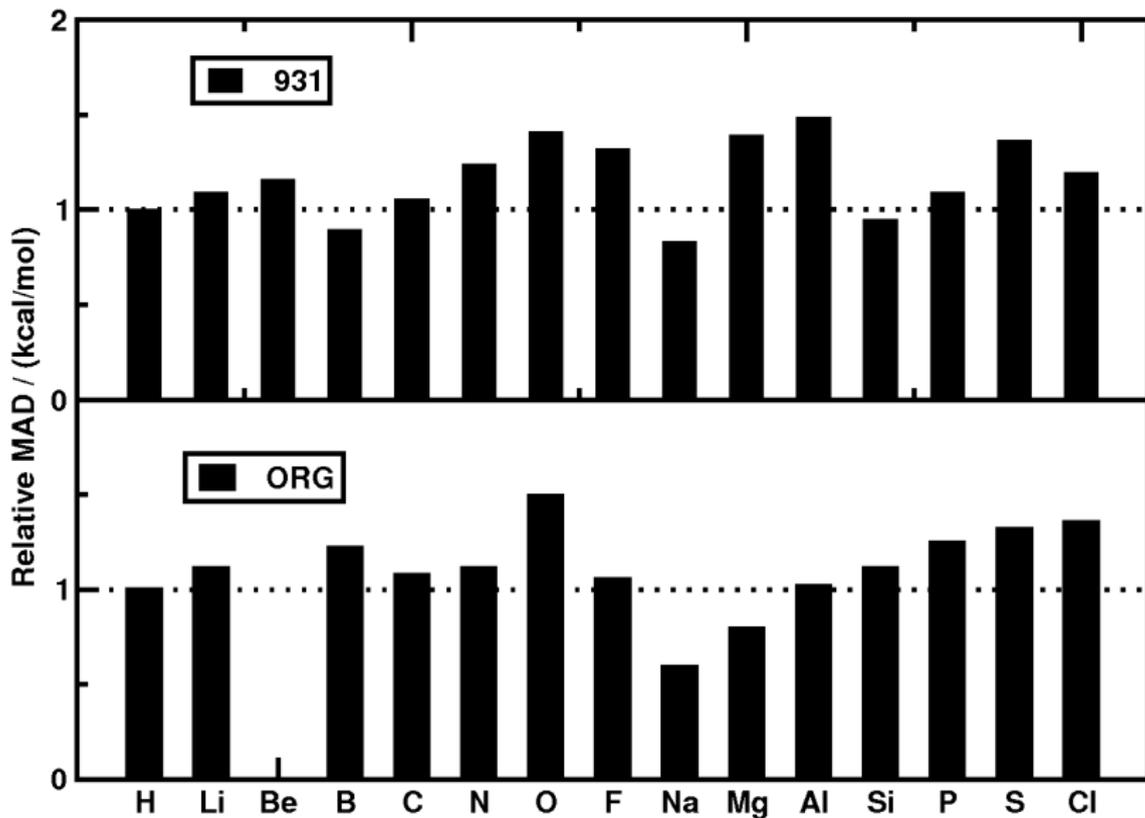
M. Korth, Stefan Grimme, JCTC 2009

DOI:10.1021/ct800511q

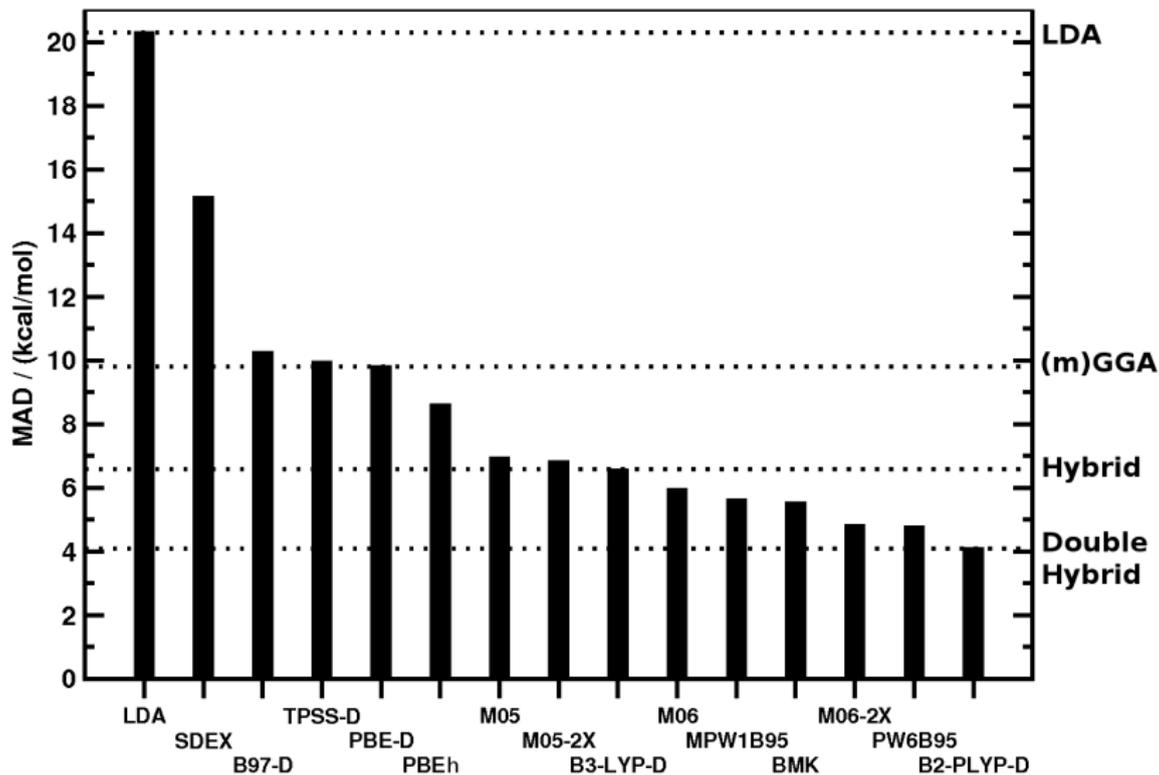
RMS over all DFT errors for each reaction



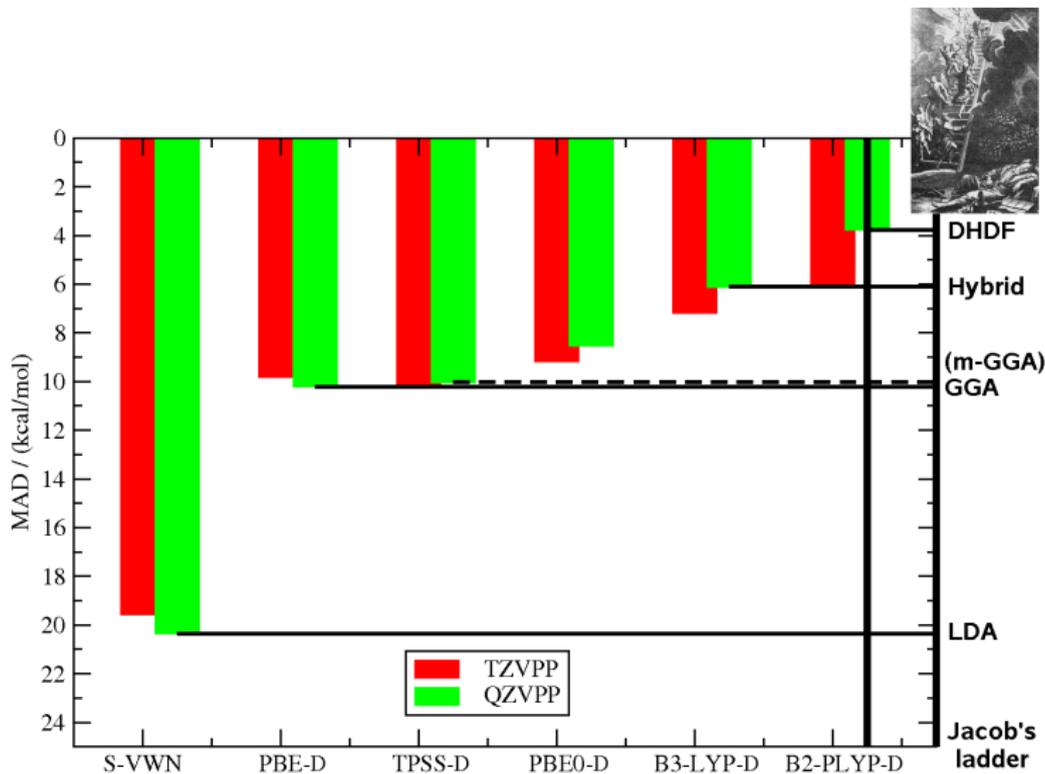
(Relative) MAD over all DFT errors for each elements



Performance of DFT methods



Perdew: Jacob's ladder of DFT development



Mindless DFT Benchmarking

- ▶ a diversity oriented approach for the preferably unbiased generation of benchmark data
- ▶ relying on systematic constraints rather than uncontrolled biases, two example benchmark sets consisting of randomly generated 'Artificial Molecules' were created
- ▶ despite restriction to small single-reference main group systems, very diverse and unusual structures were produced
- ▶ the DFT results for the MB08-165 benchmark sets perfectly fits to the Jacob's ladder of DFT development, giving further support for this metaphor
- ▶ our data assists previous claims that meta-GGAs cannot be considered a major improvement over GGAs
- ▶ state-of-the-art functionals perform significantly better than 'standard' GGAs and also Hybrids

Part IV

'Mindless' QMC Benchmarking

FNDMC for Quantum Chemistry?

Absolute energies - unbeatable!

- ▶ 'easy' to account for over 90% of the correlation energy
- ▶ several 'benchmarks' published

Atomization energies - very good!

- ▶ **Grossmann** (JCP, 2002, 117, 1434)
- ▶ G1 set / 55 Molecules
- ▶ FNDMC: MAD 2.9 kcal/mol
- ▶ CCSD(T)/aug-cc-pVQZ: MAD 2.8 kcal/mol

Reaction energies - not so clear ...

- ▶ **Manten and Lüchow** (JCP 2001, 115 , 5362)
- ▶ Test set from Helgaker group for CC / 17 entries
- ▶ FNDMC nearly as good as CCSD(T)/cc-pVTZ?

Technical details

Trial wave functions

- ▶ Slater-Jastrow type guidance functions with HF or DFT determinants and Schmidt-Moskowitz type correlation functions
- ▶ triple and quadruple- ζ (without g functions) basis sets and soft-ECPs by Burkatzki *et al.*, termed here 'BTZ' and 'BQZ'
- ▶ 'SM9' Jastrow type ($4 e_e + 3 e_n + 2 e_{en}$), parameters optimized by variance minimization

Simulation parameters

- ▶ 250-2000 work-units, each of $n \cdot 4000$ steps with an ensemble of 100 walkers and a time step of 0.005
- **not very sophisticated (but common) approach to FNDMC calculations – applicable in a 'black-box' manner**

The Helgaker test set - references

No.	reaction	cc-pVTZ	cc-pVQZ	cc-pV5Z	cc-pV(Q5)Z	experiment
1	$CH_2 + H_2 \rightarrow CH_4$	-128.3	-128.76	-128.83	-128.97	-129.9(5)
2	$C_2H_2 + H_2 \rightarrow C_2H_4$	-49.4	-49.37	-49.32	-49.23	-48.5(5)
3	$C_2H_2 + 3H_2 \rightarrow 2CH_4$	-107.7	-107.42	-107.14	-106.89	-106.5(5)
4	$N_2H_2 \rightarrow N_2 + H_2$	-42.8	-41.85	-41.41	-41.18	-41.6(0)
5	$CO + H_2 \rightarrow H_2CO$	-3.8	-4.57	-4.93	-5.16	-5.0(2)
6	$N_2 + 3H_2 \rightarrow 2NH_3$	-34.9	-37.75	-38.87	-39.46	-39.2(2)
7	$F_2 + H_2 \rightarrow 2HF$	-130.2	-134.00	-134.93	-135.47	-134.5(2)
8	$O_3 + 3H_2 \rightarrow 3H_2O$	-217.8	-223.23	-224.85	-225.77	-222.8(5)
9	$H_2CO + 2H_2 \rightarrow CH_4 + H_2O$	-57.6	-59.03	-59.51	-59.77	-60.0(2)
10	$H_2O_2 + H_2 \rightarrow 2H_2O$	-83.8	-86.34	-87.15	-87.48	-87.2(2)
11	$CO + 3H_2 \rightarrow CH_4 + H_2O$	-61.4	-63.60	-64.45	-64.93	-65.0(2)
12	$HCN + 3H_2 \rightarrow CH_4 + NH_3$	-74.8	-76.09	-76.57	-76.76	-76.4(7)
13	$HNO + 2H_2 \rightarrow H_2O + NH_3$	-102.0	-104.86	-105.89	-106.47	-106.0(2)
14	$HNC \rightarrow HCN$	-15.0	-14.78	-14.82	-14.90	-15.3(0)
15	$H_2O + F_2 \rightarrow HOF + HF$	-27.5	-28.36	-28.50	-28.67	-30.8(10)
16	$CO_2 + 4H_2 \rightarrow 2H_2O + CH_4$	-54.0	-56.85	-58.07	-58.63	-58.3(2)
17	$2CH_2 \rightarrow C_2H_4$	-198.2	-199.47	-199.83	-200.28	-201.6(7)
	MAD (to CC)	2.72	0.96	0.36	-	(0.73)
	MAD (to exp)	2.71	1.01	0.69	0.73	-

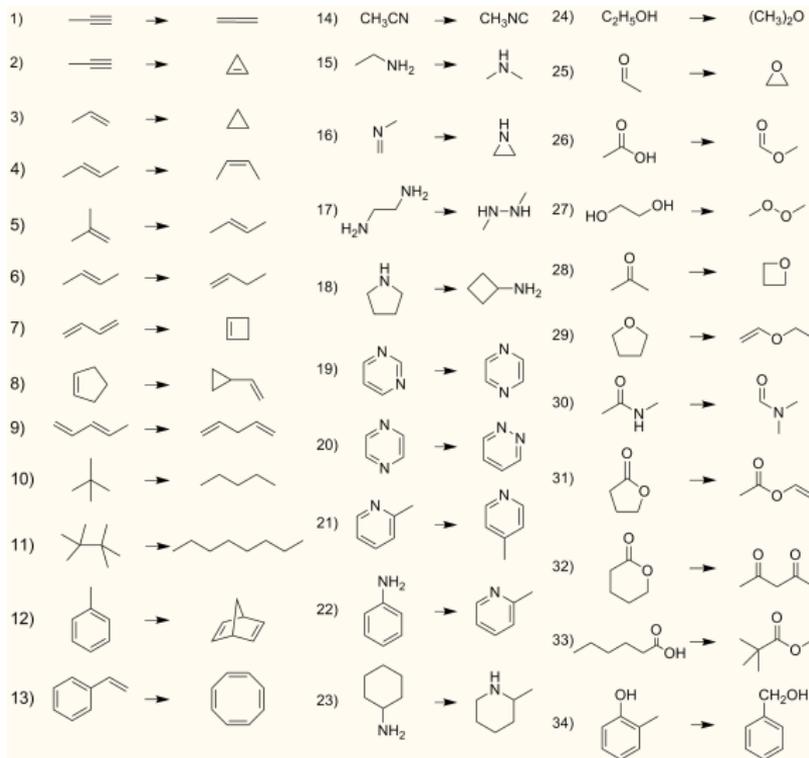
- ▶ the nice thing about Wave Function Theory methods: you clearly know which way to go ...
- ▶ cc-pV(Q5)Z fits the experiment with chemical accuracy

The Helgaker test set - FNDMC

No.	reaction	LTZ	BTZ(-f)	BTZ	BQZ	CC/(Q5)Z
1	$CH_2 + H_2 \rightarrow CH_4$	-135.0	-135.2(3)	-135.5(3)	-135.3(3)	-129.0
2	$C_2H_2 + H_2 \rightarrow C_2H_4$	-50.7	-50.3(4)	-50.4(4)	-50.5(4)	-49.2
3	$C_2H_2 + 3H_2 \rightarrow 2CH_4$	-112.6	-112.1(5)	-112.9(5)	-112.5(5)	-106.9
4	$N_2H_2 \rightarrow N_2 + H_2$	-40.7	-41.8(5)	-42.6(5)	-42.2(5)	-41.2
5	$CO + H_2 \rightarrow H_2CO$	-9.8	-9.5(5)	-8.9(5)	-9.3(5)	-5.2
6	$N_2 + 3H_2 \rightarrow 2NH_3$	-48.4	-45.1(5)	-45.2(5)	-46.0(7)	-39.5
7	$F_2 + H_2 \rightarrow 2HF$	-145.5	-143.3(6)	-144.9(5)	-144.2(5)	-135.5
(8)	$O_3 + 3H_2 \rightarrow 3H_2O$	-254.1	-247.2(7)	-248.9(7)	-247.6(7)	-225.8
9	$H_2CO + 2H_2 \rightarrow CH_4 + H_2O$	-64.4	-61.5(5)	-62.2(5)	-61.6(5)	-59.8
10	$H_2O_2 + H_2 \rightarrow 2H_2O$	-92.2	-91.3(5)	-91.5(5)	-90.9(5)	-87.5
11	$CO + 3H_2 \rightarrow CH_4 + H_2O$	-74.2	-71.0(5)	-71.1(5)	-70.9(5)	-64.9
12	$HCN + 3H_2 \rightarrow CH_4 + NH_3$	-83.9	-82.5(5)	-82.7(5)	-83.6(6)	-76.8
13	$HNO + 2H_2 \rightarrow H_2O + NH_3$	-115.8	-113.7(5)	-114.0(5)	-114.3(6)	-106.5
14	$HNC \rightarrow HCN$	-15.2	-14.7(4)	-14.7(4)	-14.9(4)	-14.9
15	$H_2O + F_2 \rightarrow HOF + HF$	-31.8	-30.5(7)	-32.0(7)	-31.7(7)	-28.7
16	$CO_2 + 4H_2 \rightarrow 2H_2O + CH_4$	-66.1	-63.1(6)	-63.0(6)	-61.7(6)	-58.6
17	$2CH_2 \rightarrow C_2H_4$	-208.1	-208.5(4)	-208.3(4)	-208.5(4)	-200.3
	MAD (to CC)	5.3	4.1	4.4	4.3	-
	MAD (to exp)	5.7	4.4	4.7	4.6	0.6

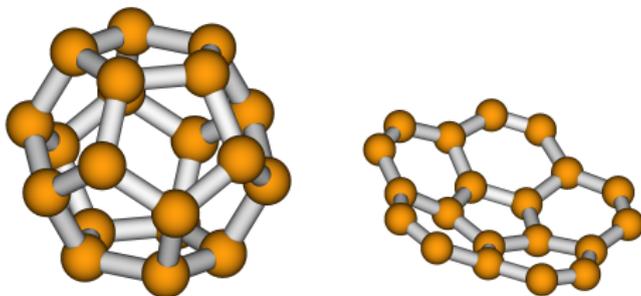
- ▶ nodes 'saturated' at TZ without f functions (for general thermochemistry)
- ▶ sizable improvement with Burkatzki *et al.* basis sets and ECPs
- ▶ FNDMC(HF/BQZ) not as accurate as CCSD(T)/cc-pVTZ

The ISO34 test set Grimme/Steimetz/Korth, JOC, 2007, 72, 2118



→ accuracy of FNDMC(HF/BQZ) comparable to MP2/TZVPP

C_{20} carbon cluster - cage and bowl isomers



Method	Bowl	Cage
MR-MP2/TZV2d1f//HF/6-31G*	0.0	29.3
MR-MP2/TZV2d1f//MP2/TZV(2d2f)	0.0	4.4
AE-FNDMC(HF-TWF)//HF/6-31G*	0.0	48.4(120)
ECP-FNDMC(HF/LQZ)//MP2/TZV(2d2f)	0.0	24.5(6)
ECP-FNDMC(HF/BQZ)//MP2/TZV(2d2f)	0.0	23.3(4)
ECP-FNDMC(DFT/LQZ)//MP2/TZV(2d2f)	0.0	20.2(6)
ECP-FNDMC(DFT/BQZ)//MP2/TZV(2d2f)	0.0	19.4(3)
CCSD(T)/CBS	0.0	11.4

- ▶ very good FNDMC results for a quite complicated case!

Benchmarking FNDMC for Quantum Chemistry

Why?

- ▶ Quantum Chemistry is not Solid State Physics
- ▶ QMC data for molecular systems is still quite limited
- ▶ no QMC equivalent of extensive DFT/WFT benchmarking
- ▶ does the existing data basis allow judgments about the performance of FNDMC for Quantum Chemistry?

What?

- ▶ use the 'unbiased' MB08-165 set
- ▶ benchmark FNDMC for thermochemistry
- ▶ first step: compare mean-field approaches, ECPs, basis sets, ...
- ▶ second step: evaluate more sophisticated FNDMC approaches

Mindless QMC Benchmarking – work in progress ...

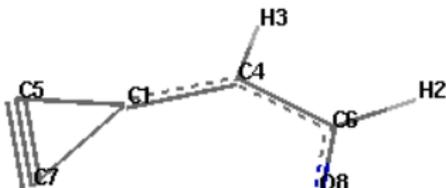
Method	MB08-165			
	MAD	MD	RMSD	Δ
FNDMC/HF-BQZ-SM9 ^a	7.0	2.0	9.5	61.8
FNDMC/DFT-BQZ-SM9 ^b	6.5	2.0	8.4	54.5
B3-LYP/QZVP	8.2	6.9	10.6	45.5
B3-LYP-D/QZVP	6.6	-2.7	8.8	49.1
CCSD(T)/cc-pVTZ	5.5	-5.3	7.0	24.6
CCSD(T)/cc-pVQZ	2.6	-2.5	3.4	11.8

[a] 161 from 165 entries [b] 159 fr. 165
statistical errors between 0.4 and 1.2 kcal/mol, on average 0.8 kcal/mol
average reaction energy of 117 kcal/mol

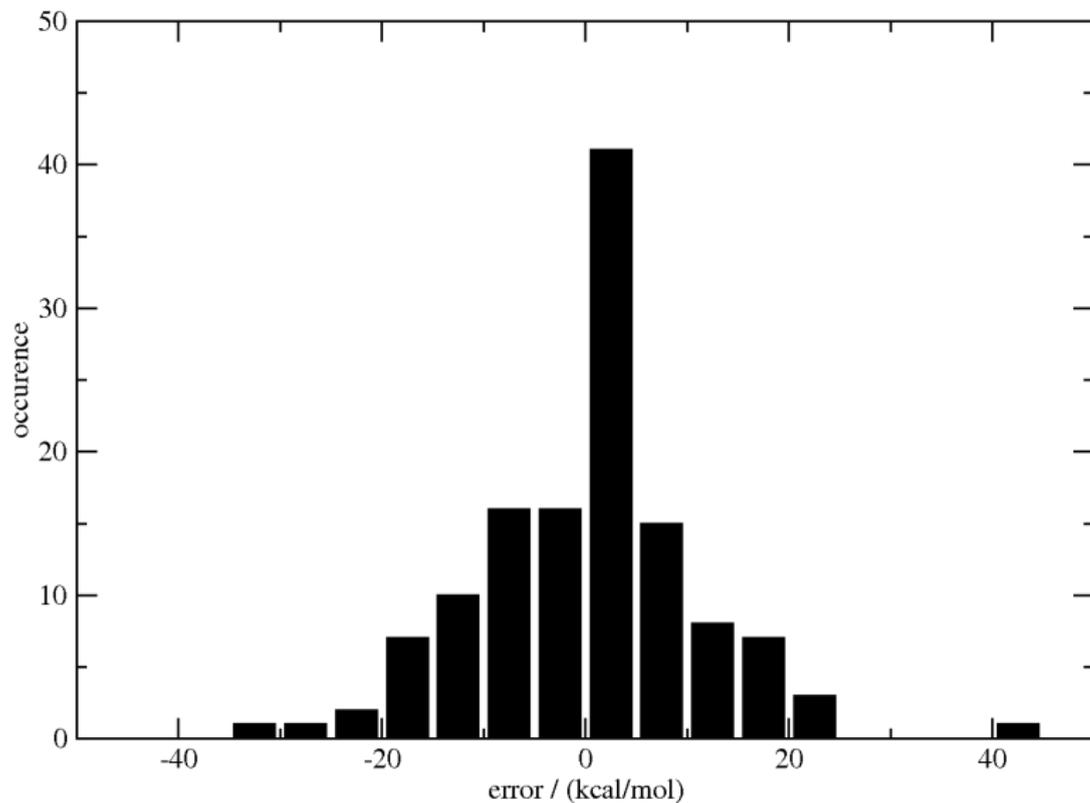
► worst case: $2 \text{ AM}_{\text{CHHCCCCO}} + 18 \text{ H}_2 \rightarrow 10 \text{ CH}_4 + \text{O}_2$

-570.6 kcal/mol AE-CCSD(T)/CBS

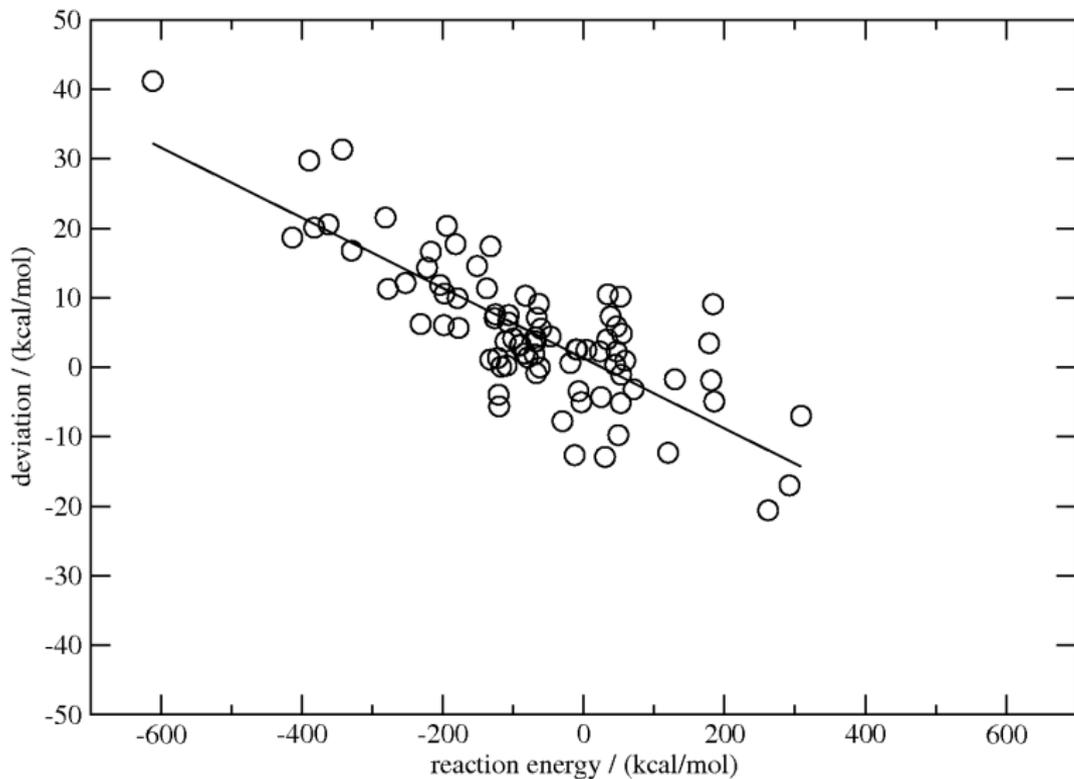
-530.2(12) kcal/mol FNDMC(HF/BQZ)



Mindless QMC Benchmarking – error distribution



Mindless QMC Benchmarking – work in progress ...



Summary

Preliminary conclusions

- ▶ nodes seem to be 'saturated' already at TZ(-f) level (for general thermochemistry)
- ▶ Burkatzki basis sets and ECPs are an improvement in most cases
- ▶ DFT-TWFs are advantageous for electronically complicated cases
- ▶ the Fixed Node Error is a serious limitation in Quantum Chemistry
- ▶ ... but seems to be quite systematic somehow ...

Further work

- ▶ extensive benchmarking of FNDMC with the MB08-165 set (mean-field approaches, basis sets, ECPs, ...)
- ▶ evaluation of more sophisticated FNDMC approaches

Acknowledgments

Cooperation partners

- ▶ Stefan Grimme, University of Münster (D)
- ▶ Arne Lüchow, RWTH Aachen University (D)
- ▶ Pavel Hobza, Czech Academy of Science Prague

Funding

- ▶ QMC@HOME was funded by the Sonderforschungsbereich 424 established by the Deutsche Forschungsgemeinschaft

And of course many thanks to ...

- ▶ ... all QMC@HOME volunteers for contributing
... **and you for your attention!**

Performance of SOTA DFT functionals

