DFT trial wave functions for molecular FNDMC calculations

Martin Korth

ESDG — 10/03/2010

FNDMC and trial wave functions

- Fixed Node Diffusion Monte Carlo (FNDMC) calculations of fermionic systems rely on nodal hypersurfaces from trial wave functions (TWF)
- this 'Fixed Node' approximation limits the final accuracy that can be obtained with FNDMC
- nevertheless often found to be very accurate even with single determinant TWFs from HF, LDA or PBE
- error cancellation for atomization and reaction energies (thermochemistry) less convincing
- Multi-determinant TWFs possible for complicated cases (but slow and oscillatory improvement reported)

DFT trial wave functions

- rather neglected topic ...
- often implicitly assumed that HF gives best possible single-determinant nodes
- LDA and PBE TWFs sometimes used, but often no significantly different performance than with HF TWFs found
- to the best of my knowledge no systematic study of DFT TWFs exists
- but even data with HF, LDA and PBE TWFs for more than one system at a time is quite hard to find
- ... in the end maybe rightfully neglected?

J. C. Grossman, *J. Chem. Phys.* **2002**, 117, 1434.

TABLE IV. Total single determinant FN-DMC energy (a.u.) of the P_2 molecule for different types of orbitals used to fill the Slater determinant.

Orbitals	Total energy
Hartree–Fock orbitals	−13.0628(1)
Natural orbitals	−13.0636(1)
LDA orbitals	−13.0652(1)
BPW91 orbitals	−13.0652(1)
B3LYP orbitals	−13.0651(1)

- big improvement for absolute energies with DFT TWFs (but very small gain for relative energies)

L. Wagner, L. Mitas *Chem. Phys. Lett.* **2003**, 370, 412.

Table 4
Binding energies (eV) by various methods

	TiO	MnO
State	$^3\Delta$	$^6\Sigma$
Hartree–Fock	2.64 [16]	−0.92 [16]
LSDA	9.12	5.88
PW91	7.45	4.79
B3LYP	6.62	3.39
VMC(HF orbitals)	6.0(1)	3.1(1)
DMC(HF orbitals)	6.3(1)	2.9(1)
DMC(B3LYP orbitals)	6.9(1)	3.4(2)
DMC(MCSCF orbitals)	6.7(2)	3.4(2)
Experiment	6.98 [17] or 6.87 (7) [19]	3.70 [17]

The DMC(MCSCF) was a multi-reference wave function with ≈20 configurations with the largest weights.

- big improvement again - rather unreliable numbers, though ...

L. Wagner, L. Mitas *Chem. Phys. Lett.* **2003**, 370, 412.

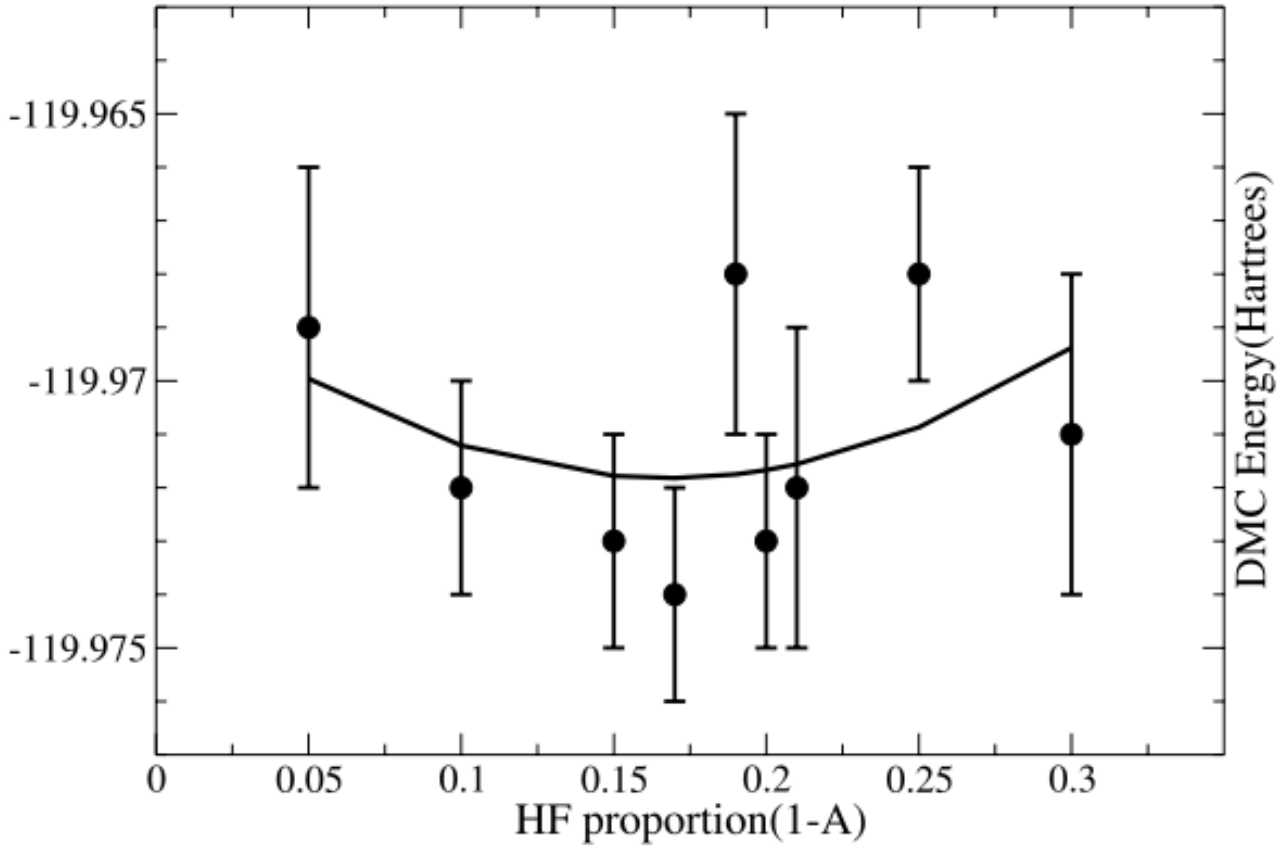


Fig. 1. MnO energy as a function of the weight of exact exchange in the B3LYP-type of GGA functional (see text). The solid line is a quadratic fit to DMC data with plotted error bars.

Reboredo/Kent *Phys. Rev. B* **2008**, 77, 245110.

A functional for nodal hypersurfaces?

- KS theory is designed for ground-state total energies and densities: The KS WF is an auxiliary quantity to generate the true ground-state density rather than an approximation to the true WF
- the non-interacting $n(r)$ is equal to the interacting $n(r)$, but no other property of the many-body ground state has to be retained
- optimal densities do not have to correspond to optimal nodes
- can we choose to retain other selected properties of the many-body ground state in a non-interacting solution?
- no results for realistic systems available so far ...

What else could be worth a try?

- meta-GGAs like TPSS (kinetic energy density τ allows to 'distinguish' between minima and maxima of the density)
- hybrid functionals like TPSSh (hybrids as the workhorse Quantum Chemistry could be advantageous at least for molecular systems)
- long-range corrected functionals like LC-ωPBE (with correct $1/r$ decay of the potential)
- HF or LHF plus TPPSc ('exact' exchange plus approximate dynamic correlation?)
- functional parametrization, e.g. with B97 ('systematic' empirical hybrid GGA allows to screen the influence of empirical DFT parameters?)

DFT TWFs for molecular FNDMC calculations

Project goals

- systematic study along Jakob's ladder with several thermochemistry test sets
- screening of empirical DFT parameters with B97
- explorative tests with alternatives like LC-wPBE

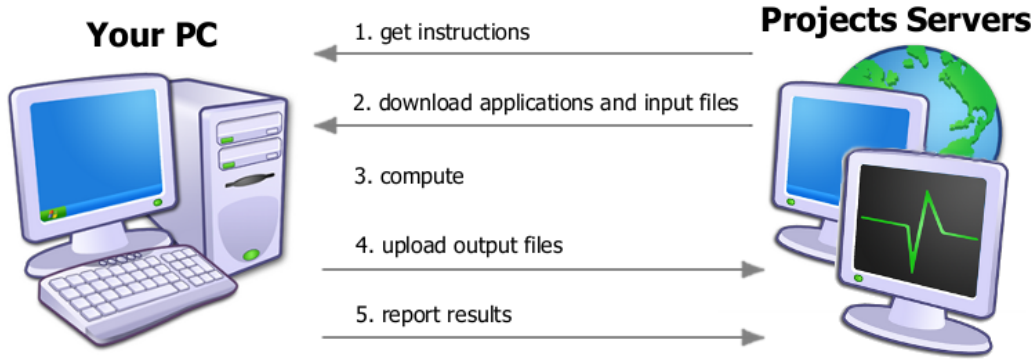
Project requirements

- 'workflow tools'
- a lot of computation time ...

... via Volunteer Computing ...

Volunteer Computing ...

- invites the public to donate spare computing time to science



... at TCM

- with QuantumFIRE alpha (QFA)

Quantum
FIRE
alpha

http://cah.tcm.phy.cam.ac.uk

How does Volunteer Computing work?

Outside

- Volunteer downloads BOINC Volunteer Computing (VC) software (standard Windows installer or Linux Bash script)
- ... types http://cah.tcm.phy.cam.ac.uk when prompted
- ... everything else works completely automatic
- ... a 'Manager' GUI gives full control if wanted

Inside

- A computer-shaped pile of second hand electronics in room 511
- ... handles clients requests and book-keeping
- ... distributes binaries and input files
- ... receives output files and status reports

CASINO for QuantumFIRE alpha

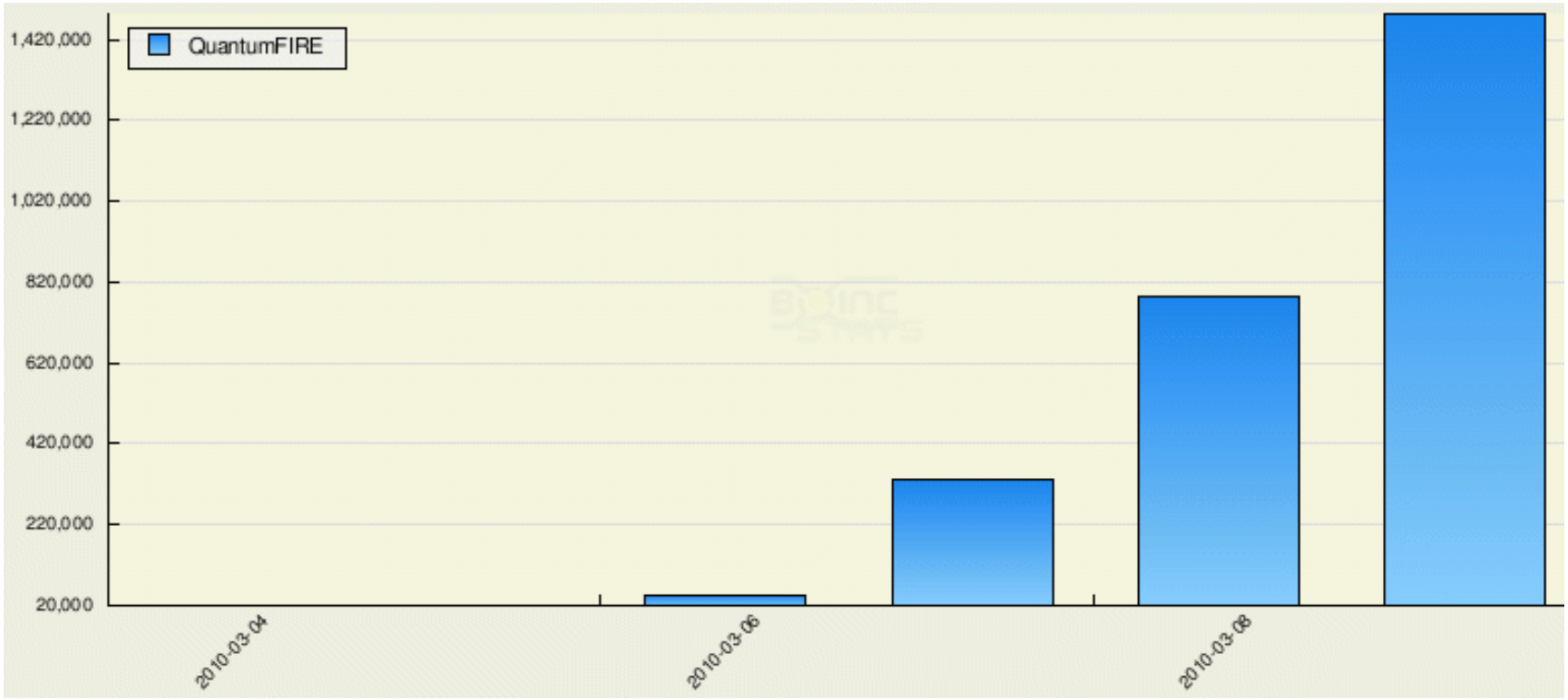
- Windows and Linux (32/64bit) binaries
- interfacing with the 'core client' (controls the application on the Volunteer's computer) included
- stable 'auto-restart' functionality (checkpoints dmc.status to get rid of dmc.history file) included
- full functionality of (non-MPI) CASINO in principle possible (though not yet realized)
- job are divided into separate calculations as 'work units' of about 5-10 hours length and returning dmc.status files are afterwards evaluated job-wise on the server

Workflow tools for QuantumFIRE alpha

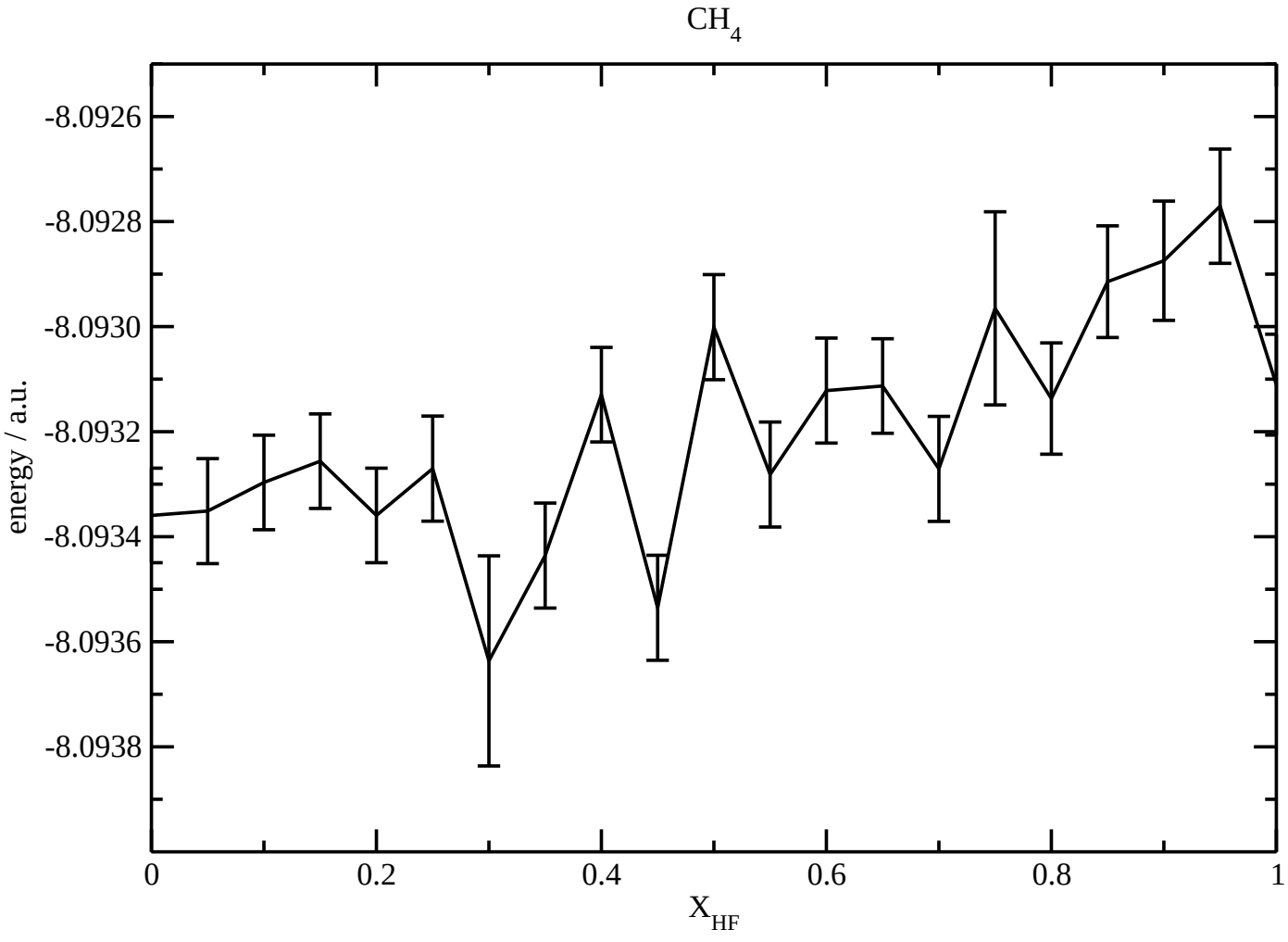
- Turbomole: a fast and flexible HF/MP2/DFT GTO code for molecular systems (http://www.turbomole.com)
- **TurboTools** to allow for a large number of QMC calculations based on Turbomole TWFs without too much human effort
  - wave function, pseudo potential converter
  - Jastrow, Backflow input generator
  - *auto_ve* (pick best Jastrow), *de* (merge dmc.status files), *auto_ex* (do τ extrapolation), *ctime* (estimate run time)
  - *prep_tm*, *prep_cas*, *prep_cah* to prepare Turbomole, CASINO-vmc_opt and VC calculations
- *prep_tm -r && prep_cas -q && prep_cah* generates everything that is needed for QFA starting from a number of geometry input files ... but no proper error handling yet
- MolproTools? Molpro is not as fast but could supply multi-determinant wave functions ...

Project status

- QFA launched 05/03/2010 (last Friday ...)
- 1103 registered users
- around 450 'highly active' compute nodes
- donating about 0.25 TeraFLOPS average computing power

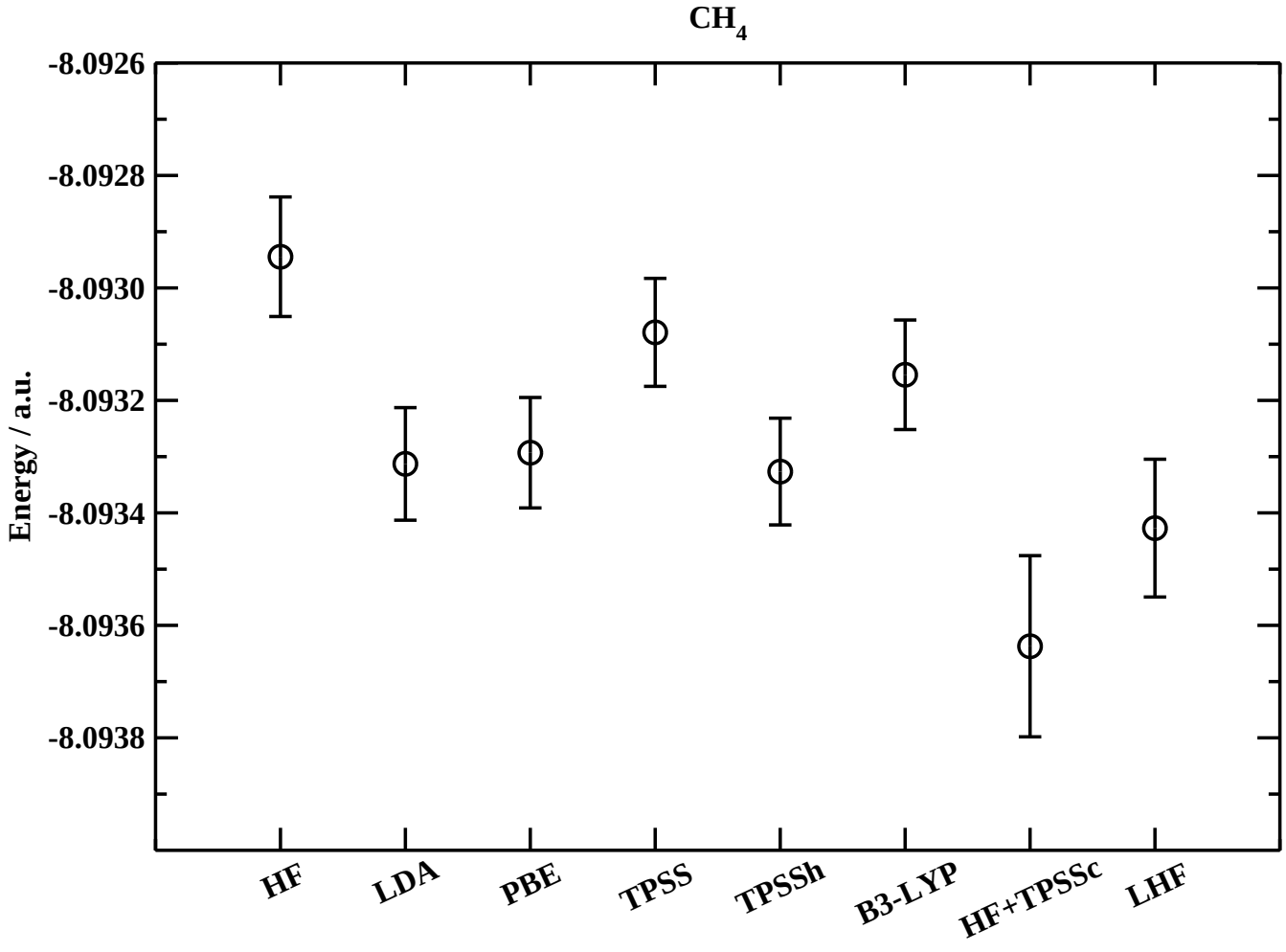


Example data



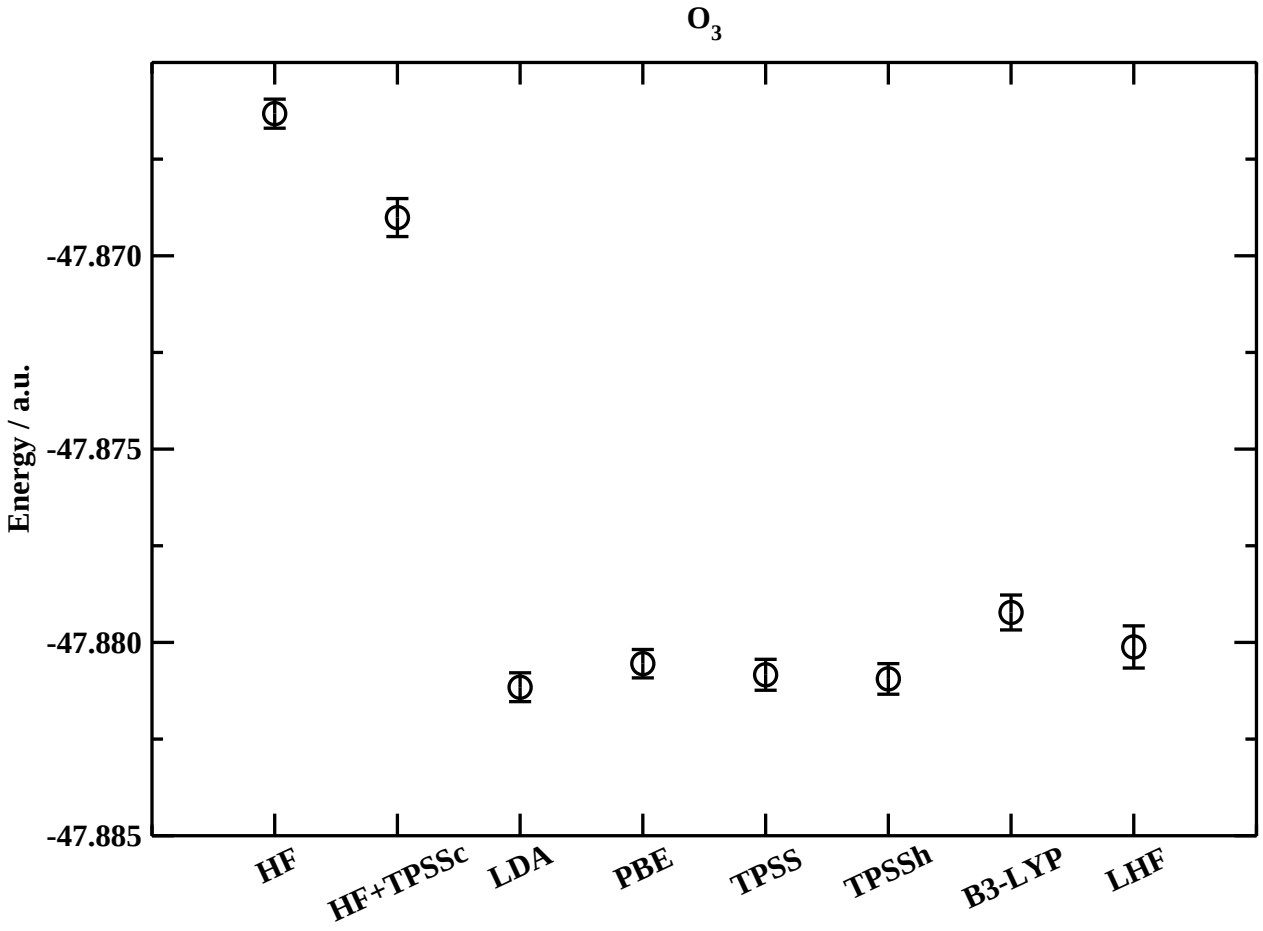
- FNDMC energy of CH4 as function of HF exchange in B97

Example data -2-



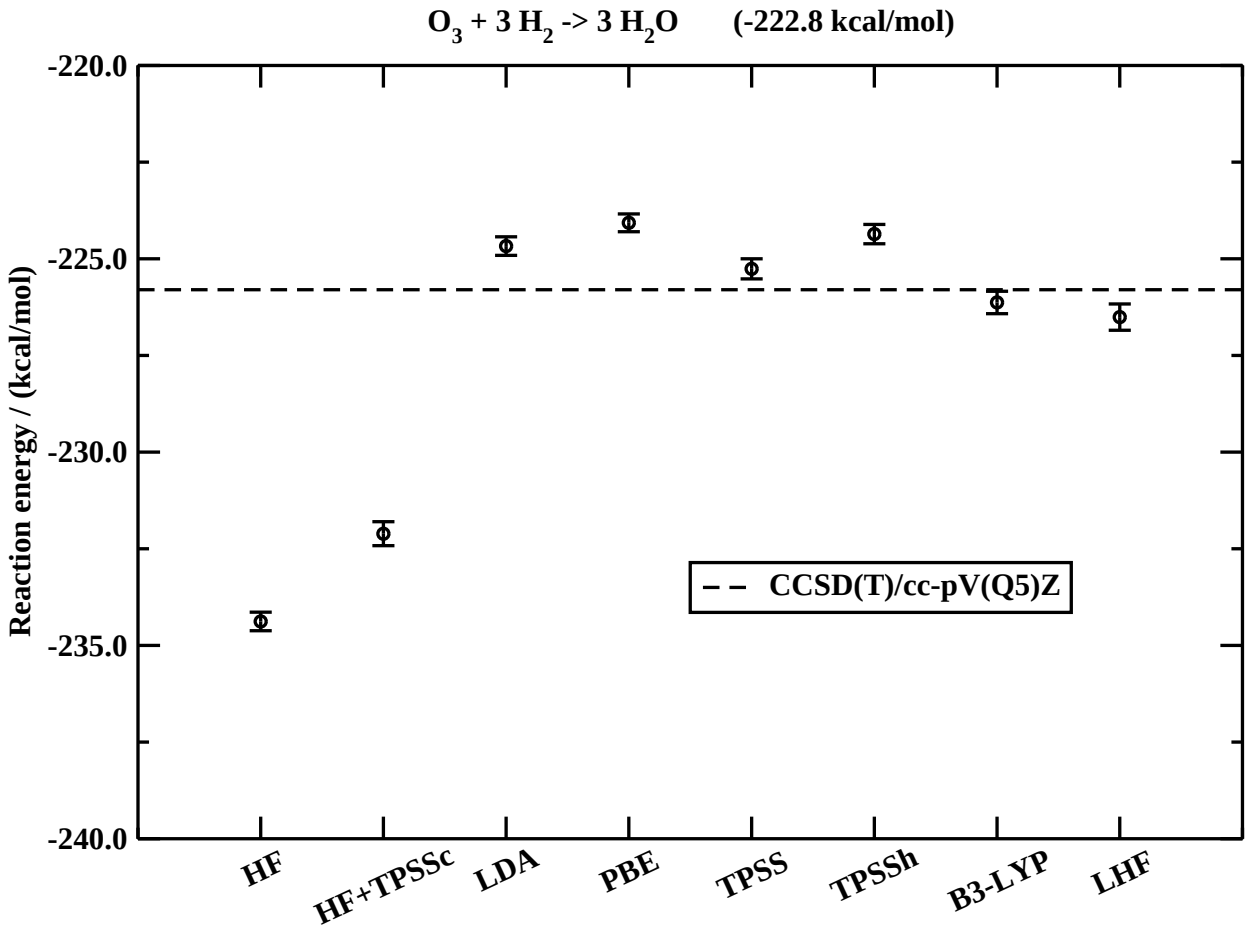
- FNDMC energy of CH_4 with different TWFs

Example data -3-



- FNDMC energy of O_3 with different TWFs

Example data -4-



- FNDMC reaction energy with different TWFs

Summary

Conclusions

- Density Functional Theory trial wave functions for Fixed Node Diffusion Monte Carlo calculations are a rather neglected topic
- QuantumFIRE alpha might shed some light on it

Outlook

- Louis on QuantumFIRE!
- Solid state QMC applications?
- Abinit on QuantumFIRE?

Acknowledgments

Money

- German National Academy of Sciences Leopoldina, Leopoldina Research Fellowship LPDS 2009-18

http://cah.tcm.phy.cam.ac.uk

People

- Mike Towler (TCM)
- David Anderson (BOINC)
- Kathryn Marks (QFA forum moderator)

DFT

Density Functional Theory

Basics

- ground state density $n(r)$ determines full Hamiltonian
- minimal principle for the energy as a functional of $n(r)$

$$E[n(r)] \geq E[n_0(r)] = E_0 \tag{1}$$

$$E[n(r)] = \int v(r)n(r)dr + F[n(r)] \tag{2}$$

$$F[n(r)] = T[n(r)] + \frac{1}{2}\int \frac{n(r)n(r')}{|r-r'|}drdr' + E_{XC}[n(r)] \tag{3}$$

- cast into the self-consistent Kohn-Sham equations

$$\left(v(r) - \frac{1}{2}\nabla^2 + \int \frac{n(r')}{|r-r'|}dr' + v_{XC}(r) - e_j\right)\psi_j(r) = 0 \tag{4}$$

$$n(r) = \sum_{j=1}^{N} |\psi_j(r)|^2 \tag{5}$$

$$v_{XC}(r) = \frac{\delta E_{XC}[n(r)]}{\delta n(r)} \tag{6}$$

Density Functional Theory

Approximations

- ground state energy

$$E = \sum e_j - \frac{1}{2}\int \frac{n(r)n(r')}{|r-r'|}drdr' - \int v_{XC}(r)n(r)dr + E_{XC}[n(r)] \tag{7}$$

- approximations for $E_{XC}[n(r)]$

$$E_{XC}^{LDA}[n(r)] = \int \epsilon_{XC}(n(r))n(r)dr \tag{8}$$

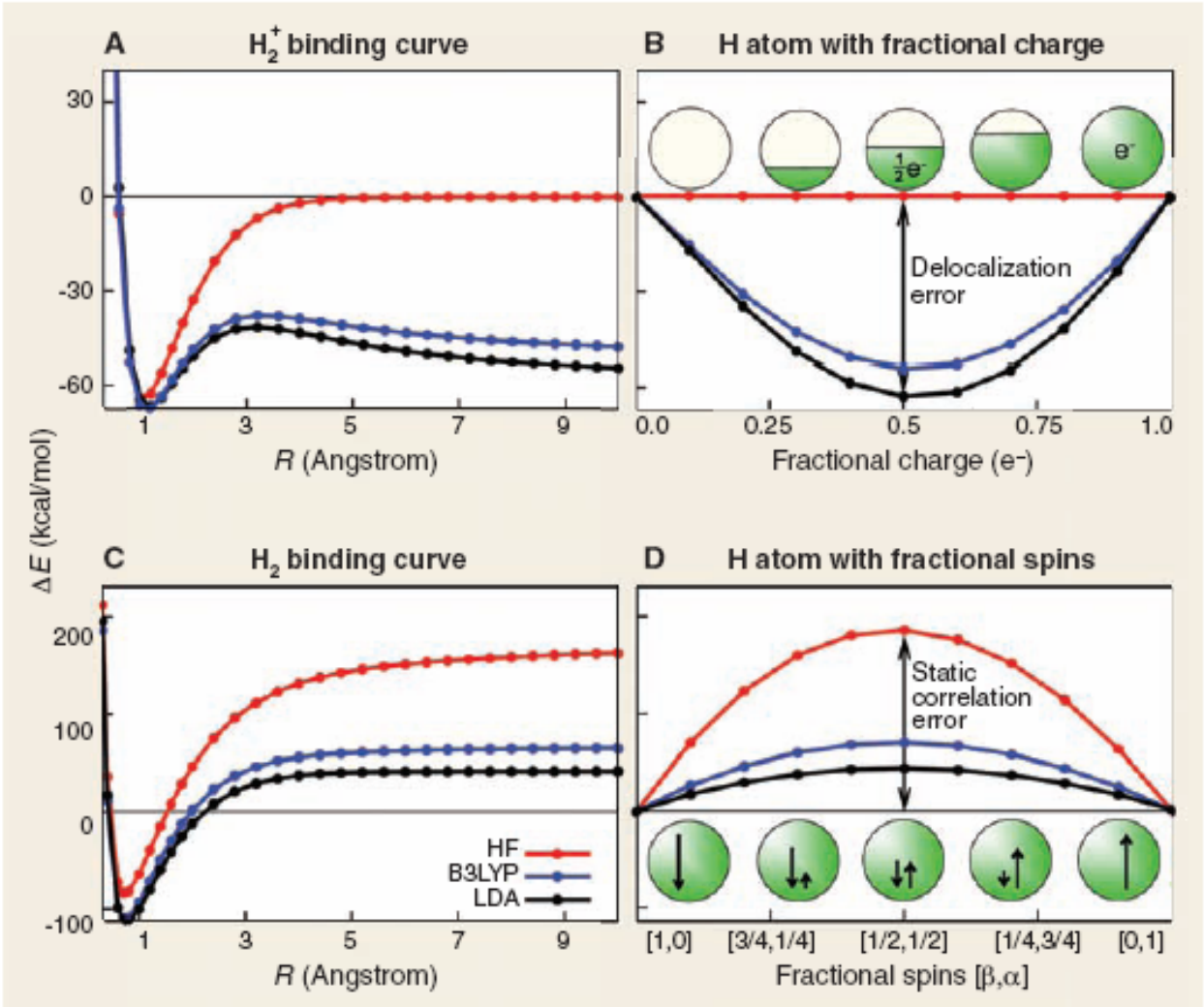
$$E_{XC}^{GGA}[n(r)] = \int f(n(r), |\nabla n(r)|)dr \tag{9}$$

$$E_{XC}^{Hybrid} = E_{XC}^{GGA} + a_0(E_X^{HF} - E_X^{GGA}) \tag{10}$$

- LDA $\rightarrow$ acceptable molecular structures
- GGAs $\rightarrow$ acceptable thermochemistry (MADs $\approx$ 6 kcal/mol)
- Hybrids $\rightarrow$ good thermochemistry (MADs $\approx$ 2 kcal/mol)

Density Functional Theory

Problems



A. J. Cohen, P Mori-Sanchez, W. Yang *Science* **2008**, 321, 792.

Use a sufficiently flexible functional: B97

Becke 1997

- systematic procedure for refining gradient corrections to get the optimum empirical Hybrid-GGA functional

$$E_{XC} = E_X^{GGA} + E_C^{GGA} + c_X E_X^{HF} \tag{11}$$

$$E_X^{GGA} = \sum \int e_X^{LDA}(\rho) g_X(s^2) dr \tag{12}$$

$$e_X^{LDA} = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{\frac{1}{3}} \rho^{\frac{4}{3}} \tag{13}$$

$$s = \frac{|\nabla \rho|}{\rho^{\frac{4}{3}}} \tag{14}$$

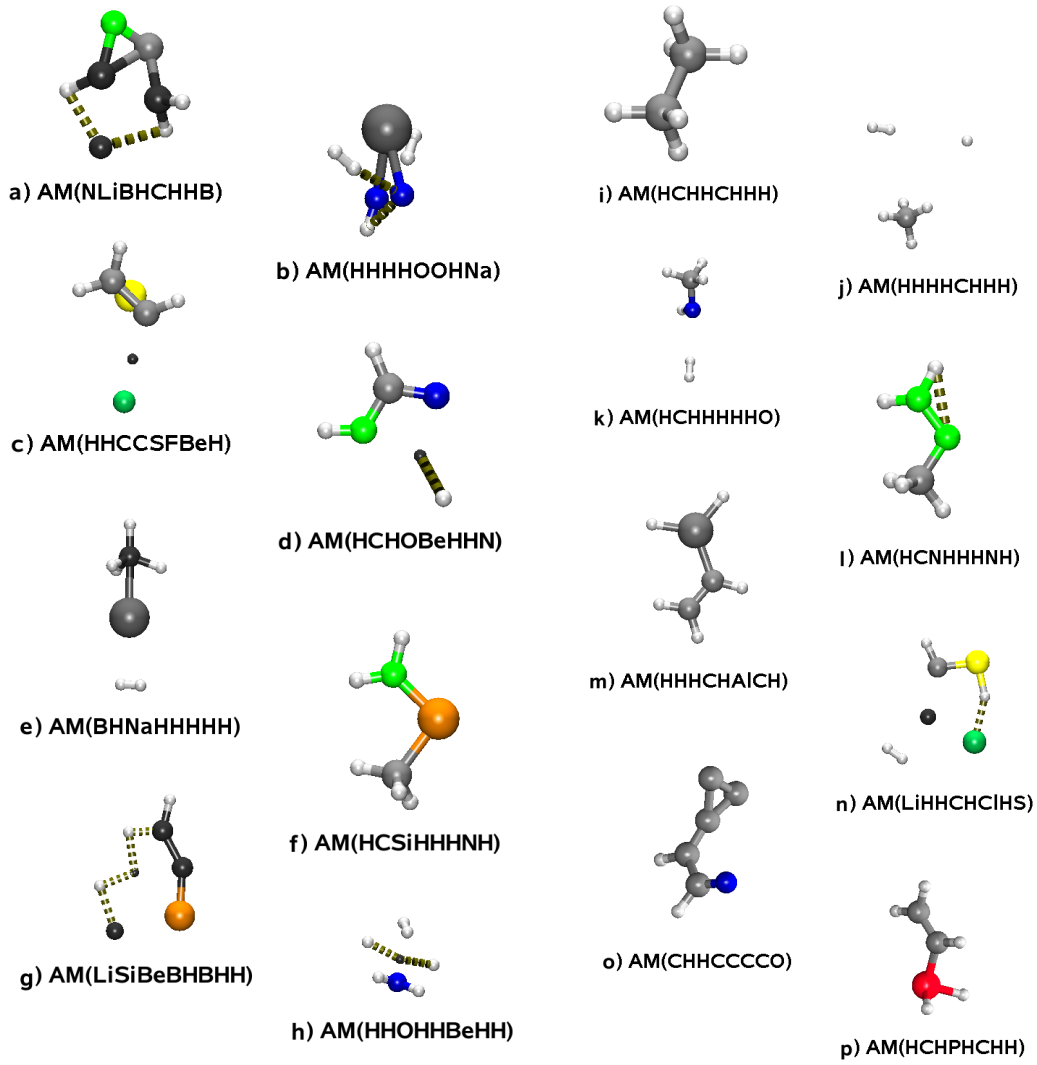
$$u_X = \frac{\gamma_X s^2}{1 + \gamma_X s^2} \tag{15}$$

$$g_X = \sum_{i=0}^{2} c_i u^i \tag{16}$$

- 10 Parameters: c_0, c_1, c_2 for $g_X, g_{C\alpha\beta}, g_{C\sigma\sigma}$ and c_X

A. Becke *J. Chem. Phys.* **1997**, 107, 8554.

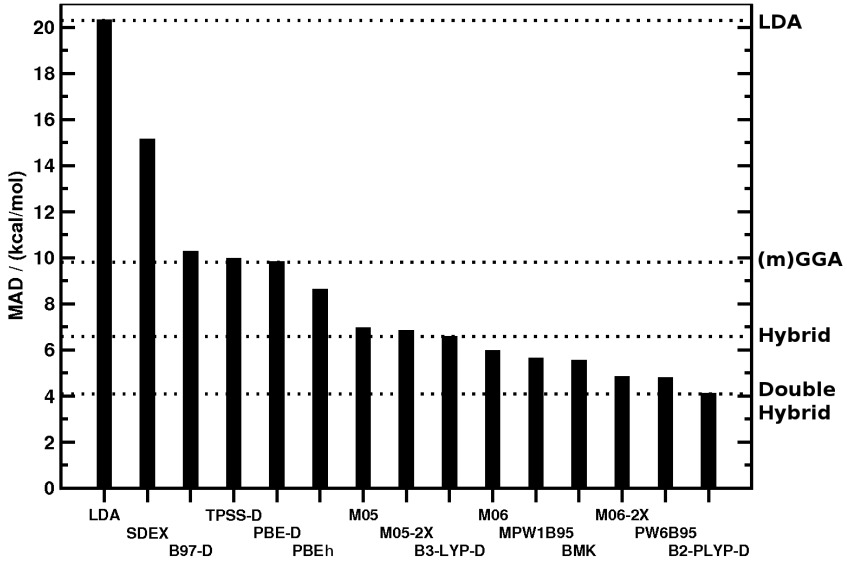
Use a fit set with enough structural diversity: MB08-165



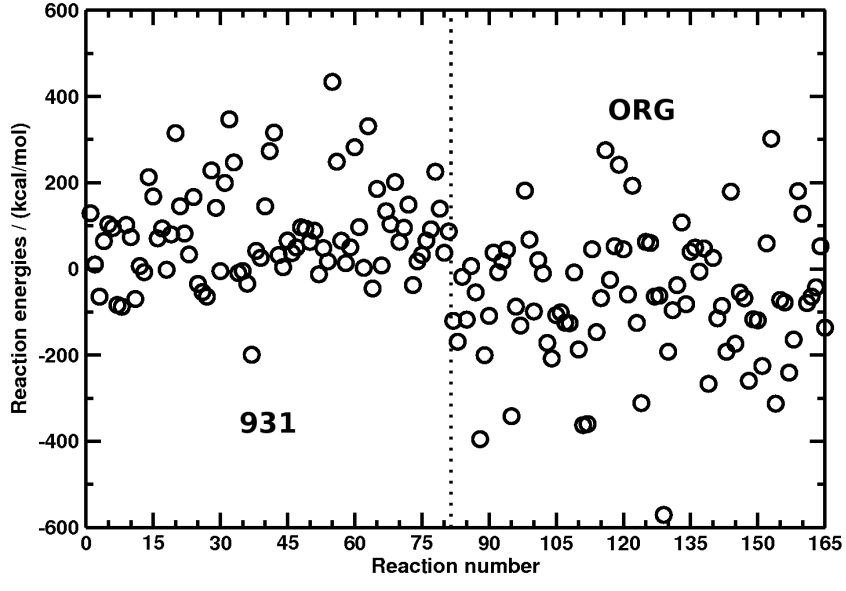
165 'Artificial Molecule' reactions with AE-CCSD(T)/CBS refs

M. Korth, S. Grimme *J. Chem. Theory Comput.* **2009**, 5, 993.

Mindless Benchmarking



DFT MAEs



Reaction energies

QAH

How does BOINC work?

Backend

- Apache web server
- MySQL database

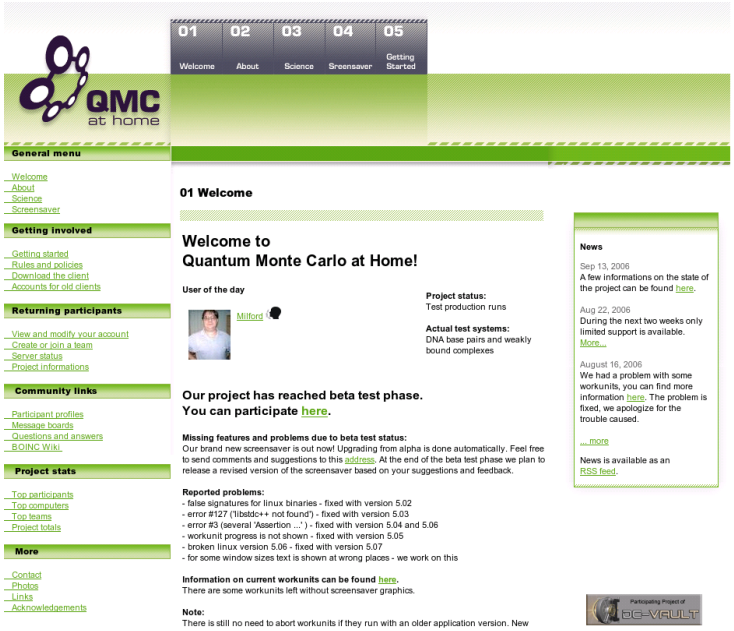
Middle-ware

- web-frontend (PHP), DB-Interface (Python)
- set of Daemons (C++/C)
  - scheduler daemons (getting data on it's way)
  - transitioner, validator (redundant computing)
  - work generator, assimilator (create and evaluate work)
- core-client: one client software for all projects!

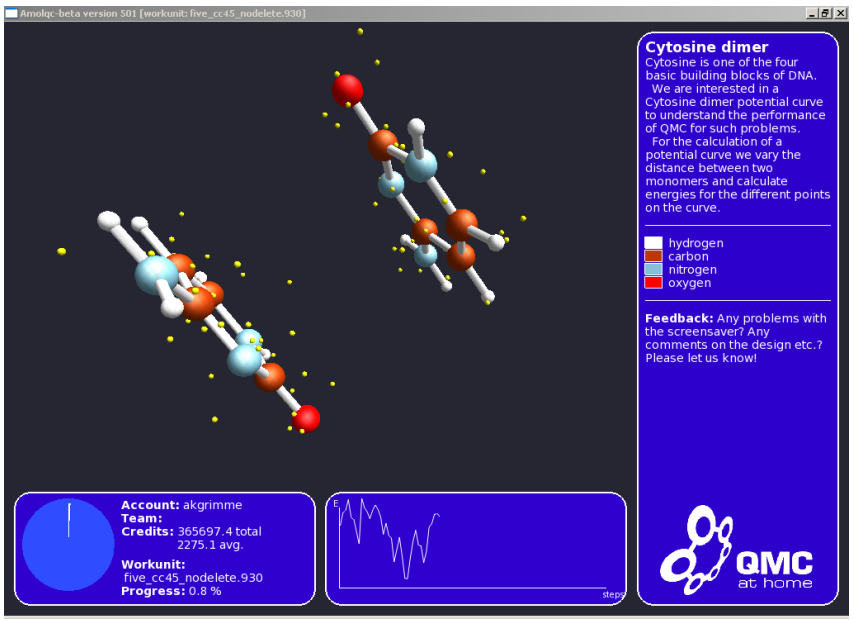
Application

- communication between core client and application via MPI-like functions (boinc_init, boinc_finish, ...)
- based on the **AmolQC QMC program by Arne Lüchow et al.**

QMC@HOME - Quantum Monte Carlo at Home

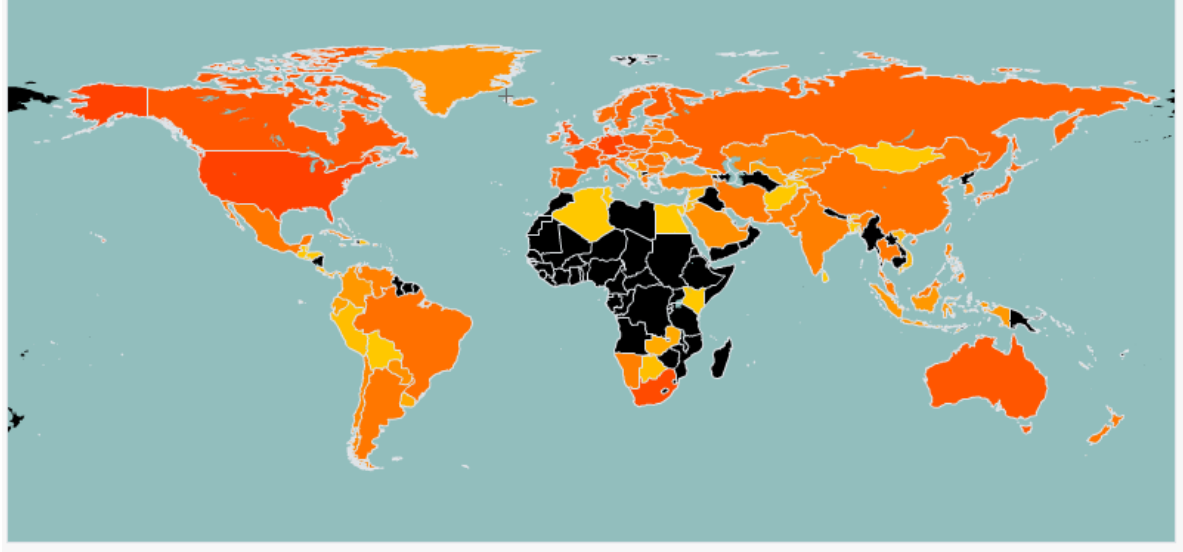


The QMC@HOME web page: http://qah.uni-muenster.de



The QMC@HOME screensaver

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!