The scaling of computational time as a function of number of processors for Quantum Monte Carlo study of CO molecule

Luiza Buimaga-Iarinca, Adrian Calborean

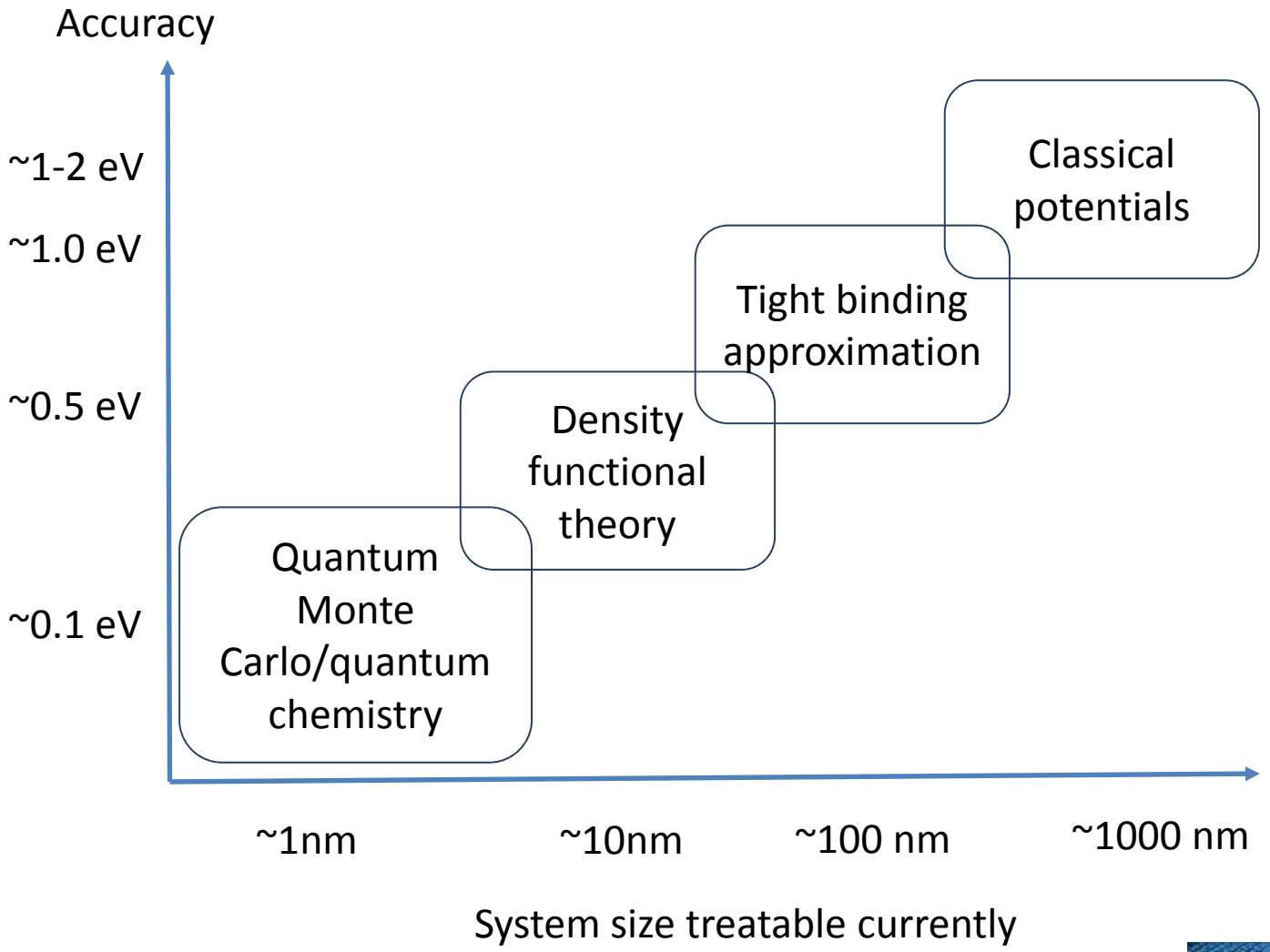
Fifth Romania Tier 2 Federation "Grid, Cloud & High Performance Computing Science“
RO-LCG 2012, 25-27 Oct , Cluj Napoca

IEEE
Advancing Technology for Humanity

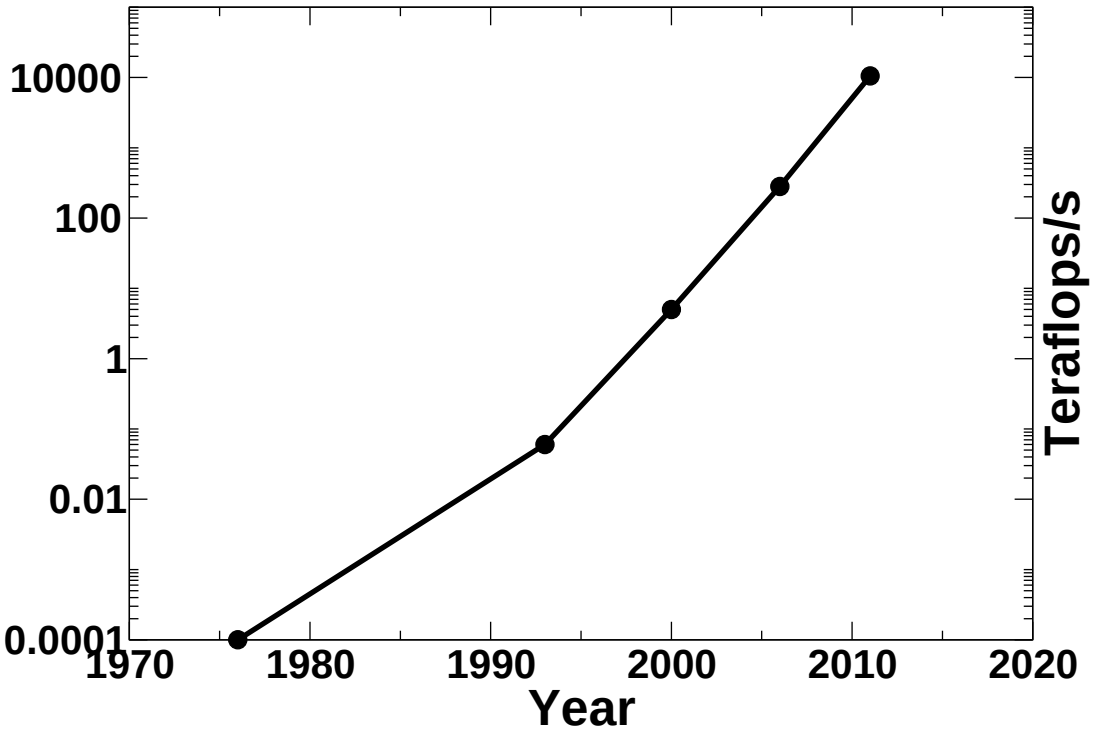
Institutul National pentru Tehnologii Izotopice si Moleculare

Structural information extracted from QM

- We need to solve Schrodinger equation for large number of electrons in periodic or non-periodic systems; this corresponds to the accurate determination of the multi-electronic wave function
- Out of the resulting wave function we can get the forces acting on atoms (i.e. theoretical equilibrium structure)
- Various observables can be extracted out the system's total wave function: vibrational properties, magnetic moments, polarizability, etc.

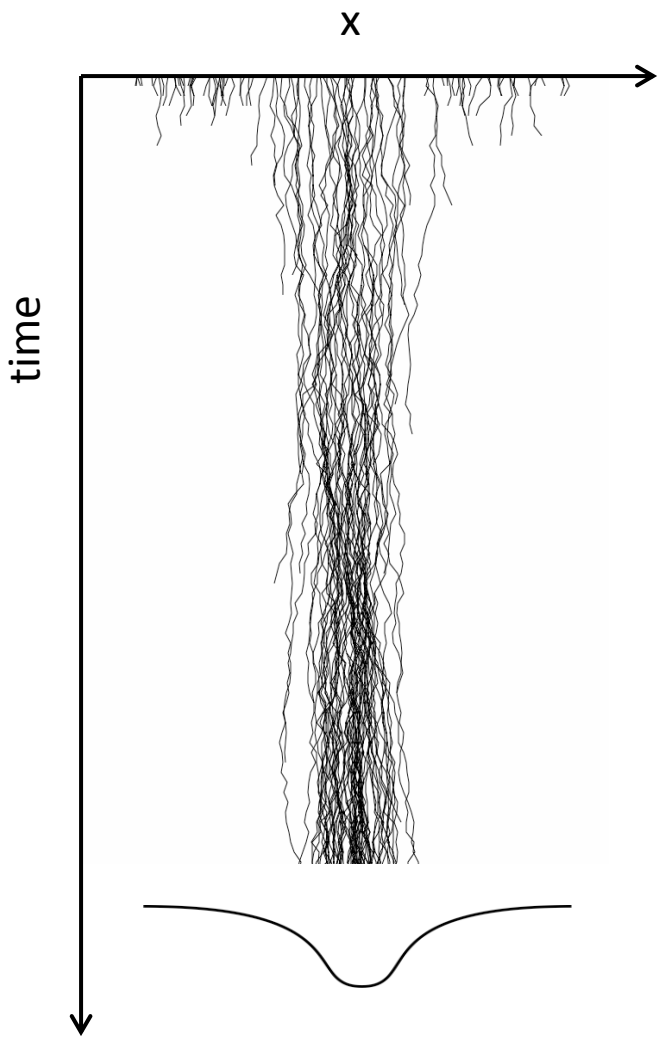

Accuracy
~1-2 eV
~1.0 eV
~0.5 eV
~0.1 eV
Classical potentials
Tight binding approximation
Density functional theory
Quantum Monte Carlo/quantum chemistry
~1nm
~10nm
~100 nm
~1000 nm
System size treatable currently

Evolution of the computing power



QMC: basic principles

- QMC gives accurate solutions
- Complex wave functions are simulated by statistical densities of model-configurations
- This ask for a large number of independent calculations
- Very efficient in parallel environments



Computational details

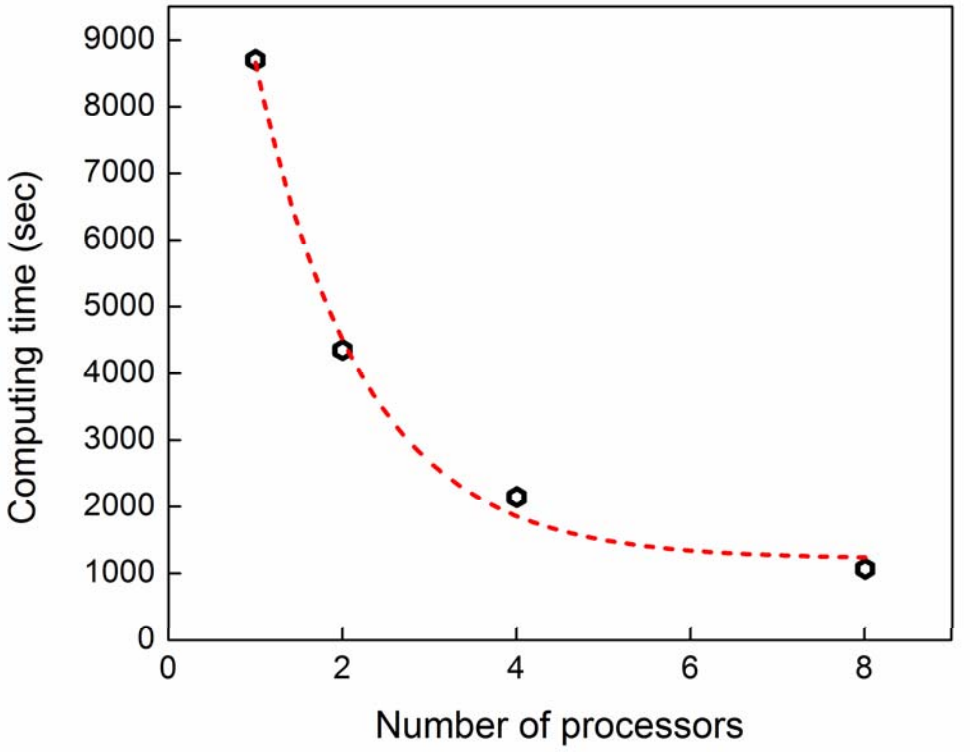
We generate our preliminary wave function by using the DFT code SIESTA; QMC calculations are performed with the diffusion Monte Carlo DMC method using the QWALK code.

The DFT single-particle states are used to construct the Slater determinant; the variational parameters, including electron-electron, electron-nucleus, and electron-electron-nucleus terms, are optimized by variance minimization.

PBE-GGA DFT calculations were carried out with triple –ζ polarization atomic orbital basis sets.

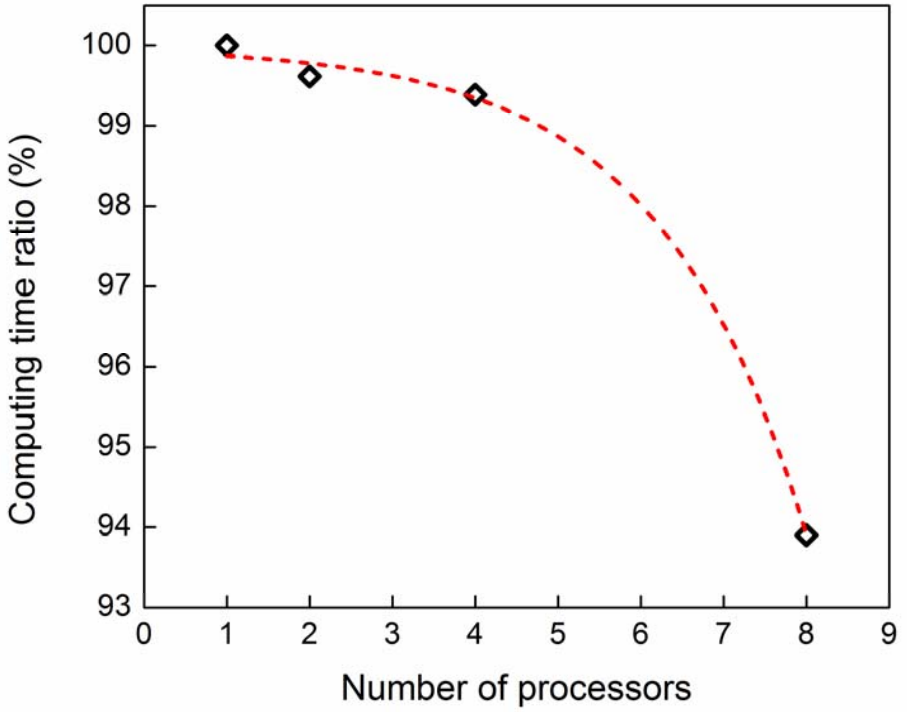
As a test-system for our calculation we use the carbon monoxide molecule (CO).

Dependence of the computing time by the number of processors



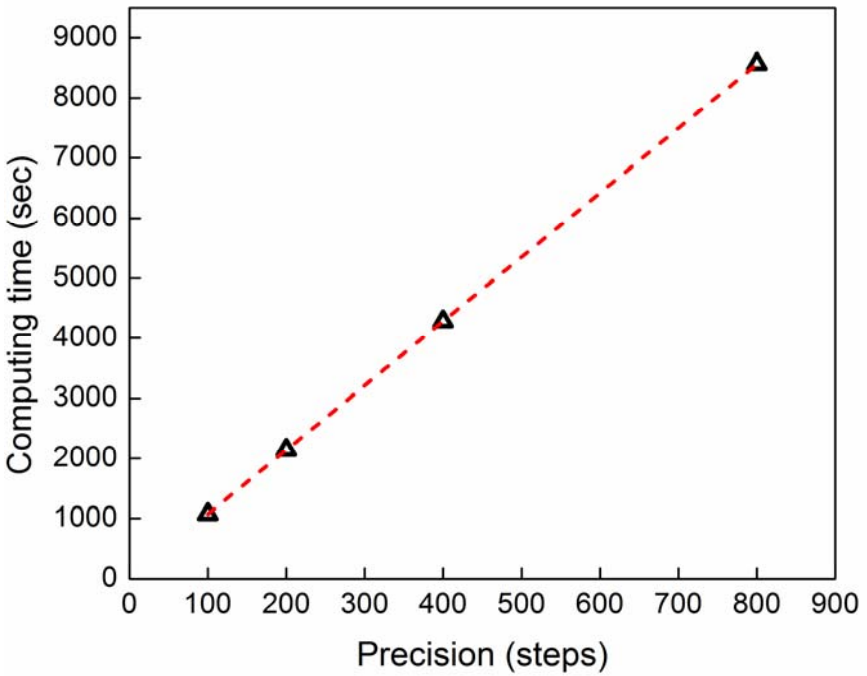
The number of QM configurations is 1000.
The number of configurations per processors is 1000/n where n is the processors number used in the QMC simulation.

Analysis of the time communication between processors



We keep constant the number of configurations per processor (i.e. 250) and we compute the total energy in QMC.

Dependence of the computing time by the number of configurations



Is represented the dependence of the computing time over the total number of configurations.
We keep constant the number of configurations per processor (250) and the total number of processors (i.e. 4 processors) and vary the number of propagation steps inside the QMC procedure.

Conclusions

Monte Carlo algorithms are known to be one of the best cases for a parallel implementation, since the communication between the execution threads is minimized in a natural way. We tested this ansatz by our first investigation.

By using this as a reference result we study the efficiency of the communication between processors in the MPI network implemented at INCDTIM.

Our results show that up to 4 processors the efficiency is superior to 99%. On the contrary for larger number of processors the efficiency drops significantly, reaching 94% for a run on 8 processors.

Acknowledgments

Investing in people!
Ph.D. scholarship, Project co-financed by the SECTORAL OPERATIONAL PROGRAM FOR HUMAN RESOURCES DEVELOPMENT 2007 - 2013

We acknowledge financial support from PN-II-ID-PCCE-2011-2-0027 BIOSENS project.

Thanks are due to INCDTIM Cluj-Napoca Data Center for providing computer facilities.