

SIMULATING BIOMOLECULES ON COMPUTERS!

PATRICK MERLOT, JON AUSTAD AND JOHANNES REKKEDAL

PATRICK.MERLOT@KJEMI.UIO.NO, JONAUS@STUDENT.MATNAT.UIO.NO, J.A.REKKEDAL@KJEMI.UIO.NO



UiO • Kjemisk institutt

PURPOSE

Most of physics has been known at the atomic scale for about a century. However simulating molecules of few thousands atoms (like important biomolecules) from first principles still requires too much computational effort. Thanks to fast and accurate linear-scaling (LS) methods in quantum chemistry, simulating such systems may soon become routine calculations, an important step for computer-aided drug design, and more generally for bridging the gap between quantum and classical physics, e.g. between the nanoworld and our macroscopic environment.

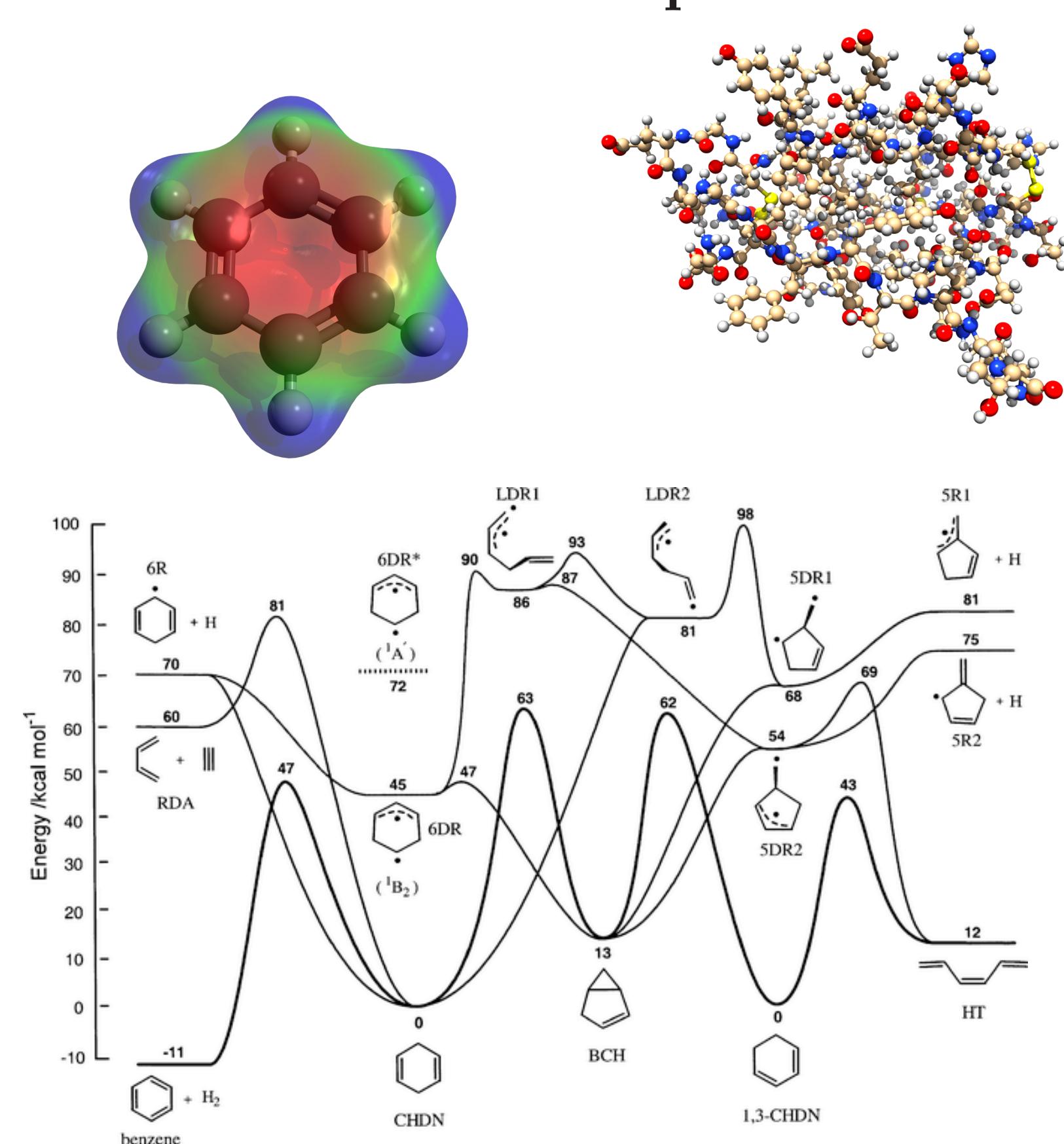
INTRODUCTION

The laws of physics at the atomic level are governed by the **Schrödinger equation**,

$$i\hbar \frac{\partial \psi}{\partial t} = \frac{\hbar^2}{2m} \nabla^2 \psi + V(\mathbf{r})\psi$$

with analytical solutions known for very few and small systems.

By approximating the solutions of this equation, computational chemistry aims at predicting molecular properties, dynamics and reactions on computers.



Today about one third of the publications in chemistry uses computation to perform verifications, interpretations of experiments or even predictions.

For molecular systems with more than one electron, a **hierarchy of quantum mechanical (QM) methods** allows for a systematic approach to the exact solution, providing reliable “measurements” or estimations when based on the simplest approximations.

CHECK THIS OUT!



Center for
Theoretical &
Computational
Chemistry
www.CTCC.no

TODAY'S LIMITATIONS

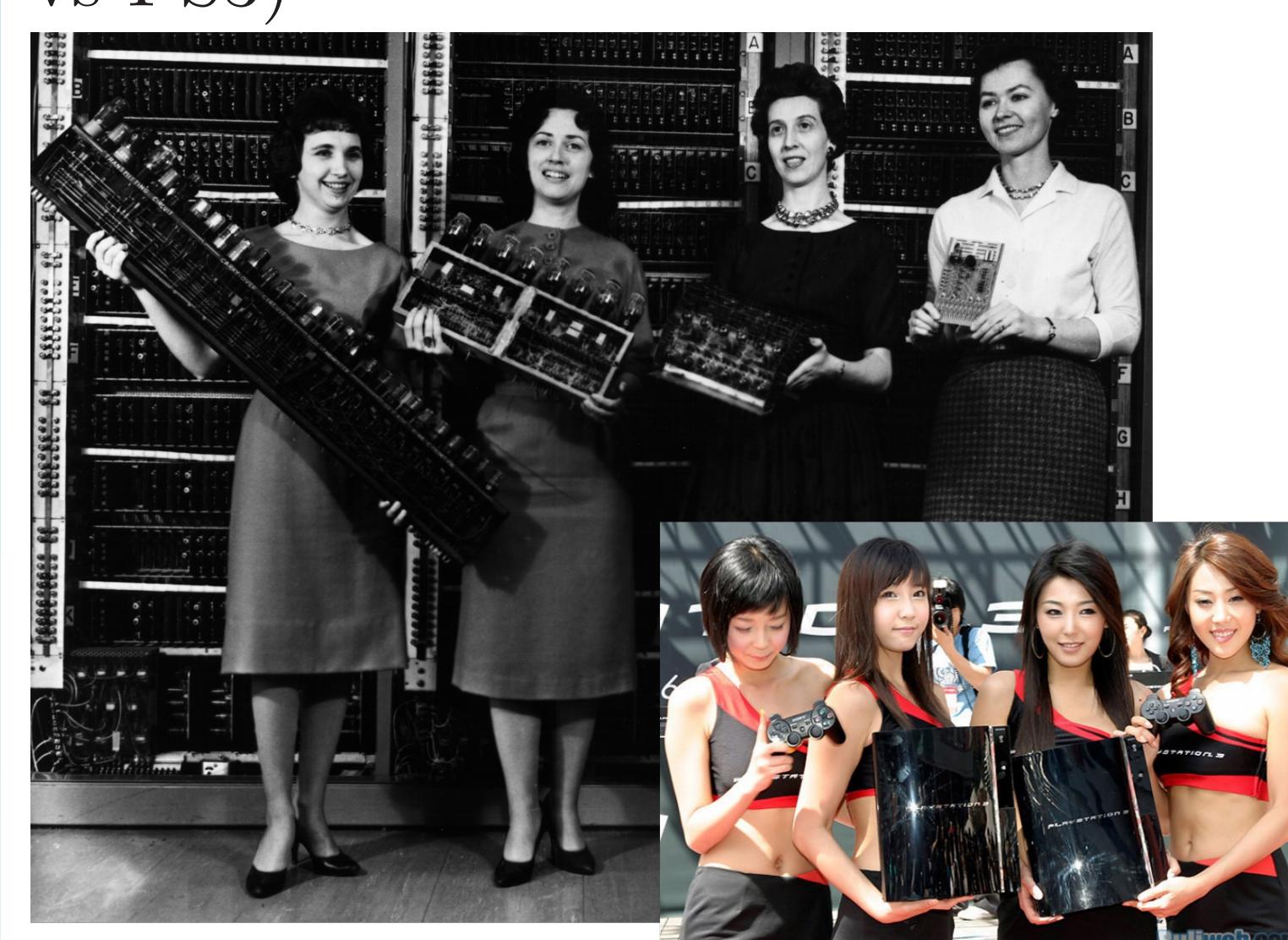
The main drawback with these methods is the strong increase in computational effort with molecular size (M).

With accurate but very expensive methods like the Coupled-Cluster (CC) method, each new digit in the energy may cost 10000 times more CPU time!
1 minute → 1 week → 200 years

Even one of the simplest approach, the Hartree-Fock (HF) method, scales conventionally as $\mathcal{O}(M^3)$.

This means that choosing another molecule that is 10 times larger than the current one increases the computational effort by a factor of 1000!

Fortunately we benefit from exciting improvements in computer power over time nowadays mainly driven by the game and IT industry. (e.g. ENIAC vs PS3)



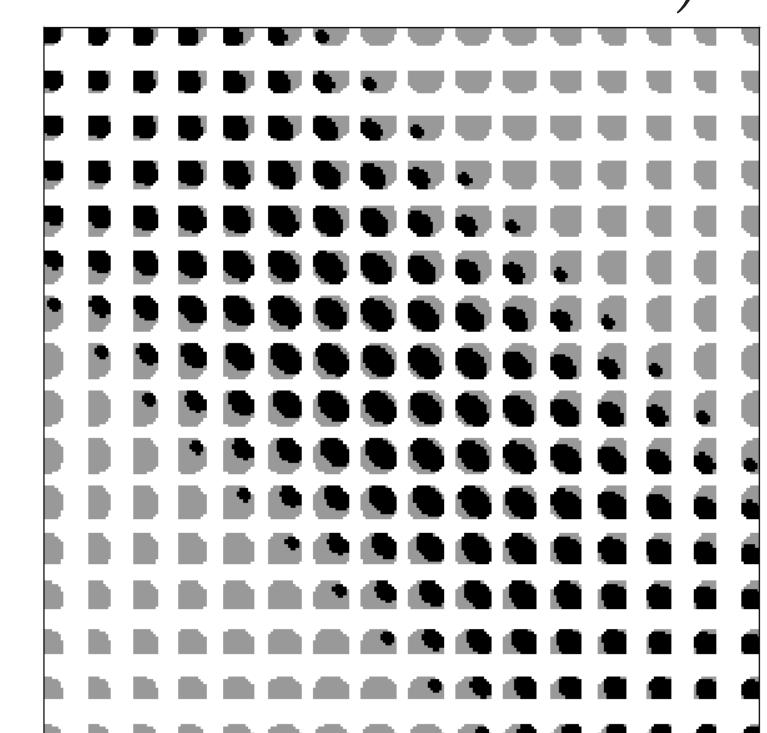
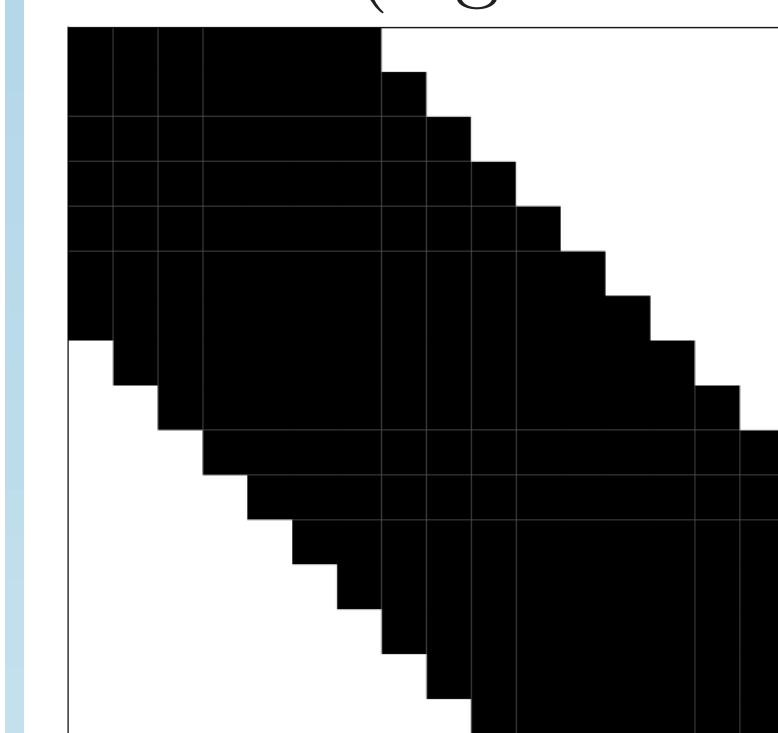
Moore's law is still astonishingly valid over the last decades, predicting that computer speed might double every 1.5 years. Then in 15 years, the computers speed might be multiplied by about a factor 1000 ($2^{10} = 1024$).

Should we wait 15 years of hardware improvements to simulate molecules only 10 times larger within the same time frame?

LINEAR-SCALING

In order to reach the size of biomolecules, **the scaling behaviour of the QM methods has to be reduced**, optimally to scale linearly with system size.

The idea is to exploit the sparsity of different matrices involved in the calculations (e.g. overlap and eri matrices).



LS techniques are already successfully developed for few methods (SCF methods): to compute the energy of a system, to optimize their geometry and to get some response properties that can be compared to experimental results (such as vibrational frequencies, NMR shieldings).

The efficient evaluation of “exact exchange in Fock-type matrices”, which is a pure quantum mechanical term, is particularly important to include for large molecular systems since without it, calculations typically fail to converge (e.g. Insulin: 787 atoms).

Linear-scaling methods also perfectly combine with new computer architectures to allow **massively parallel calculations**.



CHALLENGES

There is still great need for developing and improving linear-scaling methods. Many molecular properties remain without linear-scaling development, small “homo-lumo” gaps affect a lot the performances and the flexibility of large molecules emphasizes the importance of very fast and cheap dynamics simulations.

ACKNOWLEDGEMENTS

This work is receiving support from the Center for Theoretical and Computational Chemistry (CTCC) at the University of Oslo.