Solid state chemistry and physics

Filip Sagan, Piotr Wróbel

ACK Cyfronet AGH

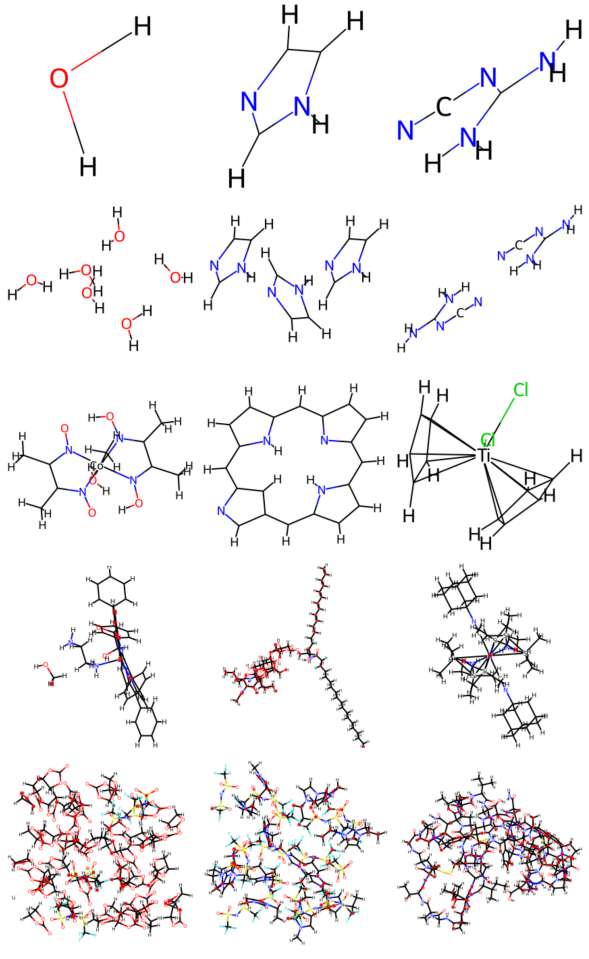
7^{th} December 2023

Purpose of tests

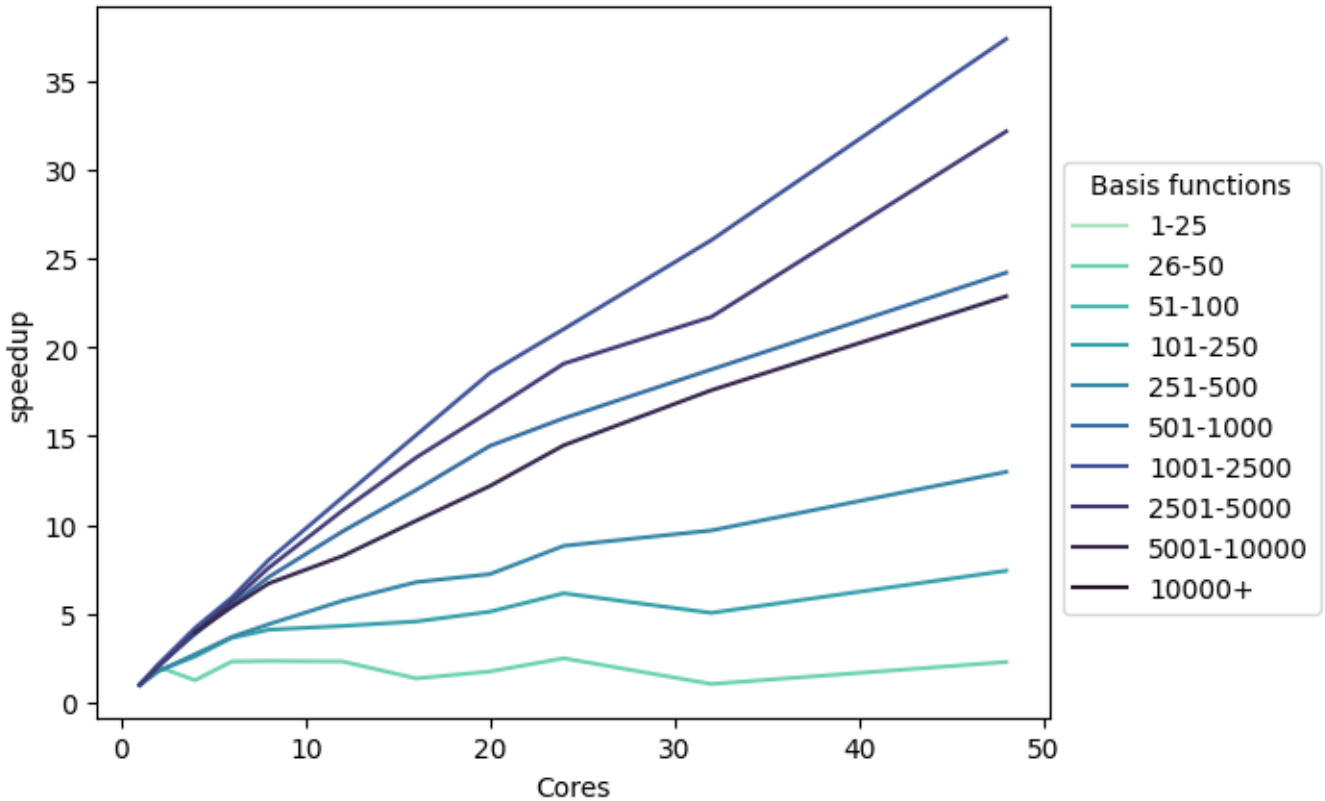
- studies of parallelisation
- description of good practices in using computational chemical software
- studies of edge cases
- measurement of performance and resource utilization on HPC clusters

Methodology of quantum chemistry software testing

- Software
  - Gaussian
  - AMS/ADF
  - ORCA
  - Molpro
  - QChem
  - Turbomole
- Methods
  - Wavefunction methods: HF, MP2
  - DFT: BLYP, B3LYP, MVS
  - Basis: common double and triple ζ basis stets: Pople, Dunning, Karlsruhe varieties
- Systems

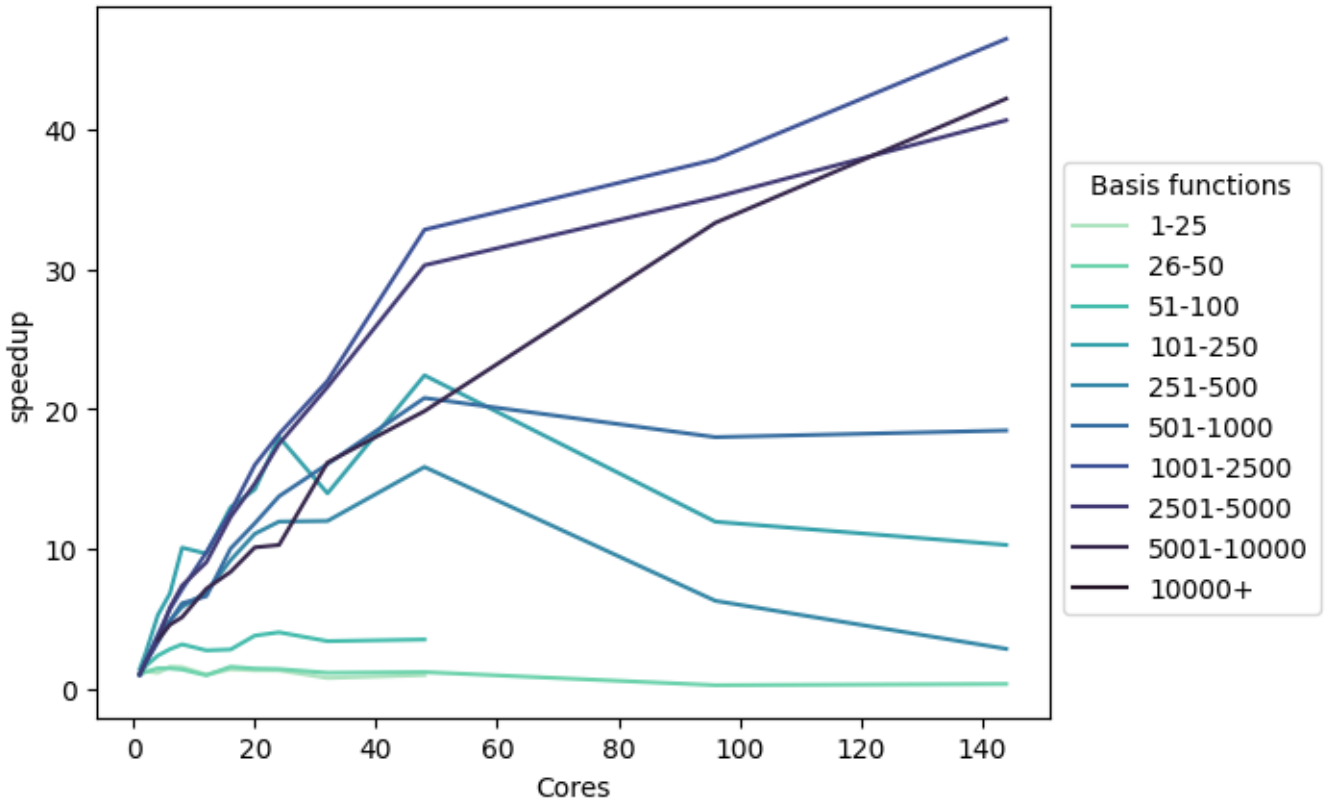


Gaussian



- Single node parallelisation
- Low demand for memory
- Great paralellisation of 500+ basis function jobs

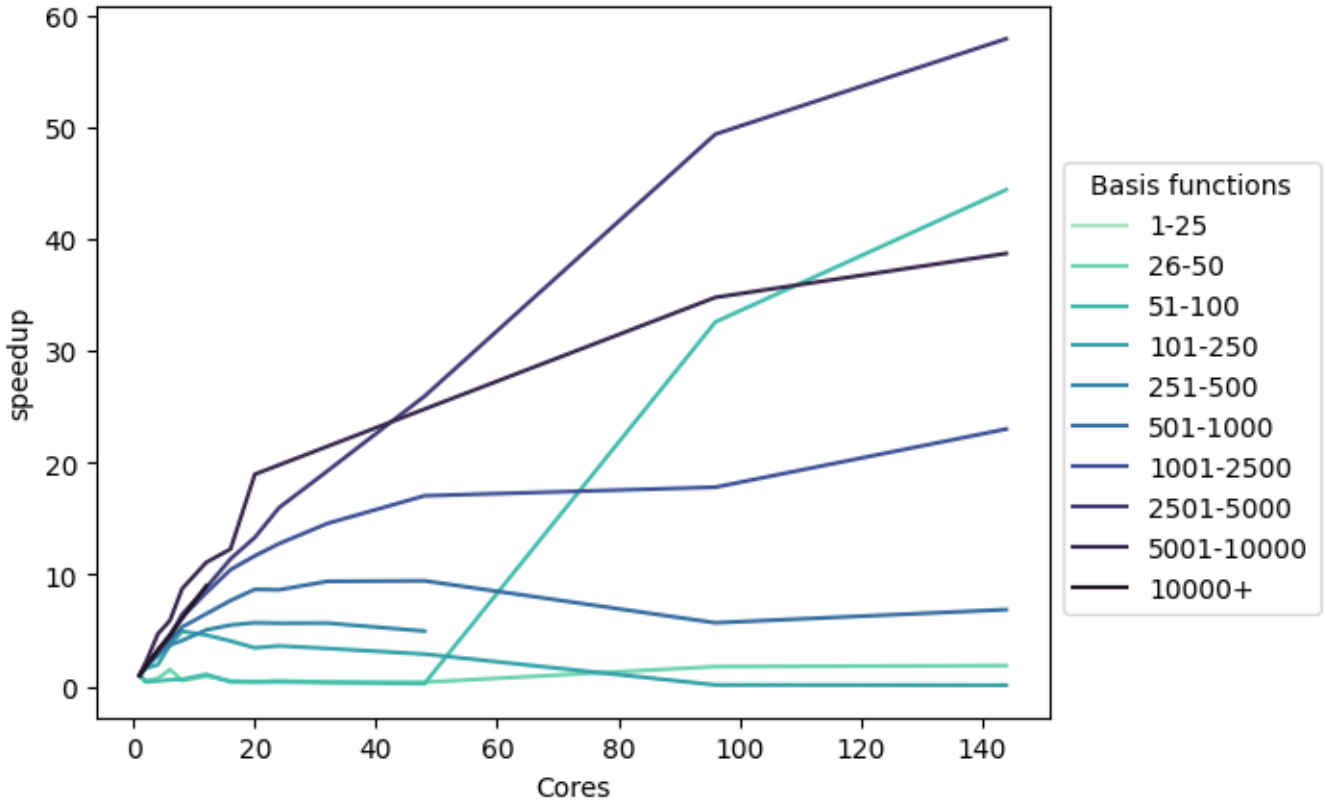
AMS/ADF



- Multiple node parallelisation
- Slater-type orbirtals
- Low demand for memory

- Efficient paralellisation of 100+ basis function jobs on one processor
- Diminishing returns for calculations on over one node and 1000+ basis functions

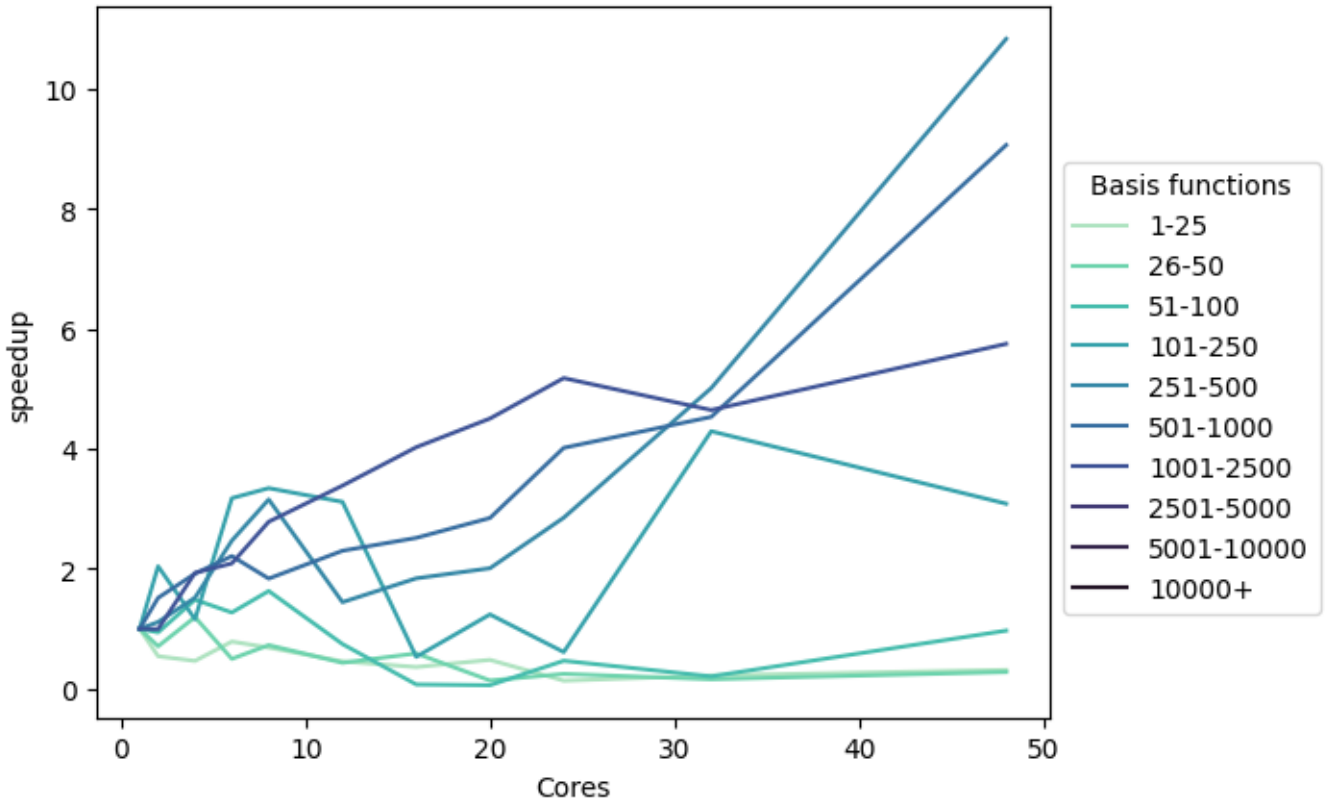
ORCA



- Multiple node parallelisation
- Large demand for memory

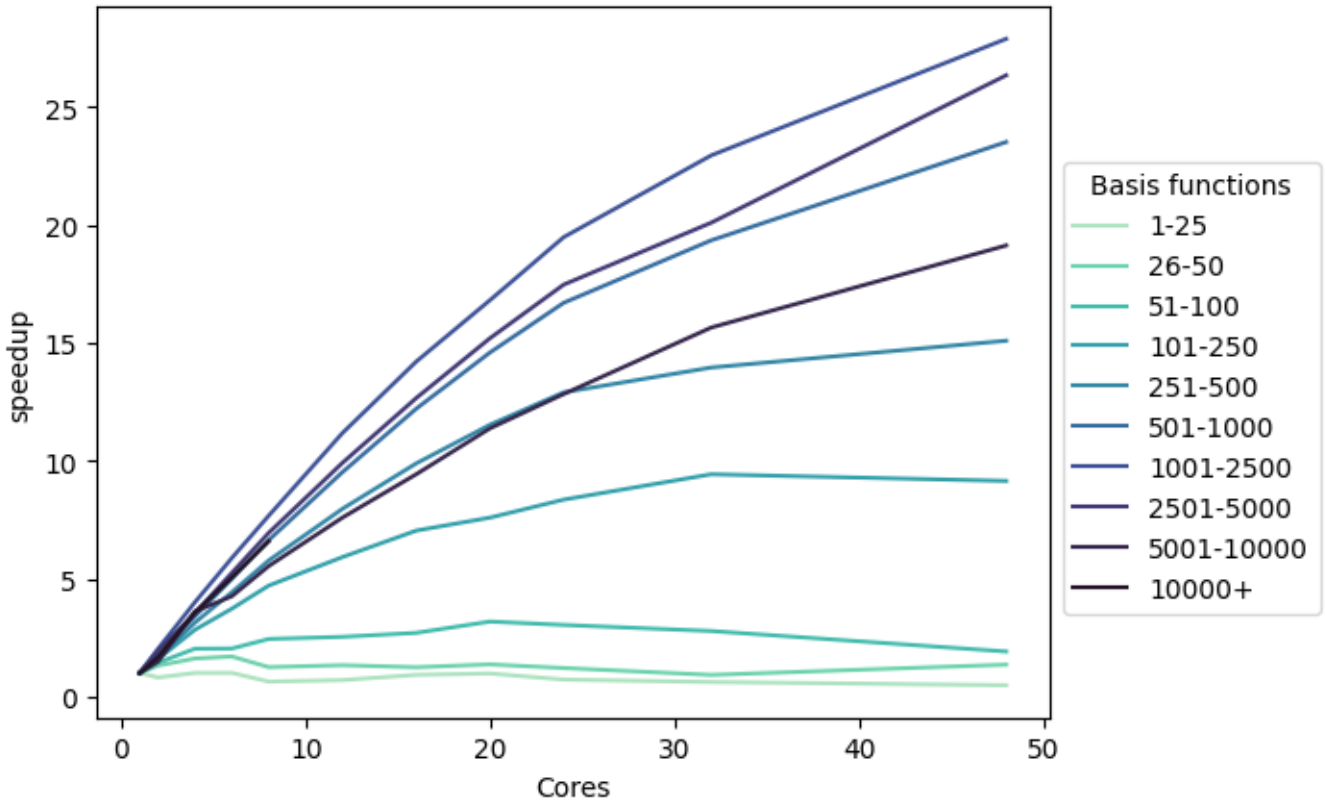
- Small (up to 500 basis functions) jobs reach reasonable perforamance at 8 cores
- Large (2500+ basis functions) jobs paralellise well on many nodes, given enough memory is provided

Molpro



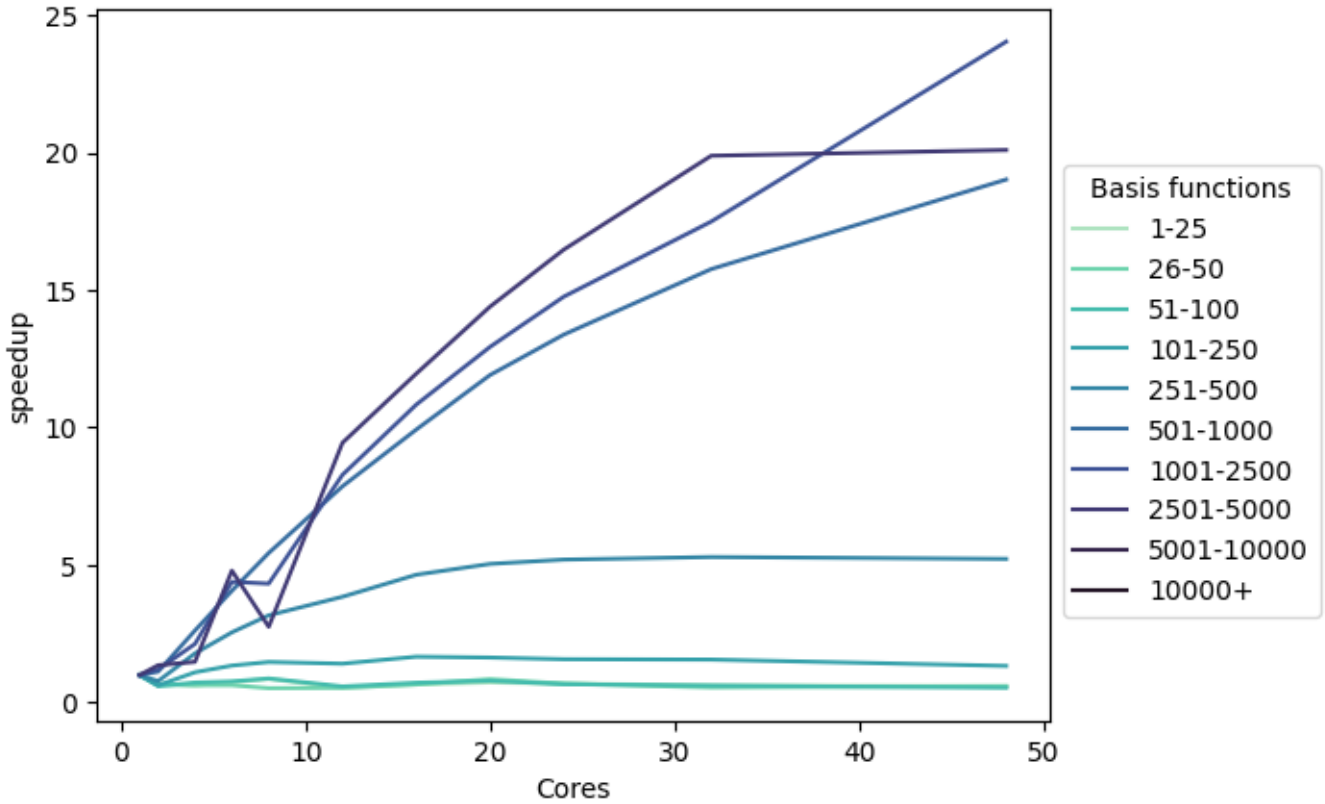
- Single node parallelisation
- Large demand for memory
- Parallelises well up to 8 cores
- Up to 400 atoms on Ares, more than 3000 basis functions runs out of memory

QChem



- Single node parallelisation
- Low demand for memory
- Great paralellisation of 500+ basis function jobs

Turbomole

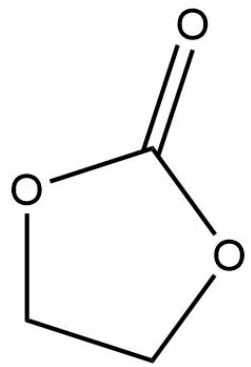


- Multiple node parallelisation
- Large demand for memory
- Efficient paralellisation of 500+ basis function jobs

Studied software

- CP2K
- CPMD
- DFTB+
- NAMD

Studied systems

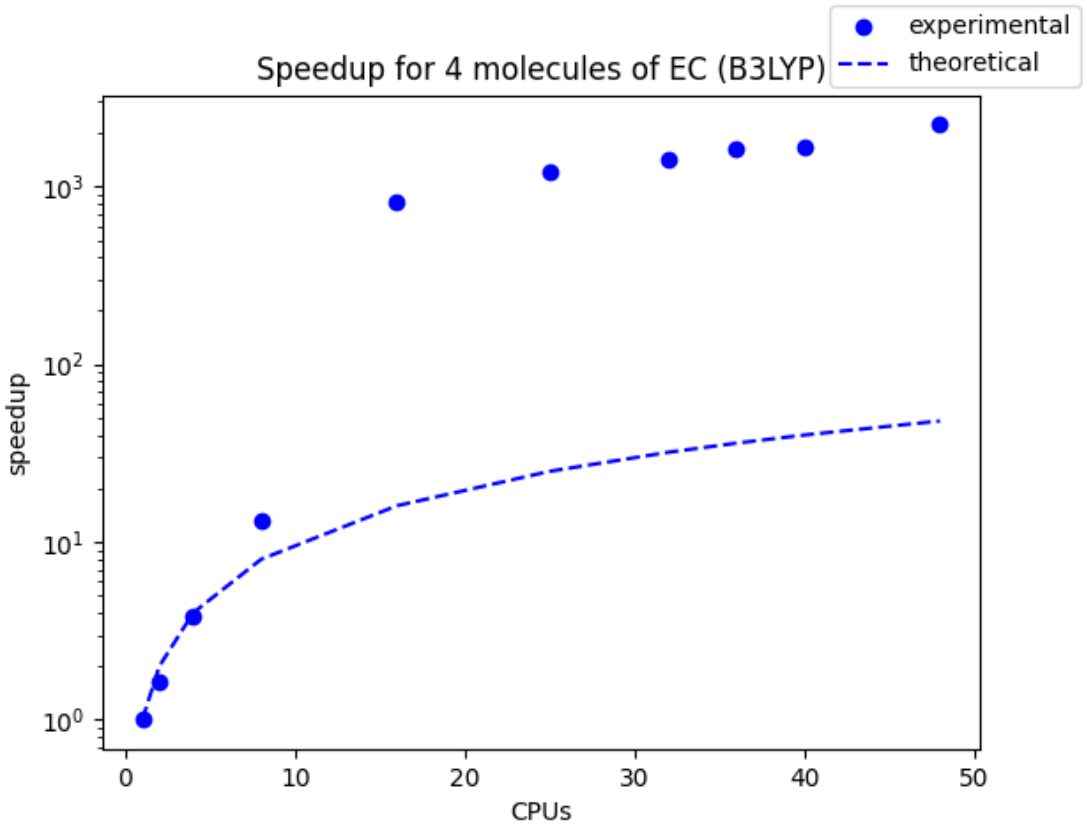


Studied systems involved ethylene carbonate (EC) boxes with different number of EC molecules, from 1 to few hundred.

Results for CP2K

For B3LYP:

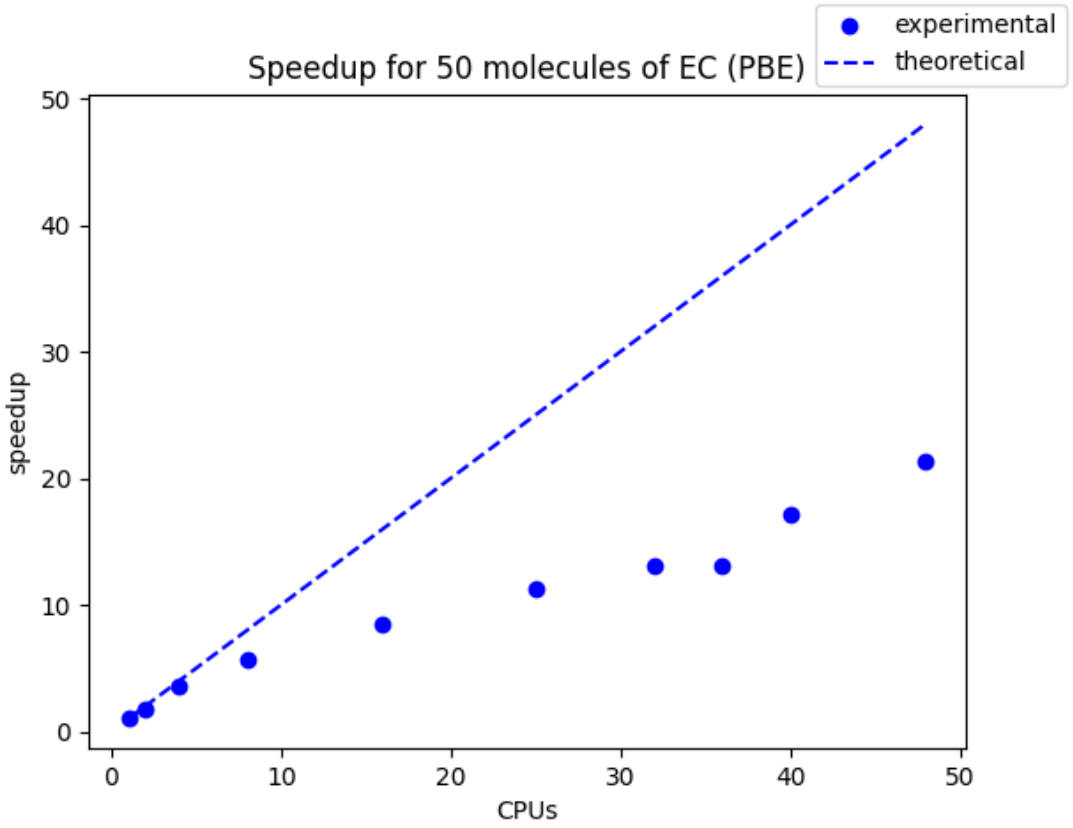
- maximal system size: 16 molecules
- very slow
- anomal behaviour with parallelisation



Results for CP2K

For PBE:

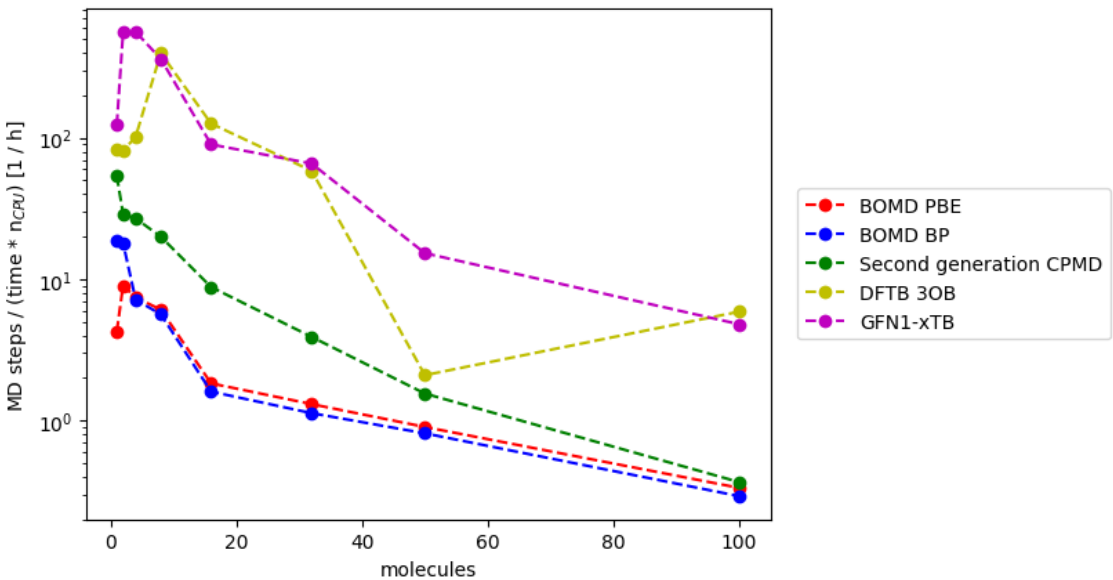
- maximal system size: 400 molecules
- reasonale speed
- typical behaviour with parallelisation



Results for CP2K

For MD:

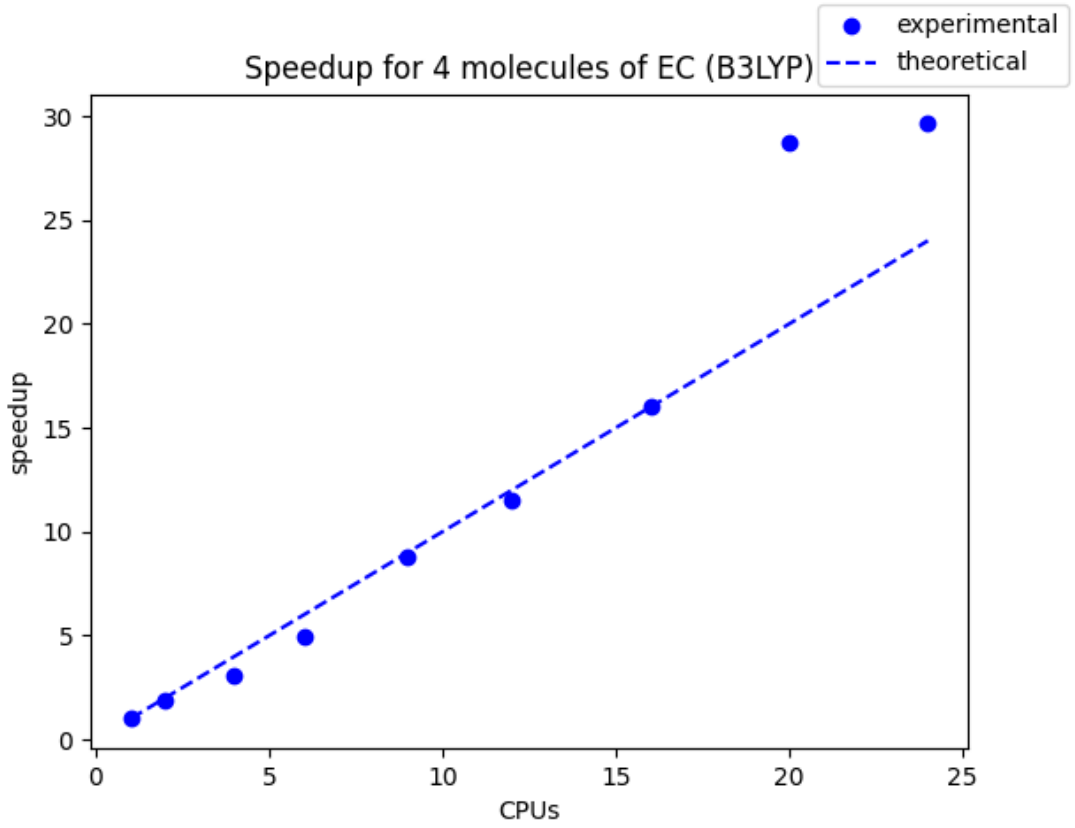
- PBE and BP are similar
- SGCPMD does not make sense for big systems
- from semi-empirical approaches GFN1-xTB is the fastest



Results for CPMD

For B3LYP:

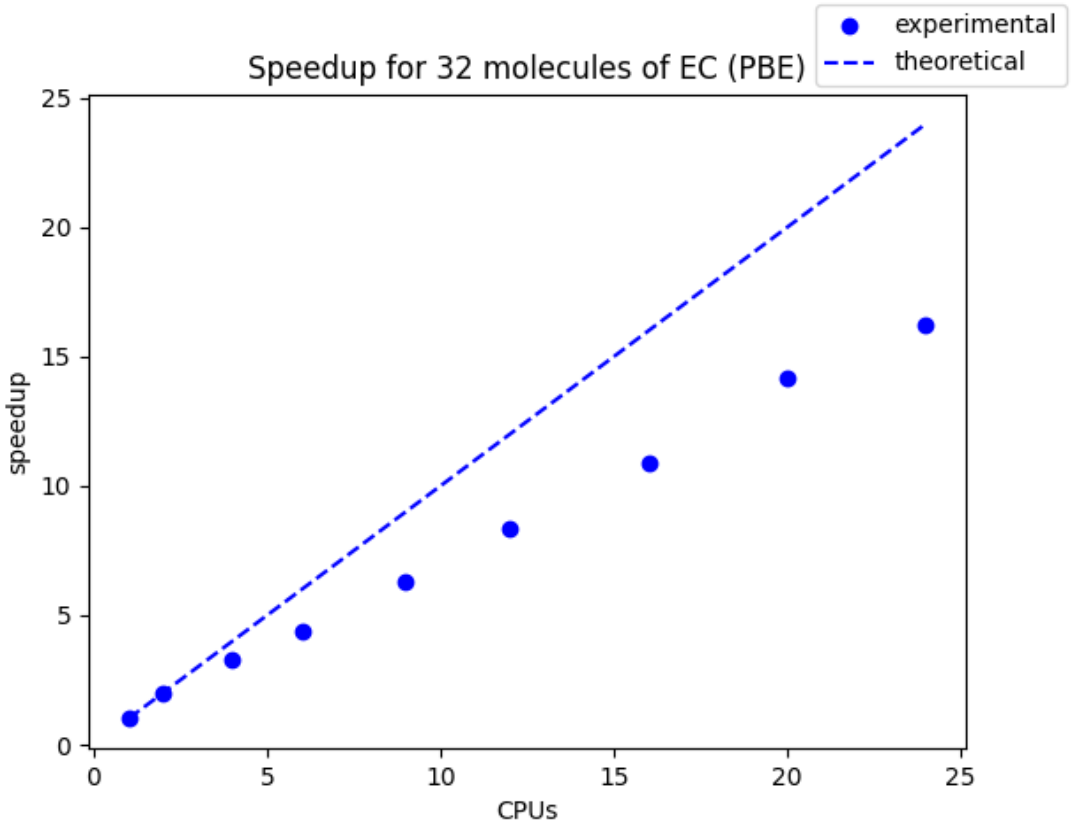
- maximal system size: 8 molecules
- very slow
- again, anomal behaviour with parallelisation



Results for CPMD

For PBE:

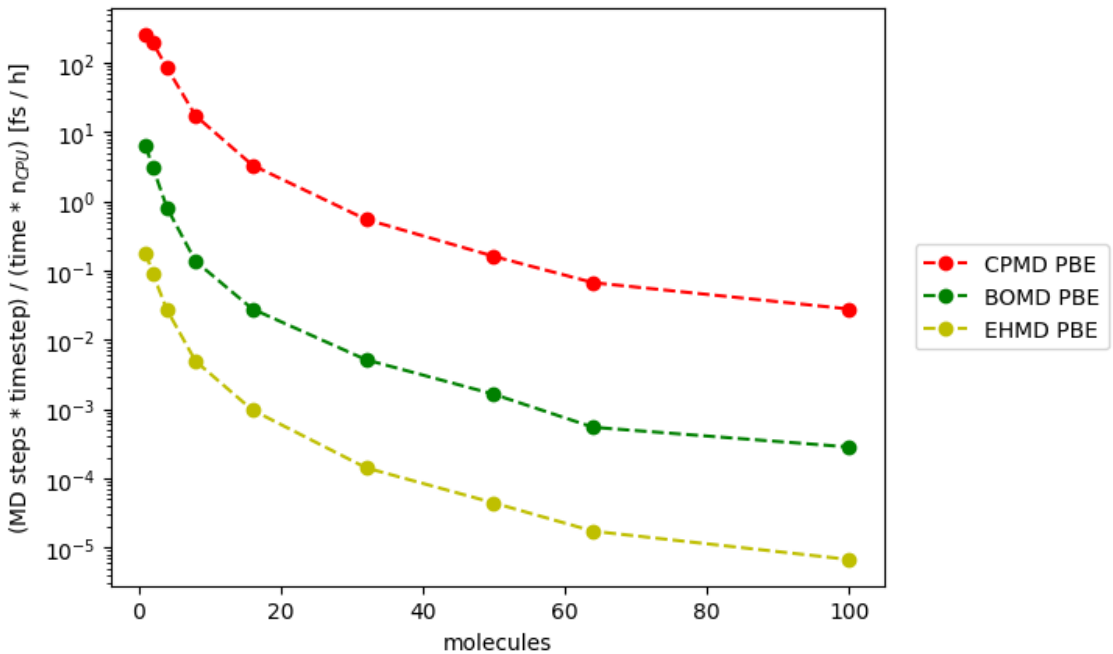
- maximal system size: 200 molecules
- reasonable speed
- again, typical parallelisation



Results for CPMD

For MD:

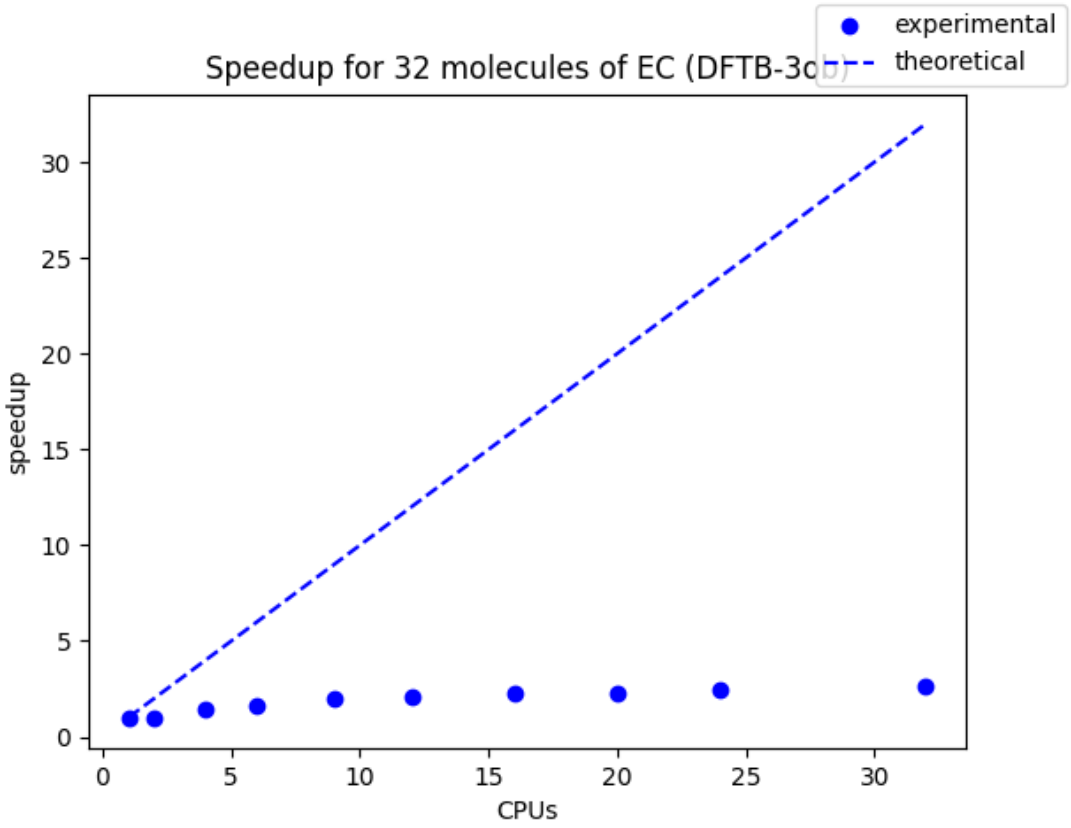
- all approaches have similar behaviour
- typical hierarchy



Results for DFTB+

For 3ob:

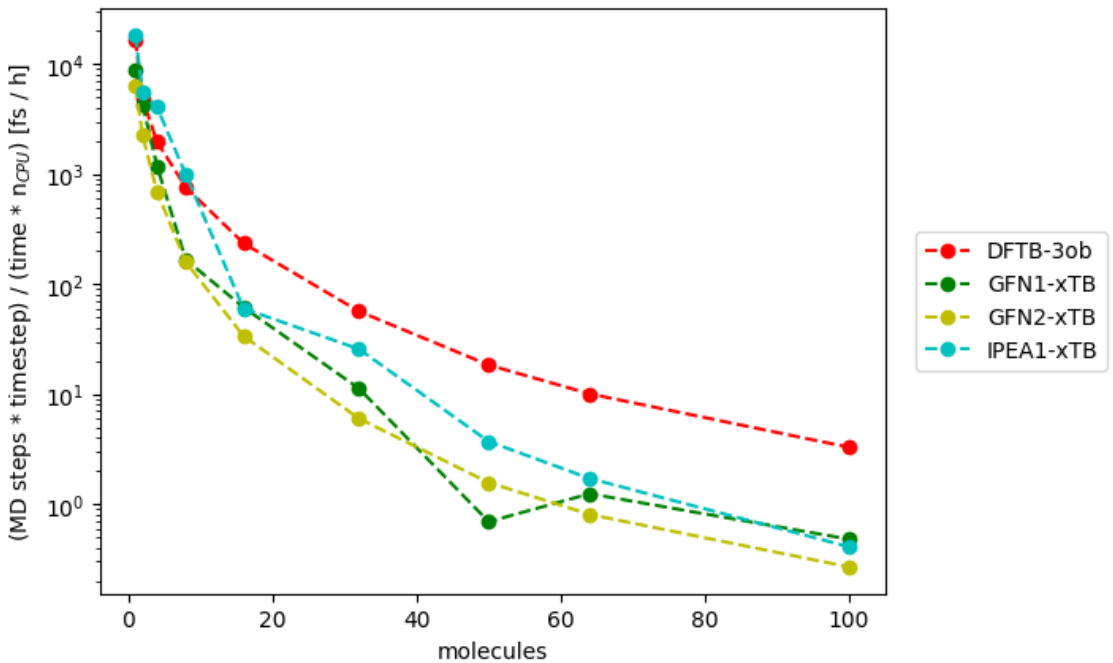
- parallelising is not so effective
- it may be problem with using MPI vs other implementations



Results for DFTB+

For MD:

- all xTB approaches have similar speed, slower than standard DFTB
- xTB has better accuracy, but this implementation has problems with memory



Summary

- parallelisation capabilities of selected software packages was studied
- maximal sizes of systems possible to study were identified
- behaviour of selected methods with changing number of CPUs was examined
- different MD approaches were compared