

CP2K-UK: Supporting Advances in Atomistic Simulation Capability

I. Bethune^{*1}, G. Gibb¹, L. Kantorovich², M. Watkins³, S. Chulkov³ and B. Slater⁴

^{*}i.bethune@epcc.ed.ac.uk



¹EPCC, The University of Edinburgh ²King's College London
³University of Lincoln ⁴University College London



Engineering and Physical Sciences Research Council

Project Overview

CP2K [1] is a freely available open-source program for atomistic simulation, best known for its implementation of the QUICKSTEP Gaussian and Plane Waves (GPW) linear-scaling Density Functional Theory (DFT) method [2]. However, CP2K provides a much wider range of capabilities, including classical pair-potentials and force-fields, QM/MM, semi-empirical Hamiltonians, hybrid density functionals and double-hybrid methods including Møller-Plesset 2nd order perturbation theory (MP2) [3], the Random Phase Approximation (RPA) [4], and GW [5]. Built on top of these Hamiltonians are an extensive set of simulation tools including Molecular Dynamics, Monte Carlo, Nudged Elastic Band, path integrals and free energy methods. Thus CP2K appeals to a wide range of users including Computational Chemists, Materials Scientists, Solid State Physicists and Biochemists. Usage of CP2K is growing, to the point where it is now the second most heavily used code on ARCHER, the UK national HPC service.

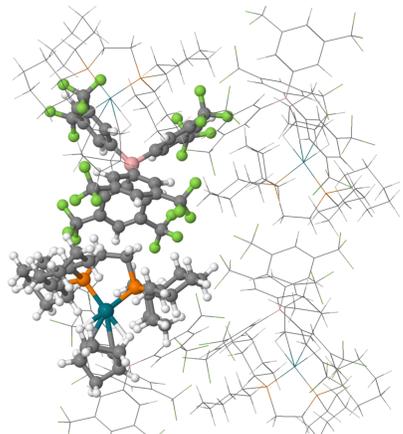


Figure 1: Computed unit cell of a sigma-alkane complex. Image from Prof. Stuart Macgregor

The CP2K-UK project, part of EPSRC's Software For The Future strategy, is a supporting the growing user community to make the best use of the capabilities CP2K has to offer. We have identified key barriers to usability, and are offering training and direct support to users to overcome there. In addition, we are building the capability of the user community to become developers of CP2K - adding functionality as required for their particular research goals.

CP2K-UK will run until 2018, and we present here some of the recent achievements of the project.

Support for Users

Networking: Our series of annual user group meetings have proved to be very popular. These include a mixture of talks from key CP2K developers, shorter talks focused on how to use particular functionality in the code, and three minute lightning talks, highlighting the breadth of research supported by CP2K. Slides from previous meetings are available at www.cp2k.org/docs#workshops

Training: The CP2K-UK project offers both through classroom-based training events, and bespoke training days with individual research groups. A group we visited in 2014 subsequently received a substantial award of CPU time through the ARCHER Resource Allocation Panel (RAP), leading to a publication [6].

Performance: CP2K offers excellent performance and scalability, but choosing the optimal set of parameters and parallelisation settings is not always straightforward. We carried out systematic benchmarking using a range of different calculation types on several HPC machines [7]. These results are available on the CP2K website along with scripts for plotting performance graphs (www.cp2k.org/performance), and we encourage users to contribute benchmarks from their own machines.

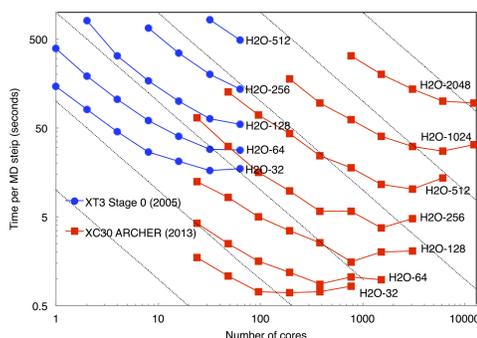


Figure 2: Scaling of DFT water benchmarks from 32 to 2048 molecules.

Tools: We have developed a visual web-based tool for editing CP2K input files, complete with tool-tip help, and automatic input validation. Within a few minutes, it is possible to create a new input from scratch, or start with an existing one and modify it, then download the resulting file and run it using CP2K. The Input Editor supports all CP2K releases since 2.5 and is available for use at:

cp2k-www.epcc.ed.ac.uk/cp2k-input-editor

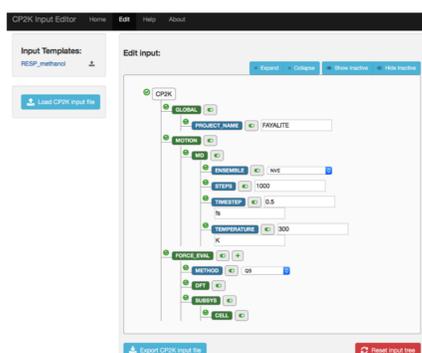


Figure 3: Creating a CP2K input using the CP2K Input Editor web tool

Support for Developers

We have worked on improving CP2K source code quality to make it easier for new developers to understand and contribute to the project. The entire CP2K code-base now contains automatically-generated 'doxygen' documentation, which can be browsed online at doxygen.cp2k.org. In addition, we have extended the CP2K automated regression test environment to cover use of the Intel Fortran compiler and MKL library, IBM's XL Fortran, OpenMPI, and macOS, and now have reliable results on these platforms.

We have also implemented several new methods in CP2K which will be of wide benefit to the community:

- Langevin Dynamics with arbitrary thermal regions [8]
- Filter matrix diagonalisation [9], a scheme where a minimal basis set is constructed dynamically during the SCF procedure which spans a new subspace, guaranteed to contain the minimum energy configuration.
- Fast linear response TDDFT with hybrid density functionals using the Auxiliary Density Matrix Method [10]
- DeltaSCF excited state calculations using the Maximum Overlap Method [11]
- Initialisation of MD velocities based on vibrational analysis to avoid long equilibration runs [12]

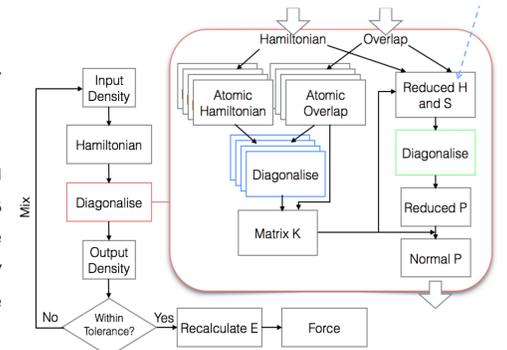


Figure 4: Schematic of the key operation of the filter matrix diagonalisation method as implemented in CP2K

Get Involved

CP2K-UK's primary purpose is to benefit the CP2K community - so if you are already using CP2K, have specific features or methods you want to find out more about, or are just starting to investigate CP2K, then we would like to hear from you! Here are some of the ways you can get involved:

- Want to get started with CP2K? Contact us to arrange a hands-on introduction to CP2K at your institution, with specific focus on your application areas.
- The CP2K website (www.cp2k.org) is now a wiki and anyone can contribute. Please make improvements, contribute tutorials based on your own experience, or ask if you think information is missing.
- Have an idea for a new feature or tool that you'd like to see implemented in CP2K? Talk to us about software development funding opportunities and training.
- If you are applying for a research grant involving CP2K, or wish to use CP2K on ARCHER, contact us to see how we can support your application - letters of support are available!
- To stay in the loop about upcoming CP2K events, we maintain a notification-only email list. Contact us to subscribe.

For any of the above, or if you have other questions about the CP2K-UK project, please email i.bethune@epcc.ed.ac.uk

- [1] Hutter *et al*, WIREs Computational Molecular Science 4(1):15-25 (2013)
- [2] VandeVondele *et al*, Comp. Phys. Comm. 167, 103 (2005)
- [3] Del Ben *et al*, JCTC 8(11) (2012)
- [4] Del Ben *et al*, Comp. Phys. Comm. 187, 120-129 (2014)
- [5] Wilhelm *et al*, JCTC 12 (2016)
- [6] Chadwick *et al*, Angew. Chem. Int. Ed., 55 (2016)
- [7] Bethune *et al*, Proceedings of the Cray User Group (2014)
- [8] L. Kantorovich, Phys. Rev. B, 78(9) (2008)
- [9] Rayson and Briddon, Phys. Rev. B, 80(20) (2009)
- [10] Guidon *et al*, JCTC 6(8) (2010)
- [11] Gilbert *et al*, J. Phys. Chem. A 122(50) (2008)
- [12] West and Estreicher, PRL 96 (2006)

CP2K-UK is supported by the Engineering and Physical Sciences Research Council [grant number EP/K038583/1]. This work used the ARCHER UK National Supercomputer Service (www.archer.ac.uk).