CP2K-UK: Supporting Advances in Atomistic Simulation Capability

Iain Bethune*[1], Lev Kantorovich[2], Lianheng Tong[2], Matt Watkins[3], Sergey Chulkov[3] and Ben Slater[4]

*ibethune@epcc.ed.ac.uk [1]EPCC, The University of Edinburgh [2] King's College London [3] University of Lincoln [4] University College London

Project Overview

CP2K [1] is a freely available (www.cp2k.org) 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 potentials, QM/MM, semi-empirical models, hybrid DFT functionals, and most recently Møller-Plesset 2nd order perturbation theory (MP2)[3] and the Random Phase Approximation (RPA)[4]. Built on top of these models 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.

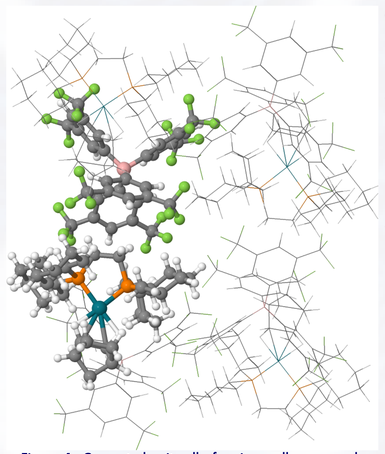

Figure 1: Computed unit cell of a sigma-alkane complex. Image courtesy Prof. Stuart Macgregor.

The CP2K-UK project, part of EPSRC's Software For the Future strategy, is supporting the growing user community to make 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 these. 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 from 2013 to 2018, and we present here some of the recent achievements of the project.

Support for Users

Networking: There is already a great deal of experience in using CP2K embedded within research groups around the UK, and part of the role of CP2K-UK is to offer opportunities for new users to make contact with existing users. 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 this year's meeting are at www.epcc.ed.ac.uk/content/cp2k-uk-workshop-2016.

Training: The CP2K-UK project offers training in the usage of CP2K, both through classroom-based training events, and bespoke training days with individual research groups. A group we visited in 2014 subsequently applied for an ARCHER 'Instant Access' grant to test out CP2K at scale, and recently received a substantial award of CPU time through the ARCHER Resource Allocation Panel (RAP), leading to a publication [5] .

Performance: CP2K offers excellent performance and scalability, but choosing the optimal set of parameters for parallelisation is not always straightforward. To help address this, we carried out systematic benchmarking using a range of different calculation types on different HPC machines, and published a paper on CP2K performance [6]. These results have been extended and made available on the CP2K website along with scripts for plotting performance graphs (www.cp2k.org/performance), and we encourage users to contribute benchmark results from their own machines .

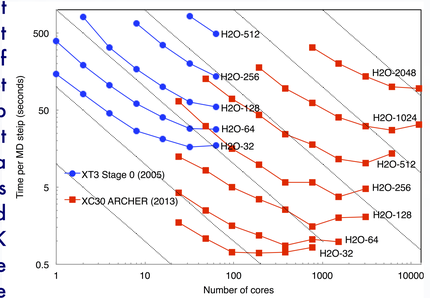

Figure 2: Scaling of water benchmarks from 32 to 2048 molecules on ARCHER (and a Cray XT3 from 2005).

Support for Developers

We have worked on improving CP2K 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, and now have reliable results on these platforms.

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

- Langevin Dynamics with arbitrary thermal regions, following the method in [7].
- Filter matrix diagonalisation [8], 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 [9].
- DeltaSCF excited state calculations using the Maximum Overlap Method [10].
- Initialisation of Molecular Dynamics Velocities based on vibrational analysis to avoid long equilibriation runs [11].

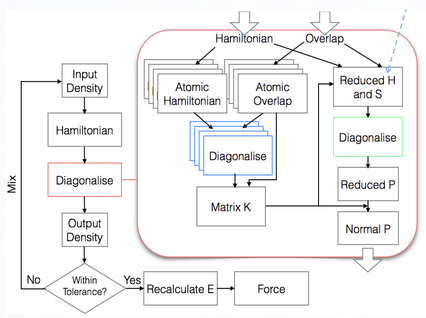

Figure 3: Schematic showing the key operations of the filter matrix diagonalisation method as implemented in CP2K.

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 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 your 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.
- 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 ibethune@epcc.ed.ac.uk

[1] Hutter *et al*, WIREs Computational Molecular Science 4(1):15-25 (2013)
[2] VandeVondele *et al*, Computer Physics Communications 167, 103 (2005)
[3] Del Ben *et al*, Journal of Chemical Theory and Computation 8(11) (2012)
[4] Del Ben *et al*, Computer Physics Communications 187, 120-129 (2014)
[5] Chadwick *et al*, Angew. Chem. Int. Ed., 55 (2016)
[6] Bethune *et al*, Proceedings of the Cray User Group (2014)
[7] L. Kantorovich, Physical Review B, 78(9) (2008)
[8] Rayson and Briddon, Physical Review B, 80(20) (2009)
[9] Guidon *et al*, Journal of Chemical Theory and Computational 6(8) (2010)
[10] Gilbert *et al*, Journal of Physical Chemistry A 122(50) (2008)
[11] West and Estreicher, Physical Review Letters 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 (http://www.archer.ac.uk).

EPSRC Engineering and Physical Sciences Research Council

KING'S College LONDON

UCL

epcc