



Scaling up and coupling adaptive kinetic Monte Carlo with large-scale DFT ARCHER2-eCSE05-4

Prof. C. Skylaris^a, G. Teobaldi^{a,b} and J. A. Purton^c

- a. School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK
- b. Scientific Computing Department, STFC-UKRI, Rutherford Appleton Laboratory, Didcot, OX11 0QX, UK
- c. Scientific Computing Department, STFC-UKRI, Daresbury Laboratory, Keckwick Lane, Warrington, WA4 4AD, UK

Abstract

An interface was written to combine the functionality of an adaptive kinetic Monte Carlo program (ACDC) and an ab initio electronic structure program (ONETEP). The merger of the two programs was motivated by the limitations in empirical potentials and allowed for the exploration of the potential energy surface of adatoms on the MgO(001) and Al(001). We examined two schemes for optimising the utilisation of the ARCHER2 platform. Due to the nature of the transition state searches (differing number of force evaluations), task farming was observed to be the most efficient. In addition, this approach is more fault tolerant and more extensible in the longer term.

Keywords: adaptive Kinetic Monte Carlo, Density Functional Theory, ion diffusion, ONETEP, ARCHER2 and eCSE.

1. Introduction

Many critical technological processes and materials rely on solid-state diffusion such as the growth of thin films (magnetic recording media, electronic semiconductor devices, LEDs, optical coatings), batteries, nuclear materials, catalysis, and solid oxide fuel cells. Therefore, there is a need to understand processes for the informed design of new devices: timescales of milliseconds are often required for industrial applications. Unfortunately, solid state diffusion is very slow and challenging to study with “standard” methods (*e.g.* molecular dynamics, **MD**). Especially at low temperatures, MD-based sampling becomes inefficient as there is insufficient thermal energy to overcome the activation energy and the configuration will remain in or close to the initial state.

Kinetic Monte Carlo (**KMC**) attempts to overcome this limitation by exploiting the fact that the long-time dynamics of this kind of system typically consists of diffusive jumps from state to state. Rather than following the trajectory through every vibrational period, these state-to-state transitions are treated directly. The result is that KMC can reach vastly longer time scales, typically seconds. The most common atomistic KMC models employ an on-lattice approximation which limits their ability to describe a system in which defects are not exactly

on the pre-defined lattice (*e.g.* interstitials) or systems which undergo significant deformation (*e.g.* materials that have been subjected to radiation damage). These atomistic KMC models also require a list of possible event mechanisms and rates (a rate catalogue) that have been determined *a priori*, through experimental and theoretical methods, or even by guessing. This is naturally a severe limitation to such a method, as atomic motion is not necessarily intuitive and can be extremely difficult to predict in advance. Furthermore, as the simulation advances, the structure of the system will change, altering the kinetics of the mechanisms. Henkelman and Jónsson [1] have proposed a variation on the KMC method, in which one builds the rate catalogue on the fly for each state (called *adaptive* KMC). This method, which is an *off-lattice* approach analogous to MD, has the key requirement of having an efficient way to search for saddle points that are connected to the current state of the system. Adaptive kinetic Monte Carlo (aKMC), as implemented in ACDC, determines the transition states and activation energies using a minimum mode following search (*e.g.* dimer method or activation relaxation technique) [2] so that neither the mechanism nor the end point of a process is known prior to the search.

Prior to this project, ACDC was only capable of calculating the energy and forces using the rigid ion model [3]. This is very fast but, as rigid ion models are often obtained from fitting to experimental properties at equilibrium conditions, of questionable accuracy transition states. Moreover, potential functions are not available for many systems (*e.g.* interfaces between metals and oxides). The primary objective of the work was to interface ACDC with the linear-scaling Density Functional Theory code ONETEP [4] and therefore to increase the accuracy of the calculated potential energy surface and broaden the scope of the aKMC.

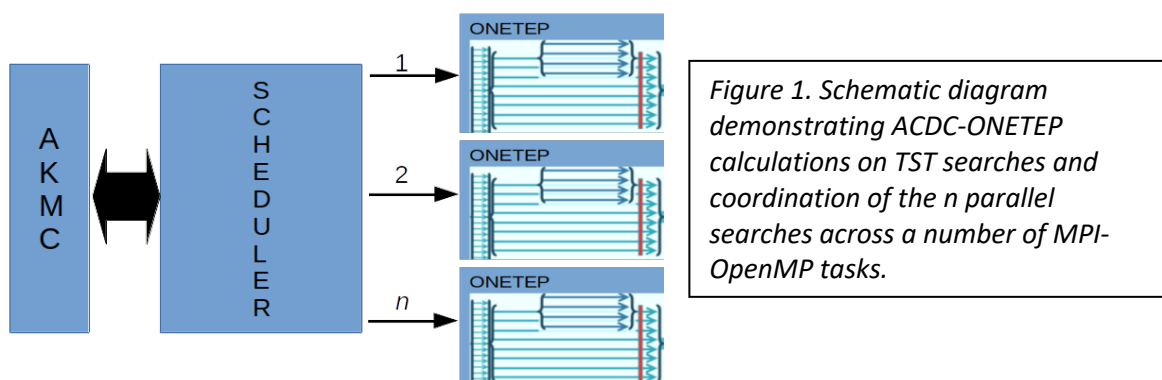
2. Methodology

To meet the primary objective described above, an interface was created so that the C++ of ACDC (written in C++) could talk to ONETEP (written in FORTRAN). The API was written to be as general as possible and easily extendable to other codes (*e.g.* DFTB+, CASTEP, BigDFT) and employs the iso-C/FORTRAN bindings. ONETEP had to be modified to receive data (*e.g.* execution commands, atomic positions) and return the energies and forces evaluated using DFT. Thus, the principal alteration to ONETEP was to turn the initialisation functions into a FORTRAN module that could manage the ab initio calculations depending on the instructions passed. In addition, a suitable MPI communicator was created in ACDC and passed to ONETEP and the memory management of FORTRAN allocatable arrays was adapted so that the ONETEP module could be repeatedly called from ACDC.

The probability of finding the lowest saddle point and the saddle point distribution is unknown and the potential energy surface can be very complex. In addition, to accurately model the diffusion processes and precisely determine the KMC clock time, it is necessary to sample a large number of possible saddle points. This may range from a few tens of saddle points to several thousands depending on the nature and size of the system. In addition, it has been estimated [2] that up to 50% of the searches fail by the negative eigenvalue turning positive or falling back into the initial energy basin. Fortunately, each search is independent of the others, and multiple transition point searches using ONETEP to determine the energy and forces can be launched simultaneously to occupy thousands, if

not, hundreds of thousands of cores. The coordination of numerous searches is not trivial due to the high rate of failure in searches and one does not know the appropriate number of workgroups at the start. Moreover, it is difficult to determine whether a search is adequately sampled (self-consistent schemes have been developed for this purpose [5]). In this project we evaluated two schemes

1. The first scheme employed a single program that instantiated a given number of “ONETEP objects” that received an appropriate MPI communicator from the main program. This approach was found to be most suitable for fast force fields searches which are relatively quick (approximately 60 seconds for 2000 atoms using 16 threads) and employ fewer nodes than ONETEP based searches. However, this method proved to be inefficient using electronic structure methods owing to the very different runtime of each transition state search. Also, the method was less fault tolerant.
2. A second approach employed a scheduler (Fig. 1) to task farm multiple searches. This method improved the load-balance for the ONETEP searches by (re-)scheduling resources and minimizing idles nodes until sufficient valid searches have been completed. Moreover, the method is more fault tolerant as a failure on a single node does not jeopardise the whole simulation, which can be restarted from a checkpoint file. Also, it is likely to be more flexible for future developments (*e.g.* potentials obtained from machine learning).



Example calculations

Using the above interface, we have examined the diffusion of atoms on the Al (001) and MgO (001) surfaces with simulation cells containing 110 and 268 atoms and exploiting between 5,120 and 10,240 cores, respectively. The number of cores used reflects the available resources and can easily be extended to a greater number of cores. The calculated activation energies from ACDC-ONETEP are in good agreement with those determined from nudged elastic band calculations using VASP [6].

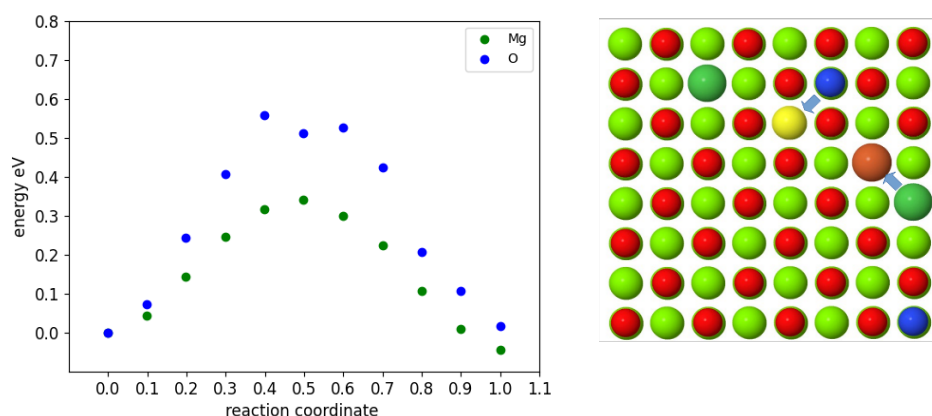


Figure 2. Calculated activation energies for the diffusion of a single oxygen and magnesium ion on the (001) surface of MgO, shown on the right. The surface oxygen and magnesium ions are shown as blue and dark green spheres respectively. The trajectory of the diffusing ions is marked by the blue arrows and the final positions of the ions are displayed in yellow (oxygen) and brown (magnesium).

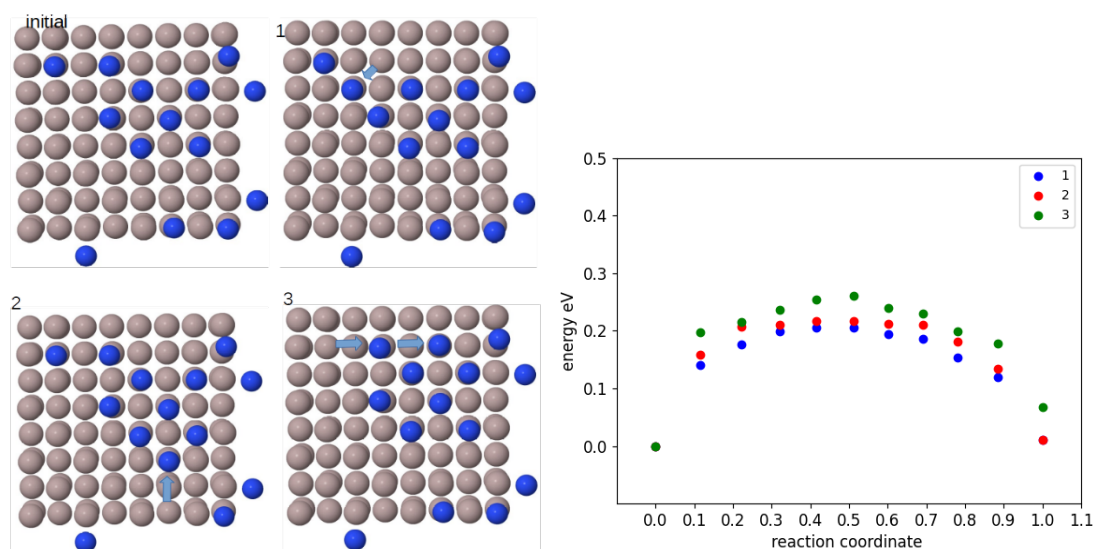


Figure 3. The initial configuration of several Al ad-atoms (blue) on the Al (001) surface. The saddle points are shown for the new configurations 1, 2 and 3. The blue arrows indicate the direction of the displacement. As self-evident, different initial states entail differently correlated ionic displacements, which are effectively impossible to anticipate without running the simulation and are strongly dependent on the accuracy of the method used to compute energies and atomic forces.

In summary, we have demonstrated that it is feasible to interface ACDC to ONETEP and calculate the activation energy of diffusing ions in solid state materials. Despite requiring large resources, it is possible to undertake adaptive Kinetic Monte Carlo simulations with a higher degree of accuracy than previously possible based on empirical interatomic potentials. The method is suitable for scaling to a very large number of cores and will benefit from the drive to exa-scale computing as well as advances in DFT-derived machine-learned interatomic potentials.

Acknowledgements.

This work was funded under the embedded CSE programme of the ARCHER2 UK National Supercomputing Service (<http://www.archer2.ac.uk>).

References

- [1] G. Henkelman and H. Jónsson, (2001) *J. Chem. Phys.*, **115**, 9657 and G. Henkelman and H. Jónsson, (2003) *Phys. Rev. Lett.*, **90**, 116101.
- [2] M. Trochet et al., *Handbook of Materials Modeling*, https://doi.org/10.1007/978-3-319-42913-7_29-1
- [3] G. V. Lewis and C. R. A. Catlow, (1985) *J. Phys. C: Solid State Phys.*, **18**, 1149
- [4] J. C. A. Prentice, ..., C-K. Skylaris (36 authors) (2020) *J. Chem. Phys.*, **152**, 174111.
- [5] D. Aristoff, S. Chill and G. Simpson, (2016) *Comm. in Appl. Math. and Comp. Sci.*, **11**, 171-186
- [6] G. Henkelman, et. al., (2005) *Phys. Rev. B*, **72**, 115437 and G. Geneste, et. al., (2005) *J. Chem. Phys.*, **122**, 174707.