

# Calculating Activation Energies in Diffusion Processes Using a Monte Carlo Approach in a Grid Environment

Mark Calleja and Martin T. Dove

Mineral Physics Group, Department of Earth Sciences, University of Cambridge,  
Downing Street, Cambridge, CB2 3EQ, United Kingdom

**Abstract.** A Monte Carlo based method due to Mishin [1] for obtaining activation energies for ionic transport processes has been implemented for empirical-potential models and used in conjunction with Condor and Globus tools. Results are shown for  $\text{Na}^+$  migrating in quartz, and  $\text{O}^{2-}$  vacancy hopping in  $\text{CaTiO}_3$ . We also describe the extensions at the Condor/Globus interface we found necessary to implement in order to facilitate transparent file transfer with Condor-G job submission.

## 1 Introduction

An understanding of the activation energies present in processes involving atomistic motions in crystalline materials is necessary if an accurate quantitative study of the kinetic properties in such systems is to be made. For example, such information is of use in determining ion-vacancy diffusion behaviour [2]. One approach to obtain an activation energy for diffusion in a crystal is via lattice energy minimisation, where an atom is pushed through the structure in steps, and the energy of the relaxed structure calculated at each step. The main disadvantage of this approach is that it is strictly only applicable for zero temperature, and therefore cannot be used when there is significant thermal disorder or where the structure changes with temperature. Moreover, for very large systems, the energy minimisation process may take a long time due to the existence of a complex and shallow energy surface. Alternatively it may be possible to study diffusion in crystalline materials using traditional molecular dynamics (MD) schemes [3], which have been successfully used for the study of atomic diffusion in fluid phases. However, these methods generally sample many nano-seconds of simulation time, which is not always sufficient to observe such processes. Some effort has been directed to modifying MD schemes to measure these quantities [4]. Temperature accelerated dynamics methods can be considered to fall in this category [5, 6]. In some cases high-temperature behaviour can be inferred albeit in an approximate way [7], such as by driving diffusion with an applied field.

Mishin suggested a scheme based on the Monte Carlo (MC) method which has been applied to the study of ionic diffusion at grain boundaries [1]. The attraction of this approach is that since it is MC based, the method parallelises trivially

and can be deployed on the resources in our mini-grid. The rest of this article is hence divided as follows: we first describe how we have implemented this method to determine the activation energies associated with moving a tracer atom from a given crystallographic site towards a target site. Next, the simulation lifecycle is described, from job submission using Condor-G to the problems we faced with automatic file retrieval from the globus gatekeeper, and how we set about solving this problem. We illustrate the work with sample results from studies of the diffusion of alkali cations in a perfect crystal of quartz and within domain walls in quartz, and of studies of oxygen vacancies in the perovskite  $\text{CaTiO}_3$ .

## 2 The Physical Model

The method studied here is a modification on the standard Metropolis Monte Carlo method [8]. In this scheme, an ion is selected for migration towards a chosen location, *e.g.* a  $\text{Na}^+$  ion migrating along a channel in quartz, or an  $\text{O}^{2-}$  ion hopping into a vacant site. The diffusing ion is marched towards its target in small steps. At every step the entire crystal is allowed to relax through the MC algorithm, except that the migrating ion is constrained to relax perpendicular to its last jump direction. This is important since the locus of equilibrium position of the jumping atom is not necessarily a straight line. In practise, the migrating ion is only forcibly marched to within  $\sim 90\%$  of its destination before being allowed to relax without constraint. This is to allow the true energy minimum to be reached, which will generally not be at the specified target location. Sample temperature is incorporated through the MC algorithm, and the structure will relax in a way that incorporates thermal fluctuations corresponding to the preset temperature.

Although this method should work independently of the potential models employed, we have used simple empirical potentials. These are much quicker, though not necessarily as accurate, as *ab initio* methods (*e.g.* see [9], although only the  $T = 0$  K case is studied by these authors). The interatomic potentials used here include the simple two-body Buckingham plus Coulombic potentials, with the following functional form for ions  $i$  and  $j$  separated by distance  $r$ :

$$V(r) = A_{ij} \exp(-b_{ij}r) - \frac{C_{ij}}{r^6} + \frac{q_i q_j}{4\pi\epsilon_0 r} \quad (1)$$

Long range interactions are handled using the well-known Ewald method [10]. Note that the method described here is not limited to two-body potentials, and our implementation supports three-body potentials (such as bond-angle dependent potentials).

After every hop of the migrating ion, the whole system is allowed to equilibrate before statistics are gathered. The number of moves per ion required for equilibration are highly system dependent, but are generally  $\sim 1000$  per ion for every migration hop. The hop distances employed for the migrating ion are typically 0.1-0.2 Å, depending on the amount of detail required from the profile.

At the end of all the parallel runs it is then possible to obtain an average energy value at each hop point. Hence note that each point is the result of two averaging processes, the first due to the statistics gathering from an individual simulation at a given hop, and the second when the values due to different simulation runs are compared for the same point. Due to the increased thermal fluctuations at high temperatures, the variance on the averages will increase for higher temperatures, meaning that if a uniform level of accuracy is required across a range of operating temperatures it will be necessary to run more configurations at the higher temperatures.

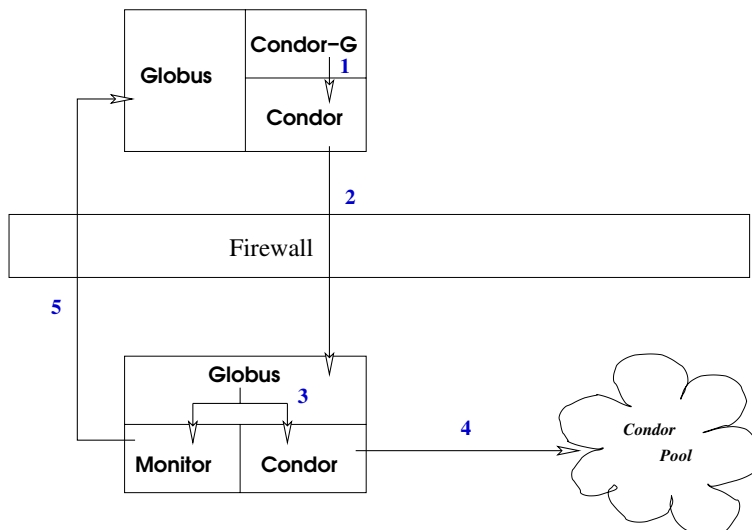
It is important to note that what we measure here is the interaction energy associated with this diffusive process, and not the free energy. To measure the latter one needs to take into account a small entropic contribution, which can be done using the method of thermodynamic integration [11]. This requires many runs at different temperatures, starting with the model  $T = 0$  K system up to the required temperature.

### 3 Job Submission on a Minigrid

Since the MC method is inherently statistical in nature, sampling the whole of phase space and not only the minimum-energy pathway, it is essential to run many calculations and obtain averages across many configurations for each step along the diffusion pathway. This approach lends itself easily to parallelisation, running different calculations on different processors.

In our studies we generally set off an identical number of initial configurations, but with different seeds for the random number generators, on a number of processors. Individual jobs were submitted to a local machine which had the required Condor [12] submission daemons running as well as a Globus installation [13]. Various versions of these tools were utilized, though as of writing they currently are 2.4.3 (Globus) and 6.6.0 (Condor). The jobs themselves were submitted using Condor-G rather than Globus's own interface since this allowed for simple workflows to be implemented using DAGMan (Directed Acyclic Graph Manager), a meta-scheduler for Condor. Figure 1 shows the scheme employed in this work. This approach was chosen since it provides a single point of access for remote facilities via the relevant globus server which helps keep firewall issues to a minimum (as opposed to, say, using Condor glideins).

The process begins with the submission of a condor job script (step 1), which specifies that the job should be run in the globus universe. The only non-standard feature of this script is the extra RSL value, `condor_home_dir`, with which we specify the directory we would like the output files to be returned to on this submitting machine. This tag is our own addition, and we'll say more about it below. The job is now submitted to the local condor pool. In step 2, the job is forwarded to the remote gatekeeper, invariably across a firewall. It is the gatekeeper's task to fork a relevant jobmanager (e.g. a condor job for an Intel/Linux machine) which submits the job to the condor pool local to the gatekeeper. For our set-up we found that the default `condor.pm` file used by the jobamanger (usually



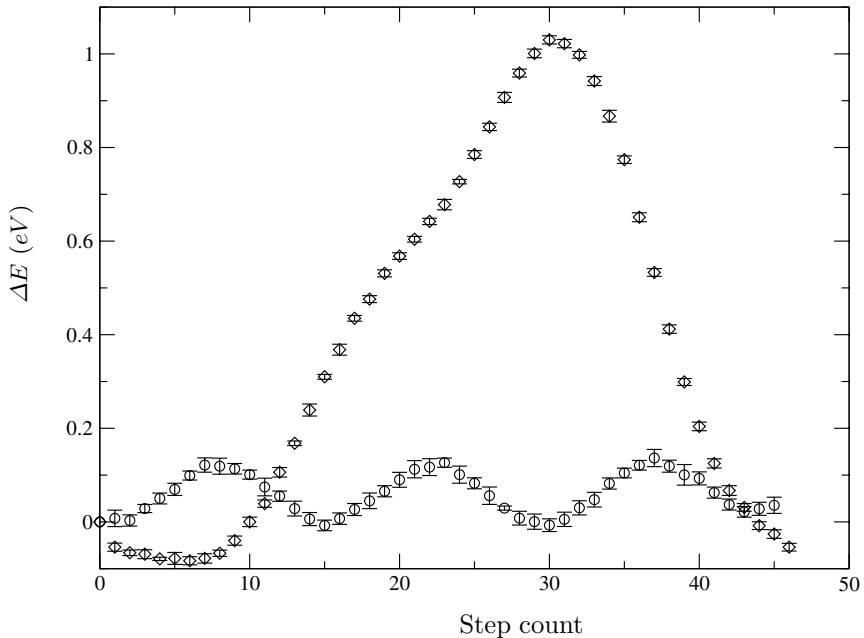
**Fig. 1.** Job lifecycle employed in this work. See text for description of workflow

located in `$GLOBUS_LOCATION/lib/perl/Globus/GRAM/JobManager`) does not provide sufficient functionality for file transfer other than `stdin`, `stdout` or `stderr`. Hence, to circumvent this we modified `condor.pm` so that it forks off a new process (step 3) which is then exec'd with a new monitoring job whose duty is to periodically check for the completion of the job in condor pool. Also, the `condor_home_dir` attribute is added to the `condor.rvf` file in `$GLOBUS_LOCATION/share/globus_gram_job_manager`. All files generated by such a job are placed in a temporary directory so that they cannot be automatically removed by condor when the job completes, or it may do so before we've had a chance to send the output files back to the submitter. In the meantime, the original job is committed to a suitable resource within the pool (step 4). After periodically sleeping, the monitoring process detects that the condor job has finished and scans the temporary directory designated for all job files, discriminating between input from output files. It then tries to send the output files back to the submitting machine in the directory given in the `condor_home_dir` value using `gsiftp` (step 5). The latter is used since the `https` socket used for initially sending input data over may now be closed and no longer available. This means that we require access through any firewall protecting the submitting machine. On the successful completion of all file transfer the monitor cleans up by removing the temporary directory and all its contents before exiting.

## 4 Results

We present some typical examples that we have studied (all at 0 GPa), with each plot being the result of eight parallel runs. Figure 2 shows the energy profile for

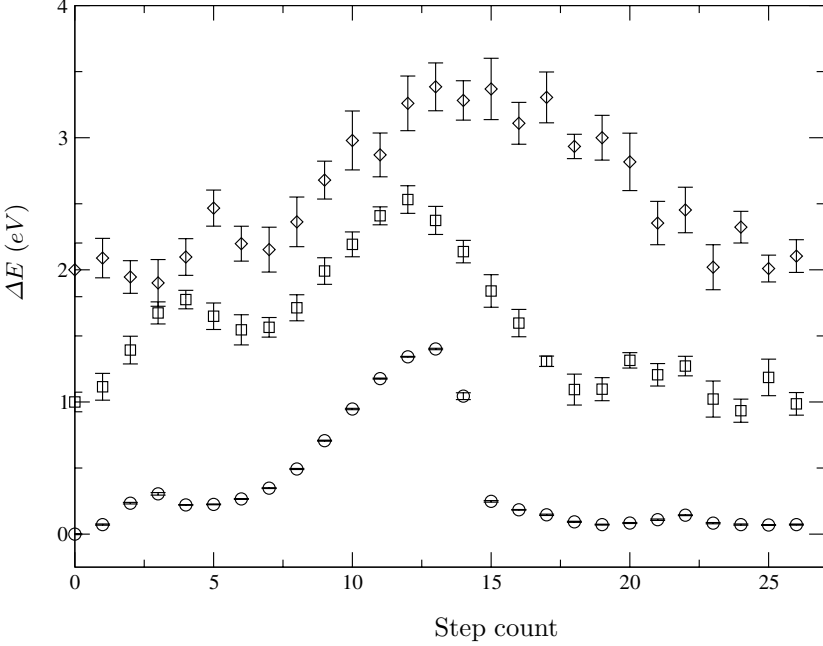
a  $\text{Na}^+$  ion migrating along the  $c$ -axis in alpha bulk and twinned quartz at 10 K (potential details are given in [14]). The abscissa denotes the hop count, and cannot readily be associated with a jump distance since the hopping atom does not necessarily follow a straight path, although the paths followed in the quartz systems studied here are relatively straight (unlike the perovskite model studied below). The ordinate gives the net change in crystal energy for the processes. Note how the three small energy barriers due to the three sets of oxygen pairs in the helical structure of the bulk can be easily distinguished. By comparison, the energy barrier seen in a twin wall is an order of magnitude greater, explaining why cationic conductivity is poisoned along this axis by twin walls [7].



**Fig. 2.** Activation energy for  $\text{Na}^+$  ion diffusing for one unit cell along [001] in quartz at 10 K. Circles are for motion in the bulk, diamonds in a wall

In Figure 3 we show a similar process, but this time the sodium ion is migrating along the [100] direction at 10, 150 and 300 K in bulk quartz. An interesting, temperature-dependent, feature is observed with the shoulder at  $\sim 0.75\text{eV}$ , which appears to develop on heating to 150 K before subsiding again on further heating to 300 K. Note the small bump at around the 23rd step in the 10 K profile; this is an example of the diffusing ion being pushed too far, so that by

this step it's being displaced away from its equilibrium position. This situation is rectified by the next step, when it is allowed to relax without any constraint.

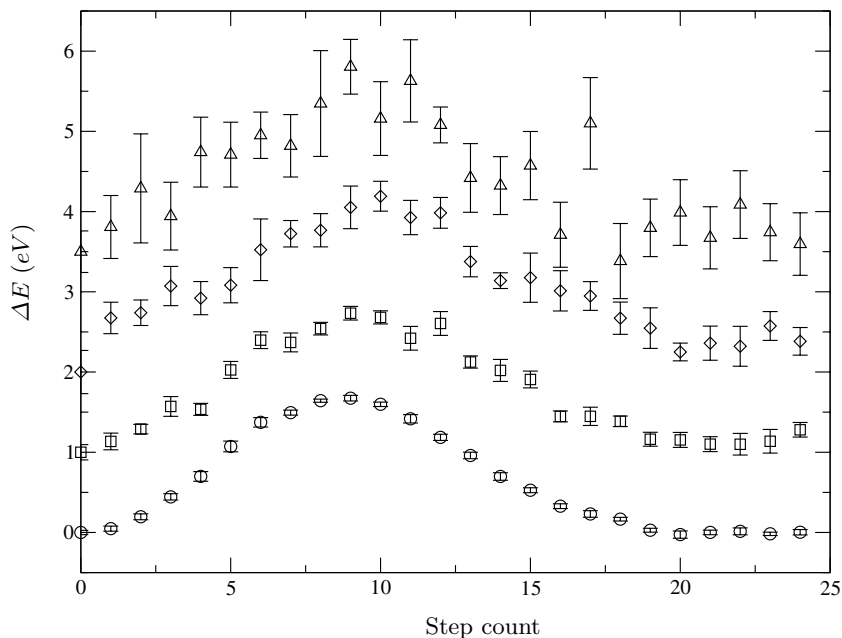


**Fig. 3.** Energy profile for  $\text{Na}^+$  ion migrating along  $[100]$  in quartz at 10 (bottom), 150 (middle) and 300 (top) K

For the next system studied we show the energetics associated with an  $\text{O}^{2-}$  ion migrating towards a nearby vacancy in bulk orthorhombic  $\text{CaTiO}_3$  for various temperatures (potential details are given in [15]). Figure 4 shows the corresponding profiles for runs at 100, 300, 500 and 1000 K. Even at the highest temperature, the saddle point energy can be identified reasonably clearly. Note the apparent weak temperature dependence, with the peak values rising with increasing  $T$ .

## 5 Discussion

A simple Monte Carlo based method due to Mishin has been shown to provide accurate activation and saddle point energies even at relatively high temperatures. Obviously the temperature range is highly system dependent, with larger



**Fig. 4.** Energy profile for  $O^{2-}$  ion migrating  $CaTiO_3$  at (from bottom to top) 100, 300, 500 and 1000 K

activation energies being easier to characterise at high  $T$  than small barriers. Furthermore, the method readily lends itself to be deployed within a grid infrastructure, though not without extending current technologies. However, we generally find that the combination of Condor, Globus and DAGMan tools provide a ready fabric for making the most of Monte Carlo simulations on a grid, with only minor modifications and extensions.

**Acknowledgements.** We would like to thank Mark Hayes (CeSC) and Rik Tyler (CCLRC) for useful discussions and the NERC (UK) (grant: Environment from the molecular level) for funding this work.

## References

1. Y. Mishin, *Defect and Diffusion Forum*, **143 - 147** (1997) 1357
2. R. J. Borg and G. J. Dienes, *An Introduction to Solid State Diffusion*, Academic Press (1988)
3. W. Smith, C. W. Yong, P. M. Rodger, *Mol. Simulat.*, **28** (2002) 385

4. S. C. Parker, N. H. de Leeuw, E. Bourova and D. J. Cooke, *Rev. Mineral Geochem.*, **42** (2001) 63
5. A. F. Voter *J. Chem. Phys.*, **106** (1997) 11
6. R. A. Miron and K. A. Fichthorn *J. Chem. Phys.*, **119** (2003) 6210
7. M. Calleja, M. T. Dove and E. K. H. Salje, *J. Phys.: Condens. Matter*, **13** (2001) 9445.
8. J. M. Thijssen, *Computational Physics*, Cambridge (1999).
9. A. I. Lichtenstein, R. O. Jones, H. Xu and P. J. Heaney, *Phys. Rev. B* **58** (1998) 6219
10. P. Ewald, *Ann. Phys.* **64** (1921) 253
11. M. C. Warren, M. T. Dove, E. R. Myers, A. Bosenick, E. L. Palin, C. I. Sainz-Diaz, B. S. Guitton and S. A. T. Redfern, *Mineral. Mag.*, **65** (2001) 221–248
12. T. Tannenbaum, D. Wright, K. Miller and M. Livny, "Condor - A Distributed Job Scheduler", in *Beowulf Cluster Computing with Linux*, MIT, (2002)
13. I. Foster and C. Kesselman, *Intl J. Supercomputer Applications*, **11(2)**, (1997) 115-128
14. G. J. Krammer, N. P. Farragher, B. W. H. van Beest and R. A. van Santen, *Phys. Rev. B* **43** (1991) 5068
15. M. Calleja, M. T. Dove and E. K. H. Salje, *J. Phys.: Cond. Matt.* **15** (2003) 2301