Nonequilibrium Rare-Event Sampling

Recommended and a commentary by Mike Allen, University of Warwick

This paper by Berryman and Schilling [1] has been in the literature for a year now, and is a valuable addition to the battery of techniques available to study nonequilibrium processes by molecular simulation; but it seems not to have attracted much attention. In the challenging field of "rare events", there are three broad target areas, of increasing degree of difficulty: (a) systems in equilibrium; (b) systems in a nonequilibrium steady state; (c) systems in a nonequilibrium, nonsteady state. The approach proposed here is aimed at the last, and perhaps most interesting category, and in tackling questions such as: what happens, at the microscopic level, to a system after a sudden quench, or shock?

For equilibrium systems, it makes sense to speak of trajectories across a free energy landscape, which may be defined in terms of a few coarse-grained order parameters. Interest lies, firstly, in the determination of residence times and rate constants, and secondly in the analysis of the reactive trajectories themselves, possibly including the construction of a simplified kinetic model. A considerable literature exists regarding the efficient sampling of reactive trajectories in this situation, for example by Transition Path Sampling [2]. The extension to driven systems which are in steady nonequilibrium states, perhaps with the aim of calculating the flux associated with an infrequent event, poses a problem for any method that relies on prior knowledge of the steady-state distribution. The situation is even more challenging if the distribution itself is changing in time.

One of the few approaches that makes no assumptions about this distribution is Forward Flux Sampling (FFS) [3]. This employs the common technique of dividing the phase space with a set of non-intersecting interfaces, level sets of an approximate "reaction coordinate", and generating trajectories that connect each interface with the next. Provided the dynamics contains a stochastic element, each starting point on a given interface can generate a whole family of equally valid ongoing partial paths. Full trajectories are constructed by combining these partial paths together. This gives the opportunity of preferentially sampling the paths, to favour the reactive ones, while at the same time taking proper account of the relative statistical weight of reactive and nonreactive trajectories. The overall trajectories, and associated statistical weights, are built up progressively from one interface to the next, in a manner reminiscent of some Monte Carlo methods for constructing polymer chain configurations [4, Chapter 11]. This can lead to a phenomenal improvement in the estimation of reaction rates, but there are some drawbacks. The paths tend to be highly correlated, with many product states having their origins in just a few starting states. The actual time taken for each partial trajectory will be different, and so the method is not well suited to tackling explicitly time-dependent problems. Also, in practice, the number and locations of the interfaces should be optimised [5], which relies on collecting some preliminary information from the runs.

Berryman and Schilling tackle the last two of these issues by abandoning the restriction that partial paths should begin and end at an interface, and instead making them of fixed duration in time. Attention is switched from the interfaces to the bins or compartments delimited by them. The number of new partial trajectories starting in each bin is calculated, at each time step, in a way that generates a sufficiently large, and constant, number of forward-going trajectories: this is the factor biasing the system towards rare-event sampling. In addition, starting points within the bins are chosen in a bias-free way that reflects their statistical weight. This effectively compensates for the fact that the paths that arrived in a given bin will have different weights, because they originated in any of several other bins. The method uses the information about the originating bin of a configuration, in the previous step, but never needs to look further back than this, because ongoing trajectories have equal statistical weights, by construction.

The algorithm is simple to implement, and the authors present arguments that it scales well as the number of bins becomes large, allowing a fine-grained description of the order parameter. The various conditional probabilities associated with partial paths may be used to discuss coarse-grained kinetic models (whether a Markovian assumption is involved or not). Naturally, as for other methods of this kind, including FFS and Transition Interface Sampling [6], it will work best if the chosen "reaction coordinate" is not too different from the "true" one. Nonetheless, the method offers exciting new prospects in the modelling of nonequilibrium processes in a wide variety of systems.

References

[1] J. T. Berryman and T. Schilling, Sampling rare events in nonequilibrium and nonstationary systems, *J. Chem. Phys.*, 2010, **133**, 244101/1–10.

[2] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Transition path sampling: throwing ropes over rough mountain passes, in the dark, *Ann. Rev. Phys. Chem.*, 2002, **53**, 291–318.

[3] R. J. Allen, D. Frenkel, and P. R. ten Wolde, Simulating rare events in equilibrium or nonequilibrium stochastic systems, *J. Chem. Phys.*, 2006, **124**, 024102/1–16.

[4] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications*, Academic Press, San Diego, 2002.

[5] F. A. Escobedo, E. E. Borrero, and J. C. Araque, Transition path sampling and forward flux sampling. Applications to biological systems, *J. Phys. Cond. Mat.*, 2009, **21**, 333101.

[6] D Moroni, T. van Erp, and P. Bolhuis, Investigating rare events by transition interface sampling, *Physica A*, 2004, **340**, 395–401.