An introduction to the study of rare events in materials using molecular-dynamics-based methods

Danny Perez

March 22, 2023

Molecular dynamics (MD), the numerical integration of atomistic equations of motion, is one of the workhorses in the computational sciences, ranging from physics, materials science, chemistry, and biology. MD simulations can be seen as numerical experiments: given a set of initial conditions and a model of the interatomic interactions, a (statistically) correct fully-resolved trajectory of the system can be generated and used to estimate a wide array of thermodynamic and dynamic properties that would be difficult (or prohibitively expensive) to obtain experimentally. This predictive power however comes at a considerable computational cost that strongly limits the space of possible simulations. While increasing the size or accuracy of simulations is greatly facilitated by the exponential growth in available computing resources, the simulation timescales have remained severely limited by intrinsically sequential nature of the integration process, even when massively parallel computers are available. This has made the study of metastable dynamics, which are extremely common in a broad range of physical conditions, especially challenging to investigate, as short simulations are often not informative of long-time behavior.

In this tutorial, I will review the basics of MD in terms of algorithms and physical models used to approximate fully quantum-mechanical solution, highlighting the commonly used parallelization strategies and motivating the so-called timescale problem of MD. I will then discuss a family of methods called Accelerated MD that have been developed to alleviate this problem for metastable systems that evolve through sequences of rare events. These methods can be seen as functional substitutes for MD, i.e., they generate statistically-correct trajectories discretized over the metastable states of the system. AMD methods will be introduced in the context of a mathematical formalism based on quasi-stationary distributions that was developed in the last decade, establishing connections with commonly used physical rate theories. I will finally discuss alternative approaches to AMD methods than can also help address the timescale problem of MD.