Statistical approaches to forcefield calibration and prediction uncertainty in molecular simulation

Fabien Cailliez[1], Pascal Pernot[1]

[1]Laboratoire de Chimie Physique, UMR 8000, CNRS / Université Paris-Sud, Orsay, France

fabien.cailliez@u-psud.fr

The great development of molecular simulation in the past decades has made it a very attractive tool for the study of condensed matter. As a matter of fact, it is now commonly used to predict thermophysical properties of fluids, both in academic studies and for industrial purposes.

The use of molecular simulation as a predictive tool requires to monitor all sources of uncertainty in the results of a simulation., among which those arising from the definition of the forcefield has long been ignored. The main reason for this is the difficulty to estimate forcefield parameters, that necessitates an extensive exploration of parameter space, incompatible until very recently with the computer time of molecular simulations.

We explored various calibration strategies and calibration models within the Bayesian framework [1] in the case of a two-parameters Lennard-Jones potential for Argon. The advantage of this system is that calibration can be done using analytical expressions. We have shown that prediction uncertainty for thermodynamical and transport properties, albeit very small, is larger than characteristic statistical simulation uncertainty [2].

For more complex systems, more parameters have to be calibrated and, in absence of analytical models, the calibration process requires to run long molecular simulations. In order to face these issues, we propose to use kriging metamodels and optimal infilling strategies to limit the number of molecular simulations to be performed during the calibration process. We have benchmarked this methodology on the water TIP4P forcefield [3].

[1] Kennedy, M., O'Hagan, A. *J. Roy. Stat. Soc. B*, 63:425-464, 2001.
[2] Cailliez, F., Pernot, P. *J. Chem. Phys.*, 134:054124, 2011.
[3] Cailliez, F., Bourasseau A., Pernot, P. *Submitted to J. Comput. Chem.*