Consistent integration of experimental and ab initio data into molecular and coarse-grained models

LUKAS VLCEK, Oak Ridge National Lab — As computer simulations are increasingly used to complement or replace experiments, highly accurate descriptions of physical systems at different time and length scales are required to achieve realistic predictions. The questions of how to objectively measure model quality in relation to reference experimental or ab initio data, and how to transition seamlessly between different levels of resolution are therefore of prime interest. To address these issues, we use the concept of statistical distance to define a measure of similarity between statistical mechanical systems, i.e., a model and its target, and show that its minimization leads to general convergence of the systems’ measurable properties. Through systematic coarse-graining, we arrive at appropriate expressions for optimization loss functions consistently incorporating microscopic ab initio data as well as macroscopic experimental data. The design of coarse-grained and multiscale models is then based on factoring the model system partition function into terms describing the system at different resolution levels. The optimization algorithm takes advantage of thermodynamic perturbation expressions for fast exploration of the model parameter space, enabling us to scan millions of parameter combinations per hour on a single CPU. The robustness and generality of the new model optimization framework and its efficient implementation are illustrated on selected examples including aqueous solutions, magnetic systems, and metal alloys.