

## **An Information-Theoretic Approach to Coarse-Graining and Multiscale Simulations**

M. Scott Shell<sup>C, S</sup>, Scott Carmichael and Aviel Chaimovich

*University of California Santa Barbara, Department of Chemical Engineering, Santa Barbara, CA, U.S.A.  
shell@engineering.ucsb.edu*

The development of accurate coarse-grained molecular models is central to many theoretical efforts in soft matter, materials, and biophysics that seek to understand large length and time scale properties of complex systems. But what is the right way to coarse-grain? How should we develop coarse models from atomically-detailed ones without making assumptions about which properties to match? How can we control errors due to coarse-graining? These fundamental ideas see multiscale approaches as not only practically useful, but in the context of a deeper philosophy that seeks systematic routes to emergent supra-atomic physics. We propose a powerful new statistical-mechanical framework for these kinds of problems based on a quantity called the relative entropy, a phase-space functional measuring the information lost upon coarse graining and hence the (inverse) fitness of a given coarse-grained model. We suggest that minimization of the relative entropy provides a universal variational principle for coarse-graining, and offers broad new strategies in multiscale simulation studies. Here, we demonstrate how this approach provides a quantitative and systematic approach to the development of coarse-grained models for a variety of different types of systems. In particular, we show that the relative entropy approach provides an a priori prediction of coarse-graining errors (i.e., errors in the properties predicted by coarse-grained models), and gives a systematic strategy for designing coarse-grained models so as to minimize them. We also describe the development of stable, efficient, and robust numerical coarse-graining algorithms based upon relative entropy minimization. In particular, these methods are able to locate optimal models for complex molecules with very high dimensional parameter spaces. Finally, we show that this new formalism is able to generate very simple but surprisingly accurate coarse-grained protein and solvation models for understanding large-scale self-assembly problems.