Title: Data-Driven Methods for Accelerating Vibrational Spectroscopy Modeling at the Medium to Large Scale

Daniel P. Tabor
Department of Chemistry, Texas A&M University

Abstract

Linear and non-linear vibrational spectroscopy can serve as crucial tools for understanding the structure and dynamics of clusters and the condensed phase. These spectra can be accurately computed from first principles; however, for larger systems, these calculations can be prohibitively expensive, particularly when anharmonic effects play a significant role. Here, we present our efforts to accelerate the prediction of vibrational spectra with data-driven methods from two different angles.

First, we describe methods for calculating the anharmonic effects of non-covalently bound clusters in an inverse spectroscopy context—the anharmonic terms are directly estimated from the spectra. These methods utilize Bayesian optimization algorithms, combined with an *a priori* local mode form of the vibrational Hamiltonian. One advantage of this technique is that these algorithms do not require extensive training data to be effective, allowing the strategy to be employed on-the-fly.

Second, we present our results on computing condensed phase spectra from coarse-grained representations in a mixed quantum-classical framework. These methods have the potential for substantial scaling advantages over their atomistic counterparts—our ultimate goal is to apply these methods to large condensed phase systems over long time scales. We will demonstrate the efficacy of different featurizations of the coarse-grained representations, sensitivity to training set composition and size, the degree to which information can be retained for the various CG representations, and the degree of transferability of the models.