Calculation of Absorption and Emission Spectra in Explicit Solvent using Large-Scale DFT and Machine Learned Interatomic Potentials

Prof. Nicholas D.M. Hine

Department of Physics, University of Warwick, United Kingdom

Theoretical spectroscopy calculations in explicit solvent require a challenging combination of high-accuracy electronic structure, large system sizes, and well-sampled molecular dynamics on the timescale of solvent reorganisations. Linear-Scaling TD-DFT approaches to LR-TDDFT can scale to thousands of atoms, but ab initio sampling is still unfeasible. Here we employ Machine-Learned Interatomic Potentials to approximate the DFT ground and LR-TDDFT excited states for clusters of solvent surrounding three isomers of a prototype sunscreen, methyl anthranilate. We develop a protocol to train accurate ML models for long-duration MD in ground and excited states. We can then sample configurations from the ground or excited nuclear ensembles, evaluate vertical transition energies, and predict absorption/emission peaks, solvatochromic shifts, Stokes' shifts and lineshapes in explicit solvent. MAE's are typically below 10nm: a dramatic improvement on implicit solvent approaches, and we can rigorously test the accuracy of the MLIP by comparing its cluster energies directly to DFT/TD-DFT. We further extend the approach by introducing Active Learning protocols based on error estimation through a committee of MLIP calculators to accelerate the learning of excitation energies.