Session: [COMP] Quantum Mechanics (Oral) or [CINF] Nanomaterials & Informatics (Oral)

Title: Enumerating the small complex inorganic universe for machine learning Authors: Stefan O. Gugler, Jon Paul Janet, Heather J. Kulik

Transition metal complexes are promising functional inorganic materials due to their wide range of tunable electronic properties. However, exhaustive enumeration of all possible ligand fields is clearly intractable due to the vast nature of chemical space. Virtual high-throughput screening with density functional theory (DFT) allows us to harvest leads with desired properties but is severely constrained by 1) long calculation times and 2) variable accuracy. More accurate correlated methods are available to address 2) but drastically worsen 1). Machine learning techniques potentially allow us to address both issues simultaneously, and our group has previously developed data-driven models based on DFT results which have highlighted the dominant role of metal-proximal (i.e. from first and second coordination shell) atoms in predicting spin state ordering, bond lengths and ionization potential of the metal center. This motivates a detailed study of the range of possible local octahedral metal environments through systematic exploration of the space of octahedral complexes made of ligands with up to two heavy organic (CNOPS) atoms. We sample the most promising compounds from this space with high-throughput DFT and assess DFT performance selectively with more accurate correlated wavefunction calculations using domain-based local pair-natural orbital coupled cluster (DLPNO-CCSD(T)). We then apply machine learning to model the difference between correlated wavefunction and DFT results in a composition-dependent manner and learn property estimates for the full space of possible core environments along with estimates of DFT reliability relative to CCSD(T).