Autonomous First-Principles Design of Transition Metal Complexes

Naveen Arunachalam

Submitted to the Department of Chemical Engineering on April 18, 2023, in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Designing novel transition metal complexes (TMCs) holds immense potential for advancing sustainability and chemical synthesis. While high-throughput virtual screening (HTVS) workflows and density functional theory (DFT) have emerged as powerful tools for discovering new TMCs, exhaustive design space exploration remains a formidable challenge due to the combinatorial growth of possible complexes with respect to ligands, metals, oxidation states, and ligand field symmetries. Machine learning models trained on databases of TMC structures and their computationally derived properties offer a promising avenue for rapidly and accurately evaluating the molecular properties of novel TMCs. However, the extrapolation error of these models in new regions of chemical space depends on the chemical space spanned by training data, making it crucial to carefully design training data to optimize the performance of HTVS workflows.

This thesis presents a comprehensive exploration of transition metal complex design, encompassing fundamental chemical insights and HTVS improvements derived from the systematic enhancement of database coverage, such as expanded coverage of ligand field symmetries, metal identities, and similarity to experimentally synthesized complexes. In addition, this work introduces interactive web-based tools for the design of TMCs and metal-organic frameworks, allowing users to explore and provide feedback on machine learning models. These tools facilitate collaboration between researchers and promote the iterative improvement of models by incorporating user feedback, leading to more effective and efficient exploration of the vast chemical space of TMCs.

Thesis Supervisor: Heather J. Kulik
Title: Associate Professor of Chemical Engineering