Mohammad Atif Faiz Afzal
Email: m27@buffalo.edu, Phone: 1-716-262-5115

Accelerated Discovery of High-Refractive-Index Polymers Using *First-Principles* Modeling, Virtual High-Throughput Screening, and Data Mining

Mohammad Atif Faiz Afzal[1], Chong Cheng[1], Johannes Hachmann[1]

[1]Department of Chemical and Biological Engineering, University at Buffalo,
The State University of New York, Buffalo, New York, United States

Abstract

Organic materials with refractive index (RI) values higher than 1.7 have attracted considerable interest in recent years due to the tremendous potential for their application in optical, optometric, and optoelectronic devices, and thus for shaping technological innovation in numerous related areas. The vast majority of organic polymers has RI values in the range of 1.3 to 1.5. The ability to tailor the molecular structure of these polymers is the key to increasing their RI values. The number of polymers that can be created by introducing different functional groups is practically infinite. The options to experimentally characterize only a fraction of these candidates are very limited.

Our work is concerned with creating predictive models for the optical properties of organic polymers, which will guide our experimentalist partners and allow them to target the most promising candidates. The present work is based on the Lorentz-Lorenz equation and thus includes the calculation of polarizability and number density values for the candidate structures. In the proposed scheme, we compute the molecular polarizability using *first-principles* electronic structure theory and the number density using the van der Waals volume method. The critical parameter in number density calculation is the packing fraction of the bulk polymer. We have devised a machine learning approach using a support vector machine on a training data set compiled from the literature to correlate the polymer structure with their packing fraction. The resulting model for this structure-property relationship is exceedingly efficient. The RI values predicted for common polymers are in very good agreement with the experimental values. We also benchmark different DFT approximations along with various basis sets for their predictive performance in this model. We demonstrate that this combination of *first-principles* and data modeling is both successful and highly economical in determining the RI values of a wide range of organic polymers.

To accelerate the development process, we cast this modeling approach into the high-throughput screening, materials informatics, and rational design framework that is developed in the group. This framework is a powerful tool and has shown to be highly promising for rapidly identifying polymer candidates with exceptional RI values as well as discovering design rules for advanced materials.