Machine Learned Electronic Structures and Optical Properties for Organic Semiconductors

Zhou Lin
Department of Chemistry, University of Massachusetts Amherst

High-throughput material and drug discovery based on density functional theory (DFT) has achieved tremendous success in recent decades. However, the power of DFT on organic semiconductors (OSC) as molecular electronic materials suffer significantly from its computational complexity and intrinsic errors. In this talk, we will discuss two of our machine learning (ML) efforts to acquire accurate electronic structures and optical properties for OSC materials with removed computational bottleneck and minimal information loss. In the first part, we will introduce a new exchange–correlation (XC) functional developed in our group, referred to as ML-ωPBE, which evaluates the molecule-specific range-separation parameter (ω) in a range-separated hybrid (RSH) functional using a stacked ensemble algorithm and composite molecular descriptor. Compared to first-principles OT-ωPBE, a well-trained ML-ωPBE reaches a mean absolute error (MAE) of 0.00504 a_0^{-1} for optimal ω's, reduces the computational cost by 2.66 orders of magnitude, and achieves comparable predictive power in optical properties to OT-ωPBE. In addition, ML-ωPBE shows a strong domain adaptation from closed shell molecules to open shell radicals. In the second part, we will apply the same algorithm and descriptor to evaluate optical properties of all synthesizable organic molecules in the QM9 database. Compared to time-dependent DFT (TDDFT) and DFT-based quantum dynamics, our algorithm reaches MAEs of 8.32 nm, 12.45 nm, and 0.10 for absorption wave lengths (λ_{abs}), fluorescence wave lengths (λ_{fl}), and fluorescence quantum yields (φ_{fl}). From both efforts, we will conclude the importance of descriptors from semi-empirical quantum chemical calculations. Our study will set the stage for developing physics-based, and data-driven computational models for high-throughput material and drug discovery.