Machine-Learned Electron Localization in Dense Hydrogen

Xiaoyu WANG (Sorbonne Université)
Julia CONTRERAS-GARCIA (Sorbonne Université)

Abstract

Machine learning offers a promising route to bypass the high computational cost associated with explicit electronic-structure calculations for real-space bonding descriptors. Here, we present a geometry-based machine-learning framework for predicting the electron localization function (ELF) of hydrides directly from atomic configurations, without requiring Kohn–Sham orbitals or kinetic-energy densities. The model is trained on first-principles datasets spanning multiple pressure regimes in random binary structures and achieves consistently high predictive accuracy, with coefficients of determination exceeding R^2>0.99. In addition to reproducing the global ELF distribution with high fidelity, the framework accurately captures the spatial organization of localization features relevant to hydrogen bonding and network formation.

To characterize the remaining discrepancies between predicted and reference ELF fields, we perform a combined real- and reciprocal-space analysis of the residual error. The residual is found to be highly structured rather than stochastic, being dominated by smooth, long-wavelength components whose correlation lengths exceed typical H–H bonding distances. The magnitude and spatial extent of these nonlocal contributions increase systematically with pressure, indicating that the most difficult aspects of ELF prediction arise from collective electronic responses extending beyond the local bonding environment. Despite representing only a small fraction of the total ELF amplitude, these components can be efficiently described using a band-limited Fourier correction formulated in the logit representation.

Importantly, ELF-derived topological descriptors remain robust against these structured residuals. In particular, networking characteristics, including connectivity patterns and critical-point topology, are preserved with high fidelity across a wide range of thermodynamic conditions, enabling the prediction of critical temperatures. Although the model is trained exclusively on binary systems, it transfers successfully to ternary hydrides.

Overall, this methodology allows rapid evaluation of ELF-derived descriptors, including networking values, bonding topology, and localization-driven structural motifs, without the need for explicit electronic-structure calculations. More broadly, these results suggest a viable path toward large-scale, geometry-based electronic-topology analysis for materials discovery and high-throughput exploration of bonding phenomena across diverse chemical environments.

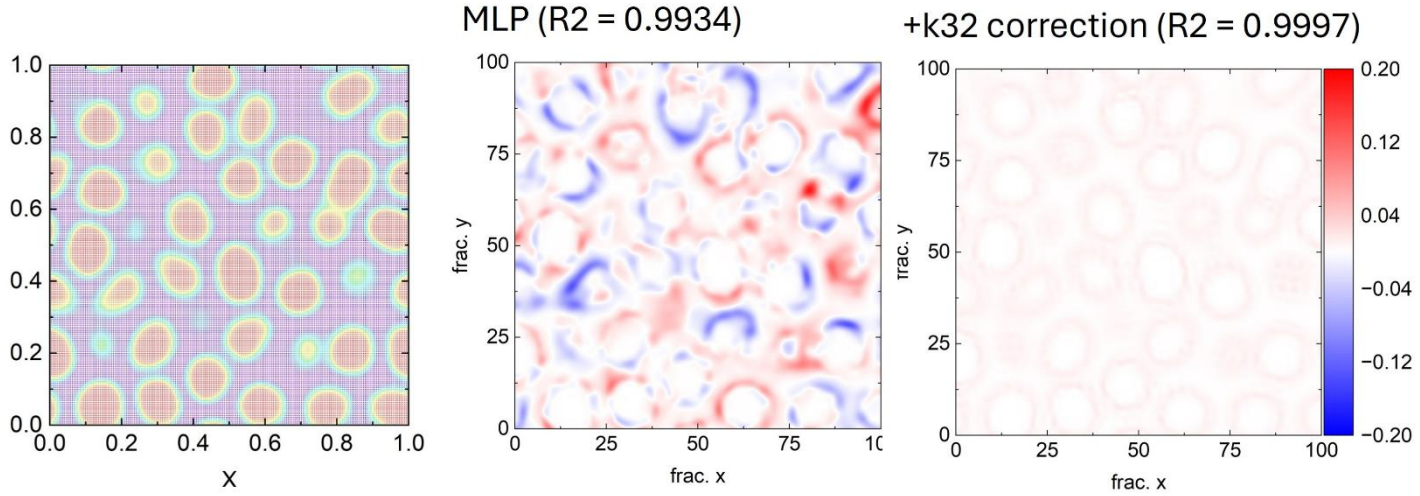

Figure 1. a) DFT-computed ELF in real space (color scale from 0 (blue) to 1 (red); fractional coordinates relative to 11~\AA, b) Real-space residual between predicted and DFT ELF with the real space predictor, c) Real-space residual between predicted and DFT ELF with the real+reciprocal space predictor.