Domain Knowledge + AI for Chemically Accurate Potentials:
Application to Diamond Surfaces

John Isaac G. Enriquez[1, 2], Yoshitada Morikawa[1], and Igor Kaganovich[2]

[1]Applied Materials and Sustainability Sciences, Princeton Plasma Physics Laboratory

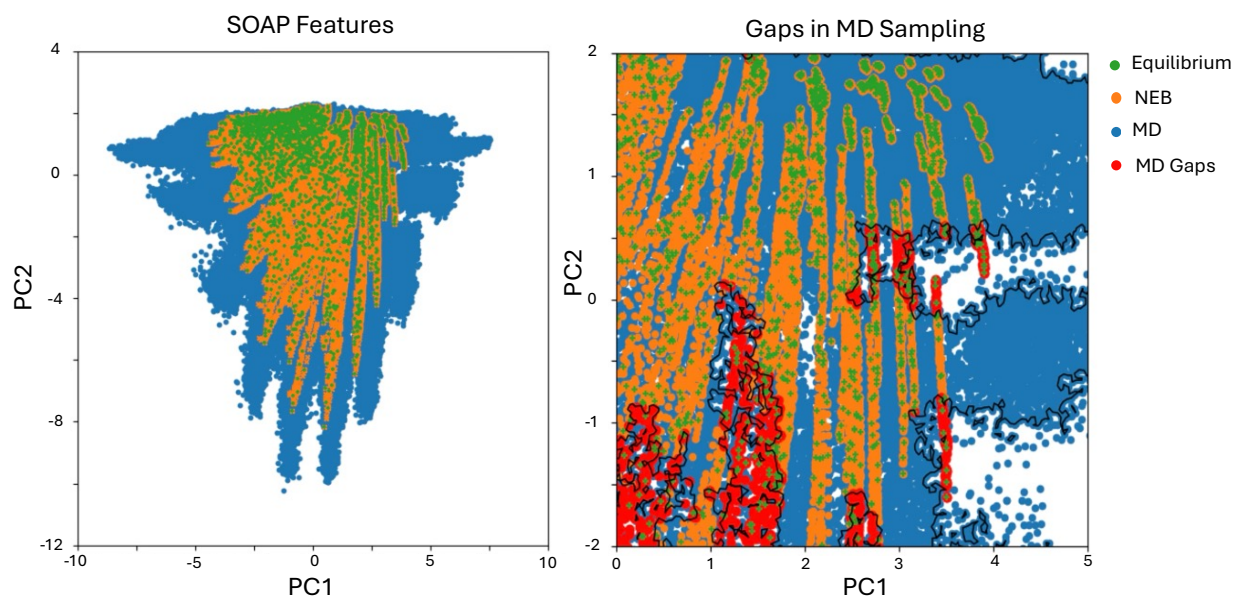[2]Graduate School of Engineering, The University of Osaka

Figure 1 | PCA analysis of structural environments sampled by MD and NEB under DIAL (Domain-Informed Active Learning).
Left: projection of SOAP (Smooth Overlap of Atomic Positions) features onto the first two principal components. Molecular dynamics (MD, blue) produces dense, blot-like clusters around equilibrium structures (green), whereas nudged elastic band (NEB, orange) traces streaks connecting equilibrium points, capturing reactive pathways. Right: zoom-in with concave hulls (black) drawn around the MD distributions. NEB points falling outside the MD hull (red) highlight reaction-relevant configurations that are missed by MD alone, demonstrating how DIAL integrates domain knowledge to enrich training datasets with chemically critical environments.