Automatic Reaction Discovery by Biasing Deep-Learned Skewed Distributions

GiovanniMaria Piccini
University Of Modena And Reggio Emilia, Modena, Italy

The automated discovery of chemical and catalytic reactions remains a fundamental challenge in computational chemistry, especially in complex systems where traditional approaches struggle to identify efficient exploration pathways. Here, we introduce Loxodynamics, a machine learning-driven method for reaction discovery based on biased molecular dynamics. The approach exploits local asymmetries in configuration space by analyzing the skewness of sampled probability distributions to determine optimal directions for low-energy barrier crossings.

At the core of Loxodynamics is the Skewencoder—a neural network Autoencoder enhanced with a skewness-based loss function—which learns reduced-dimensional representations of the system and identifies promising reaction coordinates from limited sampling. By iteratively applying a sample-and-search protocol, the method dynamically reconstructs the underlying free energy surface, capturing essential finite-temperature effects and enabling the identification of metastable states and transition pathways.

We validate Loxodynamics on a range of systems: simple two-dimensional model potentials, gas-phase reactions such as S_N2 substitution and the Diels–Alder cycloaddition, and a complex, industrially relevant catalytic transformation—ethanol dehydration in acidic chabazite zeolites under *operando* conditions. Without requiring predefined collective variables or prior mechanistic knowledge, our approach successfully uncovers the most relevant reaction pathways and rate-limiting steps.

By combining data-driven dimensionality reduction with physics-informed enhanced sampling, Loxodynamics provides a robust, generalizable framework for the automatic exploration of reaction networks. It holds particular promise for studying reactivity in fluxional, solvent-influenced, or heterogeneous catalytic environments where existing methods face substantial limitations.