Chapter 27
Machine Learning and Molecular Simulation:
A New Frontier in Quantum Dynamics

Ibtissem Jendoubi
https://orcid.org/0000-0002-2800-1748
Faculty of Sciences of Bizerte, Tunisia

Hamza Hendaoui
https://orcid.org/0000-0002-1864-1872
Faculty of Sciences of Tunis, Tunisia

Elhoucine Essefi
Faculté des sciences de Sfax, Tunisia

ABSTRACT

Understanding complex quantum systems with many interacting particles is a major physical challenge. Classical methods struggle due to the exponential increase in complexity as the number of particles increases. This study explores two promising approaches: present an overview of the state-of-the-art research related to molecular dynamics and machine learning. The first approach studies diabaticity using a neural network, which offers richer dynamic properties than atoms. By studying the hybrid calcium system, CaH2, the authors present PESs adiabatic and diabatic of the ground state and the first excited state. As scoop results presented for the first time, detailed analysis identified new approaches to molecular dynamics beyond the Born-Oppenheimer approximation. The second approach tackles the problem from a computational perspective. Machine learning algorithms, particularly interpretable methods such as neural networks (NN) with influence functions, have been explored.

DOI: 10.4018/979-8-3693-9336-9.ch027

This combination aims to achieve efficient solutions while retaining interpretability.

1. INTRODUCTION

As the computing power of computers continues to improve, machine learning methods have enabled the development of new solutions to problems in a variety of fields. It is extremely difficult for computers to simulate quantum systems with many particles (Georgescu, al., 2014; Zalka, C., 1998). There are two main reasons for this finding. First, the calculations became exponentially more difficult as the number of particles increased. This is a common problem in science called the *'curse of dimensionality*' (Köppen et al., 2000). Essentially, the vast amount of data needed to represent complex quantum systems, with their numerous variables, makes them computationally challenging to simulate. The second reason for the difficulty is even more basic and has to do with the fundamental nature of these quantum systems.

The influence of machine learning (ML) fundamentally alters our research methodologies across all scientific disciplines. Specifically, attention has been directed towards the interconnected realms of machine learning and molecular simulation. ML can be used to create models that approximate the complex interactions between molecules. These models can be considerably faster than the classical methods, allowing scientists to simulate larger systems or longer timescales.

Numerous studies have explored the convergence of machine learning and molecular simulation. Noé et al. (2020) conducted a notable study focusing on the intricate and time-intensive computations inherent in molecular simulations, highlighting their suitability for an ML revolution. Their review delved into the profound impact of the existing ML methods on these calculations. Notably, they explored the utilization of (deep) neural networks for predicting quantum-mechanical energies and forces. Furthermore, the review investigated ML applications in various aspects of molecular dynamics, including coarse-grained simulations, the derivation of free energy surfaces and kinetics, as well as generative network approaches aimed at sampling molecular equilibrium structures and computing thermodynamics (Funai & Giataganas, 2020).

Kang et al. (2020), focus on recent advances in machine learning (ML) methodologies for atomic simulations, particularly in three key aspects: generating a representative data set, increasing the complexity of ML models, and ensuring continuity of data representation. While global optimization methods are the go-to approach for building a representative data set, the stochastic surface walking method is shown to be effective in providing the desired potential energy surface (PES) sampling for both minima and transition regions. Their study demonstrates that newly devised power-type structure descriptors, when combined with a feed-forward neural network