Towards accurate universal machine learning potentials for electronic excited-state simulations

Abstract for general public

Mikołaj Martyka[1]

[1]Faculty of Chemistry, University of Warsaw

When molecules absorb light — such as sunlight or laser light — the energy can shift the molecule from the ground state into an electronic excited state. These excited states are central to how solar panels work, how our eyes perceive light, how photodynamic therapies destroy cancer cells, or how smart materials change their structure in response to illumination. Understanding and simulating excited states is therefore crucial for advancing renewable energy, molecular electronics, and next-generation materials and medicines.

However, performing these simulations with traditional quantum chemistry methods is extremely demanding. Calculations are slow and expensive, often limited to very small molecules. This project aims to change that — by using **machine learning** to create models that can predict excited-state properties *orders of magnitude faster* than conventional methods, without sacrificing accuracy.

While machine learning is already revolutionizing ground-state molecular simulations, applying it to excited states is a much greater challenge. These states are more complex and unpredictable, involving multiple interacting energy surfaces. Despite this, our group has recently developed **OMNI-P2x**, the first ever universal machine learning potential (MLP) for excited states. This model can make fast, accurate predictions of UV-Vis spectra and excited-state geometries for a wide range of molecules, and can even be adapted to simulate photochemical reactions.

In this project, we will take the next steps toward making **universal MLPs** a reliable tool for theoretical photochemistry. Our goals are to:

- Generate large, high-quality datasets specifically tailored for excited-state training;
- Design advanced neural network architectures that better respect molecular symmetry and encode electronic state information;
- Explore *delta learning* to improve prediction quality by correcting fast, approximate methods with machine learning to a higher target accuracy;
- Develop robust protocols for adapting universal models to specific simulations, such real-time simulations of photoreactions;
- Apply these models to real-world problems by rational design of molecular solar thermal energy storage (MOST) systems.

The result will be a new generation of AI-driven models that enable researchers to simulate light–matter interactions on a realistic scale and timeline — without supercomputers or specialist training. This has the potential to accelerate discovery in chemistry, materials science, and photobiology, helping build a more sustainable and energy-efficient future.