Research Project #24

Offline-to-Online Reinforcement Learning for Perovskite Stability Optimization: Benchmarking Against Bayesian Optimization in High-Throughput Experimentation (IMD-3 & JSC)

Research Project Description

Novelty and Innovation:

This project investigates whether reinforcement learning (RL) can outperform established Bayesian optimization (BO) for perovskite stability optimization in high-throughput experimentation. While BO has achieved up to 4x cost-reduction in HTE workflows through exploration-exploitation balancing [1], it suffers from two critical limitations: myopic one-step lookahead potentially leading to local optima, and inability to leverage prior knowledge from publications and thousands of historical experiments worldwide.

Recent breakthroughs in offline reinforcement learning enable training exclusively on historical data (offline) without requiring simulations or new experiments during training [2]. Offline-to-online RL grounds learning in actual experimental data from day one and leverages historic data into optimization policies guiding the experimentation (online). The core scientific question addresses whether RL's theoretical advantages (multi-step planning, long-term strategy learning, implicit prior knowledge integration) translate to practical gains in materials optimization. Through systematic head-to-head benchmarking across three methods (offline RL, pure BO, BO with priors), this work will provide the first rigorous answer defining when, if ever, RL should replace BO for experimental materials science, yielding fundamental insights regardless of outcome.

Main Tasks and Work Packages for the Doctoral Candidate:

• (Month 1 to 6) Establish performance baselines by compiling 5,000+ historical perovskite experiments from IMD-3, and public databases, benchmarking established methods (i.e. pure Bayesian optimization) to quantify current sample efficiency, convergence rates, and limitations in high-dimensional stability optimization.

• (Month 6 to 18) Develop offline-to-online RL framework using Conservative Q-Learning (state-of-the-art offline RL algorithm), training exclusively on historical data without new experiments (JSC infrastructure), developing ensemble surrogate models for uncertainty quantification as safety filters rather than primary simulators, incorporating domain-specific constraints (stoichiometry, temperature, toxicity limits).

• (Month 18 to 30) Perform systematic experimental validation through iterative closed-loop deployment in IMD-3's HTE platform: RL proposes experiment batches → surrogate filtering → human review → HTE execution → policy updates, conducting 200-500 validation experiments across 20-50 cycles, performing comparisons across all methods under identical budgets and conditions with statistical rigor (5 runs/method).

• (Month 30 to 36) Analyze cross-institute transferability and interpretability, testing whether policies trained at IMD-3 transfer to different HTE platforms (at Campus Erlangen), quantifying zero-shot performance and fine-tuning requirements, conducting ablation studies identifying which components contribute most to performance, establishing explainable AI guidelines for when RL provides advantages over BO in different scenarios.

Research Context and Embedding:

Bayesian optimization remains the dominant algorithm for high-throughput experimentation in materials science, yet fundamental questions persist about optimal exploration strategies. Recent work demonstrates two promising directions: offline reinforcement learning enabling learning from historical data without the need for simulations [3], [4], model-

based RL [2]. However, no study has rigorously compared RL against BO for experimental materials optimization with real-world validation.

IMD-3 brings critical expertise and resources to this investigation: extensive experience with BO in HTE workflows for photovoltaic materials, measurements on perovskite/OPV stability providing rich datasets for offline RL training and benchmarking studies. A newly built degradation-focused HTE infrastructure at Campus Jülich includes advanced optical perturbation techniques enabling accelerated experimental loops. Most published RL work uses simulated data or limited experimental campaigns, whereas IMD-3's comprehensive historical database provides realistic test cases matching real-world materials research complexity, enabling systematic investigation across composition spaces and prior optimization runs.

The Jülich Supercomputing Center (JSC) provides complementary expertise essential for developing RL algorithms at scale: specialized knowledge in ML and AI development, HPC infrastructure necessary for training RL agents and ensemble surrogate model, established software frameworks for offline RL implementation (Conservative Q-Learning, model-free and model-based approaches). The collaboration enables systematic comparison at scales matching our state-of-the-art autonomous laboratories [5] while maintaining rigorous experimental validation unavailable in purely computational studies.

The research addresses fundamental scientific questions with practical implications for photovoltaic materials development and broader experimental materials science. If RL achieves 2-5x sample efficiency improvement, this enables 2-year stability optimization campaigns completed in 1 year, 50% R&D cost reduction, and instant knowledge transfer between institutions through shared learned policies.

The Graduate School HITEC provides an interdisciplinary community of doctoral researchers working at the intersection of energy materials, AI/ML, and high-performance computing. Since reinforcement learning, autonomous experimentation, and materials informatics are addressed across multiple HITEC specializations, this environment offers valuable opportunities for methodological exchange and cross-disciplinary collaboration in fields of AI, infrastructure and machine-learning.

One step further, beyond photovoltaics the systematic benchmarking framework and offline-to-online RL methodology developed here could transfer to many materials science fields actively investigated at Forschungszentrum Jülich (batteries, hydrogen, catalysis, quantum materials) and is in best accordance with the FZJ strategy to make use of novel Artificial Intelligence methods for experimental acceleration and autonomous discovery platforms. The results will provide a blueprint for offline-online-type problems, and the insights are of high strategic value for the JSC. Beyond energy research, the project learnings will be transferred to other domains, e.g. by the AI consultant team in PI Stefan Kesselheims's lab at JSC.

Location of the HITEC Fellow	Forschungszentrum Jülich, Institute of Energy Materials and Devices - Photovoltaics (IMD-3), Director: Prof. Dr. Christoph Brabec https://www.fz-juelich.de/en/imd/imd-3
Partners of the HITEC Project	Forschungszentrum Jülich, Jülich Supercomputing Centre (JSC), Directors: Prof. Dr. Dr. Thomas Lippert and Prof. Dr. Kristel Michielsen https://www.fz-juelich.de/en/jsc
Specific **requirements**	The candidate has a background in physics, material sciences, electrical engineering, computer sciences
For **project specific** questions please **contact**	Prof. Dr.-Ing. Bugra Turan, IMD-3, b.turan@fz-juelich.de Prof. Dr. Stefan Kesselheim, JSC, s.kesselheim@fz-juelich.de

[1] V. Sabanza-Gil *et al.*, "Best practices for multi-fidelity Bayesian optimization in materials and molecular research," *Nat. Comput. Sci.*, vol. 5, no. 7, pp. 572–581, July 2025, doi: 10.1038/s43588-025-00822-9.

[2] R. Kidambi, A. Rajeswaran, P. Netrapalli, and T. Joachims, "MOReL : Model-Based Offline Reinforcement Learning," Mar. 03, 2021, *arXiv*: arXiv:2005.05951. doi: 10.48550/arXiv.2005.05951.

[3] A. Kumar, A. Zhou, G. Tucker, and S. Levine, "Conservative Q-Learning for Offline Reinforcement Learning," Aug. 20, 2020, *arXiv*: arXiv:2006.04779. doi: 10.48550/arXiv.2006.04779.

[4] N. Nievas, L. Espinosa-Leal, A. Pagès-Bernaus, A. Abio, L. Echeverria, and F. Bonada, "Offline Reinforcement Learning for Adaptive Control in Manufacturing Processes: A Press Hardening Case Study," *J. Comput. Inf. Sci. Eng.*, vol. 25, no. 1, p. 011004, Jan. 2025, doi: 10.1115/1.4066999.

[5] J. Zhang *et al.*, "Autonomous Optimization of Air-Processed Perovskite Solar Cell in a Multidimensional Parameter Space," *Adv. Energy Mater.*, p. 2404957, Jan. 2025, doi: 10.1002/aenm.202404957.