Project Proposal

Efficient Battery Energy Management with Learning-Based Control using DAGGER

Anup Parajuli
Athul Jose P

CPTS 570 Machine Learning
Washington State University

Problem

The core problem we are addressing is the optimization of control for energy storage systems, specifically batteries, in a typical power distribution system. Batteries play a critical role in managing energy, but efficient control becomes challenging due to the complexity of the system and the high computational costs associated with traditional Model Predictive Control (MPC). The goal is to reduce the computational burden of repeated calls to MPC while maintaining performance. The motivation for this problem stems from the increasing deployment of energy storage systems in power grids, where real-time control can significantly enhance grid stability and efficiency.

Methodology

We plan to implement the DAGGER[1] (Dataset Aggregation) algorithm to reduce the dependency on MPC. Initially, the project will involve developing an MPC to control the battery, which will serve as the expert controller. DAGGER will then be used to collect data from the MPC and train a policy that mimics its control actions. By applying this learned policy, we aim to reduce the number of direct MPC calls, thereby lowering computational costs. The DAGGER algorithm is well-suited for this task because it iteratively refines the policy, ensuring that it can generalize to a wide range of system states.

While the current project focuses on the implementation of DAGGER for this control task, uncertainty quantification (UQ) techniques are envisioned as a future extension. These techniques would be useful to further enhance the robustness of the control system by determining when to rely on the DAGGER-trained policy and when to fall back to the full MPC based on uncertainty levels.

Final Product

The final product will be a control system for battery energy storage that reduces the computational burden of MPC by using the DAGGER algorithm to handle most control decisions. The success of the project will be measured by how effectively the learned policy reduces the number of MPC calls without degrading system performance, compared to a baseline where MPC is used exclusively. This project will also set the stage for future work in uncertainty quantification, where we will quantify when the learned policy can be trusted and when MPC should be engaged.

References

[1] S. Ross, G. Gordon, and D. Bagnell, “A reduction of imitation learning and structured prediction to no-regret online learning,” in *Proceedings of the fourteenth international conference on artificial intelligence and statistics.* JMLR Workshop and Conference Proceedings, 2011, pp. 627–635.