ME 418/518 Project Proposal

Residual-Aware RL Tuning of MPC on Data-Driven Models

Erinç Ada Ceylan, Bilkent University – Mechanical Engineering
ada.ceylan@ug.bilkent.edu.tr

Project Description

Model Predictive Control (MPC) provides high-performance control by simultaneously solving an optimization problem using a predictive model of the system. However, MPC performance can degrade significantly when the model diverges from the actual dynamics of the system, which often occurs due to parameter drift, internal nonlinearities, or unmodeled disturbances. Recent reinforcement learning (RL)-based approaches attempt to adjust MPC weights online to maintain optimality under changing conditions *(Zarrouki et al., 2024; Airaldi, 2024)*. However, these methods usually assume that the predictive model itself remains reliable and do not quantify its uncertainty.

This project proposes a **Residual-Aware RL–MPC** framework, in which an RL agent dynamically tunes MPC parameters (Q, R, N) based on both the system state and a real-time model reliability indicator—the one-step residual:

$$r_t = \|x_{t+1}^{meas} - \hat{x}_{t+1}^{pred}\|.$$

Two data-driven identification methods, **DMDc** and **N4SID** will be trained using simulation data to construct predictive state-space models. The RL agent then learns to interpret the residual magnitude r_t as a confidence measure: small residuals allow aggressive tracking, while large residuals trigger conservative control. The **racecar_gym** environment (manageable and easily extensible; can be finalized later) will serve as the main simulation platform for data generation, model identification, and RL-tuned MPC evaluation. It offers a compact yet sufficiently realistic dynamic setup for this type of closed-loop experimentation. The overall methodology thus integrates identification, prediction, and residual-aware control within a single data-driven pipeline.

Motivation

The project is motivated by the increasing need for autonomous and model based controllers capable of operating under uncertainty. Classical MPC methods assume perfectly known system models, which is rarely true in practice. When the model changes or becomes inaccurate, control performance drops — and purely optimization-based control can't fix it.

This study aims to build an MPC framework that learns how much to trust its own model by simultaneously checking residuals. By explicitly feeding model residuals to an RL agent, the controller gains self-awareness of modeling errors and can react to preserve stability and performance. On a personal level, this project aligns with my interest in the intersection of control theory and the machine learning. It also allows me to apply core ME 418 concepts (DMD, N4SID, MPC and RL) in a unified and practically relevant context.

Significance

The proposed framework contributes to the emerging field of safe and learning-based control. Its significance can be summarized as follows:

- **Enhanced robustness:** By practicing residual-based feedback, the controller can automatically balance performance and safety under varying residual conditions.
- **Unified comparison:** The project provides a fair, closed-loop benchmark of DMDc and N4SID models within the same control architecture, offering insights for future data-driven control studies.
- **Real-world applicability:** The approach is transferable among real systems such as robotic manipulators, quadrotors, and process control loops, where models are uncertain.

In the long term, the findings could inform the design of autonomous controllers that monitor their own reliability, an essential step toward certifiable learning-based control in safety-critical systems.

Novelty

While RL-tuned MPC schemes have been explored before, they typically optimize MPC weights as functions of state and time, not of the model's accuracy. The novelty of this work lies in explicitly conditioning the RL policy on the model residual r_t. This residual-aware mechanism allows the control system to respond when its internal model becomes less accurate and adjust itself accordingly. Thus, the framework bridges **data-driven identification** and **learning-based control** in a way that is interpretable. Compared to existing methods *(Zarrouki et al., 2024; Airaldi, 2024)*, this project emphasizes model reliability as a first order signal rather than an implicit assumption. The outcome is a self correcting controller capable of maintaining stable performance even as its predictive model degrades.

Key References:

- B. Zarrouki, M. Spanakakis, and J. Betz, "A Safe RL-driven Weights-Varying MPC for Autonomous Vehicle Motion Control," *arXiv:2402.02624*, 2024.
- F. Airaldi, "mpcrl: Reinforcement Learning with Model Predictive Control," TU Delft, 2024.
- J. N. Kutz et al., *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems*, SIAM, 2016.