Model Predictive Control with Gaussian Processes

A Case Study as the CERN SPS

Ferdinand Ferber[1,2], Nicky Bruchon[1], Simon Hirlaender[1,2], and Verena Kain[1]

[1]CERN
[2]University of Salzburg

Background

Slow Extraction Spill

Most physics experiments at CERN are *fixed target experiments*, where subatomic particles are accelerated to very high energies and then are steered towards a stationary target made out of solid, liquid, or gas. Instead of delivering the particle beam in a short burst (as done in the LHC), the beam is usually extracted slowly to provide a continuous stream of particles. At CERN, several *transfer lines* connect the *Super Proton Synchrotron* (SPS) with the fixed target experiments hosted in the North Experimental Area (NA).

Power Supply Ripples

Power supply ripples are periodic voltage fluctuations in a DC power supply, caused by residual AC components after rectification or by dynamic current draw from nonlinear devices. These ripples appear as harmonics of 50 Hz and higher in the extracted beam intensity, leading to unwanted fluctuations in the spill, which negatively affects the experiments.

Previous Work

Power supply ripples have always been a problem in the SPS. Before the *Long Shutdown 2* (LS2) in 2021, the fluctuations around 50 Hz, 100 Hz, and 150 Hz were corrected by injecting a suitable voltage modulation in the main quadrupole circuit QF at adjustable phases and amplitudes, which needed infrequent manual adjustments. After the upgrades performed during LS2, those manual interventions became infeasible and were replaced by numerical optimization using the BOBYQA algorithm. However, only the suppression of spill ripples at around 50 Hz was within the requirements of the experiments. Therefore, in 2023, Adaptive Bayesian Optimization (ABO) with Empty Bucket Channeling (EBO) was deployed, which in particular improved the quality around the 100 Hz frequency. One drawback of ABO is the slow speed with which it can react to changes in the environment.

Our Contribution

We present an implementation of the GP-MPC approach that is compatible with the stable-baselines3 API and test it on the problem of correcting power supply ripples in the slow extraction spill from CERN's SPS. To propagate uncertainty through time, we implement an uncertainty-aware kernel instead of using moment matching.

GP-MPC

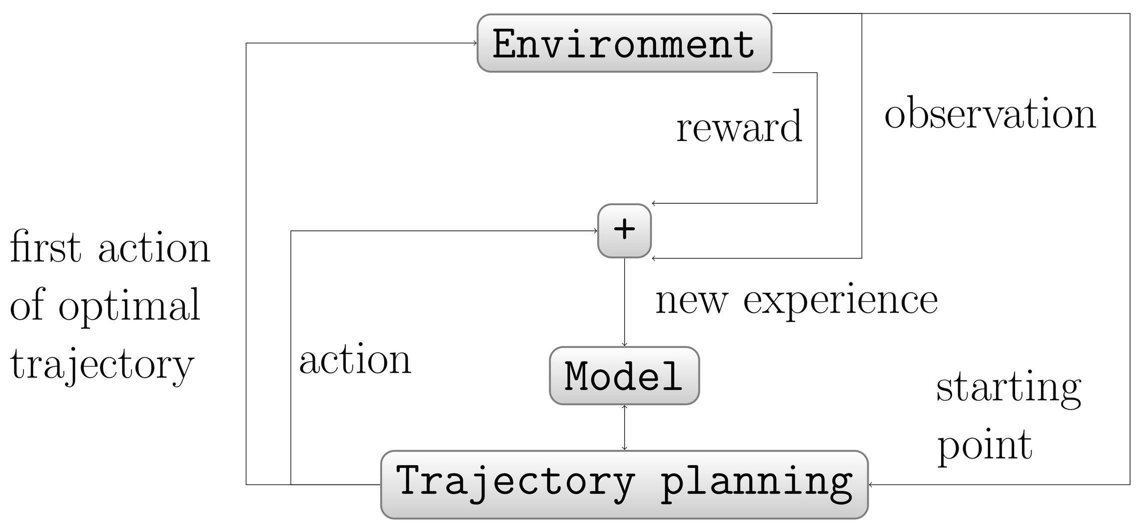


Figure 1: Schematic overview over the use of MPC in an RL setting.

We try to solve a sequential optimization problem that is based on a *Markov Decision Process*. The agent interacts with the environment, observes its resulting state, and receives feedback in the form of rewards or penalties. Its goal is to maximize the long-term cumulative returns. With *Model Predictive Control* (MPC), the agent keeps a model of the world and uses it to plan the optimal sequence of actions over a finite decision horizon. After doing the first action of the optimal trajectory, the internal model is updated according to the actual behavior of the environment. Since the model improved, the optimal trajectory is re-planned. When a probabilistic model is used, the agent is able to quantify its uncertainty when planning. This allows the agent to avoid exploring suboptimal trajectories, even when being optimistic.

One choice for such a probabilistic model is a *Gaussian Process*, which can be thought of as a function that, upon input x, does not return a precise point $f(x)$, but instead a "fuzzy point" specified by a mean $\mu(x)$ and a variance $\sigma^2(x)$, in a way, such that for any finite choice $x_1, \ldots, x_n$ of inputs, the joint distribution of the outputs is a multivariate normal distribution. A Gaussian process is fully determined by its *mean function* and its *kernel function*. The choice of a kernel heavily influences the shape of the possible functions that can be approximated. If the kernel contains *hyperparameters*, they need to be fitted, usually by maximizing the marginal likelihood of the model.

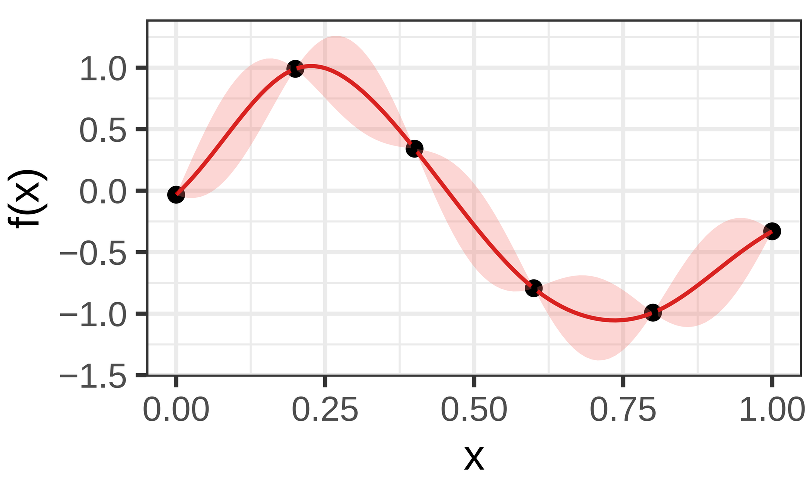


Figure 2: A Gaussian Process conditioned on six random points of a sine curve. We show the mean prediction and the confidence region.

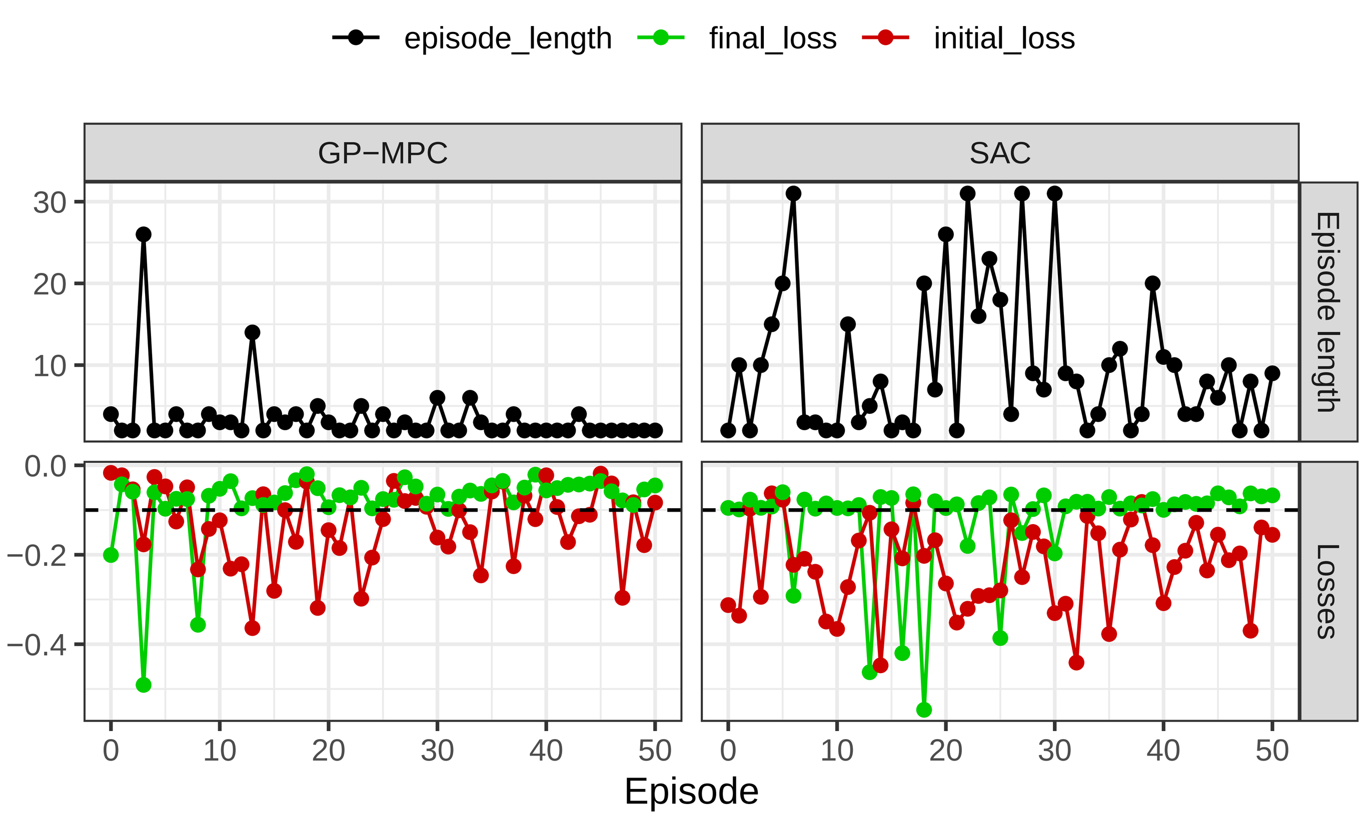
Experimental Setup

We test our GP-MPC implementation directly on the SPS. Because the rate at which the power supply ripples change is relatively slow, we increase the difficulty by injecting random ripples at multiples of 50 Hz in one of the quadrupole power converters and run the GP-MPC agent on a second power converter on the same circuit, where it tries to inject a voltage modulation that suppresses the resulting ripples.

Results

The GP-MPC agent was trained with just 300 interactions and was able to solve the problem almost every time, even during training. We tested different settings and found that the Expected Improvement acquisition function works better than the Upper Confidence Bound acquisition function for this problem. A *Soft Actor Critic* (SAC) agent tested on the same problem needed about 40 episodes until it was able to solve the problem reliably.

Findings



Conclusions

GP-MPC proved to be a viable alternative to existing RL algorithms for the problem of power supply ripple suppression in the transfer line between the SPS and the fixed target experiments in the NA. Although traditional RL algorithms like SAC are able to solve the problem, they require a substantial amount of interaction with the environment until they learn a reliable policy. Our GP-MPC implementation required only a few hundred samples to solve the problem. Additionally, once the agent is trained, it can be used in a mode where it behaves risk-averse and avoids trajectories with high uncertainty.

Future Research

More Uncertainty-Aware Kernels

One of our main questions was if it is possible to use an uncertainty-aware kernel in the GP to propagate uncertainty through time. Previous implementations of GP-MPC used methods like moment matching or approximations with Taylor series to evaluate a GP on an uncertain input. Our results indicate that uncertainty-aware kernels can be used as an elegant alternative. However, we only implemented one such kernel that is based on the traditional RBF kernel and it would be beneficial to have more choice.

Constraints

Another addition to previous GP-MPC implementations is the availability of the *Expected Improvement Acquisition Function*. This acquisition function has the advantage that it is non-negative, which makes it easy to incorporate constraints into the trajectory planning by multiplying the expected improvement with the probability that no constraint violation will occur.

References and acknowledgements

This internship was generously supported by a scholarship from the rectorate of the PLUS Salzburg.

References

[1] A. Girard et al. 2003: *Learning a Gaussian Process Model with Uncertain Inputs.*

[2] V. Kain et al. 2024: "Slow extracted spill ripple control in the CERN SPS using adaptive Bayesian optimisation".

[3] S. Kamthe et al. 2018: "Data-Efficient Reinforcement Learning with Probabilistic Model Predictive Control".

LaTeX TikZposter