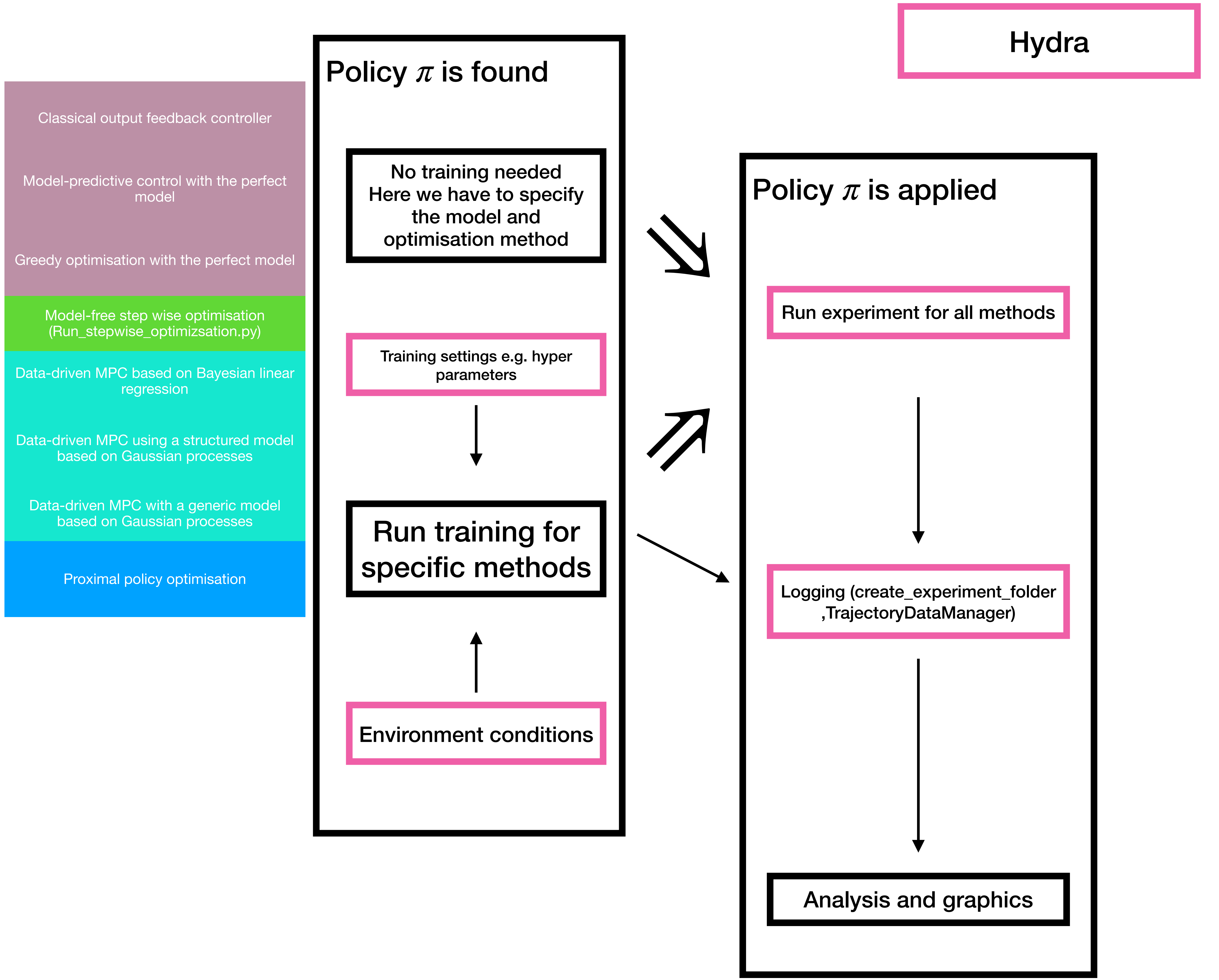
Hydra
Classical output feedback controller
Model-predictive control with the perfect model
Greedy optimisation with the perfect model
Model-free step wise optimisation (Run_stepwise_optimizsation.py)
Data-driven MPC based on Bayesian linear regression
Data-driven MPC using a structured model based on Gaussian processes
Data-driven MPC with a generic model based on Gaussian processes
Proximal policy optimisation
Policy π is found
No training needed Here we have to specify the model and optimisation method
Training settings e.g. hyper parameters
Run training for specific methods
Environment conditions
Policy π is applied
Run experiment for all methods
Logging (create_experiment_folder ,TrajectoryDataManager)
Analysis and graphics

Thesis

Bridging Classical Control and Reinforcement Learning for Delayed Outcome Optimisation demonstrated on CERN AWAKE electron line

This thesis explores advanced strategies for optimal control in systems with delayed consequences, using beam steering in the AWAKE electron line at CERN as a benchmark. We formulate the task as a constrained optimization problem within a continuous, primarily linear Markov Decision Process (MDP), incorporating measured system parameters and realistic termination criteria. A wide range of approaches is implemented and compared, including classical response matrix inversion, control-theoretic methods, reinforcement learning, and a novel problem specific structured model-based technique.

While classical methods like matrix inversion offer accurate convergence, they fail to account for delayed effects and are sensitive to noise. Control-theoretic approaches, such as Model Predictive Control (MPC), leverage known dynamics and handle delays effectively when models are available. Data-driven methods, including Proximal Policy Optimization (PPO), adapt to uncertainty and non-linearities but require large amounts of data. Structured Gaussian Processes (GP)-MPC bridges both paradigms by learning system dynamics using GPs while respecting the problem's causal structure, significantly improving robustness and sample efficiency.

Our experiments highlight key performance differences, particularly in how each method handles delayed outcomes, noise, slight non-linearities, model errors and structural assumptions. We find that exploiting the causal structure of the problem provides a notable advantage, and that method choice ultimately involves trade-offs between adaptability, data efficiency, and computational cost. These findings offer guidance for applying advanced control strategies in high-dimensional, partially structured environments.

Structure and timing of the thesis

The problem - already done

Thesis - easy:
- Motivation - why is my problem interesting? - already done
- How to approach the problem methodically? - at the end or during breaks
- <u>Experiments to probe the ideas - most time consuming</u>
- <u>Results and discussion - most time consuming</u>
- <u>Summary and outlook - at the end</u>

Paper - higher priority
General path follows the poster + extra studies on modifying the dynamics of the system (robust MPC is still missing)
Starting on 25th of April - reproducing the results of the poster and be critical
Make a logical structure for Hydra has to be set up
- <u>Experiments to probe the ideas - most time consuming</u>
- <u>Results and discussion - most time consuming</u>
- <u>Summary and outlook - at the end</u>

- Motivation - why is my problem interesting? - already done
- How to approach the problem methodically? - at the end or during breaks