IDALAB

PARIS
LODRON
UNIVERSITÄT
SALZBURG

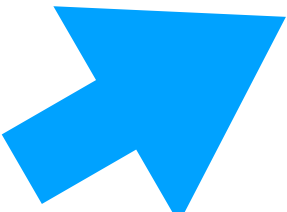
Simon Hirländer

Tutorial RL4AA

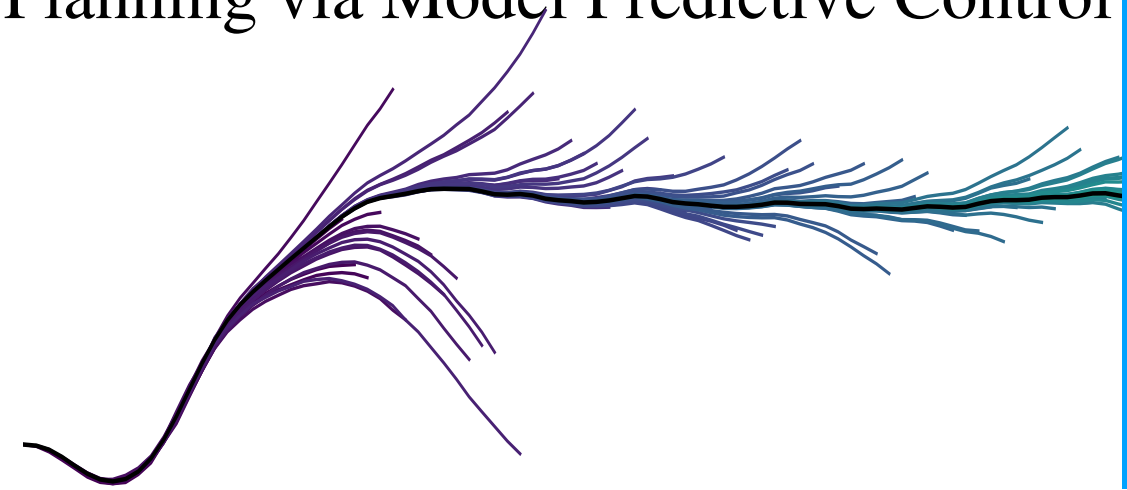
GP-MPC the BO of RL

- Setup the dynamics-reward model
- Use PMP to obtain sparse optimization with gradient information
- Choose optimization algorithm
- Consider safety (constraints)
- Set up training

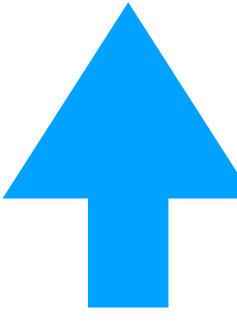
47

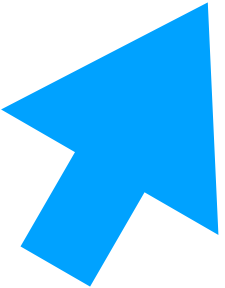


Planning via Model Predictive Control



Optimization



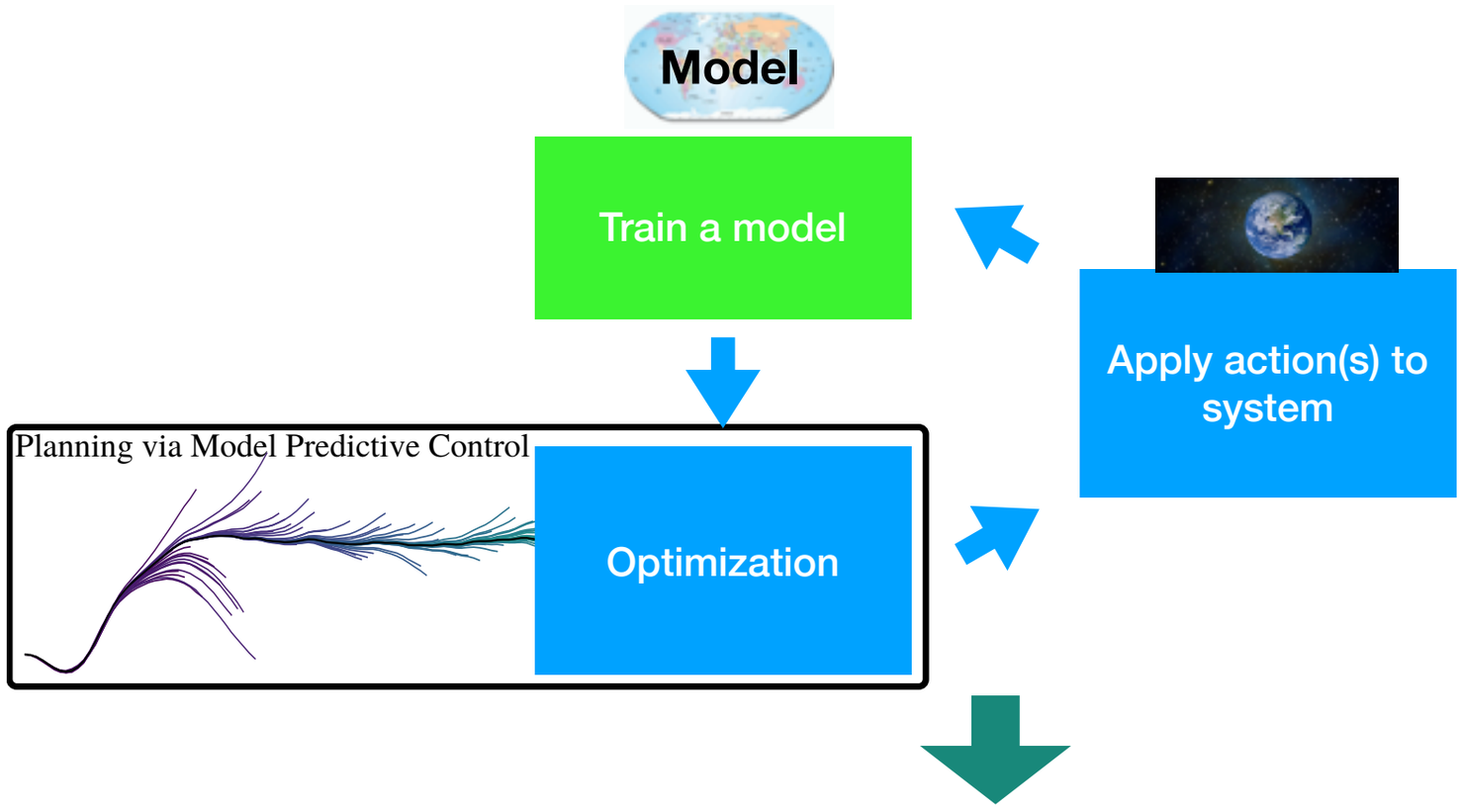


Train a model

Apply action(s) to system

Model (GP)

GP-MPC the BO of RL



- Setup the dynamics-reward model
- Use PMP to obtain sparse optimization with gradient information
- Choose optimization algorithm
- Consider safety (constraints)
- Set up training

We don’t know the model

- Learn the model from data:
  - ➡ Aleatoric uncertainties
  - ➡ Epistemic uncertainties - minimise model bias
- Gaussian processes (GPs) are used assuming $\mathbf{s}_{t+1} = \mathbf{f}(\mathbf{s}_t, \mathbf{a}_t, \omega_t)$ and $\omega_t \sim \mathcal{N}(0,\sigma)$
- Include if needed the emitted reward
- Use RBF Kernel - allow for analytical propagation of uncertainties
- Standard GPs training: evidence maximization