Learning-based Optimisation of Particle Accelerators Under Partial Observability Without Real-World Training

5th DESY ML Round Table

Jan Kaiser, Oliver Stein and Annika Eichler
Hamburg, 25 November 2022

Motivation

High-impact real-world application of reinforcement learning

Challenges of RL for Real-world Applications

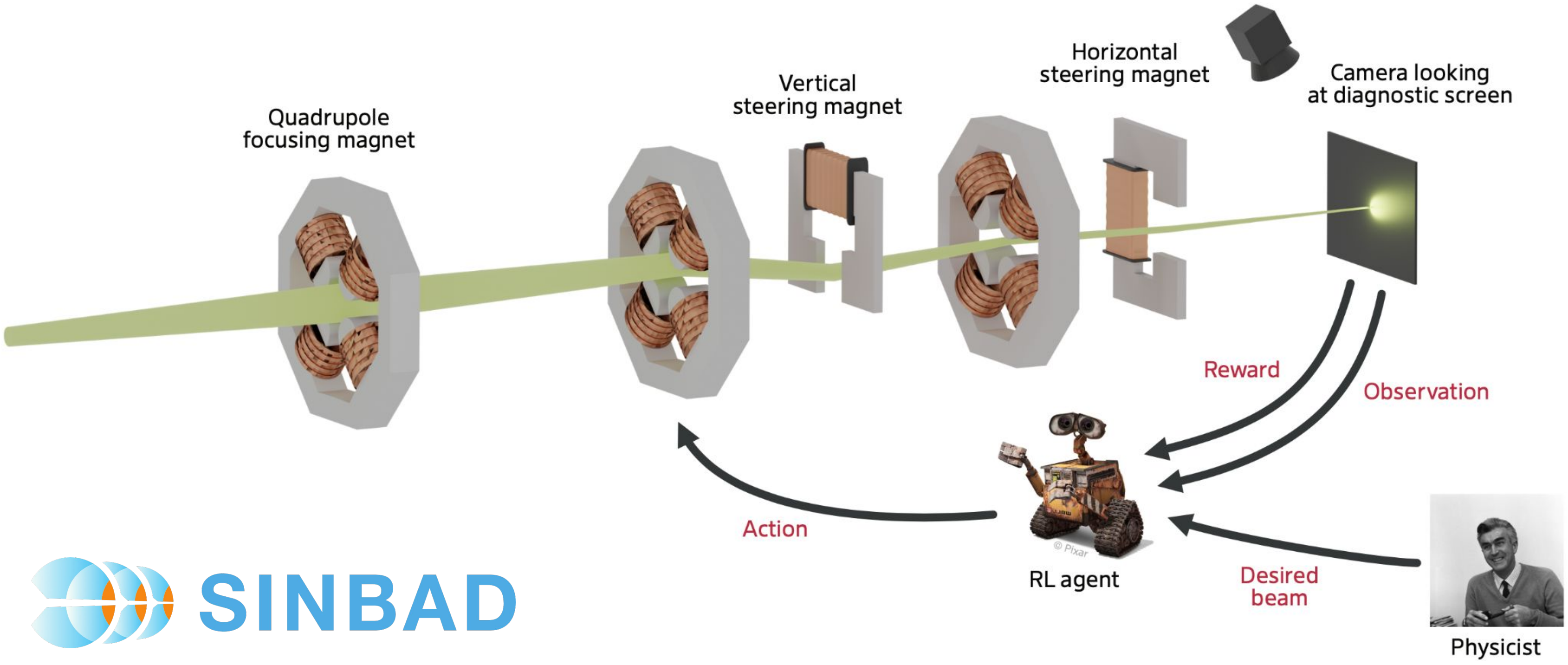
- **High cost-per-sample** can make training in the real world infeasible
- **Simulations** might not be available, have limited accuracy or are expensive to compute
- Disturbances in the real world such as **noisy** observations or action effects
- **Safety** concerns
- **Partial-observability** is often more prominent than in academic examples

Reinforcement Learning for Particle Accelerators

- Highly advanced machines used in a large **variety of applications** in physics, engineering and medicine
  - e.g. physics discovery, vaccine and drug development, cancer treatment, material development
- **Sparse availability** and high operating costs
- **Tuning** to desired configurations a **long time** which cannot be used for useful operations
- Tuning is still largely done **by hand** as automation of these complex tasks has proven difficult
- Final tunes are **difficult to repeat** and strongly depend on operator experience

RL Task

Transverse beam parameter optimisation at the ARES accelerator



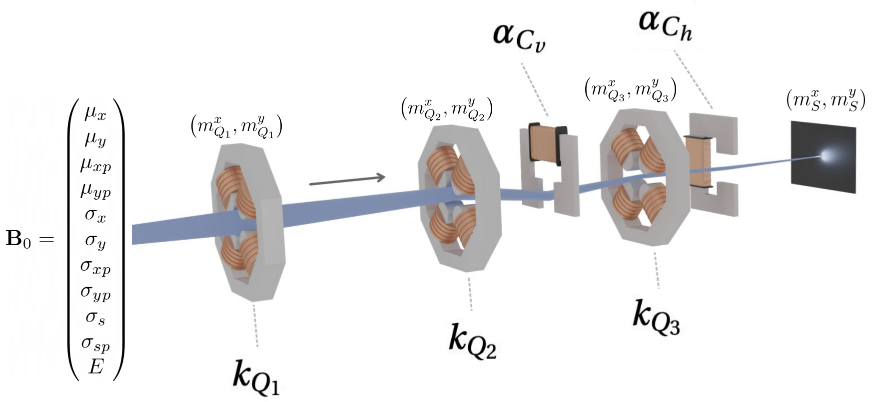
SINBAD

Real-world Challenges at ARES

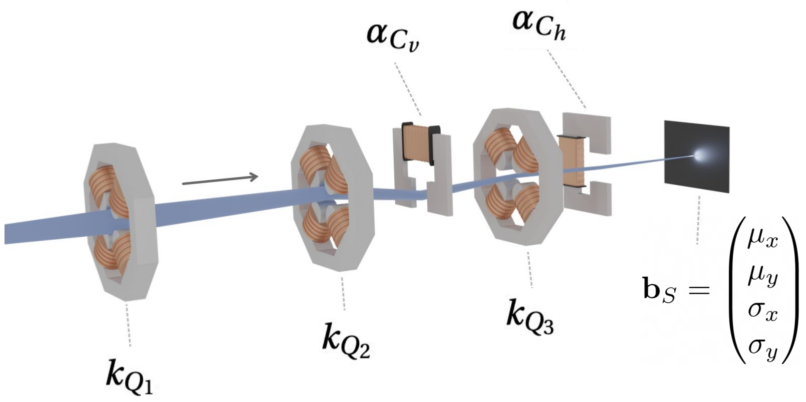
How the task at ARES reflects other real-world tasks

- **Partial observability** (9 observables vs. at least 26-dimensional state-space)
    - No diagnostics available for incoming beam and magnet/screen misalignments
- **Noisy** observations and disturbed actions
    - Magnet hysteresis changes effects of magnet, screen is noisy and has limited accuracy, other environmental disturbances affect accelerator notably
- Training in the real-world is prohibitively **expensive**
    - Collecting just one sample takes 10-20 seconds, beam time is a sparse commodity and there exists only one accelerator exactly like ARES. In addition there are machine safety concerns during training.

Fully-observable state s_t



Partially-observable observation o_t



Design Decisions

Solving the RL task at ARES

Training

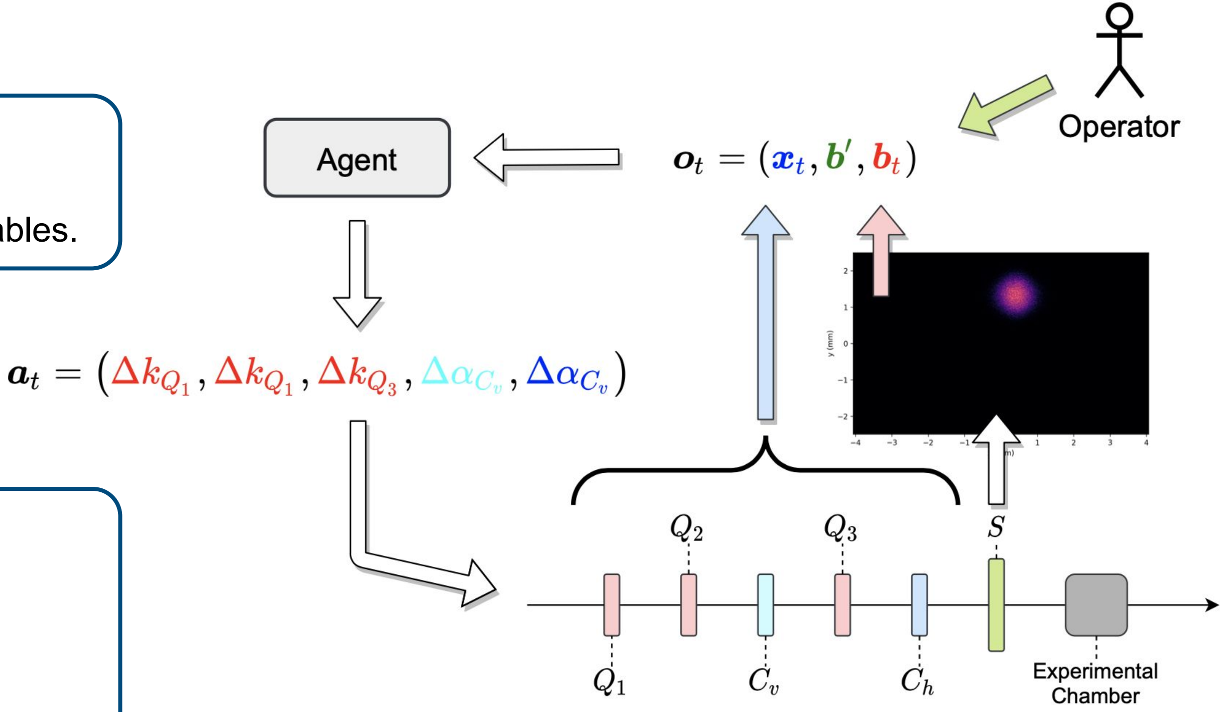
Train in **simulation** with **domain randomisation** over unknown variables.

Reward

$$O(\boldsymbol{x}_t) = \ln \sum_{p \in \boldsymbol{b}_t, p' \in \boldsymbol{b}'} w_p \left| p - p' \right|$$

$$\hat{R}(\boldsymbol{s}_t, \boldsymbol{a}_t) = O(\boldsymbol{x}_t) - O(\boldsymbol{x}_{t+1})$$

$$R(\boldsymbol{s}_t, \boldsymbol{a}_t) = \begin{cases} \hat{R}(\boldsymbol{s}_t, \boldsymbol{a}_t) & \text{if } \hat{R}(\boldsymbol{s}_t, \boldsymbol{a}_t) > 0 \\ 2 \cdot \hat{R}(\boldsymbol{s}_t, \boldsymbol{a}_t) & \text{otherwise.} \end{cases}$$



$$\boldsymbol{o}_t = (\boldsymbol{x}_t, \boldsymbol{b}', \boldsymbol{b}_t)$$

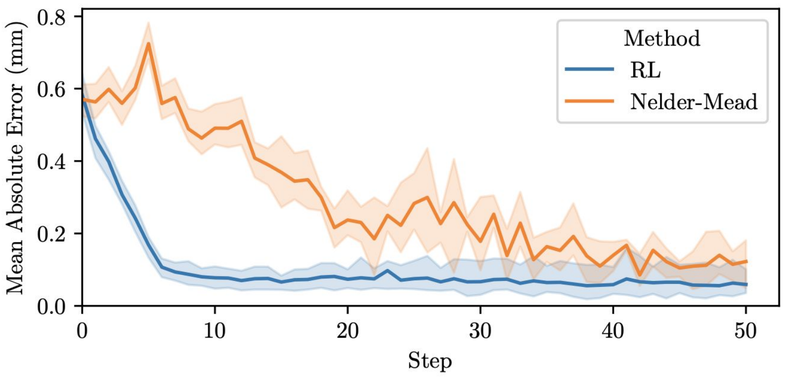
$$\boldsymbol{a}_t = \left(\Delta k_{Q_1}, \Delta k_{Q_1}, \Delta k_{Q_3}, \Delta \alpha_{C_v}, \Delta \alpha_{C_v}\right)$$

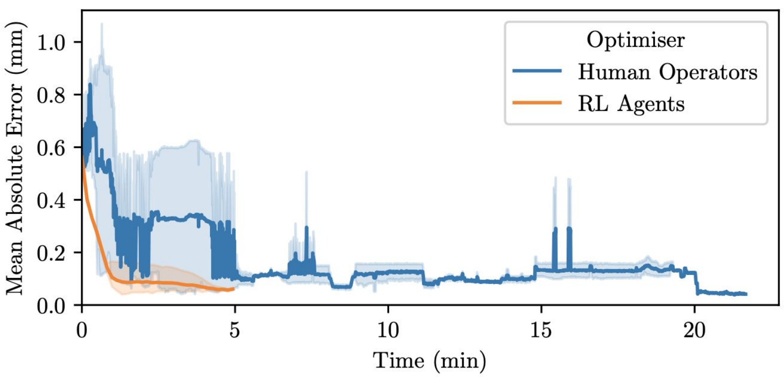
Results

Evaluation in simulation and on the real accelerator

- The **sim2real transfer** was successful!
- Trained agents achieve final beams in line with those achieved by the best-performing **black-box optimisation algorithm** and **expert human operators**.
- RL agents performed the optimisation around **10x faster** than competitive optimisation algorithm and up to **4x faster** than human experts.

Algorithm	MAE Median (mm)	Convergence Median (Steps)
Do Nothing	1.122	0
Zero	0.588	1
FDF	0.699	1
Random	0.267	101
Powell	0.259	119
COBYLA	0.105	34
Nelder-Mead	0.007	112
Bayesian	0.081	101
Ours	0.008	7
Ours (Machine)	0.036	12





Thank you!

Contact

DESY. Deutsches
Elektronen-Synchrotron

www.desy.de

Jan Kaiser, Oliver Stein & Annika Eichler
Machine Beam Control (MSK)
jan.kaiser@desy.de | oliver.stein@desy.de | annika.eichler@desy.de