Applications and Opportunities for Machine Control

V. Kain for the CERN ATS ML Community Forum

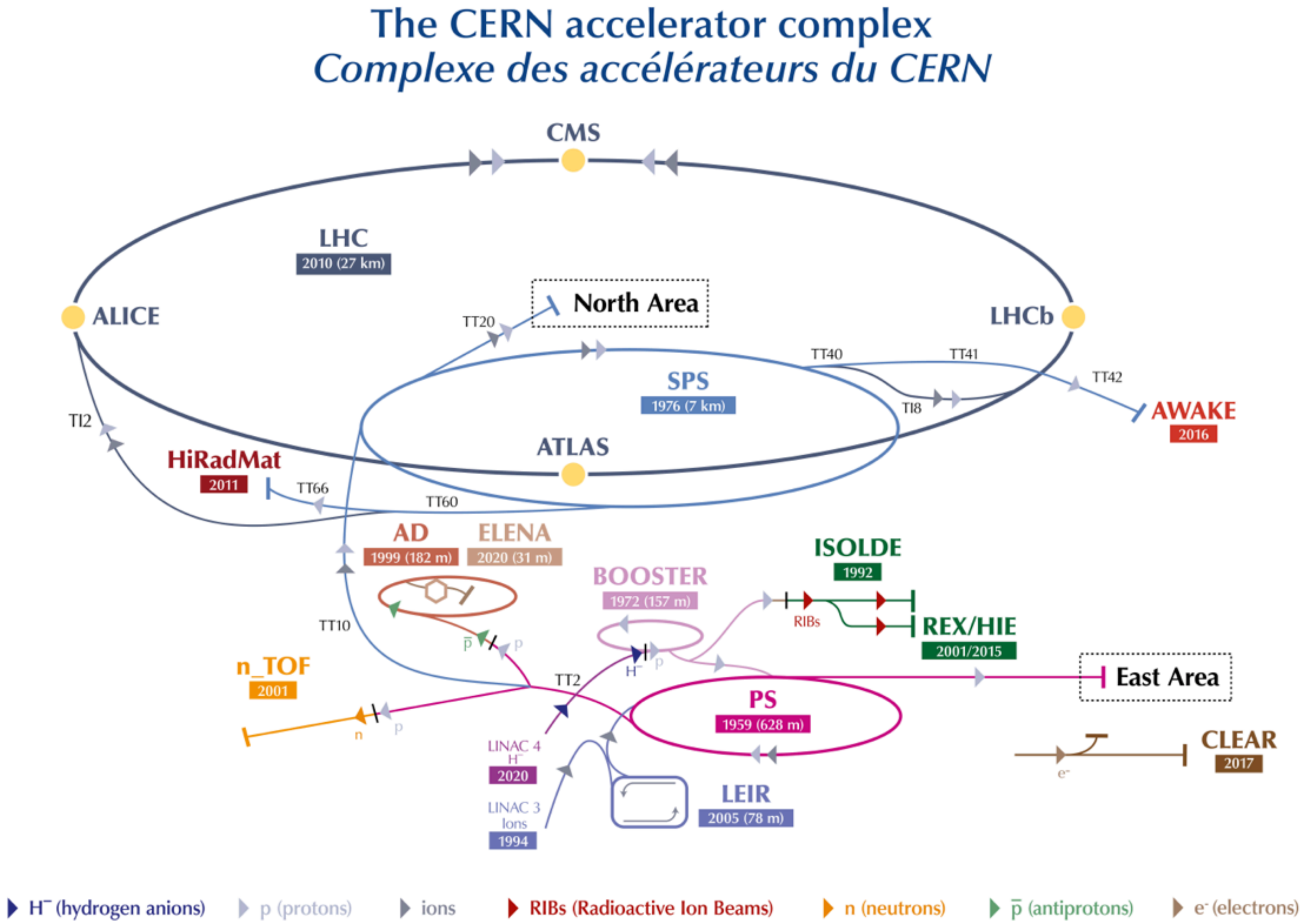
Disclaimer

Will focus on the approach at the CERN accelerators

CERN accelerators are not investing (yet) into fast ML or ML on chip → more about applications than opportunities

Most ML is currently done from remote, far away from the actual hardware → distributed systems (e.g. 1000 pick-ups around 27 km of circumference)

The CERN accelerator complex



Physics models in the control room

Beam dynamics equation of motion in static magnetic and RF fields linearised and solved
→ **closed form solutions used as models in the control room to control**

- ★ mean energy, energy spread, beam size, orbit,...collective motion of particles,...

→ **global parameters:** $B\rho$ or p, the tunes Q_x, Q_y, Q_s and tune spreads through additional global parameters e.g. chromaticity $Q'_{x,y}$

→ Use **high level physics parameters** to control accelerators instead of direct hardware parameters: i.e. normalised magnetic fields instead of currents.

- ★ e.g. dipole magnet's control parameter: bending angle change $\Delta x'$ or $\Delta y'$ instead of current in power supply.
- ★ needs hardware to physics parameter translation: e.g.. *transfer function* $Bl \rightarrow I$ for every magnetic circuit

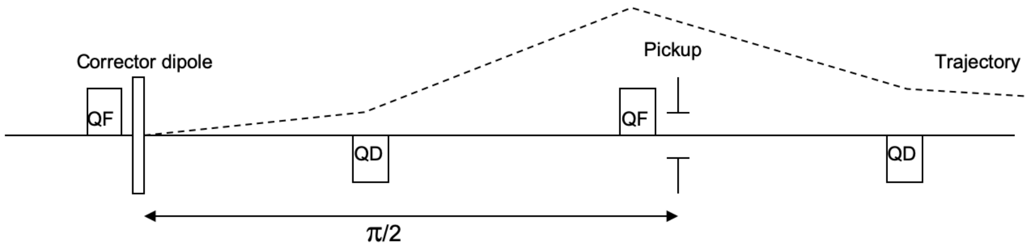
[1] D. Jacquet et al, LSA- The high level application software of the LHC and its performance during the first 3 years of operation , ICALEPS 2013

Physics models in the control room

Build parameter models, store transfer functions in controls DB, pre-calculate the settings of accelerate according to uploaded "optics", injection/extraction momentum,...

This allows: model-based one-shot correction of imperfections

- ★ E.g. trajectory correction



- ★ Calculate response R, with R^{-1} settings for correctors for given $(\Delta x_1, \Delta x_2, \ldots \Delta x_m)$

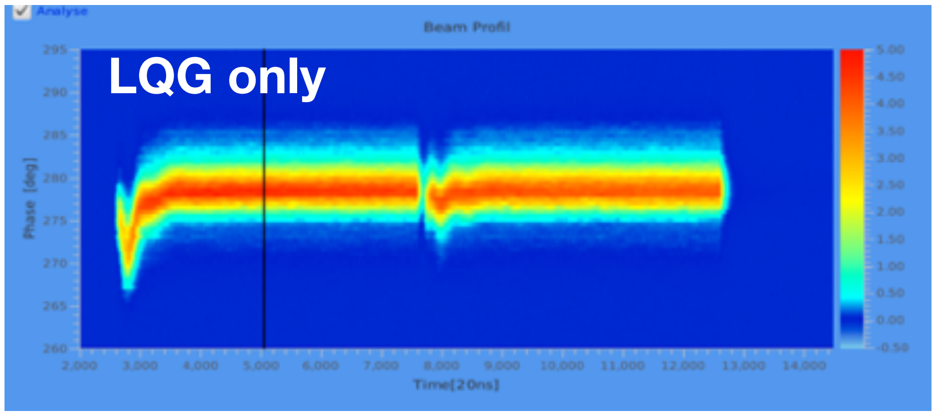
$$R \begin{pmatrix} \Delta x'_1 \\ \Delta x'_2 \\ \cdots \\ \Delta x'_n \end{pmatrix} = \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \cdots \\ \Delta x_m \end{pmatrix}$$

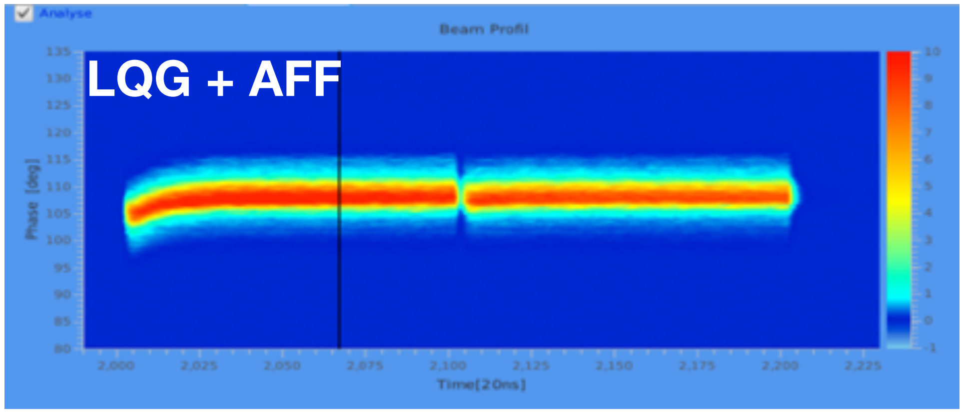
  - ✳ R is linear for our machines (i.e. matrix) → SVD algorithm

Example: Controlling the cavities at LINAC4

- ★ Strong field perturbations from beam loading
  - ✳ e.g. CCDTL1 cavity 1 MV induced by 25 mA DC current
- ★ Needs to be compensated along the 600 μs long beam pulse injected into the PS booster.
  - ✳ PI controller not fast enough
  - ✳ → model-predictive control with LQG controller ~ 900 ns response (Kalman predictor response ~ 200 ns) on FPGA
  - ✳ Adaptive feedforward (iterative learning controller) on FEC for optimum performance: pulse per pulse adjustment

P. Baudrenghien et al





Limitations of classical approach

★ Need model → some of the accelerators in the chain are 1/2 century old

★ Need beam instrumentation to define where and what the beam is, as input to correction algorithm

★ Models can be very complex → days/weeks of simulation. Inverse models?

★ Models are rarely complete → as-built machine models?

- ✳ But also: drifts,…

Common approach still: operators on shift that manually tune

Since 2018: towards numerical optimisation and ML techniques

ML and Advanced Algorithms?

Machine Learning has many applications for accelerators (from accelerator design to offline data analysis,…)

Will focus here on ML and Advanced Algorithms to

- ★ speed up commissioning and
- ★ speed up configuration switch times
- ★ shorten turn-around time
- ★ stabilise performance containing drifts with clever algorithms

→ ML to complement classical models with “learned" models

→ ML for enhanced diagnostics

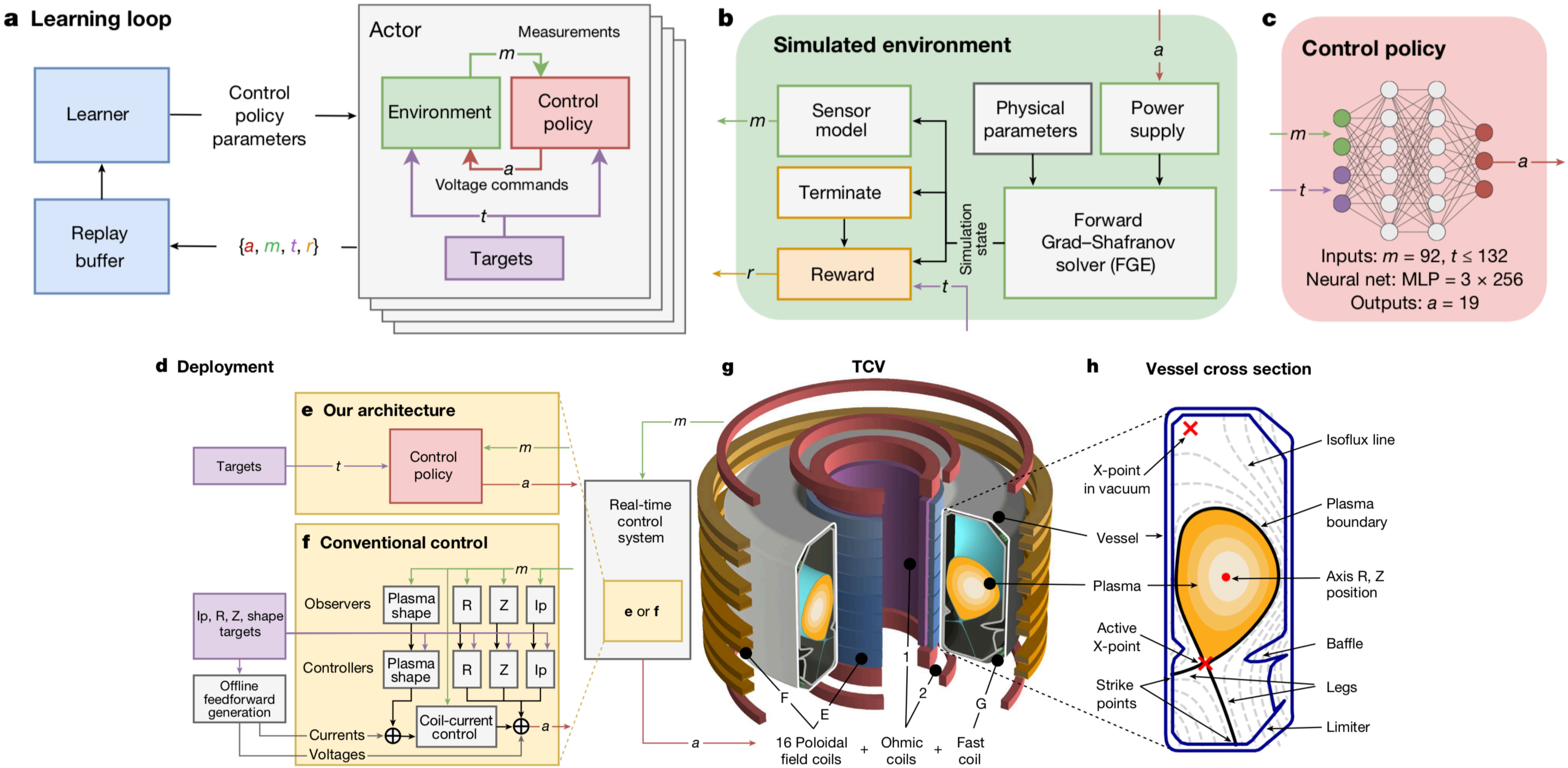
→ ML to automate → towards self-driving accelerators

Deepmind and Tokamak control

Magnetic control of tokamak plasmas through deep reinforcement learning

https://doi.org/10.1038/s41586-021-04301-9

Time-varying, non-linear, multi-variate control problem



Example from DESY

ARES Our Testbed

Small research accelerator at DESY's SINBAD facility

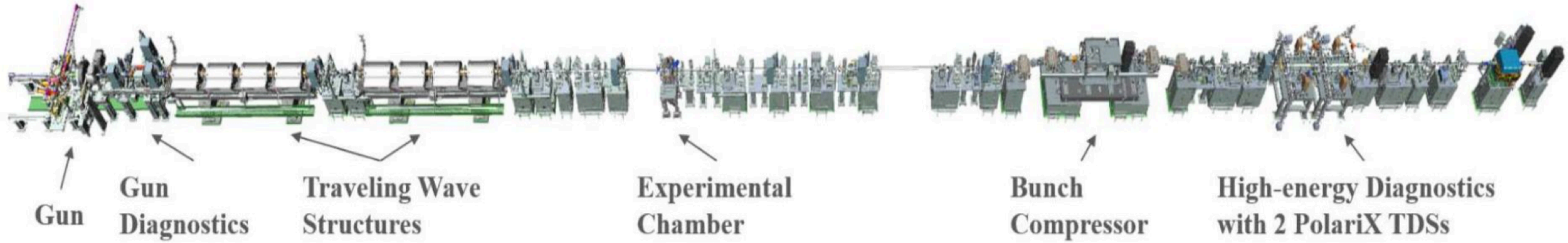
ARES a small (~60m) linear electron accelerator.

Goal: Characterize ultrashort e- bunches for applications (advanced & compact long. diag. and accelerating structures development, FLASH radiotherapy, etc.).

ARES is an easy accessible accelerator posing as a testbed for testing new developments.

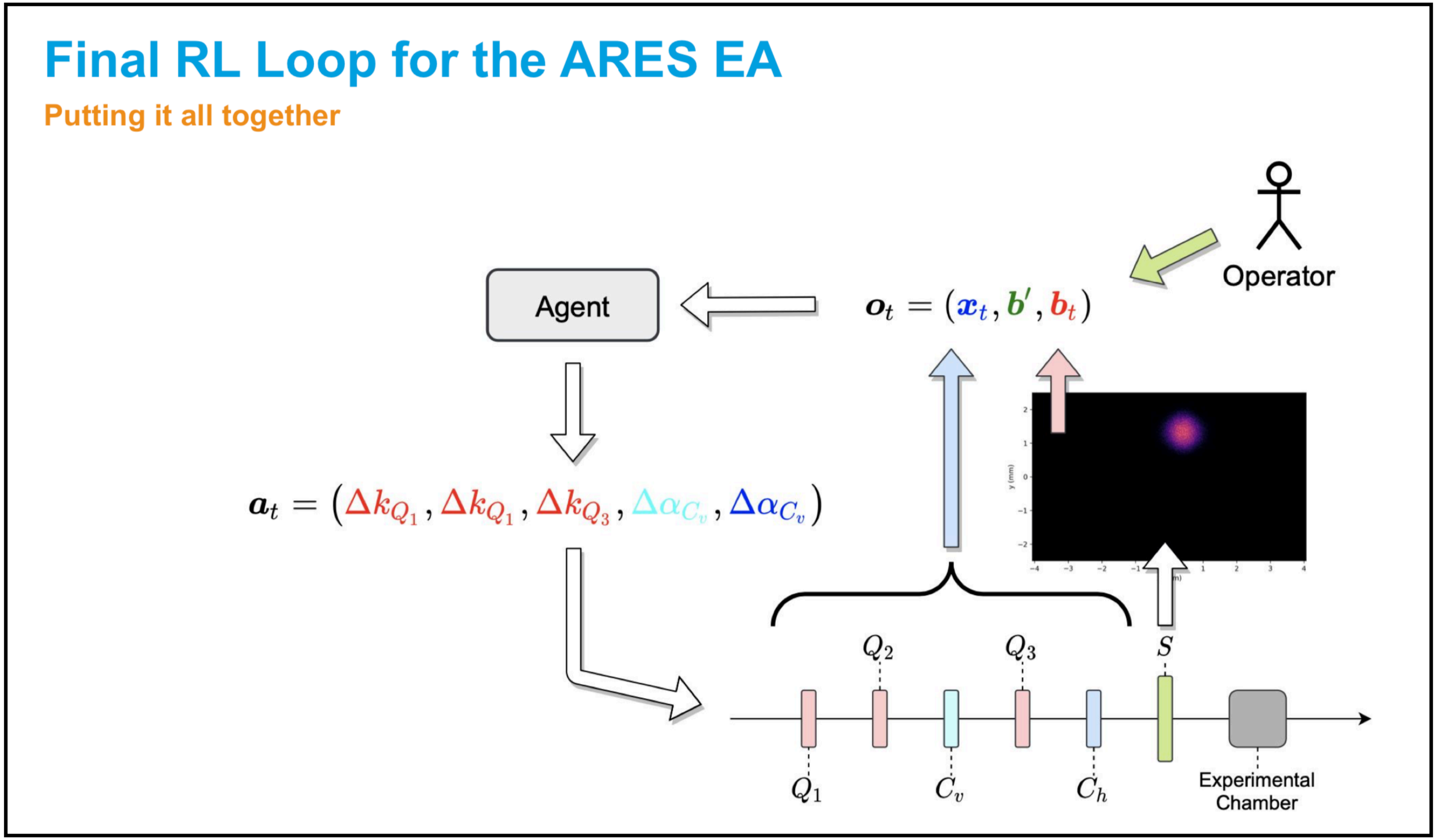
Frequent (almost weekly) beam times at ARES with close collaboration with the ARES team.

Properties	Target Values	Status
Charge	0.1 - 100 pC	0.1 – 50 pC
Momentum	50 -100 Mev/c	50 – 156 MeV/c
Momentum Spread	1,00E-04	1,00E-04
Transverse emittance	< 0.8 π.mm.mrad	≈ 0.4 π.mm.mrad
Duration	Sub-fs to ≈ 10 fs	≈ 40 fs



O. Stein, J. Kaiser, A. Eichler

Example from DESY



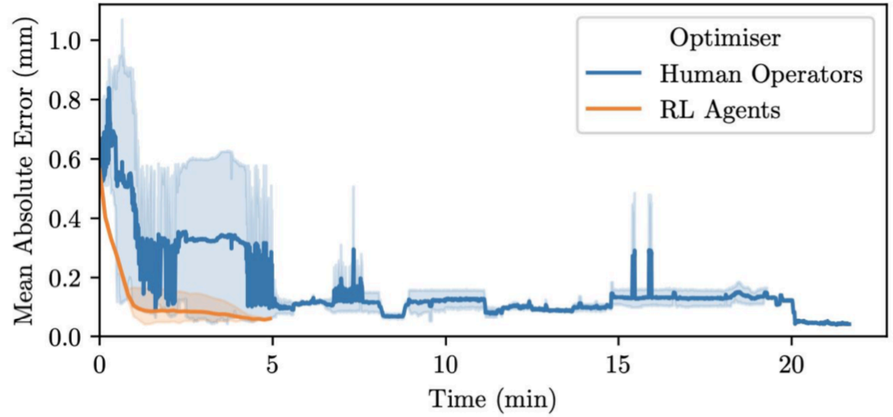
O. Stein, J. Kaiser, A. Eichler

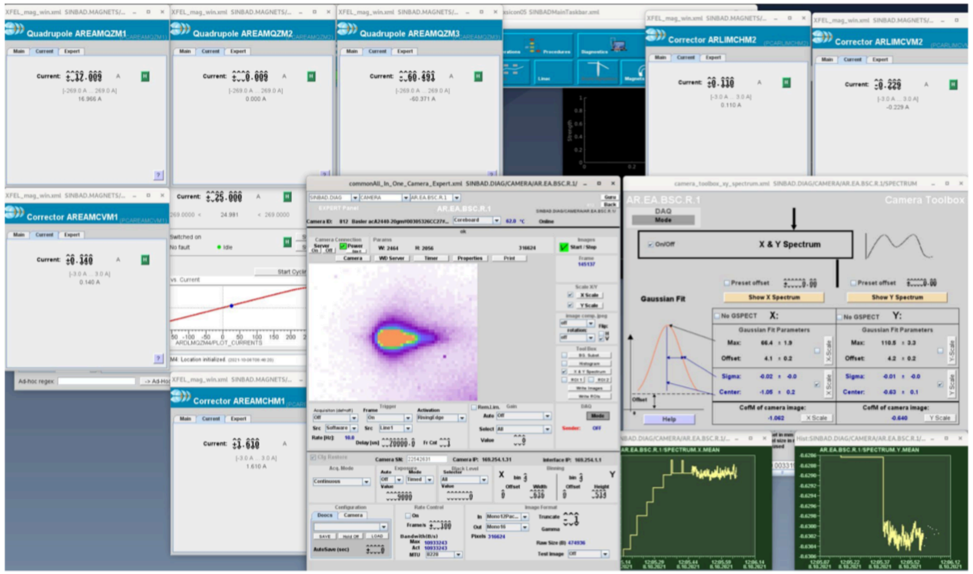
Example from DESY

Human vs. Machine

Can John Connor beat Polished Donkey?

- Two experienced human operators were tasked to solve the same problem as three sibling RL agents.
- Initially RL agents converge slightly faster than operators. Operators take a long time to fine tune to a competitive result but eventually achieve better result than RL agents.
- RL agent speed severely limited by accelerator hardware speed in ARES EA. Likely much faster on other hardware.
- RL agent performance much more consistent than operator performance.
- Both follow sensible strategies but humans simplify (use only two quadrupoles and focus on one parameter at a time) as opposed to RL agents who are agnostic to simplicity of solutions.





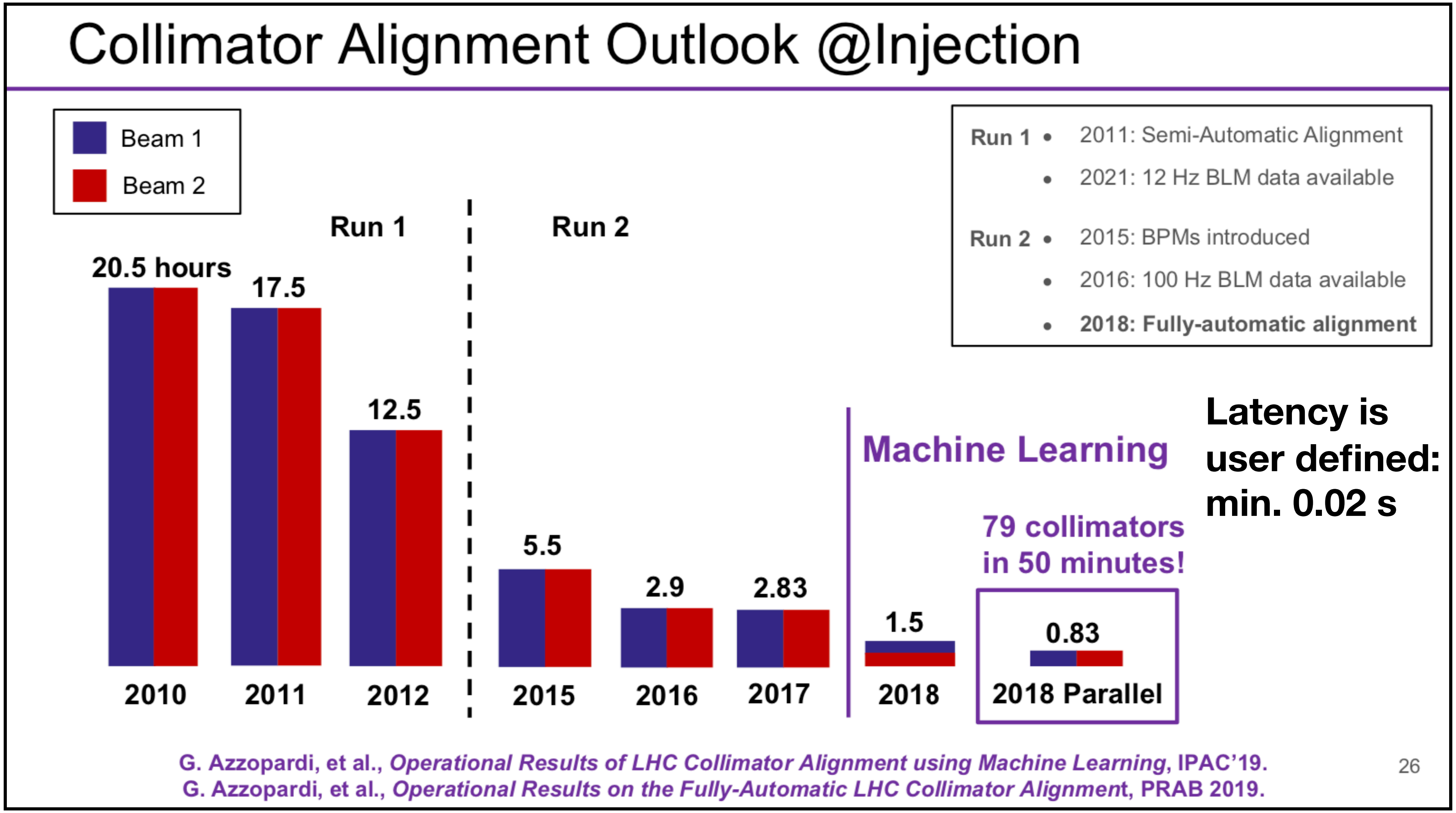
Operators took time ahead of timing them to set up well, though.

O. Stein, J. Kaiser, A. Eichler

Back to CERN: Speed up commissioning

Almost 100 collimators in the LHC…

SVM for beam loss signal classification on front end computer.



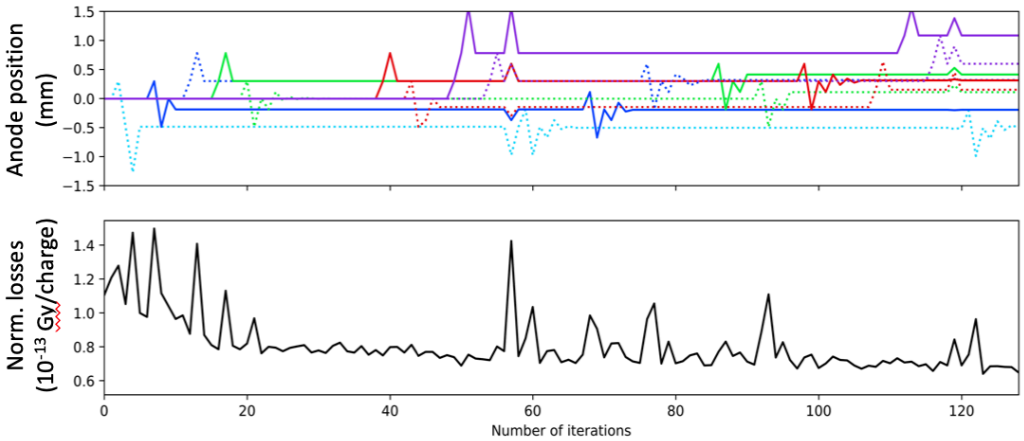
G. Azzopardi et al.

Speed up commissioning… again alignment

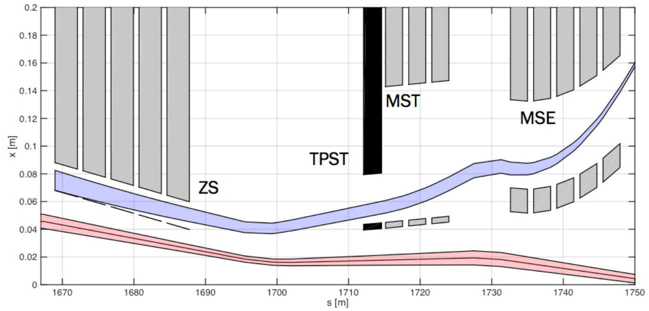
Alignment of electro-static septum, SPS, 9 DOF

November 2018: algorithm POWELL

- ✳ Before numerical optimisation for alignment: ~ 8 h

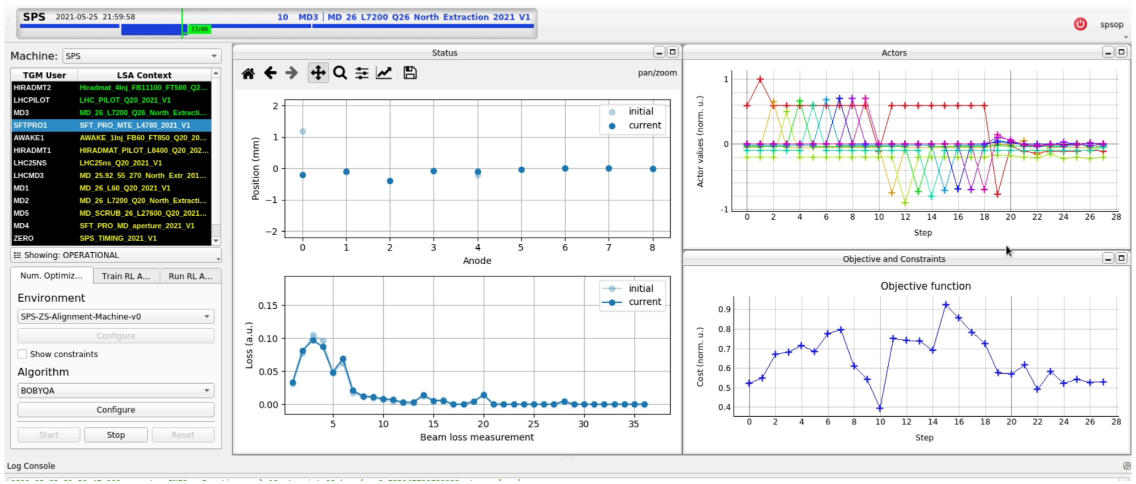


Circulating (red, ±3σ) and extracted (blue) horizontal beam envelopes and apertures in the LSS2 extraction region.



2018
~ 130 iterations.
~ 45 minutes

2021 BOBYQA and generic optimisation framework



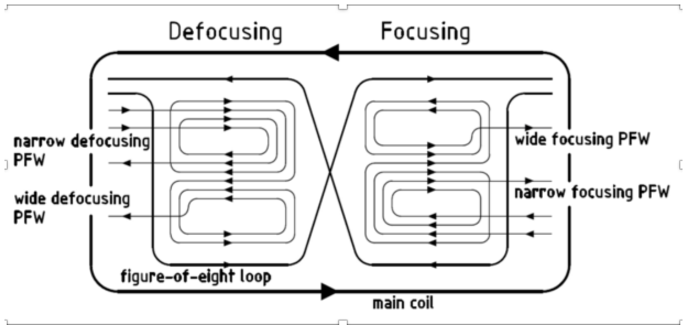
2021
~ 30 iterations.

Speed up commissioning: use models

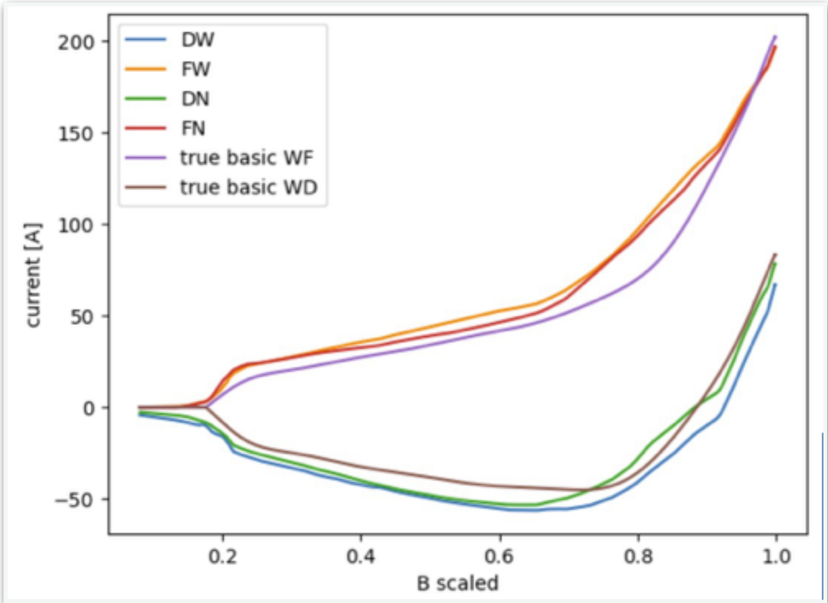
Build models from data → regression with ML

Example: **Modelling “Pole Face Windings" control PS**

- Control for $\Delta Q_{h,v}$ and $\Delta Q'_{h,v}$ available from polynomial fits
- From data learn neural network F^* for “generation" of current functions for desired $Q_{h,v}$ and $Q'_{h,v}$ function for given B function
- Will be cycle-by-cycle feedforward correction on most abstract level → settings server (LSA server)



$$F^* \begin{pmatrix} Q_x \\ Q_y \\ Q'_x \\ Q'_y \\ B \end{pmatrix} = \begin{pmatrix} I_{DN} \\ I_{FN} \\ I_{DW} \\ I_{FW} \\ I_{8L} \end{pmatrix}$$

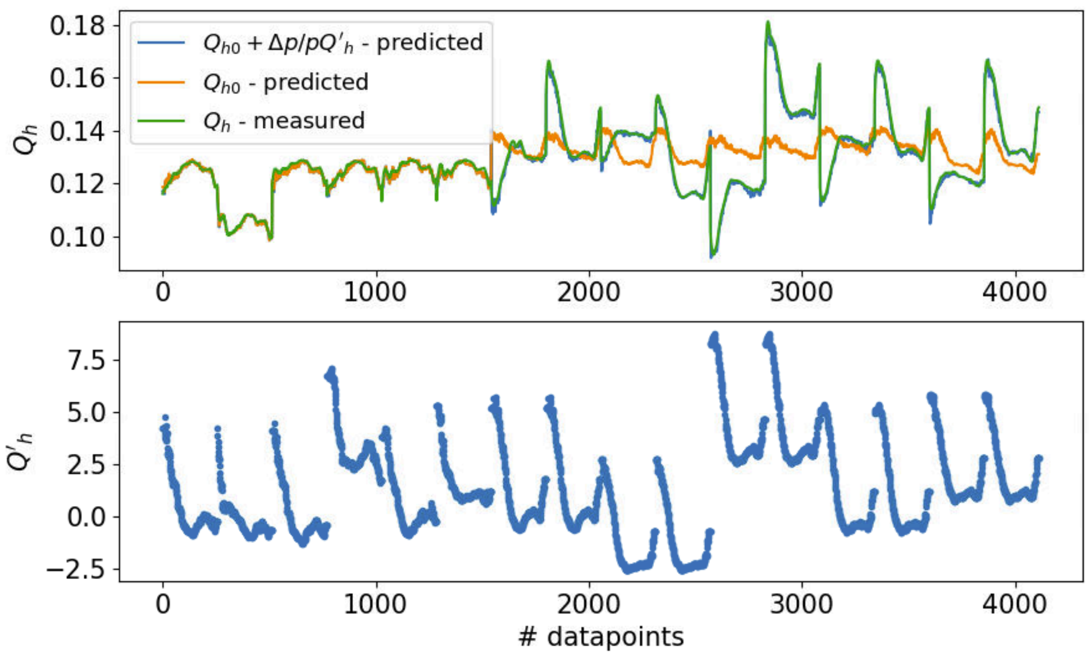


Speed up commissioning: use models

Example: **Modelling effect of eddy current on $Q_{h,v}$ and $Q'_{h,v}$ in the SPS; goal: generate sextuple and quadrupole circuit settings for given $B, \dot{B}, Q_{h,v}, Q'_{h,v}$**

- Similar to PFW modelling in the PS
- $\rightarrow Q_{h,v}$ data taken through acceleration for different $\dfrac{\Delta p}{p}$ with PI

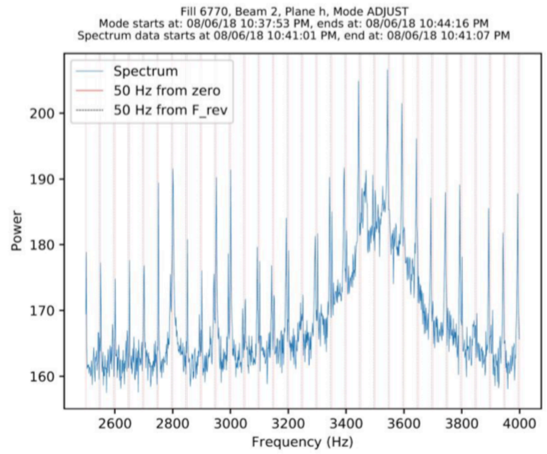
loss function $\mathscr{L} = \sqrt{\left[Q_{h_{true}} - (Q_{\beta h} + \frac{\Delta p}{p} Q'_h)\right]^2 + \left[Q_{v_{true}} - (Q_{\beta v} + \frac{\Delta p}{p} Q'_v)\right]^2}$



$$F\begin{pmatrix} k_{QF} \\ k_{QD} \\ k_{SF1} \\ k_{SD1} \\ k_{SF2} \\ k_{SD2} \\ k_{S3} \\ B \\ \dot{B} \end{pmatrix} = \begin{pmatrix} Q_{\beta h} \\ Q_{\beta v} \\ Q'_h \\ Q'_v \end{pmatrix}$$

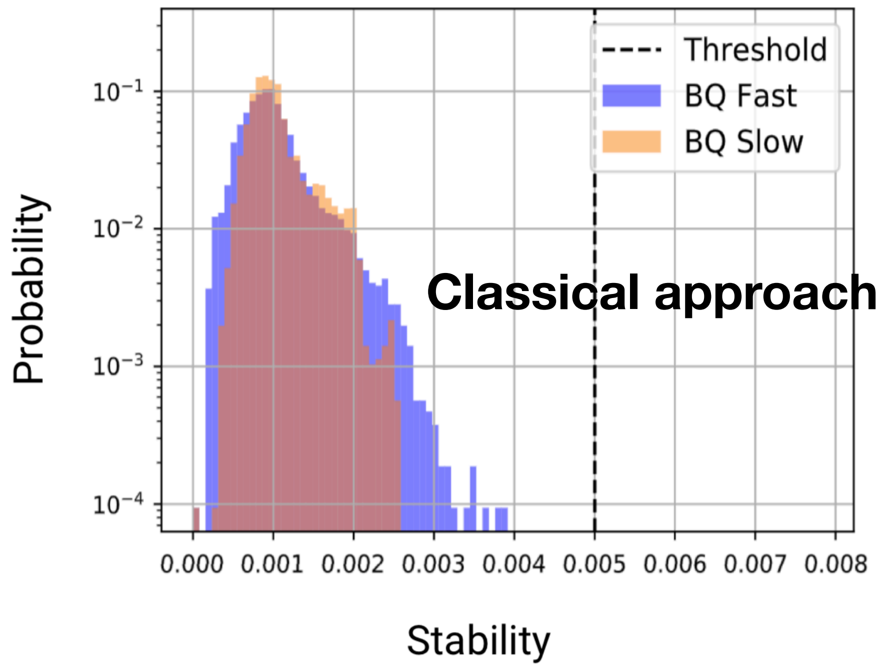
Efficient operation: interpreting the LHC tune signal

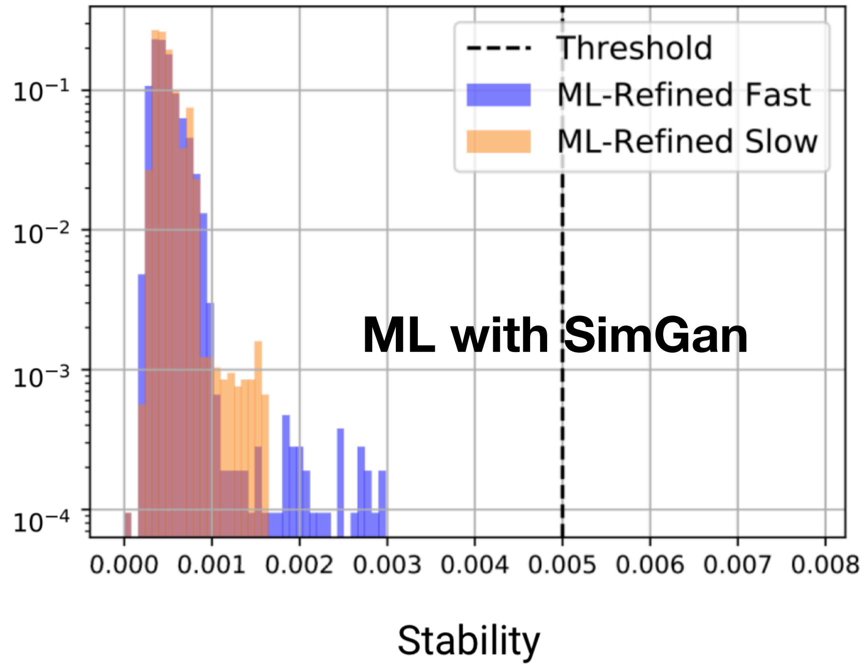
Tune estimation algorithm on BBQ spectra that does not get fooled by 50 Hz noise. Signal available at 1 Hz.



Refined Approach II

- What about the QFB stability metric?





- ML-Refined gave the most stable estimates from all tune estimation systems attempted

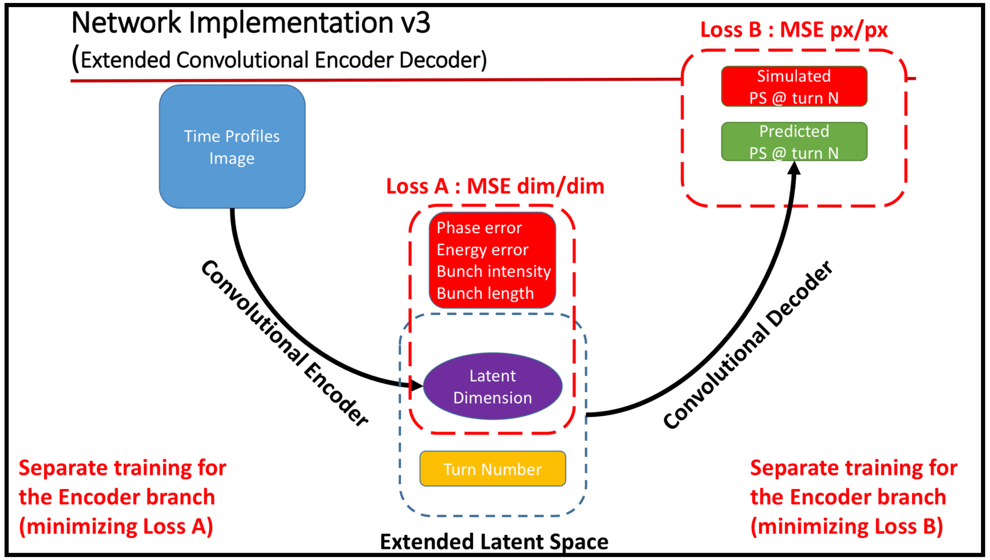
L. Grecht

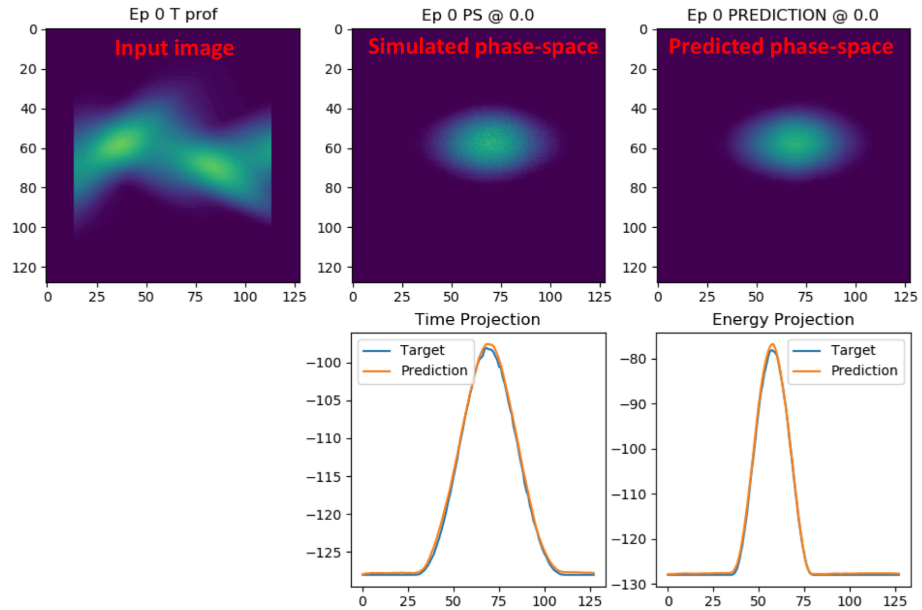
Efficient operation: LHC bunch-by-bunch parameters from AI tomography

Longitudinal beam parameters at injection from fit of longitudinal profiles and tomography.

→ in the LHC online only possible for single bunch; too time consuming

… unless one uses ML.



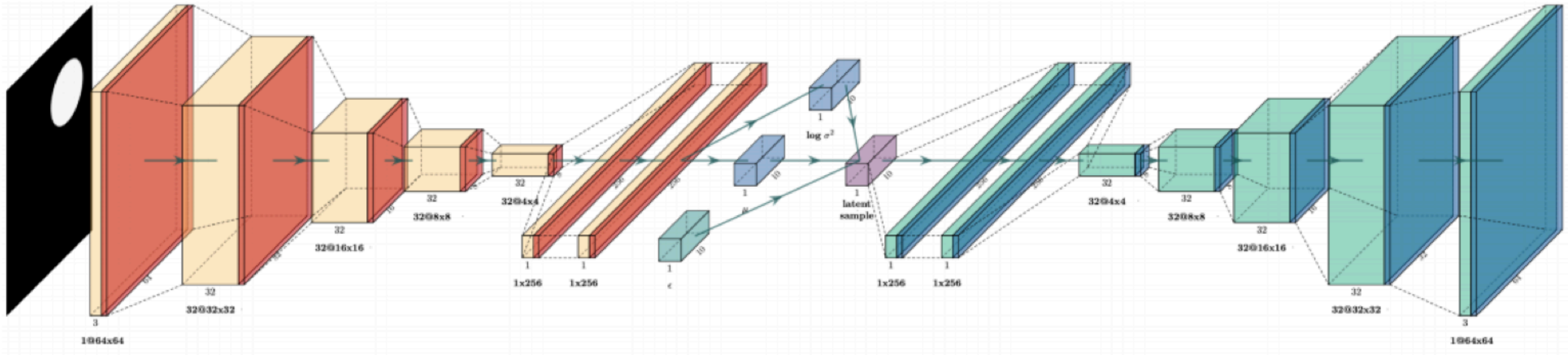


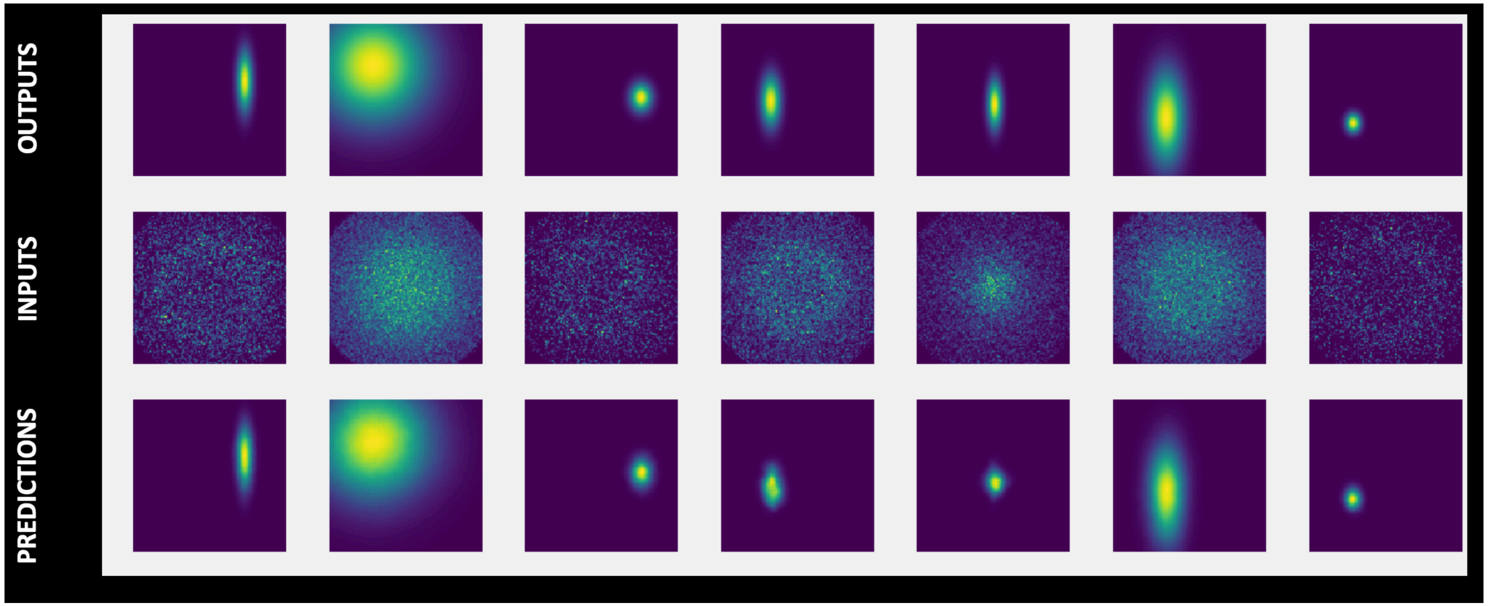
Once per injection: ~ 30 s

G. Trad and T. Argyropoulos

Efficient operation: more diagnostics…

Variational Auto-encoders for radiation hard **Optical Fibre Imaging** → next generation beam profile monitors?



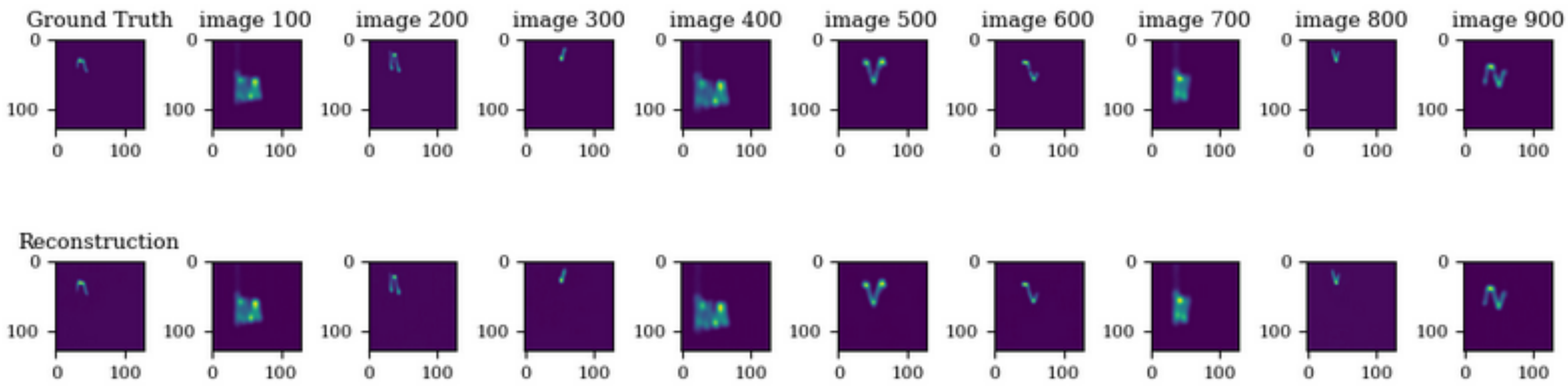


G. Trad

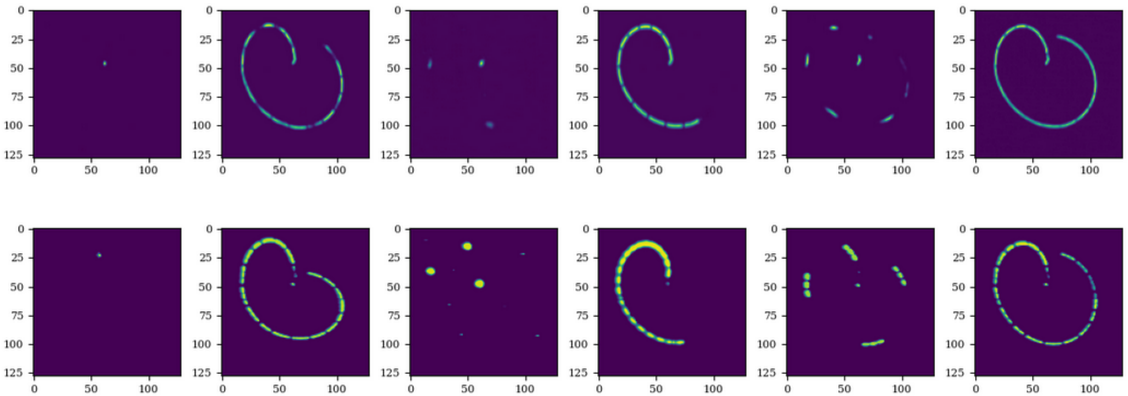
Detecting faults early: interpreting beam dump pattern

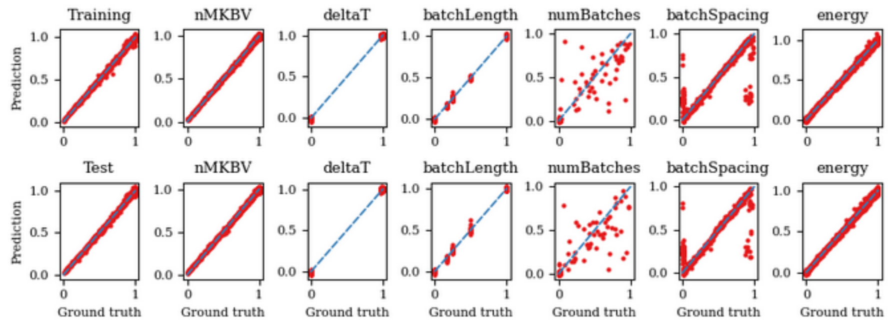
Classify dump kicker failures from the beam dump pattern images. SPS and LHC

Model results in simulated data



Model trained on simulations and applied on real data...and extract physical information about the system from images





F. Velotti and B. Goddard

Stable and flexible operation: Predicting timeseries data → Hysteresis control

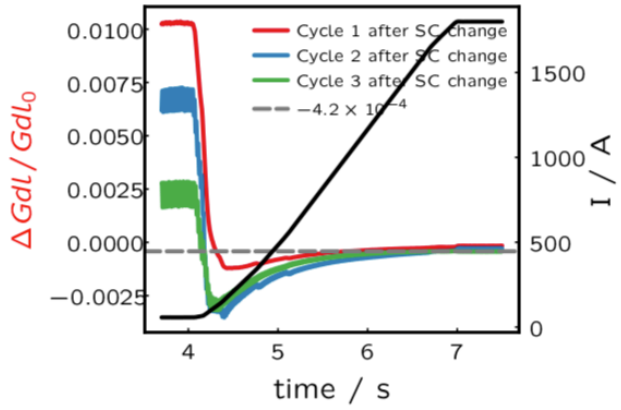
Big issue for multi-cycling machines (e.g. no LHC dedicated filling)

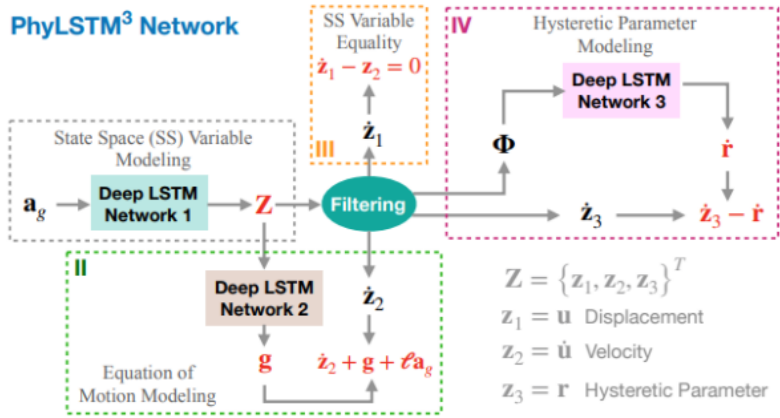
Example: predicting hysteresis of SPS main quadrupoles

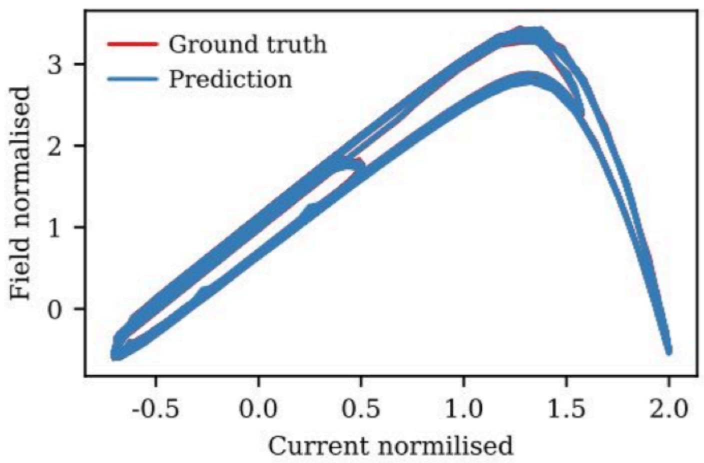
Solution with sufficient precision: Physics inspired LSTMs for field prediction; field measured in lab

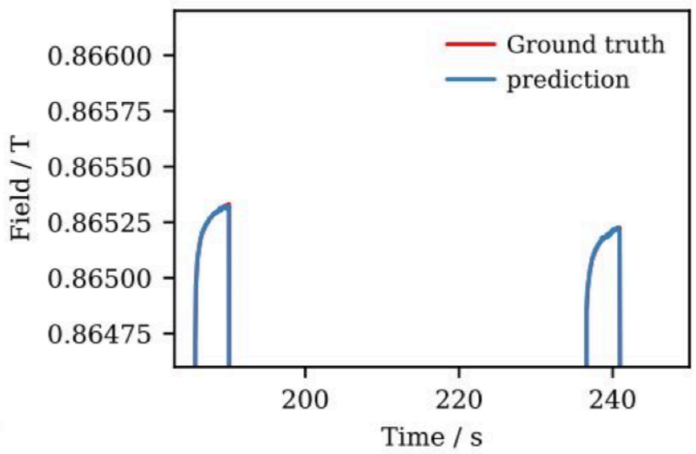
To be tested during run of 2022.

$$\mathcal{L} = \sum_{n=1}^{N} \frac{1}{N} \left(\alpha \left\| y - \bar{y} \right\|_2^2 + \beta \left\| \dot{y} - \dot{\bar{y}} \right\|_2^2 + \gamma \left\| \ddot{\bar{y}} - NN(x, \dot{y}) \right\|_2^2 \right)$$



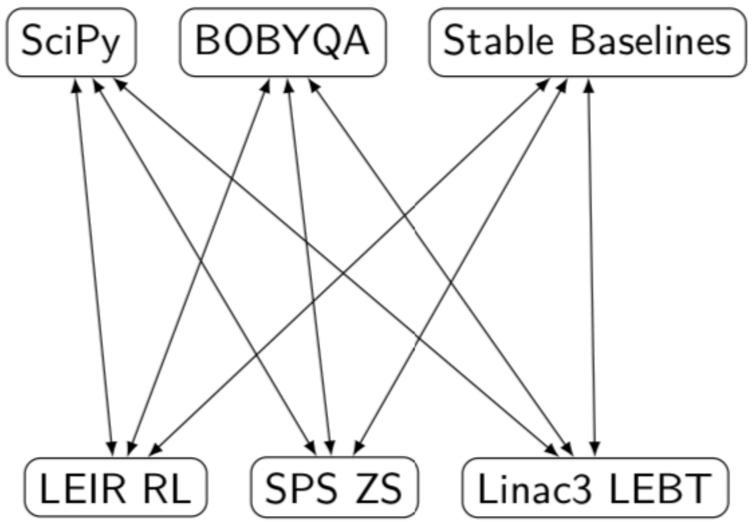


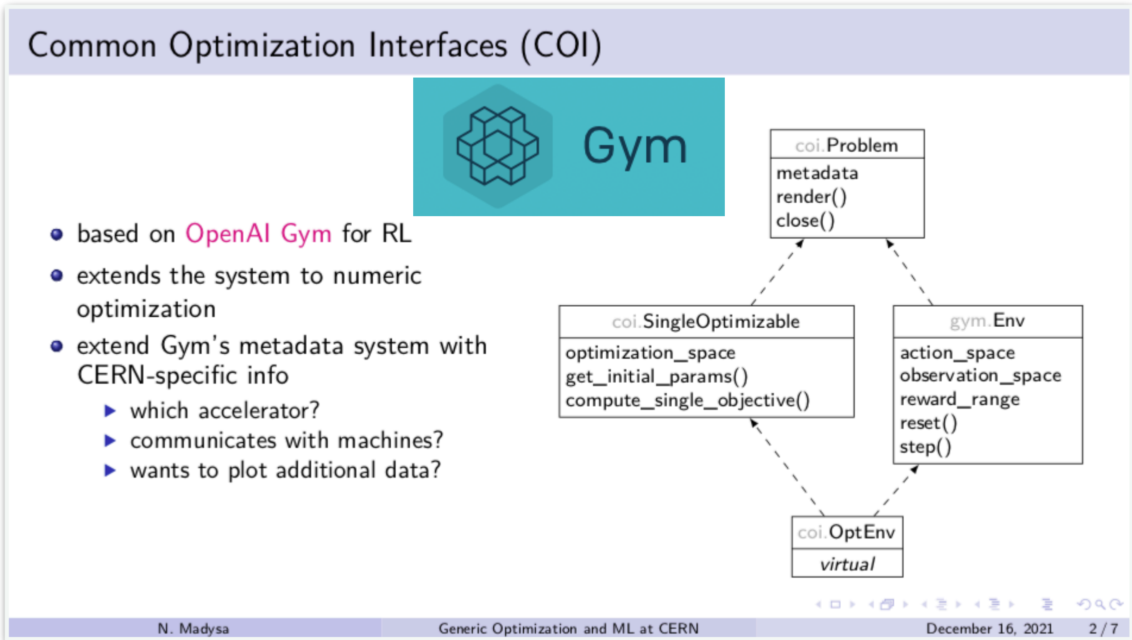


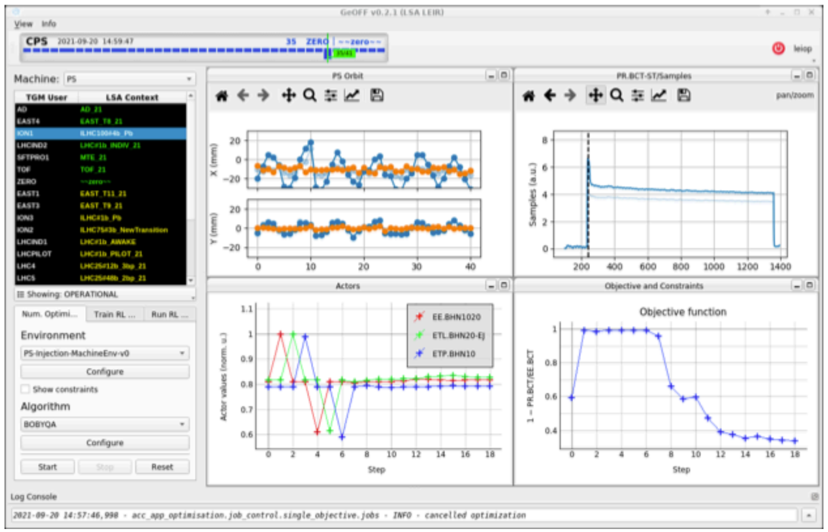


Make numerical optimisation plug&play

Offer GUI, hide algorithm complexity → separate optimisation problem definition from algorithm







CERN GeOF (Generic Optimisation Framework):

→ optimisation problems interfaces + GUI

→ GUI to deal with control aspects, e.g. RBAC

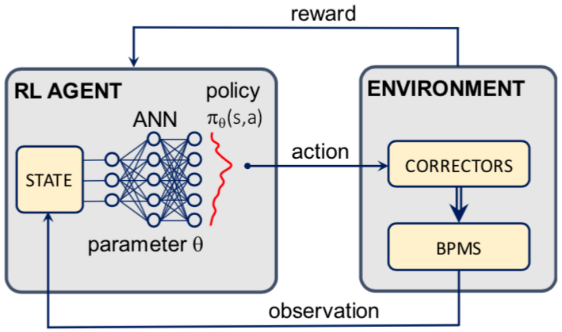
→ offers basic default plotting

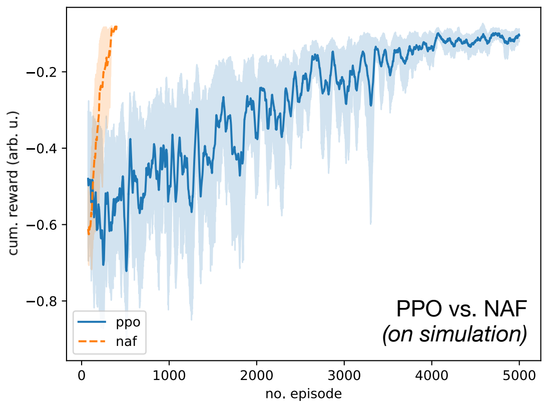
→ allows to integrate custom plotting

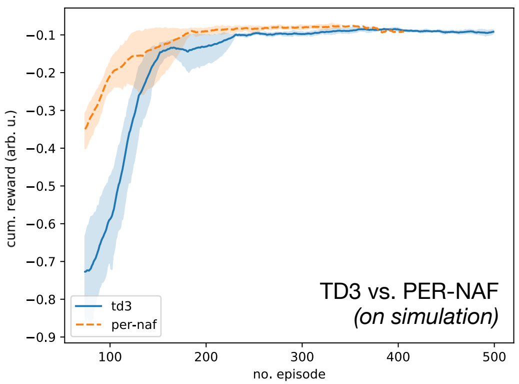
→ not only numerical optimisation, but also RL

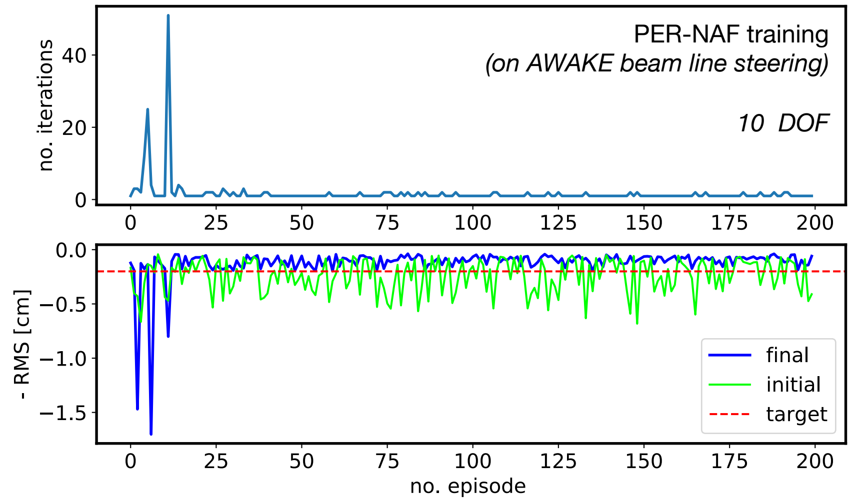
Reinforcement Learning (RL)

- **Numerical optimisation needs exploration at every deployment**
- **With RL, after training, exploration phase is reduced to a minimum**
  → one iteration in the best case
  - ◉ Agent learns underlying **dynamics of the problem**
  - ◉ **Key:** sample efficiency, cont. action







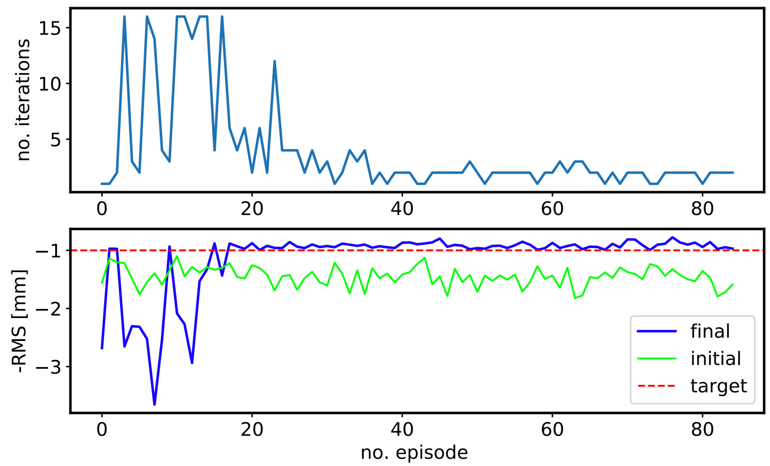


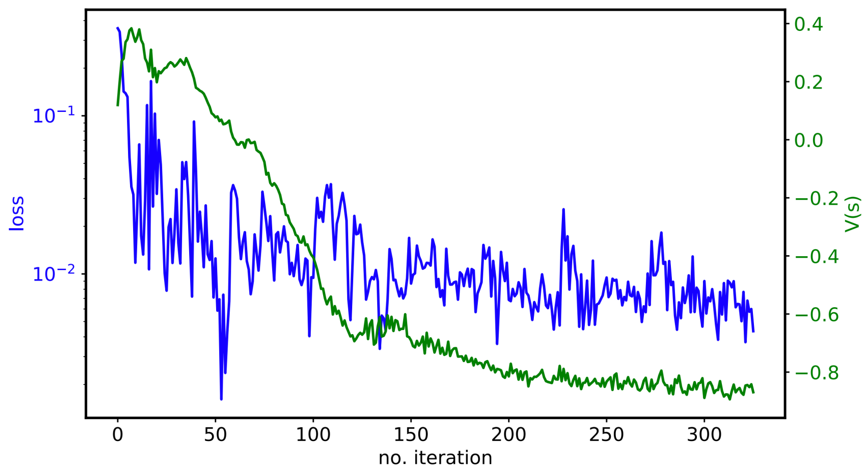
→ **Controllers like with model-predictive control**

Reinforcement Learning as part of the Toolkit

Real world Application: Linac4 H^- commissioning run (16 DOF)

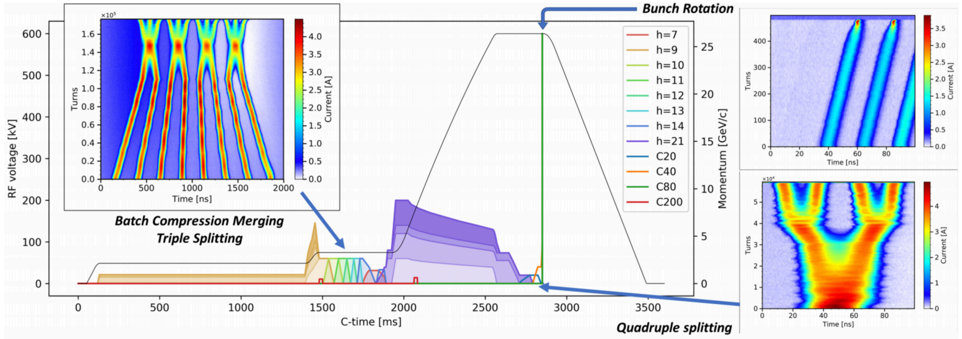
- Similarly to AWAKE beam line, using **NAF with *Prioritised Experience Replay***
- **Hyper-parameters:** quite some tuning necessary to avoid running into machine safety constraints during exploration
- After training for **~25 episodes** (~125 iterations), agent starts correcting well
- Corrects to **better than 1 mm RMS** within 3 iterations
- Value function converges to -0.85, corresponding to **0.85 mm RMS**





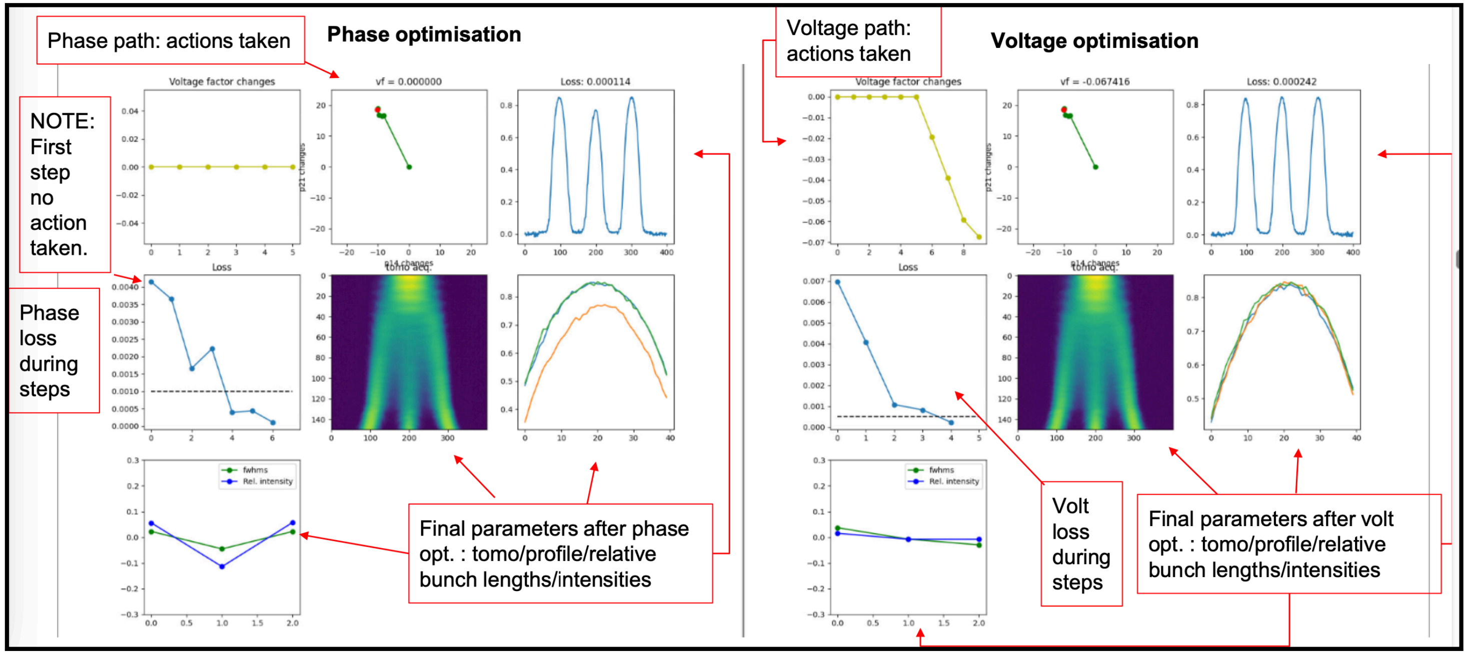
RL for RF manipulations in PS

- **LHC beam production** requires quadruple splitting at 26 GeV/c in PS
- **RF phase errors** introduce spread in bunch-by-bunch intensity and emittance



- **Train RL agent to correct RF phase and voltage to produce uniform splitting**
- Sample-efficiency a concern for training on PS directly
  - ★ Use BLonD code to produce simulation data set and create RL environment on top
  - ★ Train agent till convergence (~50k iterations)
  - ★ RL algorithm: Soft Actor-Critic (SAC); multi-agent algorithm using CNN to define initial set point

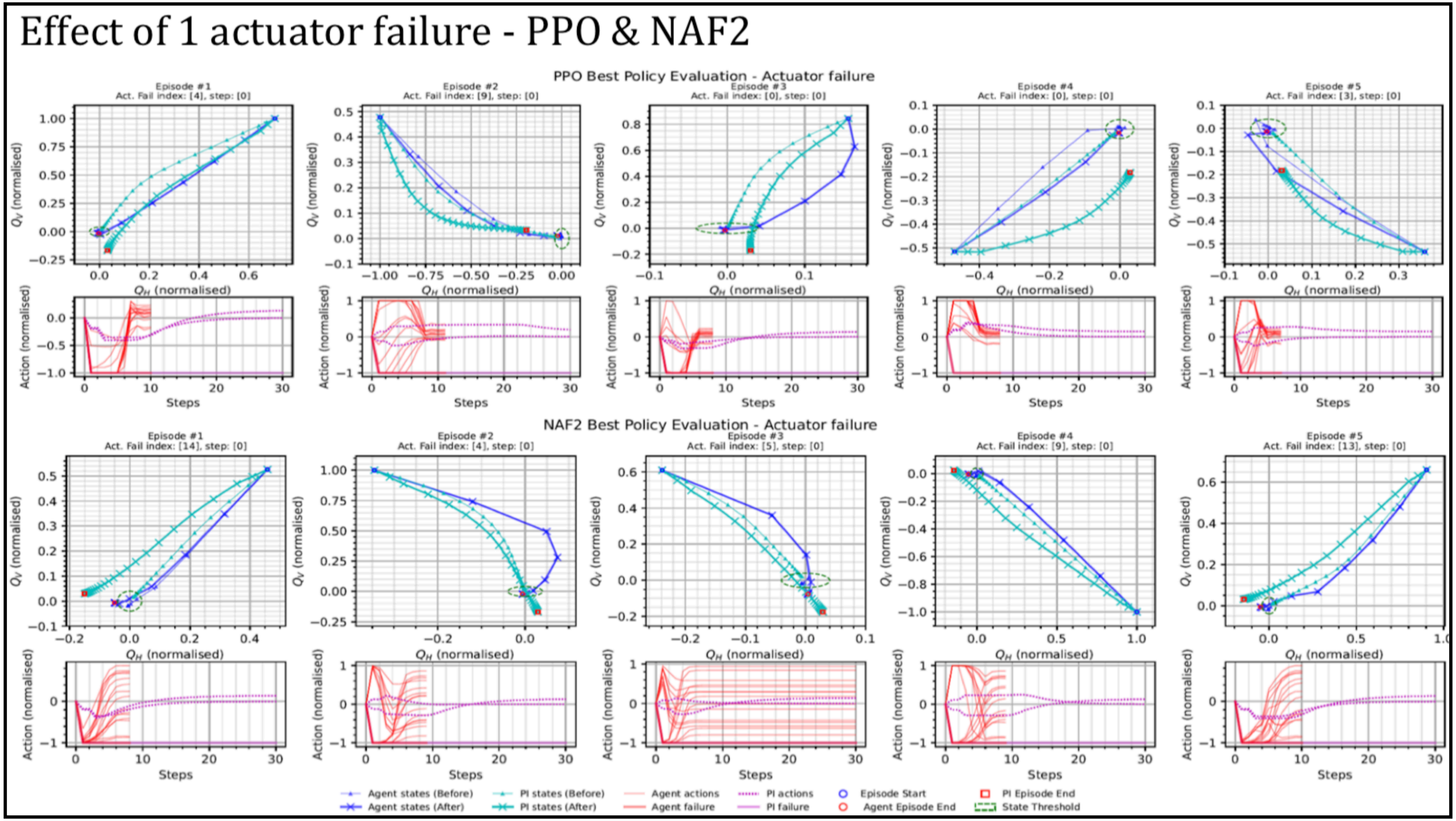
J. Wulff and A. Lasheen



Successful Deployment in the control room

RL for tune control in the LHC?

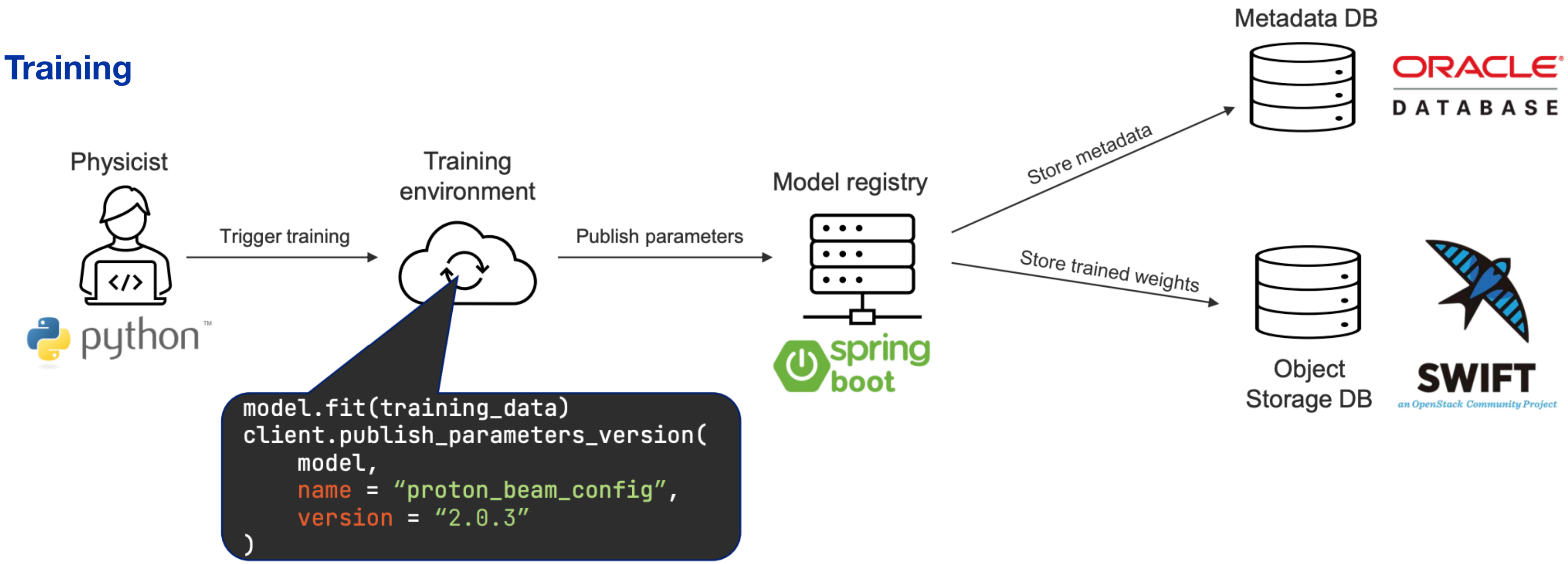
RL more robust controller than PI controller for tune feedback in the LHC? **Current orbit feedback running at 0.1 - 1 Hz**



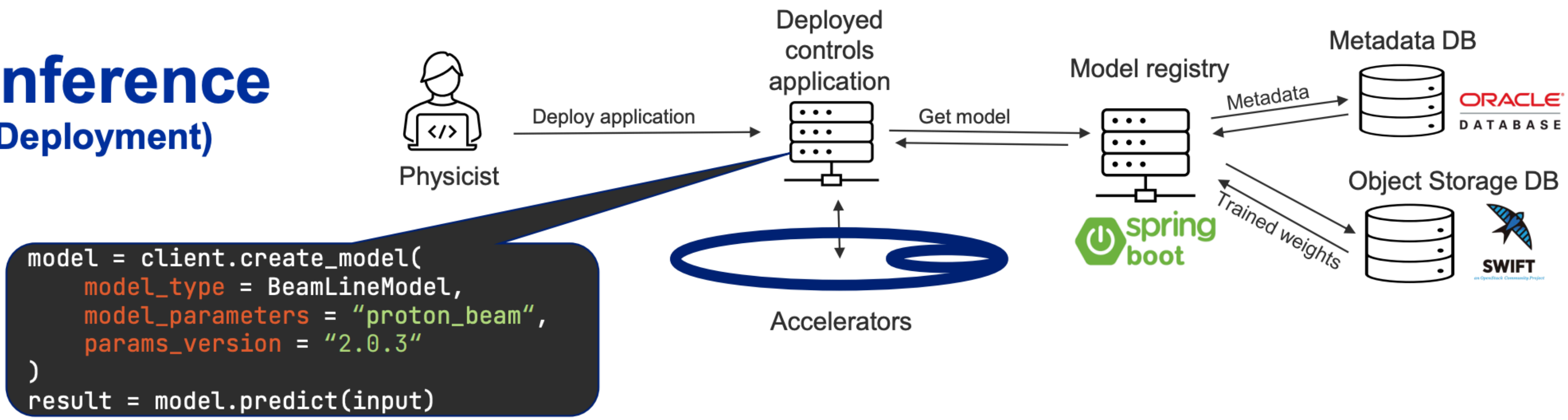
L. Grecht

Infrastructure: Machine Learning Platform

Training



Inference
(Deployment)



J.-B. De Martel et al., *Machine Learning Platform - Deploying and Managing Models in the CERN Control System*, Dec 2021

Conclusion

ML will be one of the paths to further increase the efficiency and flexibility of CERN's accelerators.

- ★ Huge potential for efficiency at low cost - e.g. ML solution very often inexpensive
- ★ Already many successful implementations available
- ★ Additional push with recent issues with cost of energy

- ★ No ML on edge devices or chip yet for the CERN accelerator control
  - ✳ Potential for fast model-predictive control for RF systems and transverse damper if using learned models or other algorithms (e.g. RL)