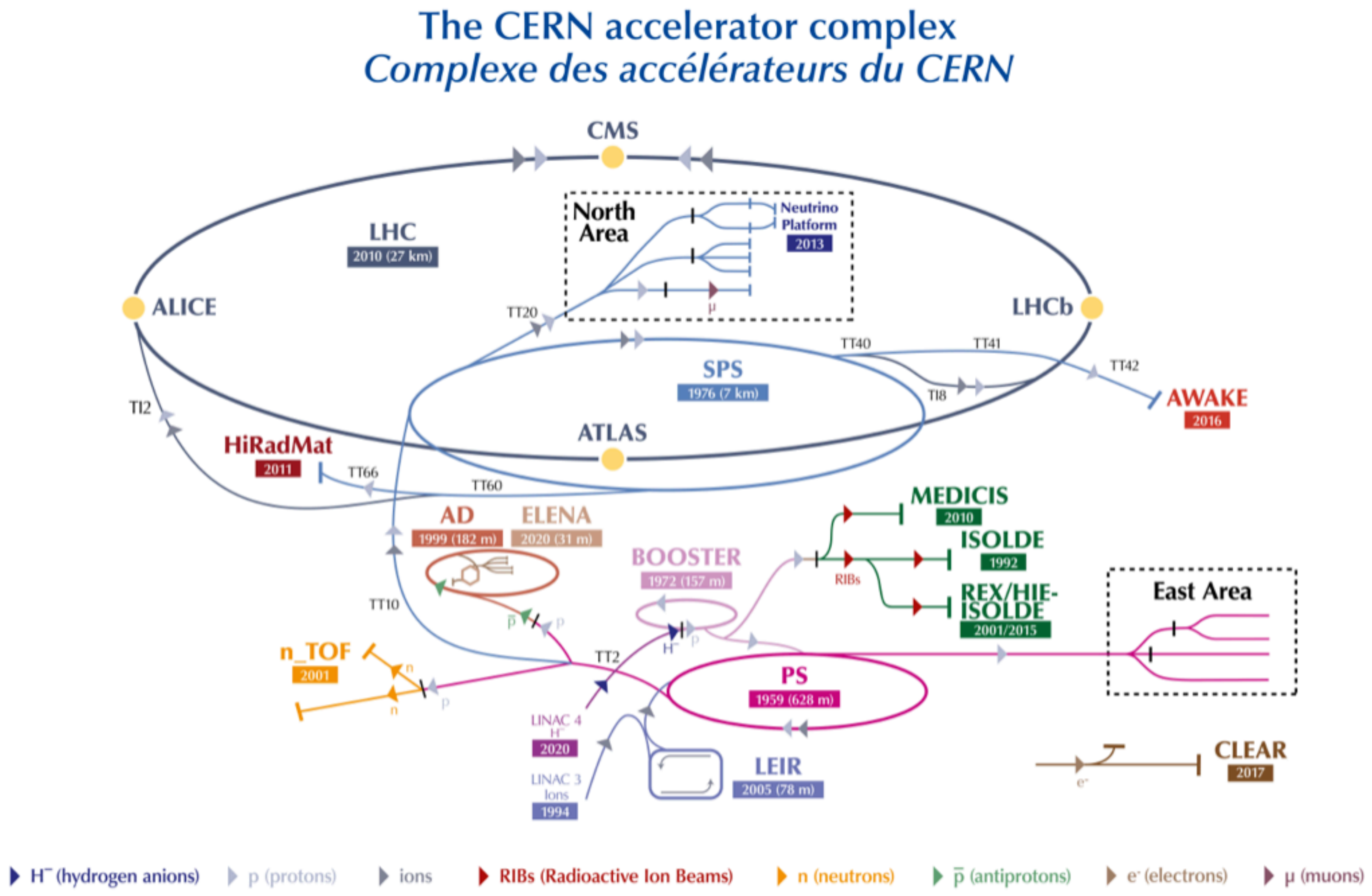
Reinforcement Learning at CERN's accelerators

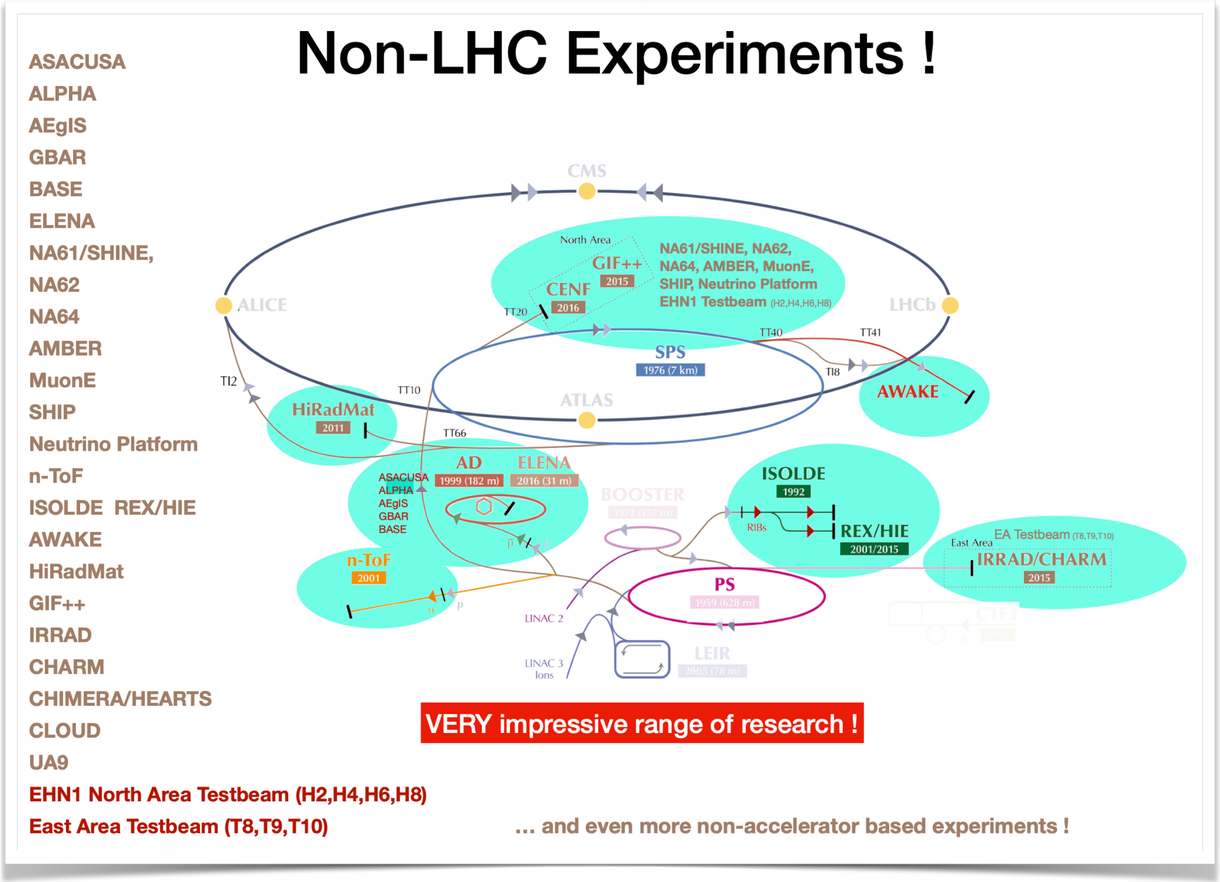
V. Kain, N. Bruchon, S. Hirlander, N. Madysa, B. Rodriguez Mateos, M. Schenk, M. Remta, F. Velotti, J. Wulff

The CERN accelerator complex



LHC and non-LHC physics

From last week's Chamonix Accelerator Performance workshop: **many, many different beams!**



Flexibility comes with a price...

Summary talk Injector and Experimental Facility Workshop (IEF) 2021

2. Address reproducibility and availability

- Availability OK, under control of Groups. **Reproducibility** is critical concern with increasing flexibility and multi-destination operation
- Transmission problems and instability in beam delivery in many locations. "Need more time in 2022"→have to ensure this is there (add in schedule?) #A
- Addressing reproducibility relies on many factors including equipment, accelerator modelling and high-level controls approach

Other input from IEF'21

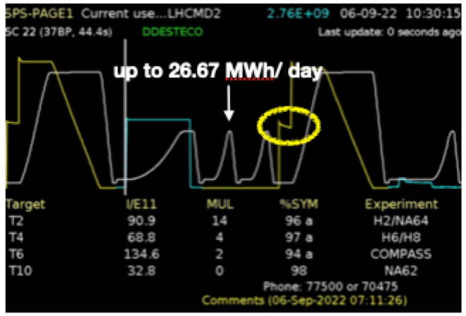
→ Current **beam scheduling** has severe impact on resources needed to run accelerators and on efficiency

- ✳ Statistics: 20-100 clicks to change supercycle = 2-25 min; 40-60 times/24 h

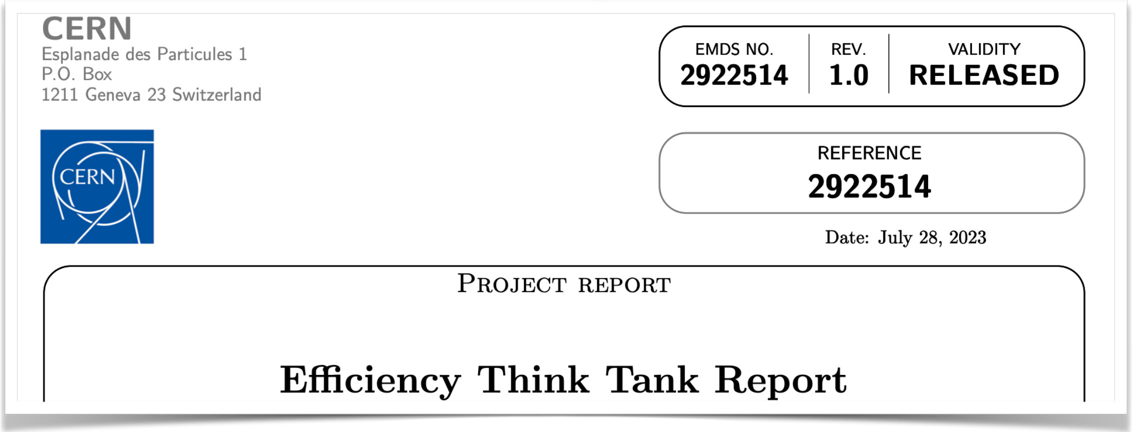
Input from JAPW'22

→ **Hysteresis** is severe limitation for efficiency and flexibility in most machines, current mitigation methods wasting energy

- ✳ ~ 15 % of yearly cost of SPS fixed target cycle for "waste" cycles and quasi-degauss Cycle MD1



7 recommendations → Automating exploitation


CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

1. **Hysteresis compensation**
2. **Automatic and dynamic beam scheduling**
3. **Automatic LHC filling**
4. **Auto-pilots**
5. **Automatic fault analysis, recovery and prevention**
6. **Automatic testing and sequencing**
7. **Automatic parameter optimisation**

7 recommendations → Automating exploitation

CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

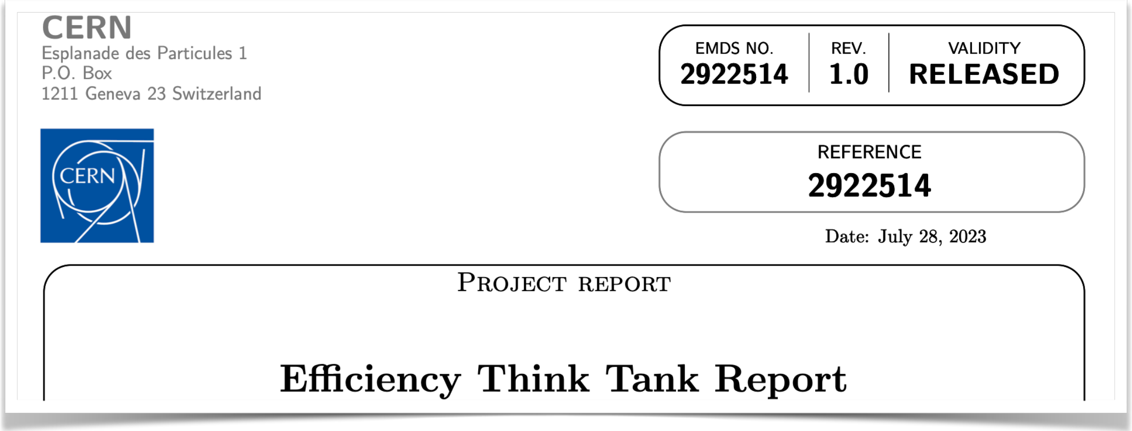
1. **Hysteresis compensation**
2. **Automatic and dynamic beam scheduling**
3. **Automatic LHC filling**
4. **Auto-pilots**
5. **Automatic fault analysis, recovery and prevention**

→ **Fully automated standard physics operation**

6. **Automatic testing and sequencing**
7. **Automatic parameter optimisation**

→ **Goal: reduce commissioning time by 50 %**

7 recommendations → Automating exploitation


CERN
Esplanade des Particules 1
P.O. Box
1211 Geneva 23 Switzerland

EMDS NO.	REV.	VALIDITY
2922514	1.0	RELEASED

REFERENCE
2922514

Date: July 28, 2023

PROJECT REPORT

Efficiency Think Tank Report

AI Hysteresis compensation
2. Automatic and dynamic beam scheduling
AI Automatic LHC filling
AI Auto-pilots
AI Automatic fault analysis, recovery and prevention

→ Fully automated standard physics operation

6. Automatic testing and sequencing
AI Automatic parameter optimisation

→ Goal: reduce commissioning time by 50 %

Efficient Particle Accelerators (EPA) project

Approved in autumn 2023 after pre-study in Efficiency Think Tank (ETT)

10 work packages: ETT recommendations and controls infrastructure evolution.

	WP0 Project Management Lead: V. Kain Deputy: A. Huschauer	
WP1 Dynamic Beam Scheduling Lead: F. Irannejad	WP2 Automated LHC Filling Lead: A. Huschauer, G. Trad	WP3 Automated Parameter Control & Optimisation Lead: M. Schenk
WP4 Hysteresis Compensation Lead: C. Petrone	WP5 Next Generation Sequencer Lead: R. Gorbonosov	WP6 Efficient Settings Management Lead: M. Hostettler
WP7 Automated Equipment Testing Lead: A. Calia, J.C. Garnier	WP8 Automate Equipment Lead: F. Velotti, K. Papastergiou	WP9 Data Processing Framework Lead: M. Sobieszek

Efficient Particle Accelerators (EPA) project

Approved in autumn 2023 after pre-study in Efficiency Think Tank (ETT): project length 5 years

10 work packages: ETT recommendations and controls infrastructure evolution.

WP0
Project Management
Lead: V. Kain
Deputy: A. Huschauer

WP1
Dynamic Beam Scheduling
Lead: F. Irannejad

WP2
Automated LHC Filling
Lead: A. Huschauer, G. Trad

WP3
Automated Parameter Control & Optimisation
Lead: M. Schenk

WP4
Hysteresis Compensation
Lead: C. Petrone

WP5
Next Generation Sequencer
Lead: R. Gorbonosov

WP6
Efficient Settings Management
Lead: M. Hostettler

WP7
Automated Equipment Testing
Lead: A. Calia, J.C. Garnier

WP8
Automate Equipment
Lead: F. Velotti, K. Papastergiou

WP9
Data Processing Framework
Lead: M. Sobieszek

Automation Infrastructure - readiness

Many classical automation concepts came from the LHC → injectors

- Sequencer, high level parameter control
- **EPA WP5 (Next Generation Sequencer) & WP6 (Efficient Settings Management) to ensure evolution for new requirements**

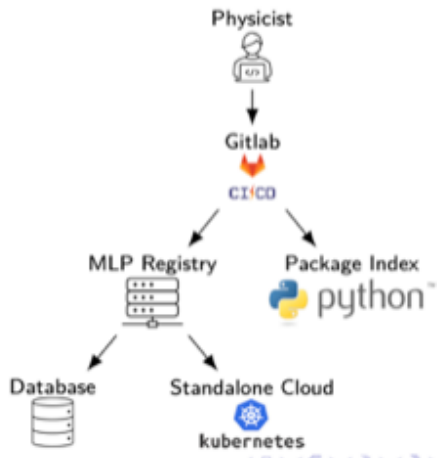
Automation Infrastructure - readiness

Many classical automation concepts came from the LHC → injectors

- ✳ Sequencer, high level parameter control
- ✳ **EPA WP5 (Next Generation Sequencer) & WP6 (Efficient Settings Management) to ensure evolution for new requirements**

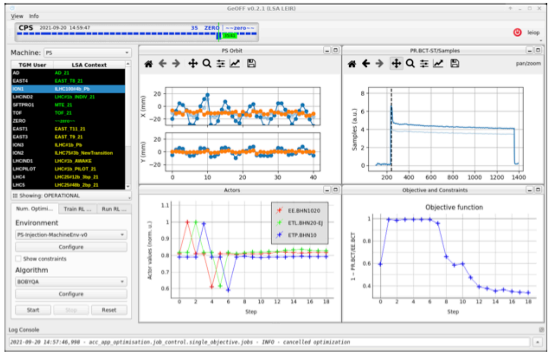
Since LHC: preparing for **automation including AI/ML** - injectors on forefront

- **Acc-Py** (Accelerating Python): unlocked the potential of Python in CERN ATS including control rooms
  - ✳ Python distribution, Python Package Index, release of applications to centrally managed deployment location
- **UCAP:** Unified Controls Acquisition and Processing ("Virtual Device Service") → servers on-the-fly in JAVA or Python
  - ✳ Provides infrastructure to run "transformations" and event building
  - ✳ Expect **evolution with EPA WP9** (Data Processing Framework)
- **MLP** (Machine Learning Platform): store and share AI models between users and applications of different languages



Automation Infrastructure - readiness

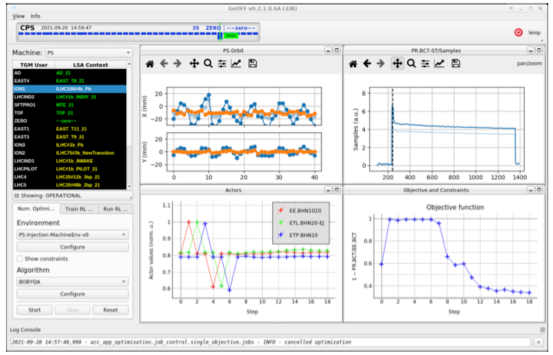
- **Generic Optimisation Framework GeOFF**
  - Manual scans and grid scans are inefficient for multi-parameter problems → optimisation algorithms
  - GeOFF = easy and flexible parameter optimisation in the control room
  - To date: > 20 parameter optimisation problems automated across complex



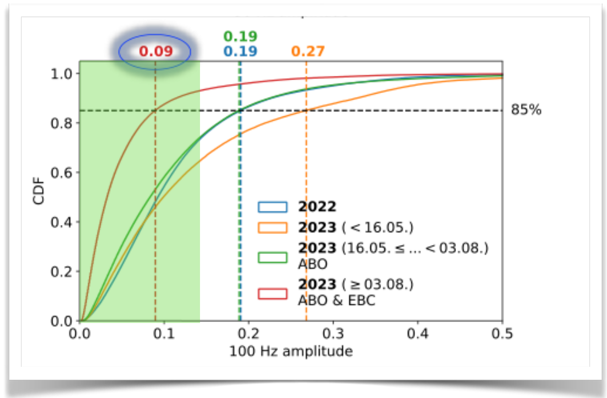
Automation Infrastructure - readiness

Generic Optimisation Framework GeOFF

- Manual scans and grid scans are inefficient for multi-parameter problems → optimisation algorithms
- GeOFF = easy and flexible parameter optimisation in the control room
- To date: > 20 parameter optimisation problems automated across complex



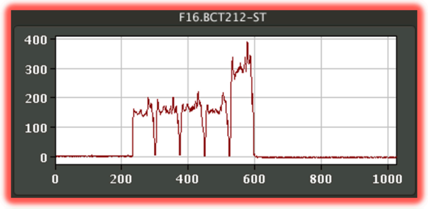
100 Hz content of NA spill with ABO and EBC

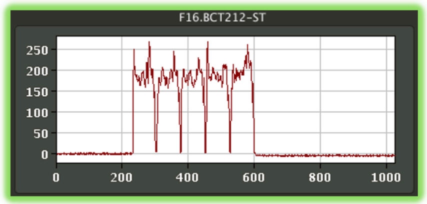


Optimisation framework for auto-pilots

- GeOFF on UCAP → acc-geoff4ucap released in summer 2023.
- Operational: $n \times 50$ Hz control for NA spill with **GPUs on UCAP**
- **EPA WP3** (Automated Parameter Control & Optimisation) to implement **in 2024:**
  - automated PS2SPS steering
  - MTE efficiency drift stabilisation

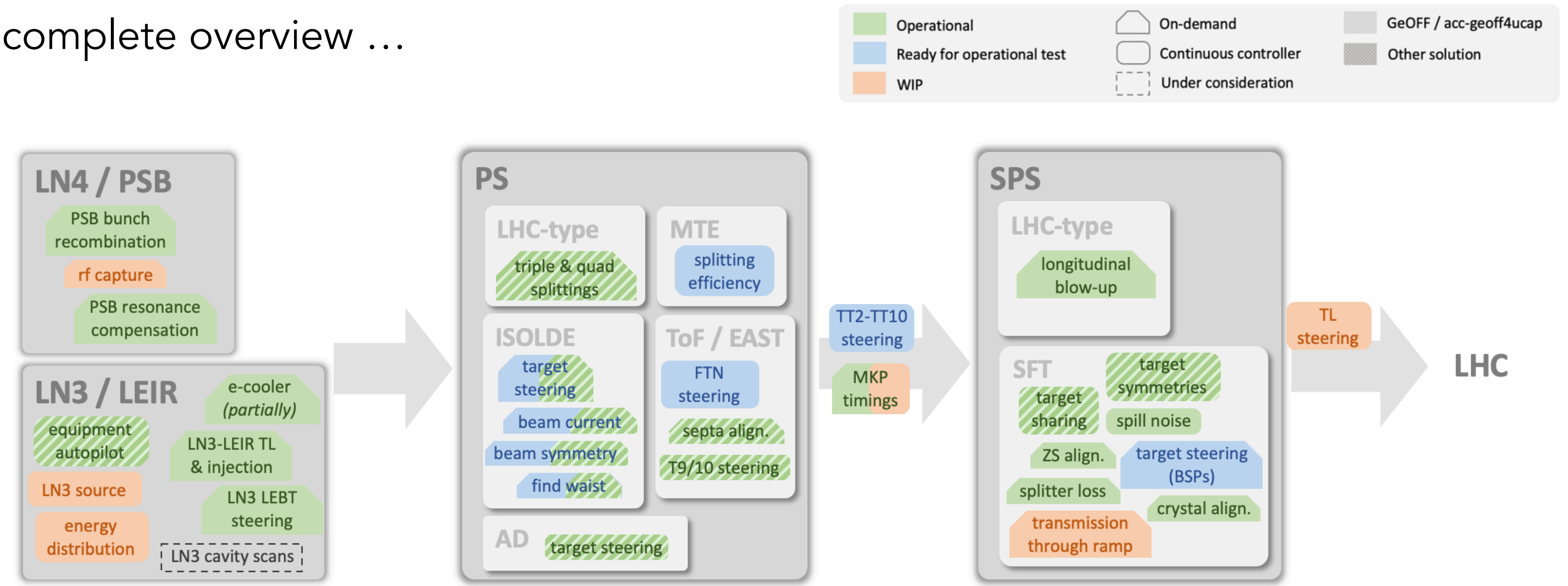
MTE island intensities to be equal with core





Status: Auto-pilots, optimisers,...

An incomplete overview …



Courtesy M. Schenk

Status 2023: many optimisers and auto-pilots used operational, many added in 2023

Trends 2024: on-demand → continuous (UCAP) | some new auto-pilots

Until end of run 3: automation of all typical optimisation and continuous control problems

What about RL as auto-pilot?

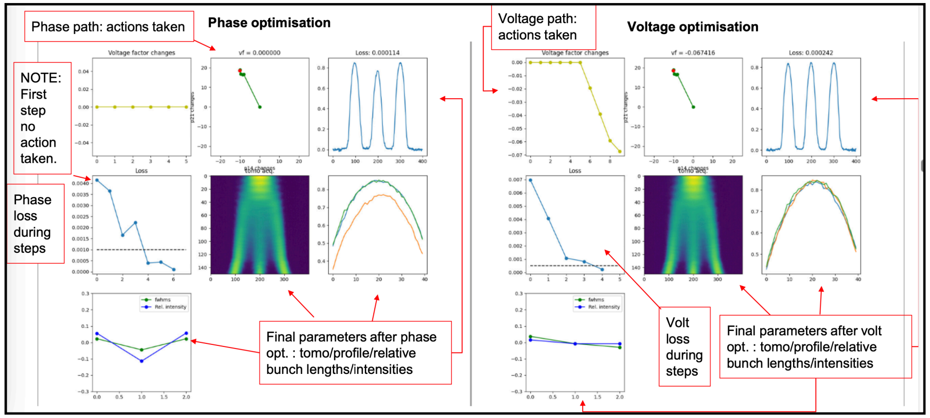
RL in the control room

RL agent to correct **RF phase and voltage** to produce uniform RF splitting in PS for LHC beams

★**Trained in simulation and successfully transferred to control room → fully operational**

★RL algorithm: Soft Actor-Critic (SAC); multi-agent algorithm using CNN to define initial set point

★Next step: from on-demand to continuous: → UCAP



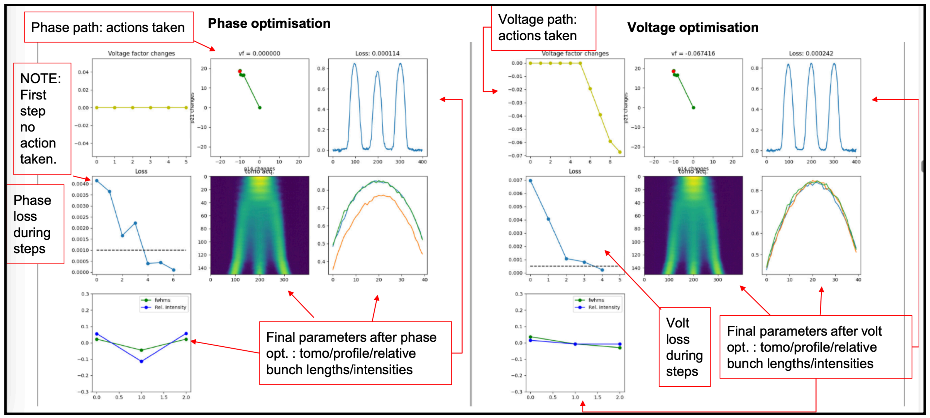
RL in the control room

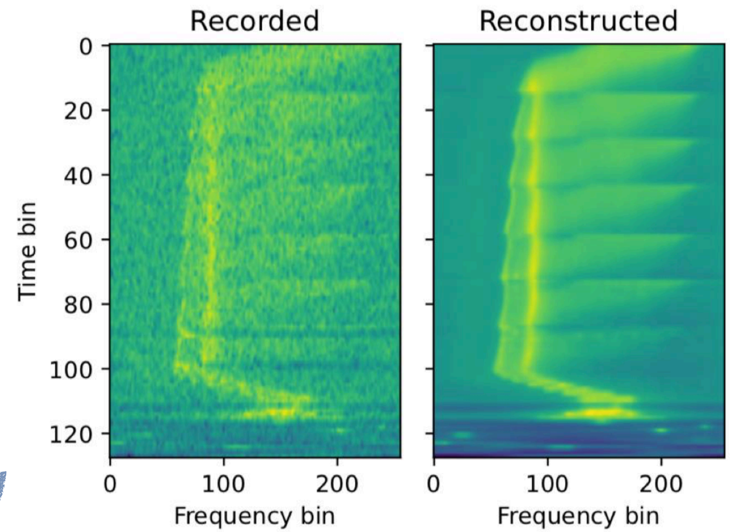
RL agent to correct **RF phase and voltage** to produce uniform RF splitting in PS for LHC beams

★**Trained on simulation and successfully transferred to control room → fully operational**

★RL algorithm: Soft Actor-Critic (SAC); multi-agent algorithm using CNN to define initial set point

★Next step: from on-demand to continuous: → UCAP



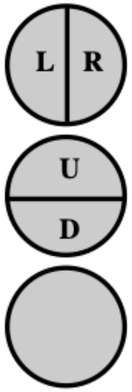


PhD ongoing for: control r**amping and debunching cavity in LINAC3** for optimal injection efficiency into LEIR, based on **Schottky spectrum. Trained on data-driven dynamics**

RL in the control room

Work in progress: RL to steer DC beams in the CERN TT20 transfer line using split-foil secondary emission monitors (BSPs).

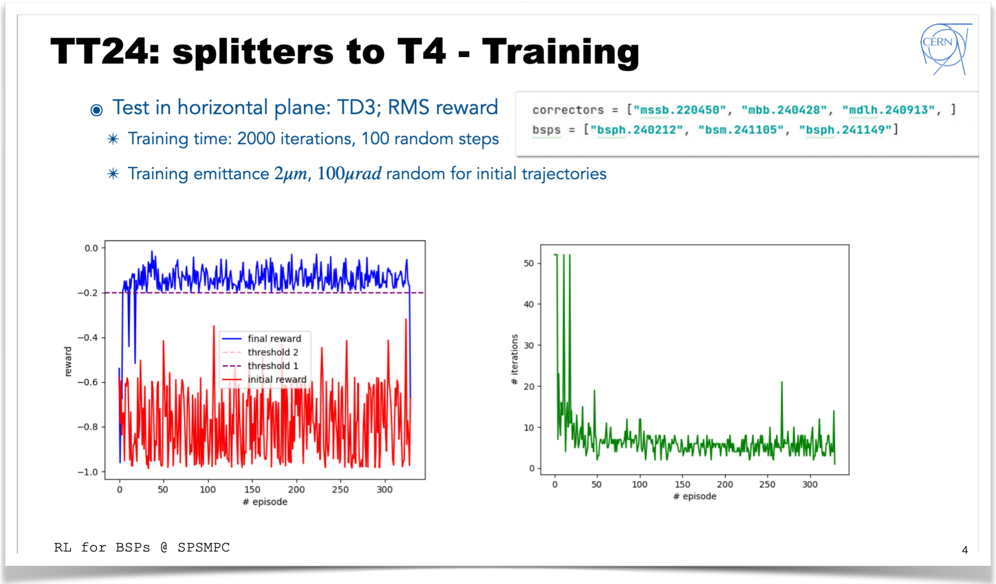
- RL state $\vec{s}$: $[(I_1 - I_2)_i]$ for each monitor; all intensities are normalised.
- Our metric: symmetries per monitor: $S = \sqrt{1 - \frac{|I_1 - I_2|}{I_1 + I_2}}$; Goal: S > 0.8



BSPH-BSMH
Horizontal beam position

BSPV-BSMV
Vertical beam position

BSI
Beam intensity



Also tested for TT23: 10 DOF
Also tested: different distributions

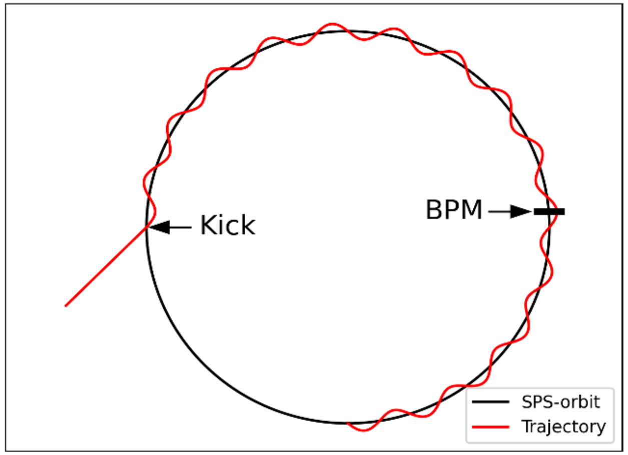
→ Trained on simulation!

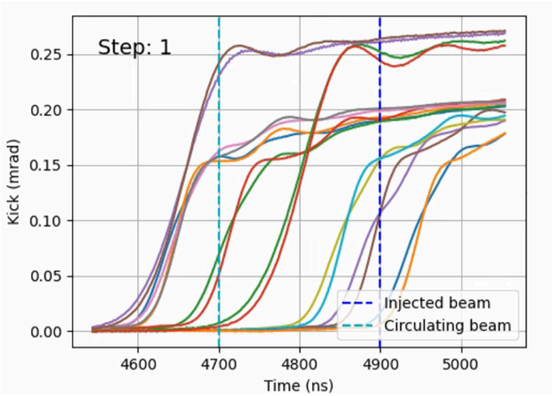
Test of transfer foreseen for startup 2024.

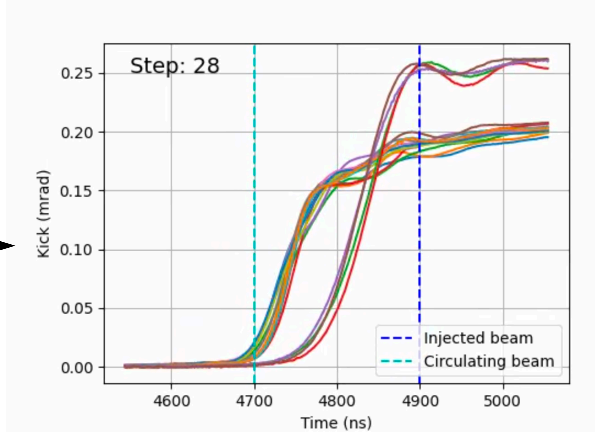
RL in the control room

Ready for transfer test: Adjusting the fine delays of SPS injection kicker with RL

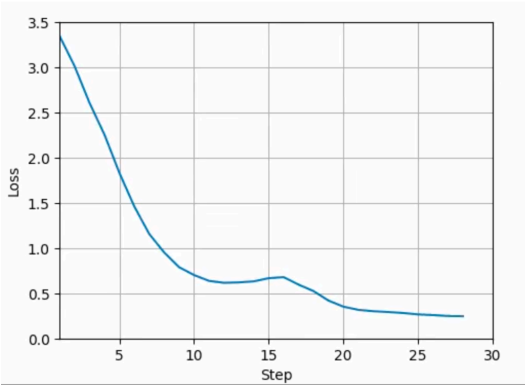
Trained on data-driven dynamics model: PPO





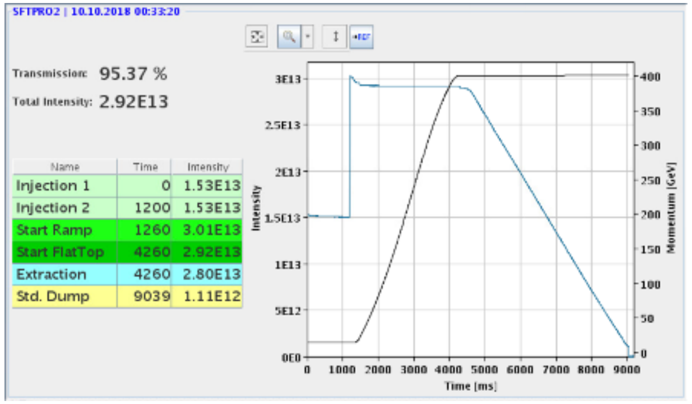


$$r = -\text{ loss } = -\left(x_{inj}^2 + x_{circ}^2 + (x_{inj} - x_{circ})^2\right)$$



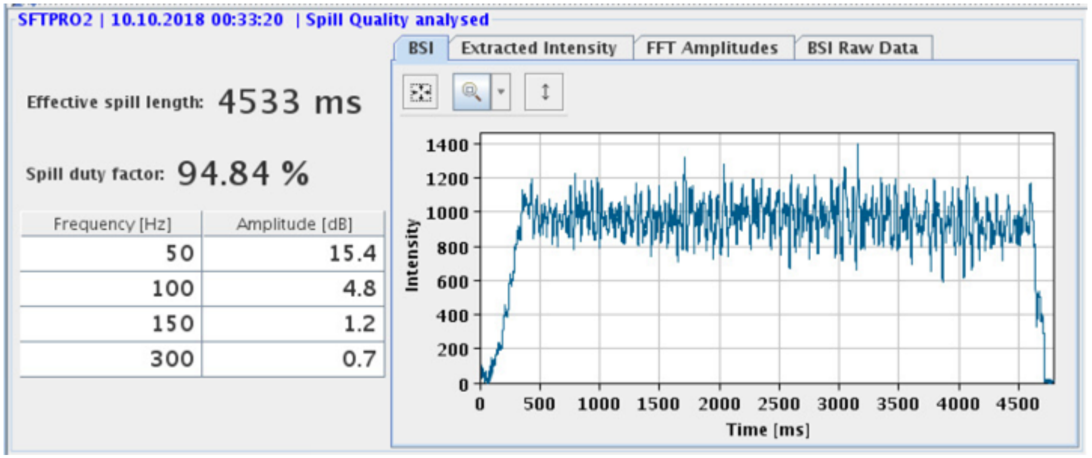
RL in the control room (or not)

Controlling the $n \times 50$ Hz noise in the slow extracted spill to the North Area Experimental Hall.



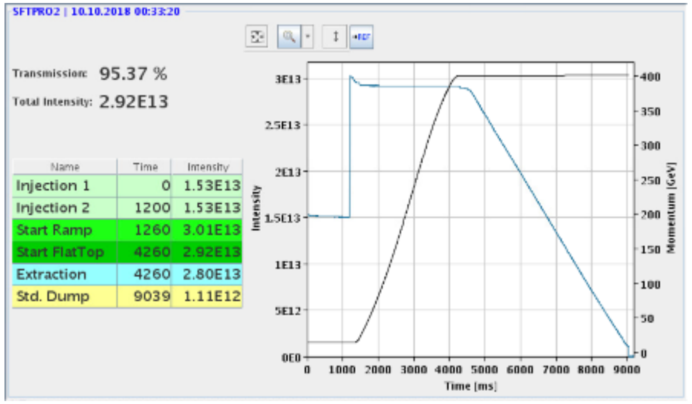
Pile-up constraints by experiments:

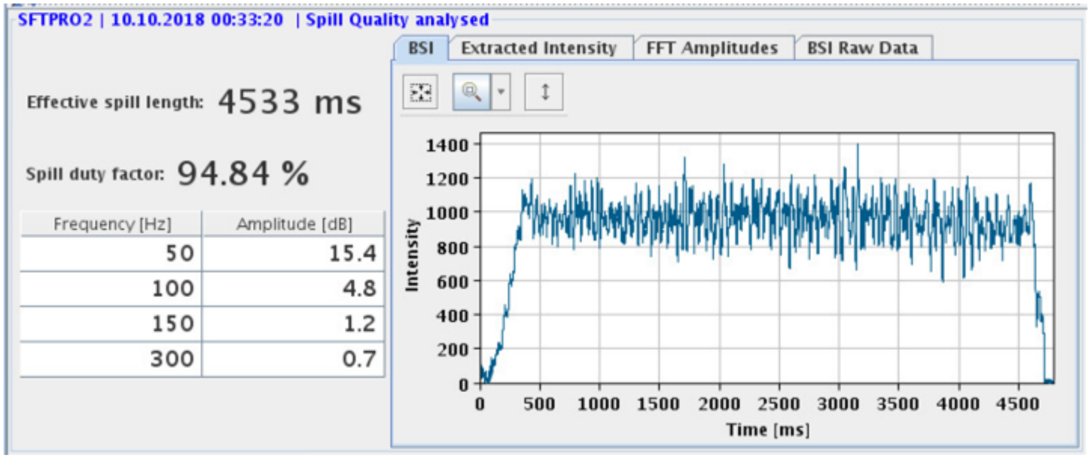
$\rightarrow n \times 50$ Hz norm. amplitudes < 0.15
for $> 85\,\%$



RL in the control room (or not)

Controlling the $n \times 50$ Hz noise in the slow extracted spill to the North Area Experimental Hall.





September 2019

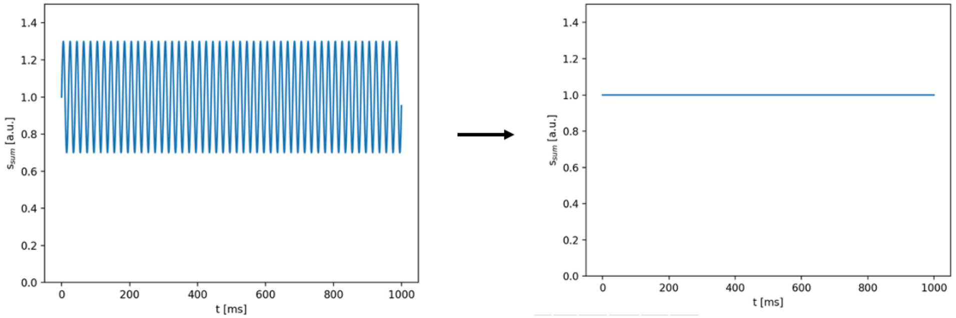
RL for $n \times 50$ Hz noise control?

Simulated environment:

$$\vec{s} = [A_{spill}, \phi_{spill}, A_{corr}, \phi_{corr}]$$

$$r = -\sqrt{A_{noise}^2 + A_{corr}^2 + 2A_{noise}A_{corr}\cos\Delta\phi}$$

Spill monitor signal (2kHz) for 1 s SHiP cycle



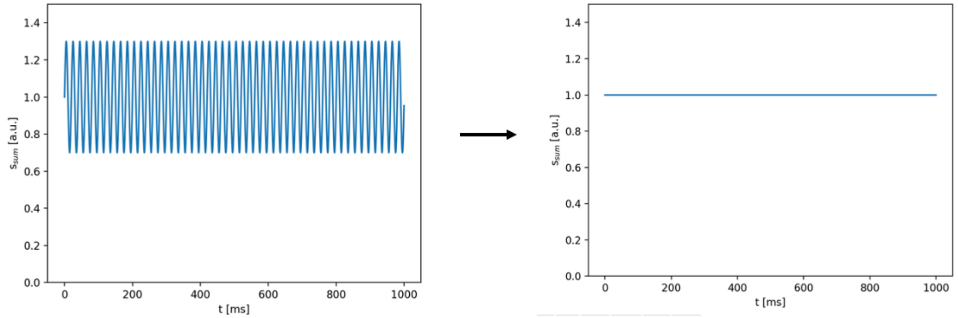
RL for $n \times 50$ Hz noise control?

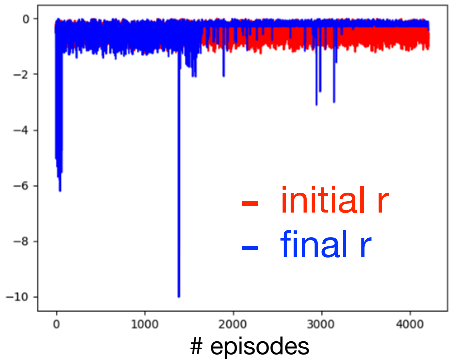
Simulated environment:

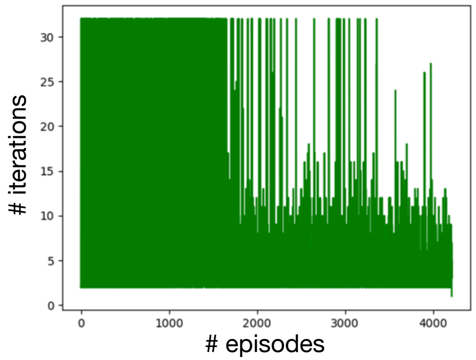
$$\vec{s} = [A_{spill}, \phi_{spill}, A_{corr}, \phi_{corr}]$$

$$r = -\sqrt{A_{noise}^2 + A_{corr}^2 + 2A_{noise}A_{corr}\cos\Delta\phi}$$

Spill monitor signal (2kHz) for 1 s SHiP cycle







→ TD3/SAC actually learn to optimise it.
Could work as controller…
But….

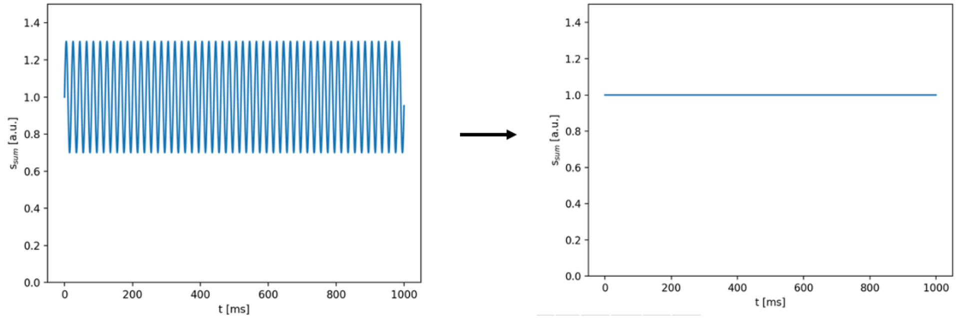
RL for $n \times 50$ Hz noise control?

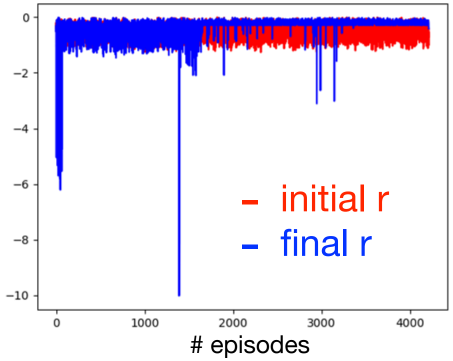
Simulated environment:

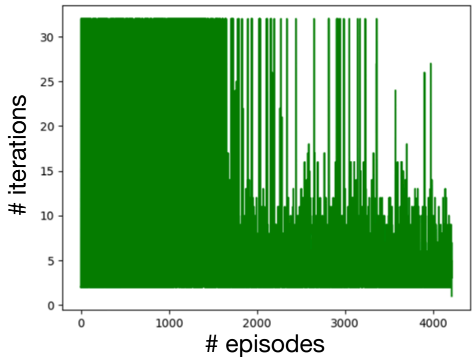
$$\vec{s} = [A_{spill}, \phi_{spill}, A_{corr}, \phi_{corr}]$$

$$r = -\sqrt{A_{noise}^2 + A_{corr}^2 + 2A_{noise}A_{corr}\cos\Delta\phi}$$

Spill monitor signal (2kHz) for 1 s SHiP cycle







→ **TD3/SAC actually learn to optimise it. Could work as controller… But….**

→ Can we transfer? How does V_{QF} translate to A_{corr}?

→ Training on the machine takes too long. Also, how to change A_{noise}, ϕ_{noise}?

So what we did instead...

…as auto-launch numerical optimisation and analytic solution did not work well either.

Adaptive Bayesian Optimisation

Idea: build Gaussian Process for timeseries prediction with *SpectralMixtureKernel* $S(t, t')$

Gaussian Process Kernels for Pattern Discovery and Extrapolation

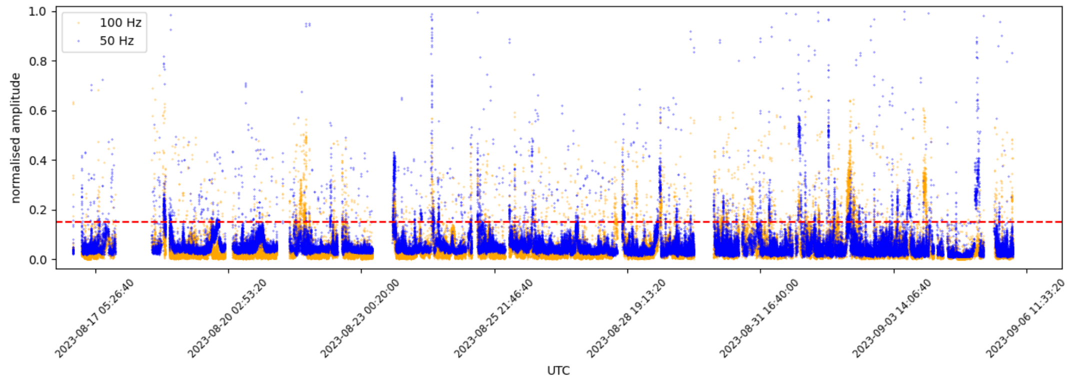
Andrew Gordon Wilson AGW38@CAM.AC.UK
Department of Engineering, University of Cambridge, Cambridge, UK
Ryan Prescott Adams RPA@SEAS.HARVARD.EDU
School of Engineering and Applied Sciences, Harvard University, Cambridge, USA

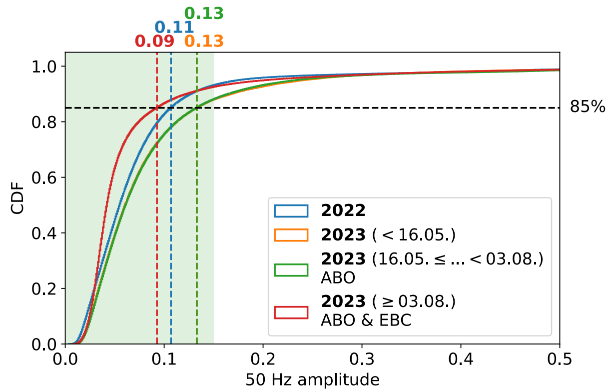
→ add one dimension in problem space: t to predict $t+1$ into future

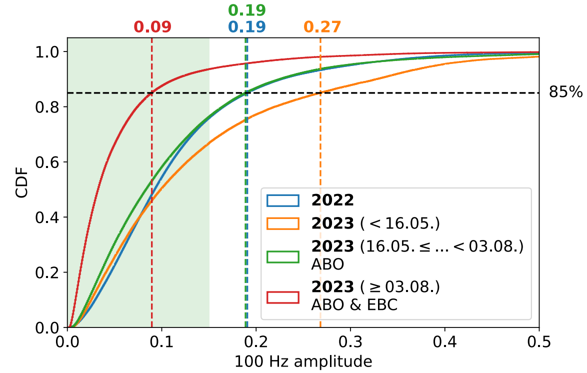
→ GP with composite kernel: the kernel that is currently used:

$$\sigma^2 \times S(t, t') \times RBF(x, x')$$

Example: a couple of weeks during August 2023
→ ABO tracks well. Some issues: controller lock-up due to shared GPU; "exploration" spikes → 2024 proximal biasing







RL in the control room: observations

- Few RL based controllers compared to many optimisation problems solved with black-box optimisation algorithms*

RL in the control room: observations

- Few RL based controllers compared to many optimisation problems solved with black-box optimisation algorithms*

- None of the RL algos in the control room were trained on the machine!
  - → **simulation**
  - Data-driven dynamics → dynamics "easier" to learn than policy from cold. Offline RL?
    - E.g. no exploration issues,…

RL in the control room: observations

- Few RL based controllers compared to many optimisation problems solved with black-box optimisation algorithms*

- None of the RL algos in the control room were trained on the machine!
  - → **simulation**
  - Data-driven dynamics → dynamics "easier" to learn than policy from cold. Offline RL?
    - E.g. no exploration issues,…

- Reasons?
  - States, intuition for dynamics (often) missing.
    - How far do we get with POMDPs?
  - RL not sample-efficient enough
    - Solutions are → GP-MPC, MBRL including physics,…
    - Offline RL?

RL in the control room: observations

- Few RL based controllers compared to many optimisation problems solved with black-box optimisation algorithms*

- None of the RL algos in the control room were trained on the machine!
  - → **simulation**
  - Data-driven dynamics → dynamics "easier" to learn than policy from cold. Offline RL?
    - E.g. no exploration issues,…

- Reasons?
  - States, intuition for dynamics (often) missing.
    - How far do we get with POMDPs?
  - RL not sample-efficient enough
    - Solutions are → GP-MPC, MBRL including physics,…
    - Offline RL?

PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

CrossMark
click for updates

Article submitted to journal

Subject Areas:

Physics-informed Dyna-Style Model-Based Deep Reinforcement Learning for Dynamic Control

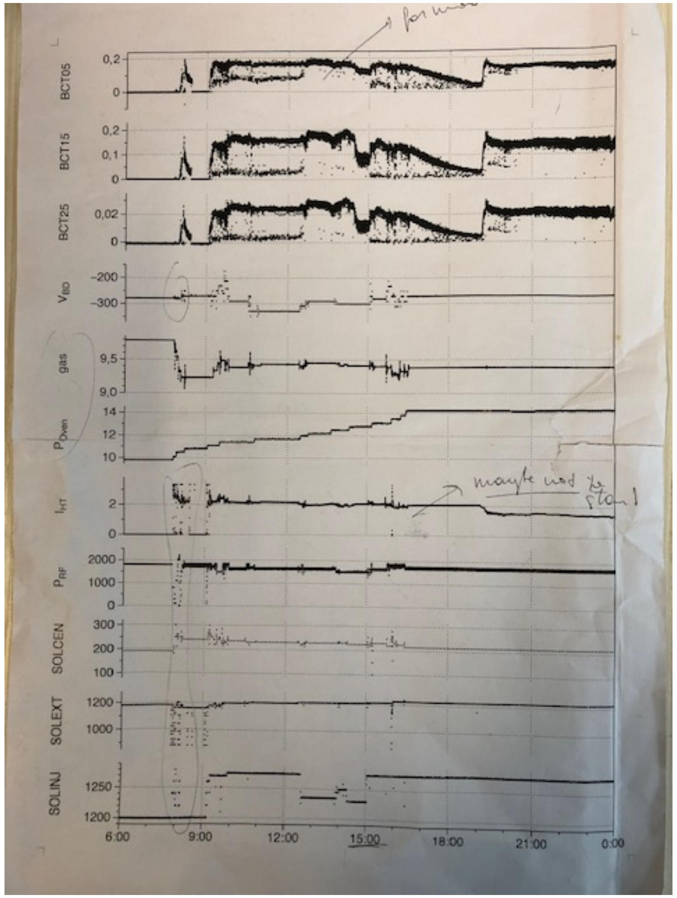
Xin-Yang Liu[1] and Jian-Xun Wang[1]

[1]Department of Aerospace & Mechanical Engineering, College of Engineering, University of Notre Dame, Notre Dame, IN, USA

→ Proposal for LINAC3 energy distribution control

Offline RL?

Collab with Simon: Tuning/stabilising the LINAC3 Pb^{54+} source

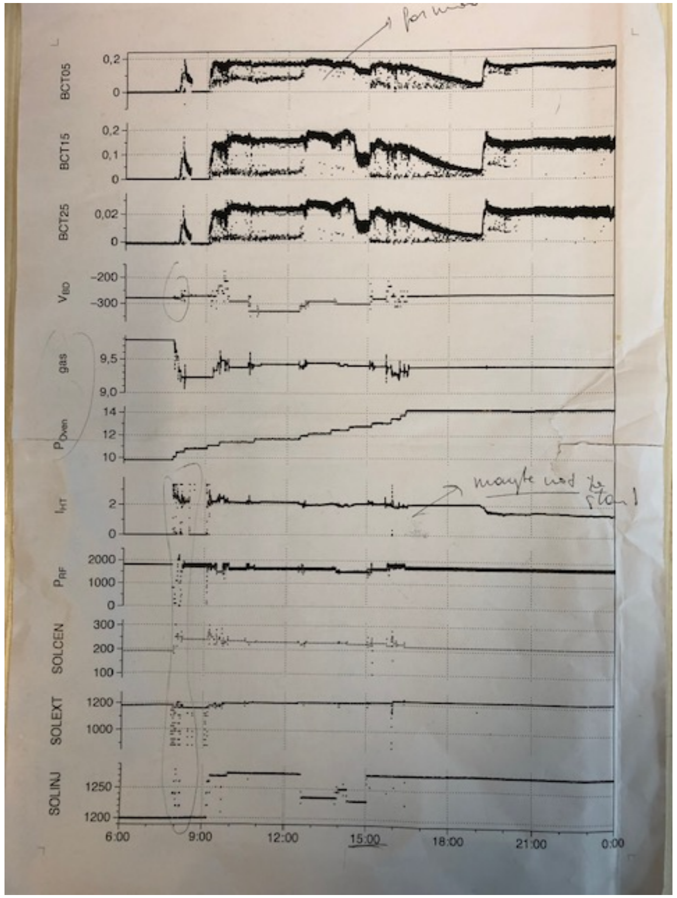


Offline RL?

Collab with Simon: Tuning/stabilising the LINAC3 Pb^{54+} source

LINAC3 source specialist wish list:

- behaviour cloning from historic data
- ideally zero-shot transfer, safe exploration otherwise
- no continuous control: only correct when necessary
- Additional remark: **response with delay!**
  - Different delays for different parameters

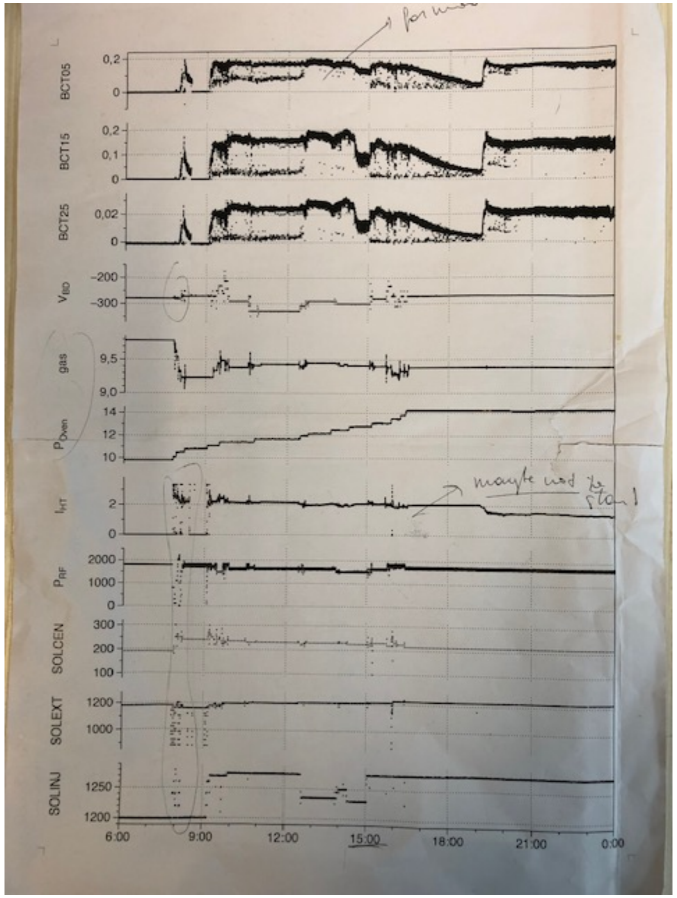


Offline RL?

Collab with Simon: Tuning/stabilising the LINAC3 Pb^{54+} source

LINAC3 source specialist wish list:

- behaviour cloning from historic data
- ideally zero-shot transfer, safe exploration otherwise
- no continuous control: only correct when necessary
- Additional remark: **response with delay!**
  - Different delays for different parameters

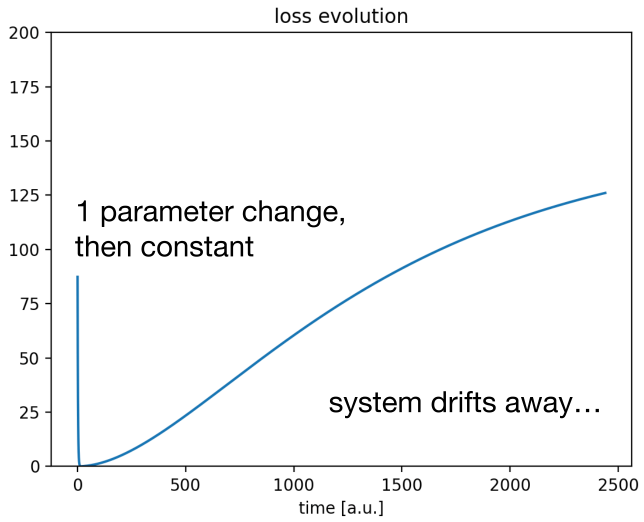


Simon & Verena: How to solve this?

- no obvious state information
  - Maybe absolute settings
- no useful simulation
- (not enough data to train “a" model)
- How to deal with different delays of actors?

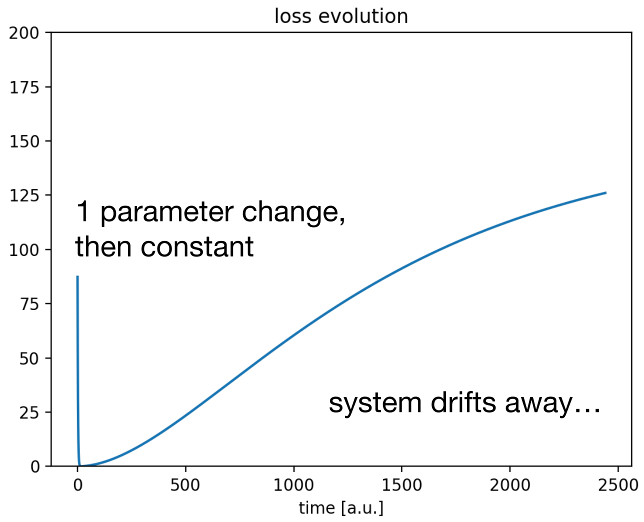
LINAC3 source: Offline RL - not yet

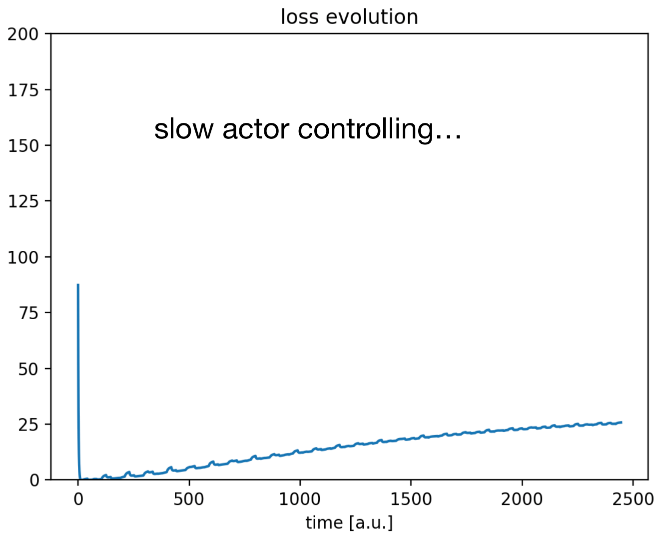
- Building **toy simulation** of dynamic system that changes with time
- Different actors, with delayed response; actors depend on each other
- Currently all response functions convex
- → to test algorithm type!



LINAC3 source: Offline RL - not yet

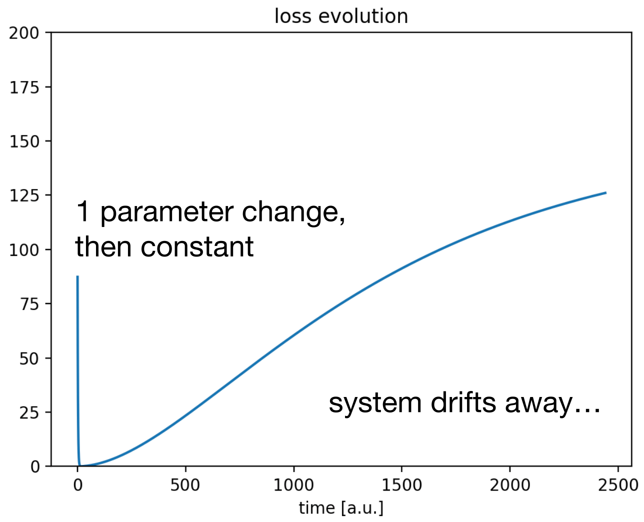
- Building **toy simulation** of dynamic system that changes with time
- Different actors, with delayed response; actors depend on each other
- Currently all response functions convex
- → to test algorithm type!

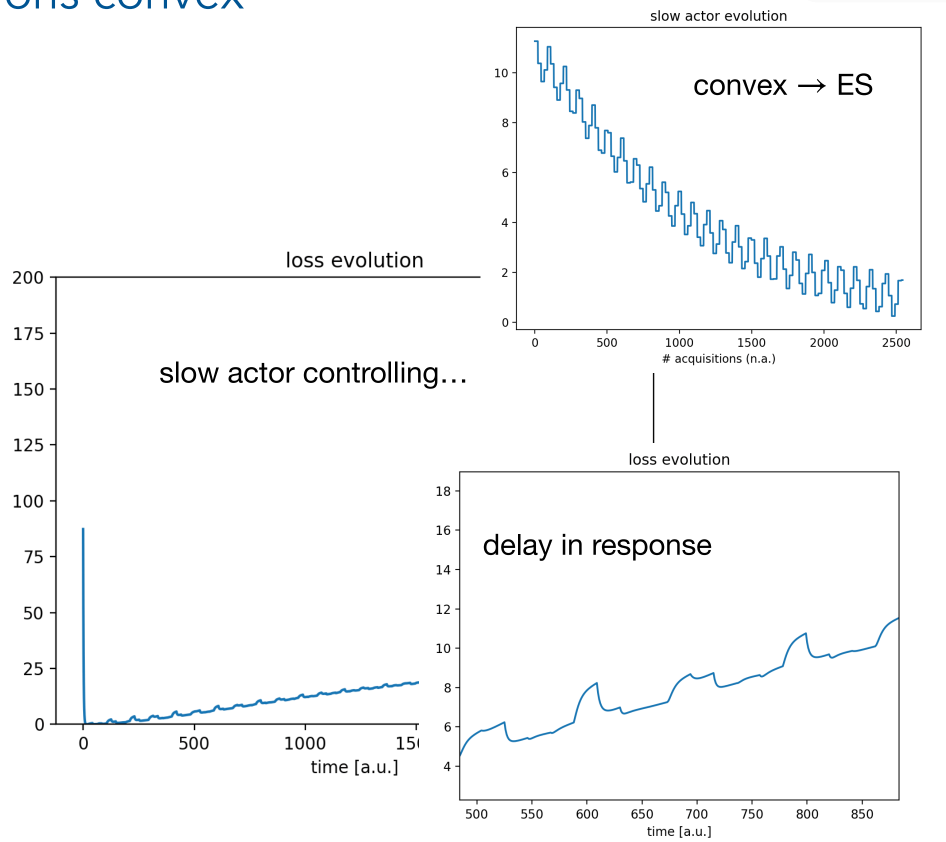




LINAC3 source: Offline RL - not yet

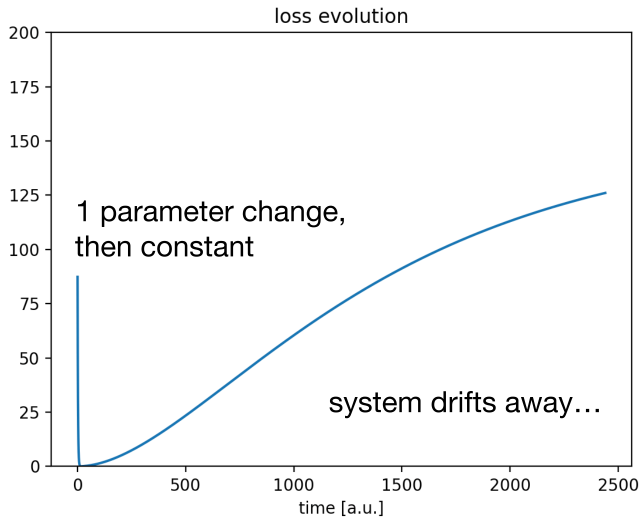
- Building **toy simulation** of dynamic system that changes with time
- Different actors, with delayed response; actors depend on each other
- Currently all response functions convex
- → to test algorithm type!

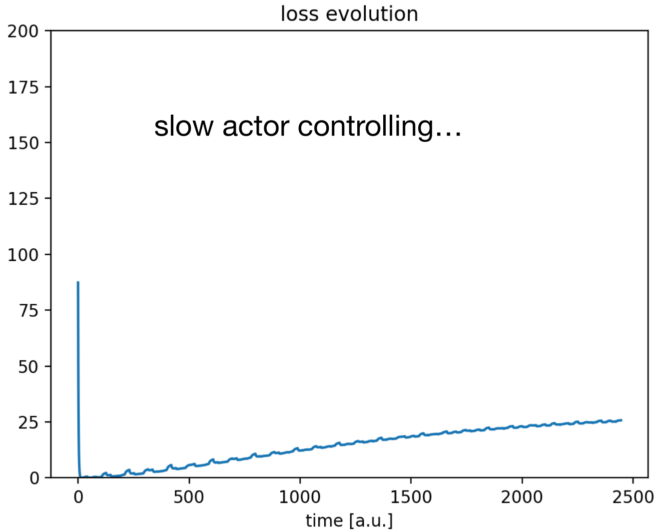


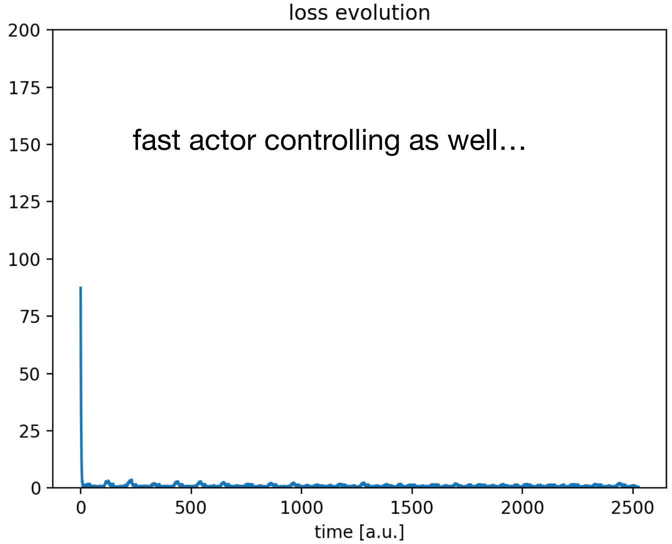


LINAC3 source: Offline RL - not yet

- Building **toy simulation** of dynamic system that changes with time
- Different actors, with delayed response; actors depend on each other
- Currently all response functions convex
- → to test algorithm type!







→ simulations!!

Conclusion

Efficient Particle Accelerators (EPA) project has recently been put in place

to improve on various efficiency limiting aspects through **automation** and also **improved modelling**.

- Including AI/ML techniques at scale

The project was given 5 years to be ready with improvements for the HL-LHC era

Conclusion

Efficient Particle Accelerators (EPA) project has recently been put in place

to improve on various efficiency limiting aspects through **automation** and also **improved modelling**.

- Including AI/ML techniques at scale

The project was given 5 years to be ready with improvements for the HL-LHC era

EPA WP3: **Reinforcement Learning will be part of the new exploitation paradigm.**

Conclusion

Efficient Particle Accelerators (EPA) project has recently been put in place

to improve on various efficiency limiting aspects through **automation** and also **improved modelling**.

- Including AI/ML techniques at scale

The project was given 5 years to be ready with improvements for the HL-LHC era

EPA WP3: **Reinforcement Learning will be part of the new exploitation paradigm.**

Key ingredient for RL to be adopted → simulations of all processes and accelerators.

- Sim2real transfer: Need to get more experience with domain adaptation, adaptive agents
- Offline RL: data-driven dynamics if easily obtainable from historic data also of interest