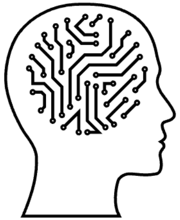
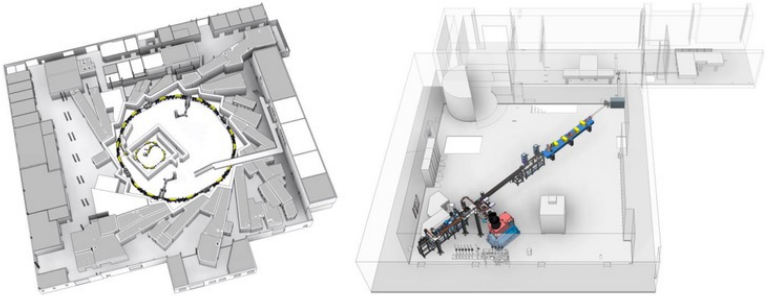
Karlsruhe Institute of Technology

RL4AA'24
5-7 Feb. 2024



Breakthrough in real-time control with reinforcement learning on hardware at KARA

A. Santamaria Garcia, L. Scomparin, E. Blomley, E. Bründermann, M. Caselle,
A. Kopmann, J. L. Steinmann, C. Xu, J. Becker, A.-S. Müller



Two proof-of-principle studies at KARA implementing RL "at the edge"

L. Scomparin

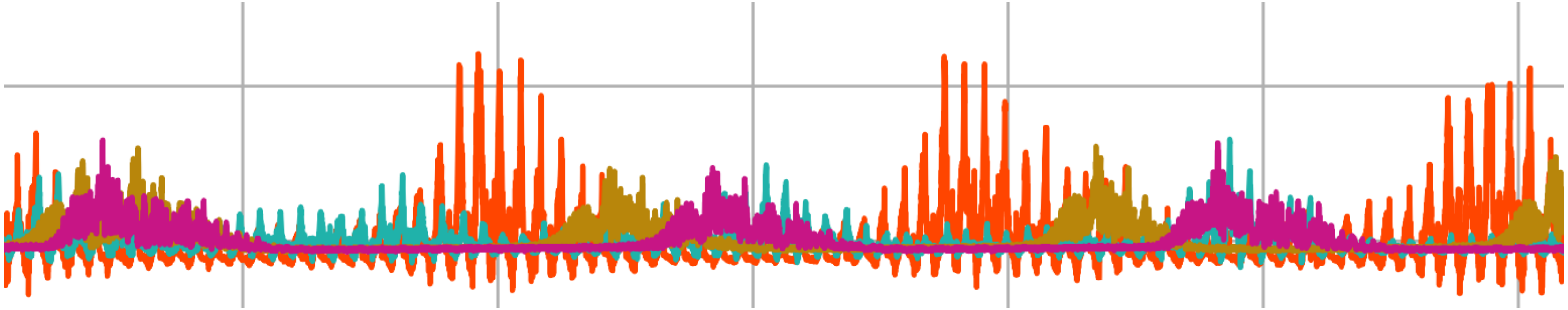
Step 1: trying the concept in a simple problem
Damping of induced betatron oscillations

Step 2: using the validated concept for more difficult endeavors…
Control of the microbunching instability

First RL deployment in accelerators with purely **online training** and **running on hardware** at very low latencies!

Motivation:

Control of coherent synchrotron radiation emission

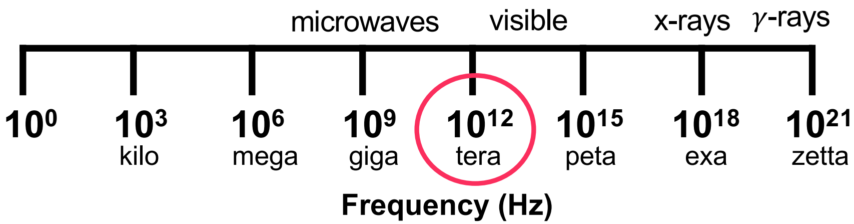


The THz frequency

Great scientific potential!

0.3 THz ↔ 3 THz
3.3 ps ↔ 0.3 ps
1 mm ↔ 100 μm

Between microwave and far infrared ("THz gap")



- Material characterization
- Lower-energy solution to X-rays (non-ionizing)
- High-resolution images of the interior of solid objects (security screening, artwork analysis)
- Submillimeter astronomy (chemical abundances and cooling mechanisms of molecular clouds)

Same scale as…

- Frequency of rotation of small molecules
- Duration of collisions between gas molecules at room temp.
- Peak frequency of blackbody-like emission of galaxies
- Oscillation of gaseous and solid-state plasmas
- Frequency of resonance of electrons in semiconductors
- Frequency of superconducting energy gaps
- Frequency of vibration of biologically-relevant collective modes of proteins
- …

["THz techniques" E. Bründermann et al.]

The THz frequency

Great scientific potential!

Common desiderata:

- High peak fields
- Coverage to higher frequencies with coherent broadband sources
- Full pulse-shaping
- Excellent source stability

↳ Accelerator-based sources provide THz radiation with high **brightness**, **power**, and **repetition rate**

More generally:

Non-destructive

Single-shot, bunch-by-bunch diagnostics

Constant delivery of high quality, intense and stable photon beams to a variety of beamlines

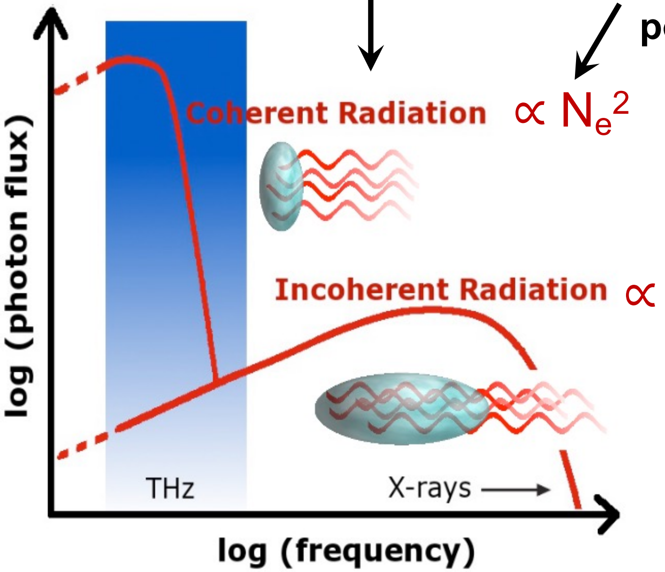
Feedback

"Accelerator-Based THz Radiation Sources", A.-S. Müller & M. Schwarz

Coherent synchrotron radiation (CSR)

Short bunch

Increased radiation power!



*N = number of electrons

Courtesy of Prof. A.-S. Müller

$$\frac{dI}{d\omega} = [N_e + N_e(N_e - 1)F(\omega)]\frac{dI_0}{d\omega}$$

- $\frac{dI}{d\omega}$: Synchrotron radiation spectral intensity
- N_e: Incoherent radiation
- $N_e(N_e - 1)$: Coherent radiation
- $F(\omega)$: Form factor
- $\frac{dI_0}{d\omega}$: Single particle spectrum

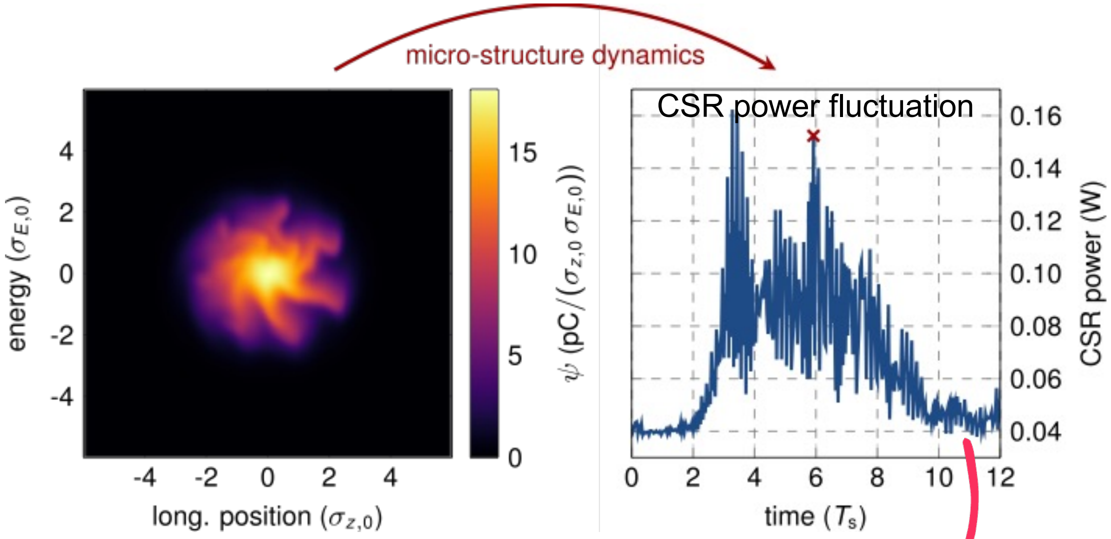
$$F(\omega, \vec{n}) = \left| \int \rho(\vec{r}) e^{i\omega\, \vec{n}.\vec{r}/c}\, \mathrm{d}^3\vec{r} \right|^2$$

Highly dependent on the **shape** of the generating **charge distributions**

["Accelerator-Based THz Radiation Sources", A.-S. Müller & M. Schwarz]

The microbunching (MBI) instability

Control for stable, enhanced, or damped CSR



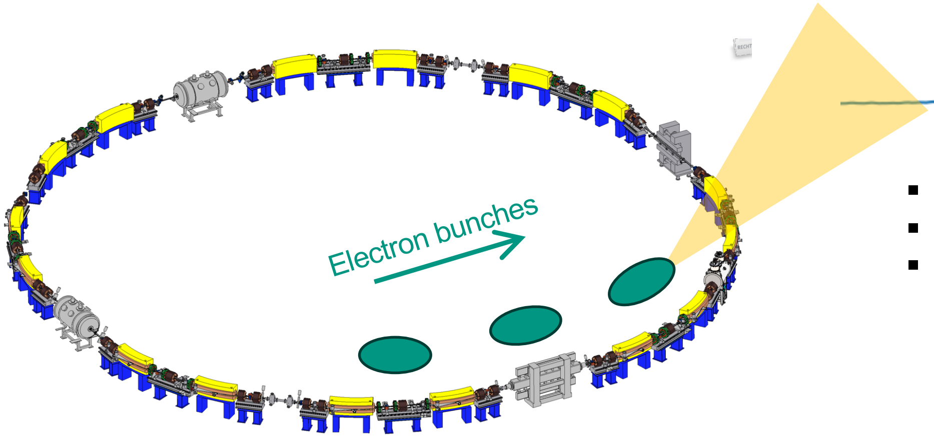
We would like a high average, low variance emission

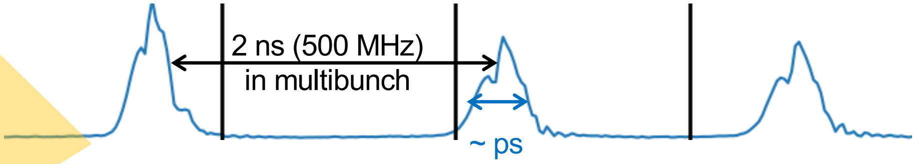
RL considerations

- How to influence the instability? (actions)
- How fast does the action need to be to influence a physical phenomenon?
- How fast can we detect THz radiation? (observable)
- Can we achieve the required latency?

[“Micro-Bunching Control at Electron Storage Rings with Reinforcement Learning”, T.Boltz]

Timescale perspective and requirements



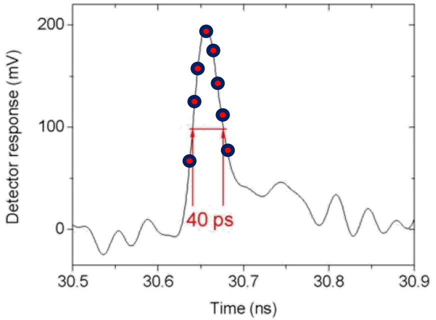


- Revolution frequency: 2.7 MHz (T_{rev} = 370 ns)
- Synchrotron frequency: 7-9 kHz (T_{sync} = 110-143 μs)
- → 300-400 T_{rev} ~ T_{sync}

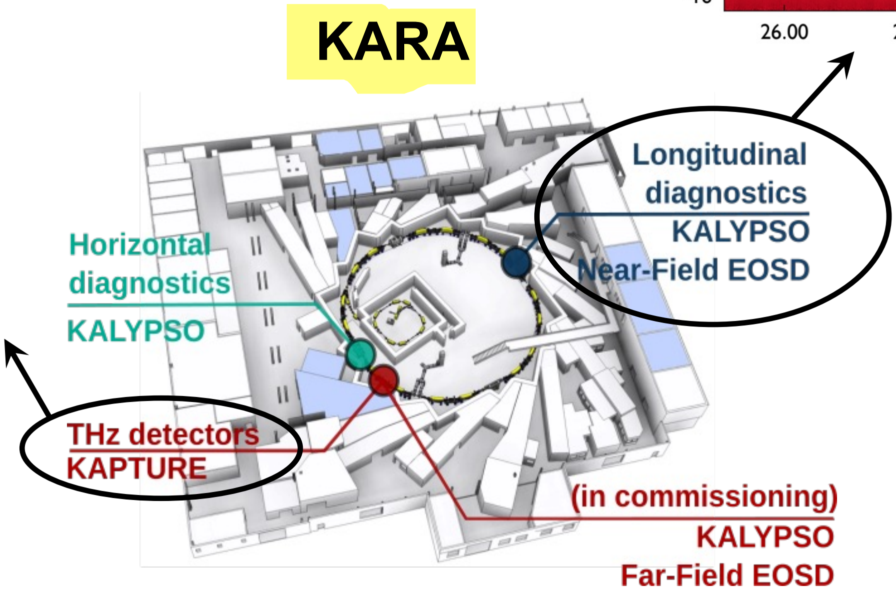
- We want to be able to detect light every **2 ns**
- We want to be able to detect light pulses in the **ps order**
- We want to be able to act every **~ T_{sync} (110-143 μs)**

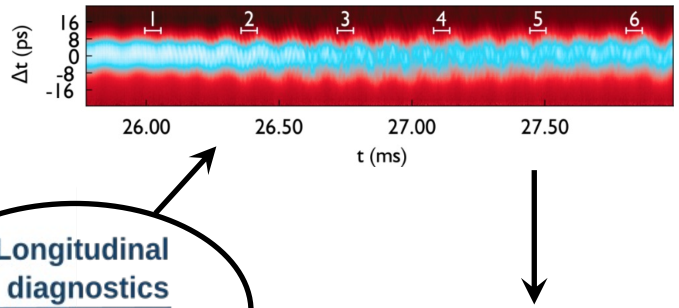
Real-time, high-repetition data acquisition

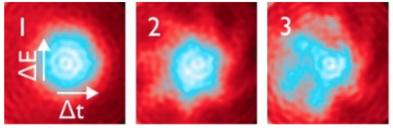
State-of-the-art detectors



Turn by turn sampling (2.7 MHz)
1024 samples







Phase space density reconstruction

[“High throughput data streaming of individual longitudinal electron bunch profiles”, S. Funkner]

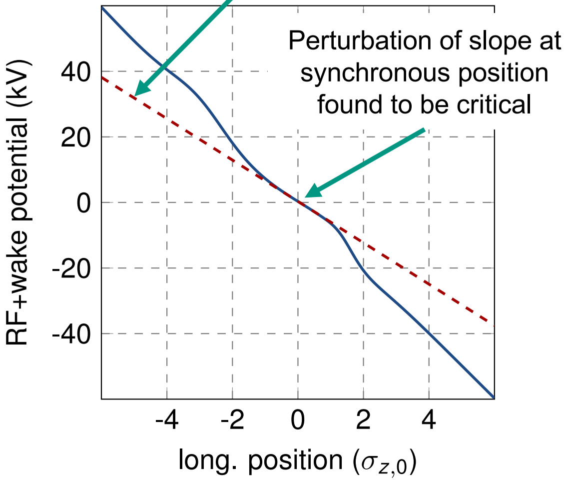
[Revealing the dynamics of ultrarelativistic non-equilibrium many-electron systems with phase space tomography]

[“KAPTURE-2. A picosecond sampling system for individual THz pulses with high repetition rate”, M. Caselle]

Influencing the instability

CSR self interaction

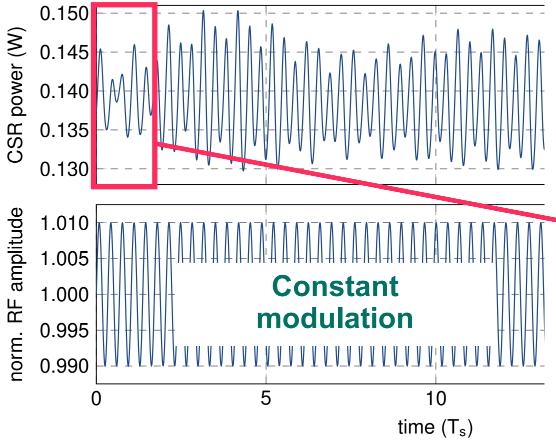
$$V_{eff}(q) = V_{RF}(q) + V_{CSR}(q)$$



Compensate the effect of the CSR perturbation by **modulating the RF voltage (amplitude)**

$$V_{RF} = \hat{V}(t)\sin(2\pi f_{RF} t)$$

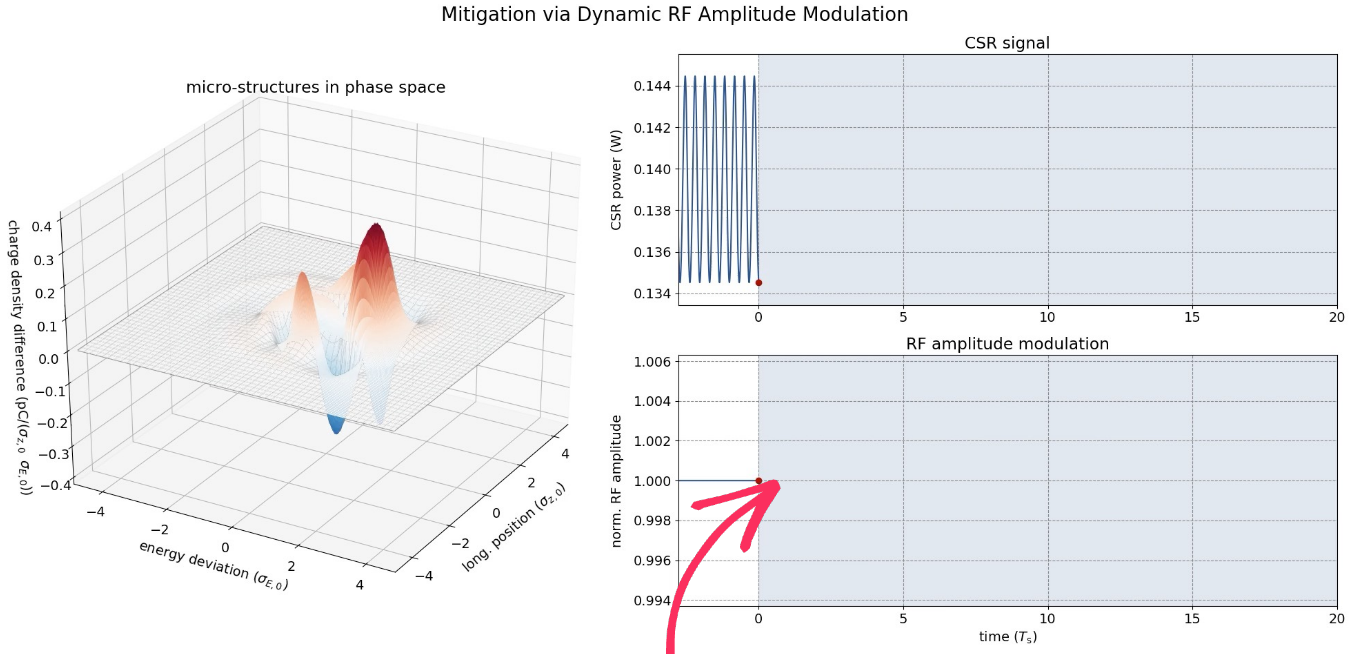
$$\hat{V}(t) = \hat{V}_0 + A_{mod}\sin(2\pi f_{mod} + \varphi_{mod})$$



Initial damping, but quickly out of sync…we need **dynamic control**!

Courtesy of T. Boltz

Influencing the instability (simulation)



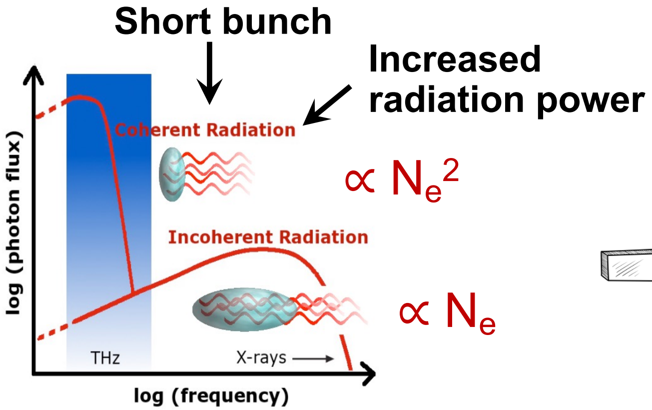
High average, low variance CSR!

$$V_{RF} = \hat{V}(t)\sin(2\pi f_{RF} t) \qquad \hat{V}(t) = \hat{V}_0 + A_{mod}\sin(2\pi f_{mod} + \varphi_{mod})$$

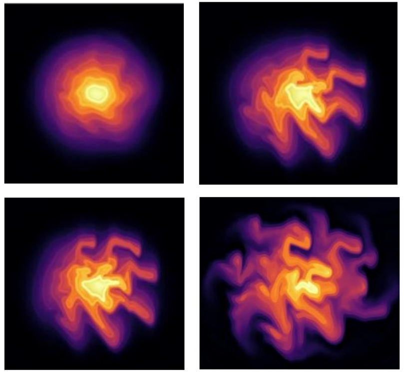
Courtesy of T. Boltz

Simulation done with Inovesa, Vlasov-Fokker-Plack solver developed at KIT

Control of the microbunching instablity with RL at KARA



Low-α_c optics → MBI

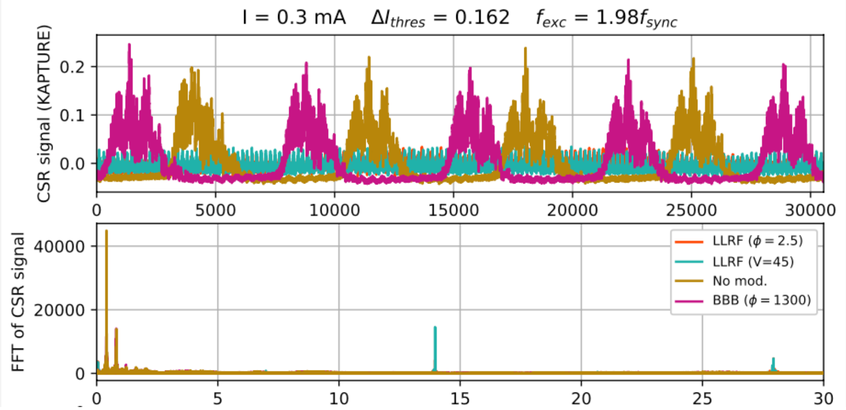


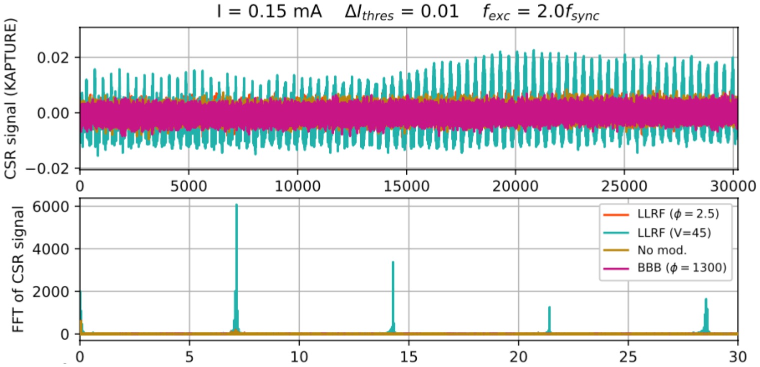
Bursting can be controlled with RF modulations

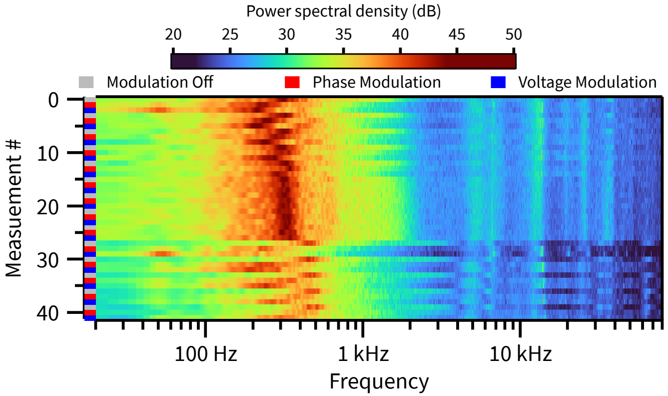
- Bunch-by-bunch feedback system
- Accelerating RF cavities

Detailed actuator study was carried out
[A. Santamaria Garcia et al, IPAC23-WEPA018]

How to influence the instability at KARA?



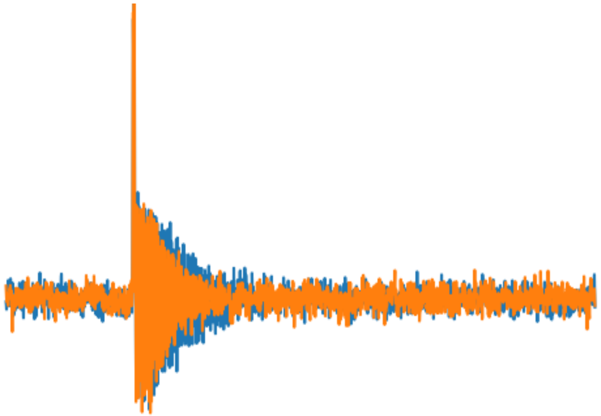




- Only the LLRF system could influence the MBI
- The LLRF needed to be modified by the manufacturer to accept continuous input signals

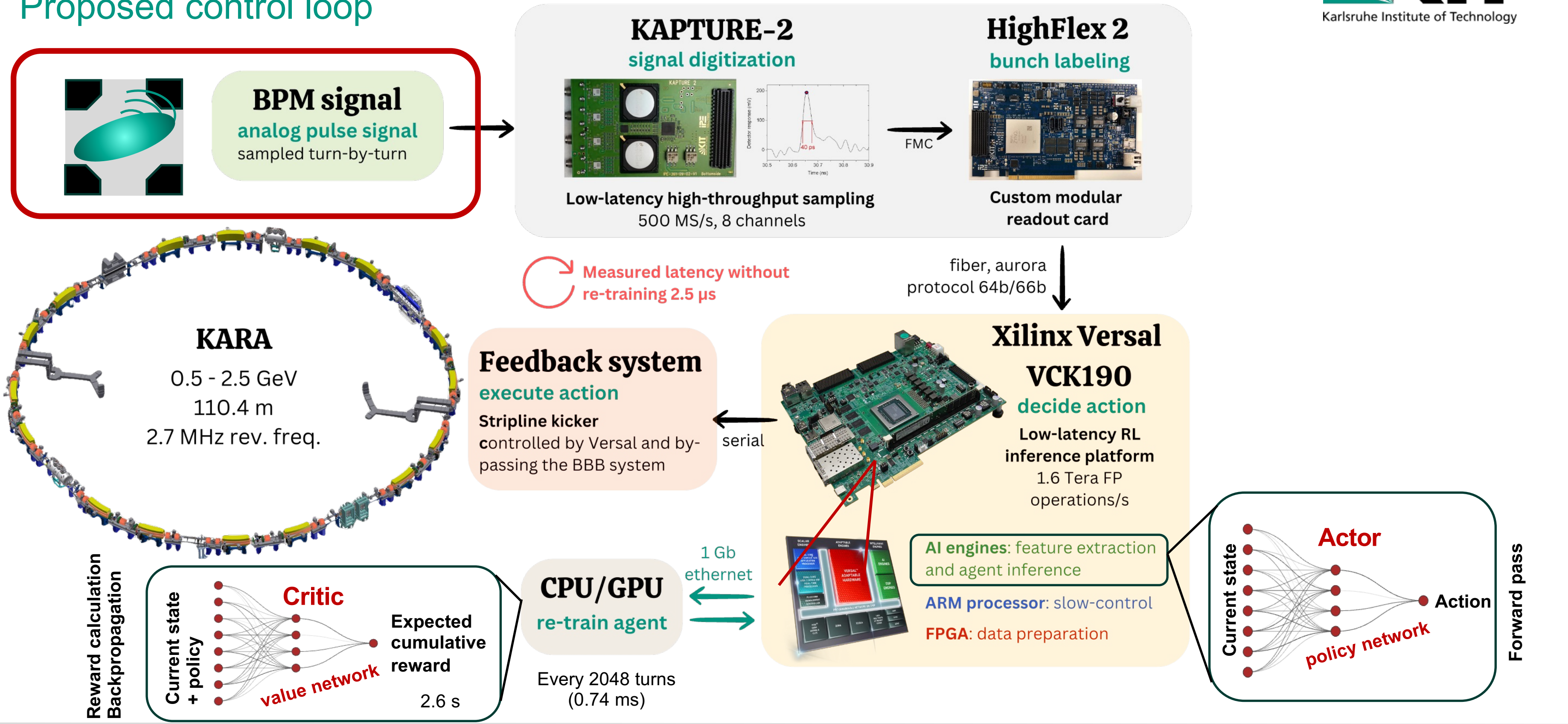
A simple case of online RL:

damping of induced betatron oscillations

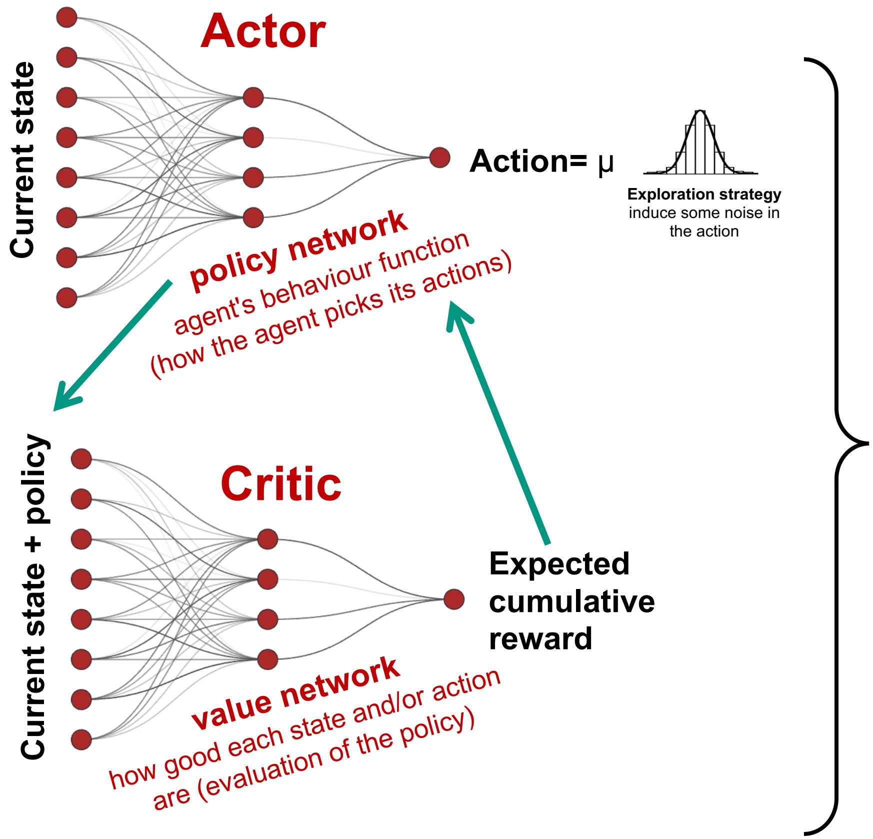


Damping of transverse oscillations

Proposed control loop



Actor-critic RL



Training

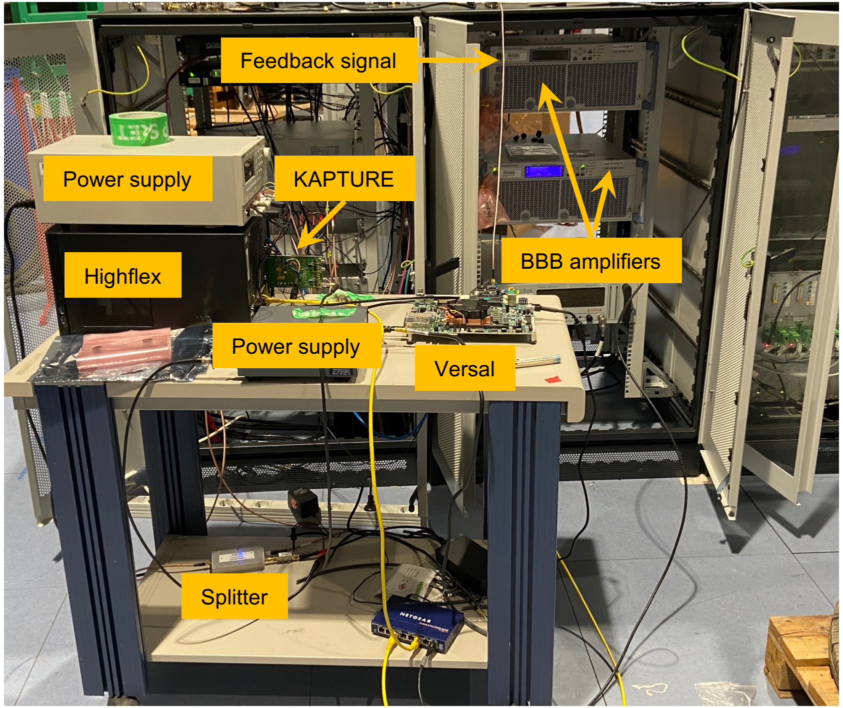
- The weights of both networks are updated iteratively using gradient descent
- The **critic** assesses how good the current policy is in terms of expected rewards, and aims at minimizing the difference between the expected value and the actual observed rewards
  - ➢ Calculation of the observed reward only needed at backpropagation time
- The **actor** updates updates its policy to maximize the expected rewards

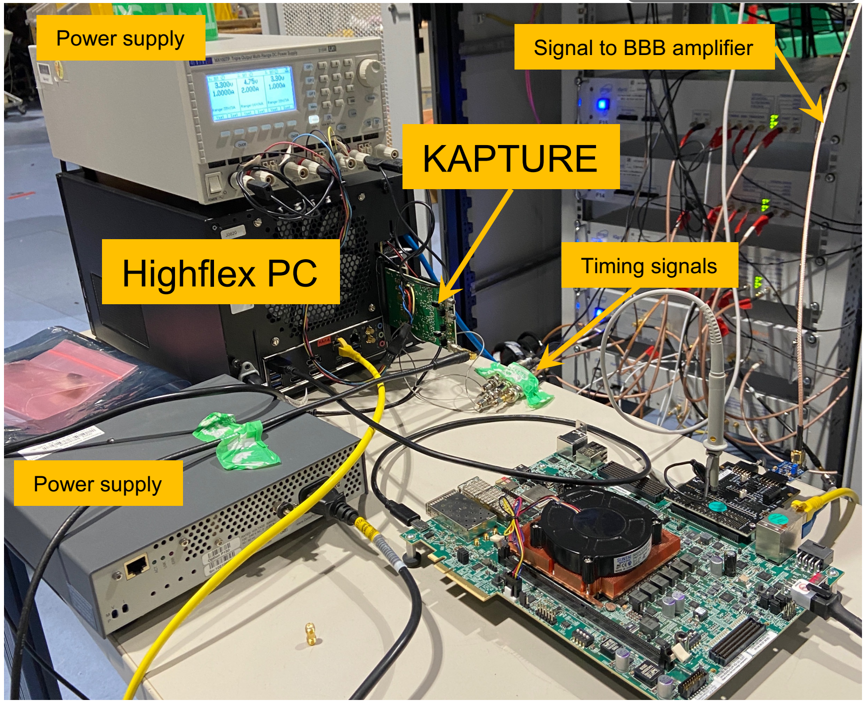
This happens “offline” using the experience accumulated

Damping of transverse oscillations

Hardware setup at KARA

L. Scomparin





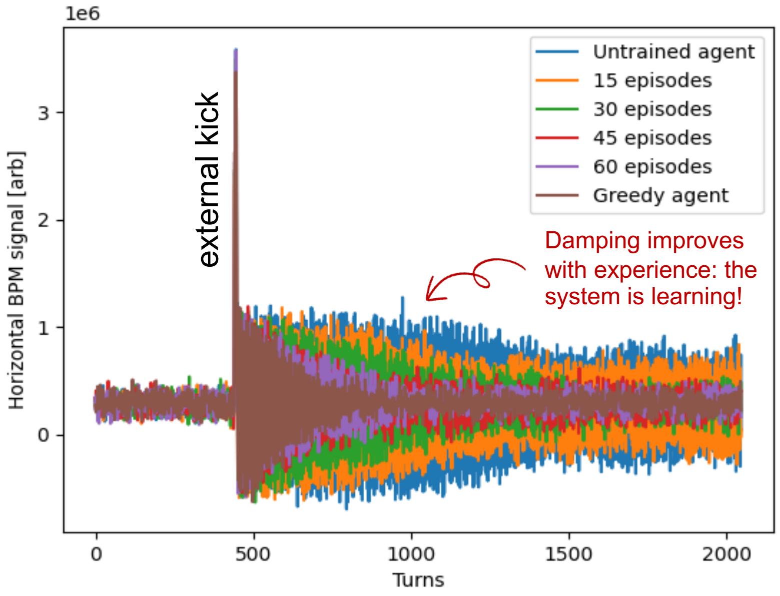
Damping of transverse oscillations

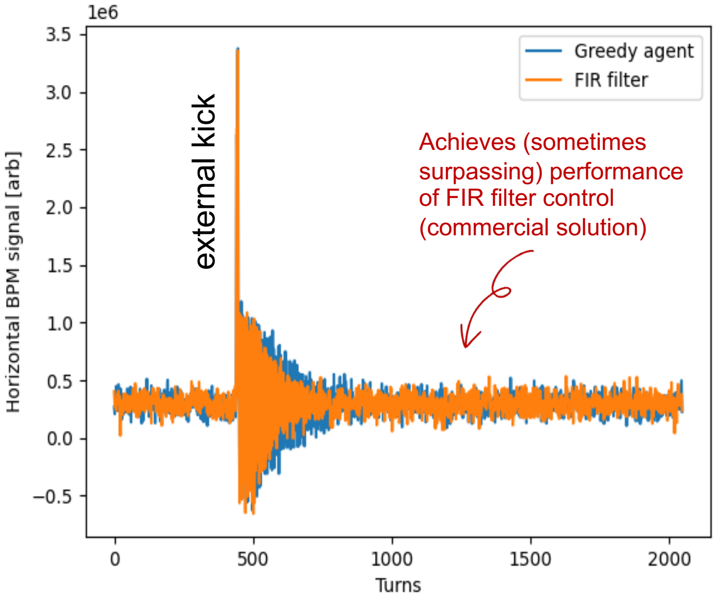
RL implementation

Karlsruhe Institute of Technology
L. Scomparin

- **Algorithm**: Vanilla PPO from Stable Baselines3
- **Actor & critic architecture**: 8-16-1
- **Reward**: metric of the beam position (low as possible)
- **Observation**: last 8 BPM samples

- **Strategy:**
  1. Agent acts during 2048 turns (0.74 ms)
  2. Agent stops and is re-trained in a CPU (~2.6 s)
  3. New weights are sent to Versal board and agent starts again



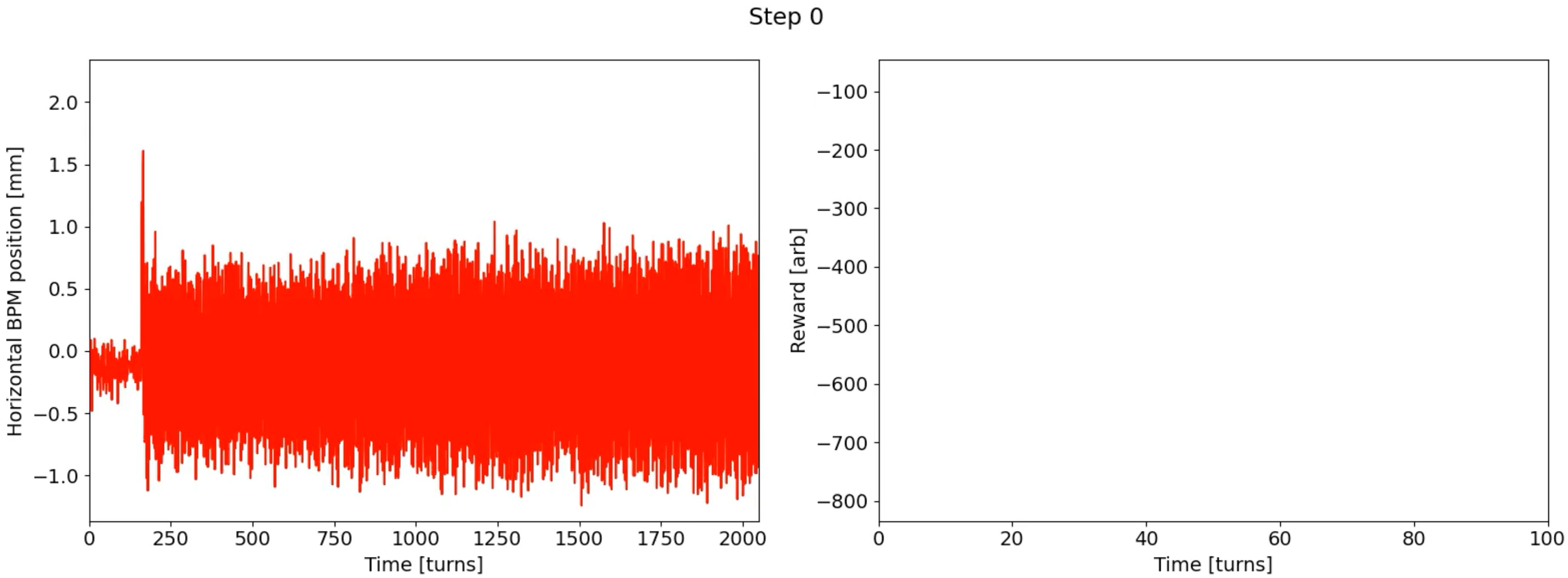


Damping of transverse oscillations

Time evolution per episode

L. Scomparin

Learns by pure interaction with the accelerator

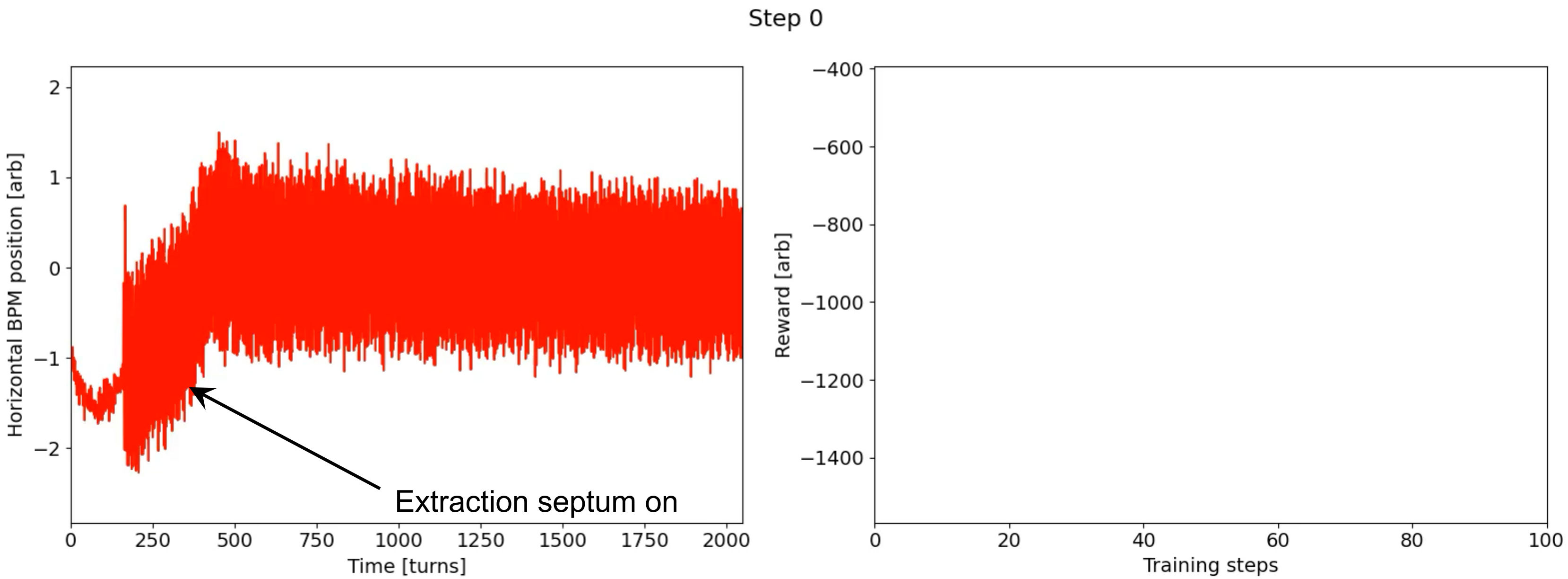


Damping of transverse oscillations

Time evolution per episode

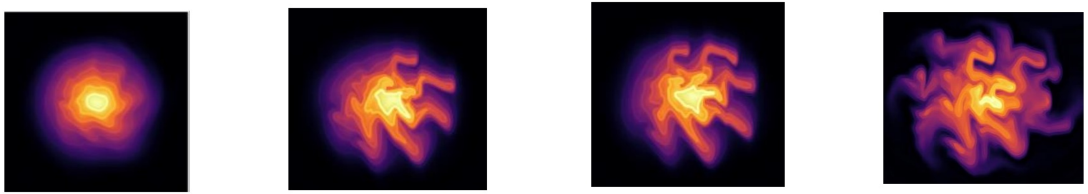
L. Scomparin

Learns by pure interaction with the accelerator



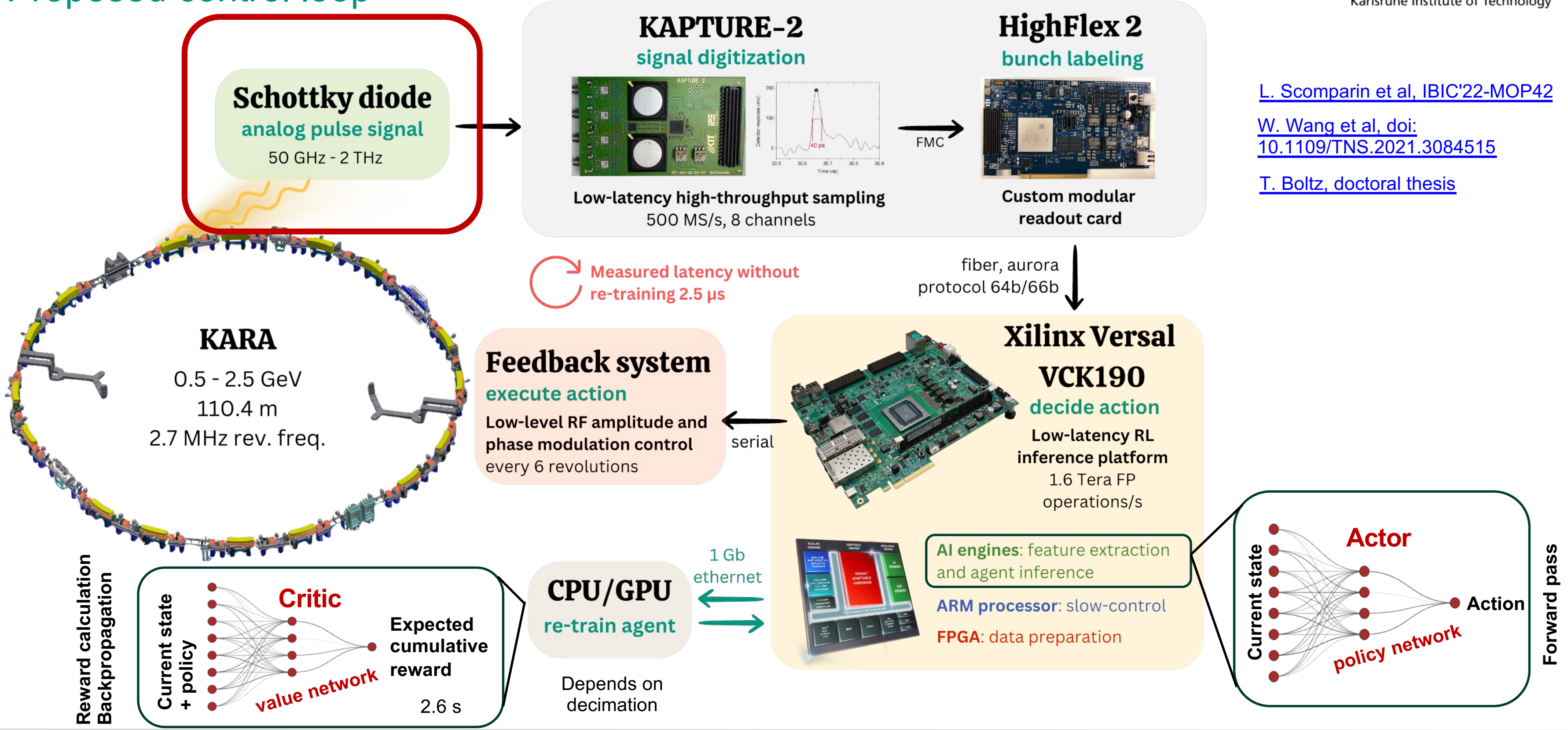
A more complicated case of online RL:

control of the microbunching instability



Control of the microbunching instablity with RL

Proposed control loop



L. Scomparin et al, IBIC'22-MOP42

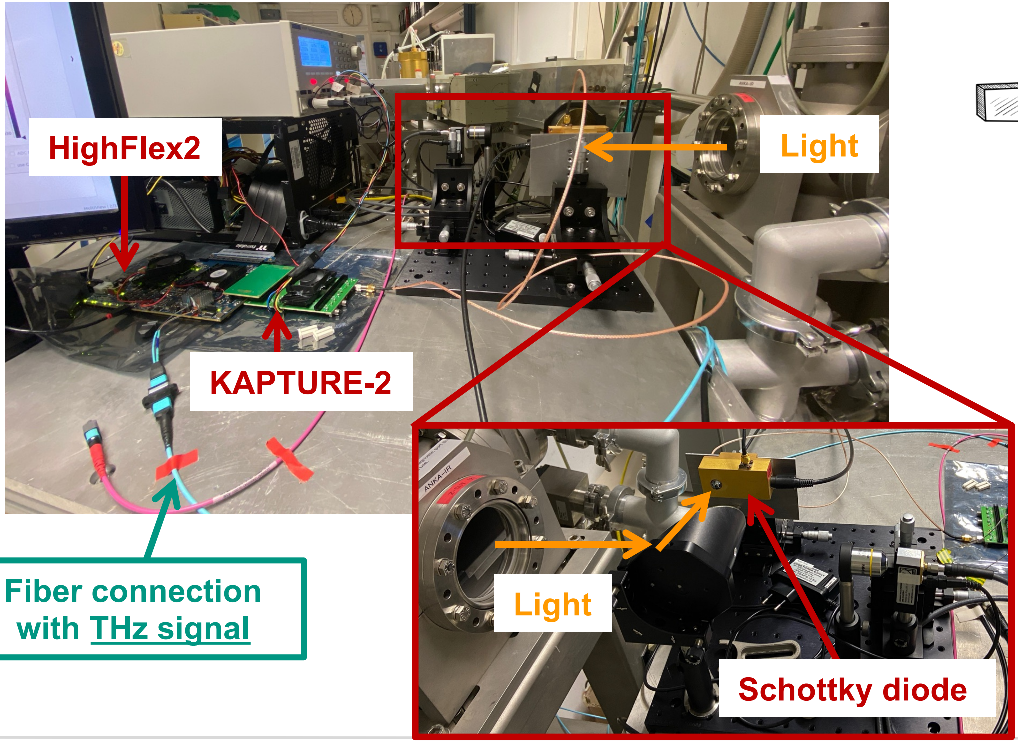
W. Wang et al, doi: 10.1109/TNS.2021.3084515

T. Boltz, doctoral thesis

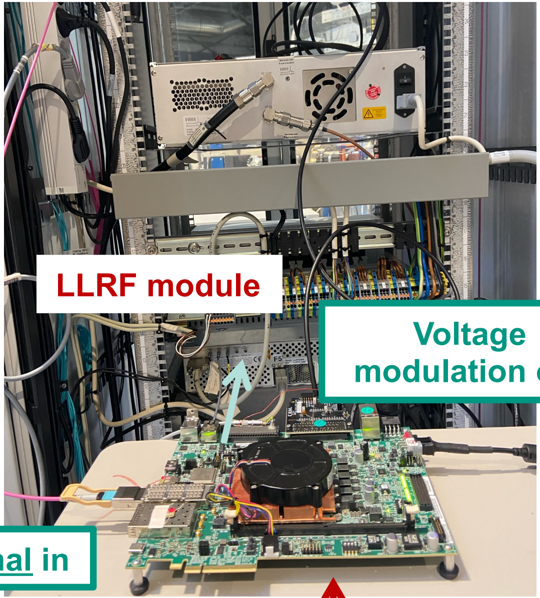
Control of the microbunching instablity with RL

Hardware setup at KARA

At the beamline



In the KARA ring



Control of the microbunching instablity with RL

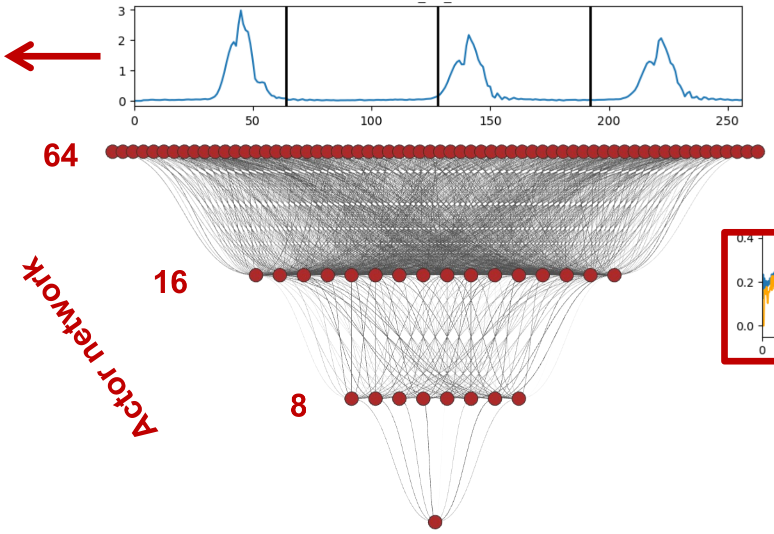
RL implementation

- **Algorithm**: vanilla PPO from Stable Baselines3

Observation

Last 64 THz signal samples (decimated)

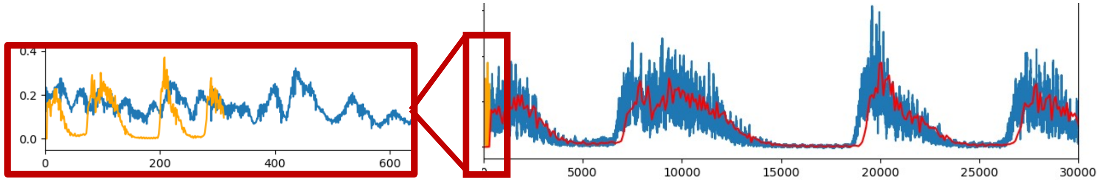
Circular buffer (keeps last 64 samples in memory)



Action (voltage value)

Interpolated (smoothed) to avoid sharp steps (cavity interlocks)

- **Decimation (two stages):**
  - Controls timescale agent "sees"
    - Rate of action = decimation x T_{rev}
  - Makes infererence & training easier (smaller networks)
  - We decimate 16 x 6 = 96 (take a sample every 96 revolutions)
    - We act every 96 x T_{rev} ~ 28 kHz ~ 0.25 x T_{sync} ~ 36 μs
    - We use 440 samples per second



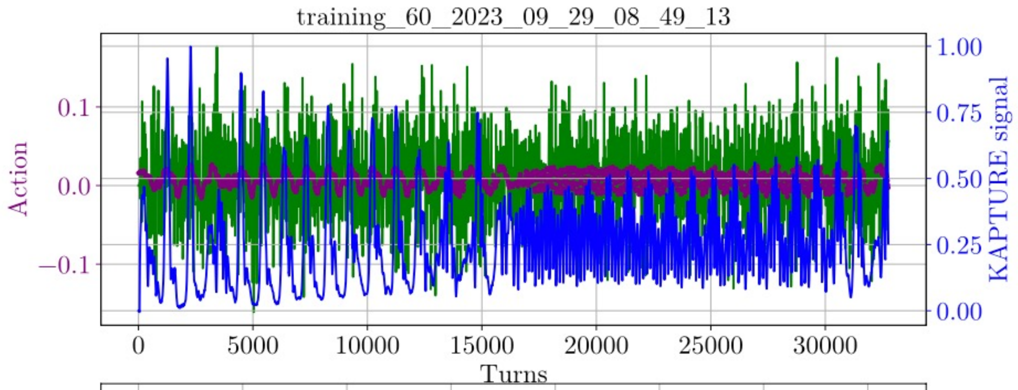
Filtering to remove spurious content + decimation

- **Strategy:**
  1. Agent acts during 2048 steps (samples of decimated signal)
  2. Agent stops and is re-trained in a CPU (takes ~2.6 s)
     - We train every (2048 x 96) T_{rev} = 509 T_{sync}
  3. New weights are sent to Versal board and agent starts again

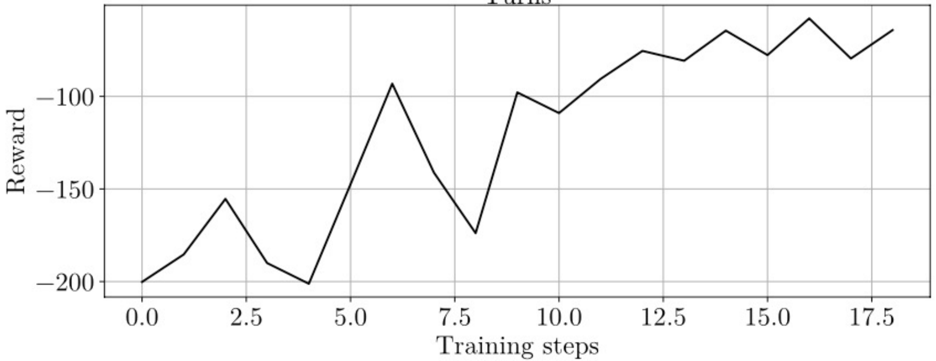
Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC



Energy = 1.3
Current = 0.4199
Alpha = 0.00836
RF voltage = 767.29
fsync = 9.555
Decimation = 16
Ramping steps = 240



-((rollout_buffer.T[2]-0.3)**2)

Purple: mean of action distribution
(policy output, non cumulated)

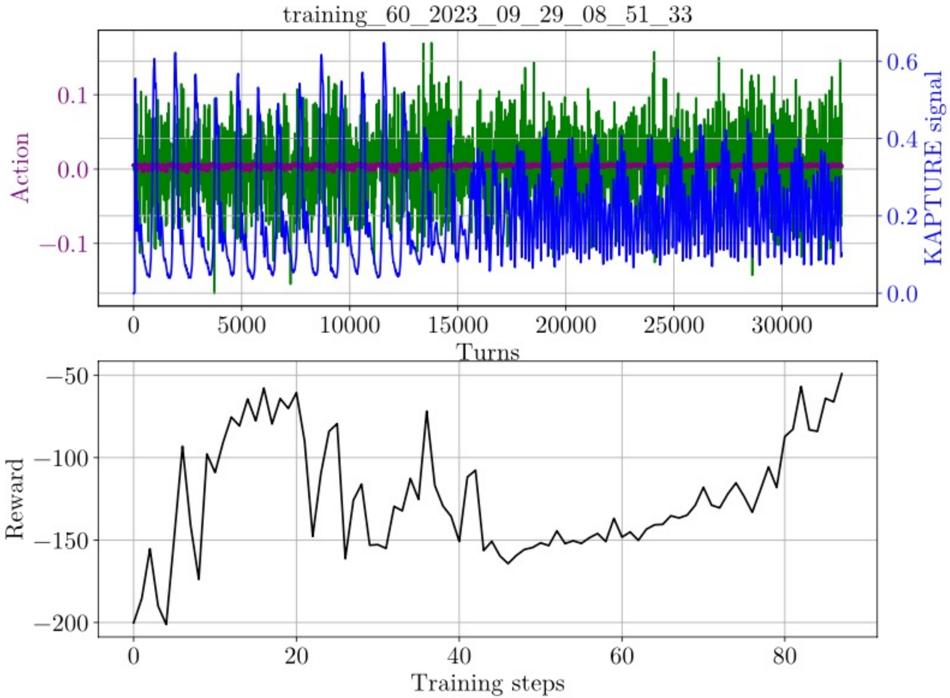
Green: action taken from distribution

Real action that reaches the cavity: ?

Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC



Energy = 1.3
Current = 0.4051
Alpha = 0.00836
RF voltage = 767.11
fsync = 8.859
Decimation = 16
Ramping steps = 240

-((rollout_buffer.T[2]-0.3)**2)

Purple: mean of action distribution
(policy output, non cumulated)

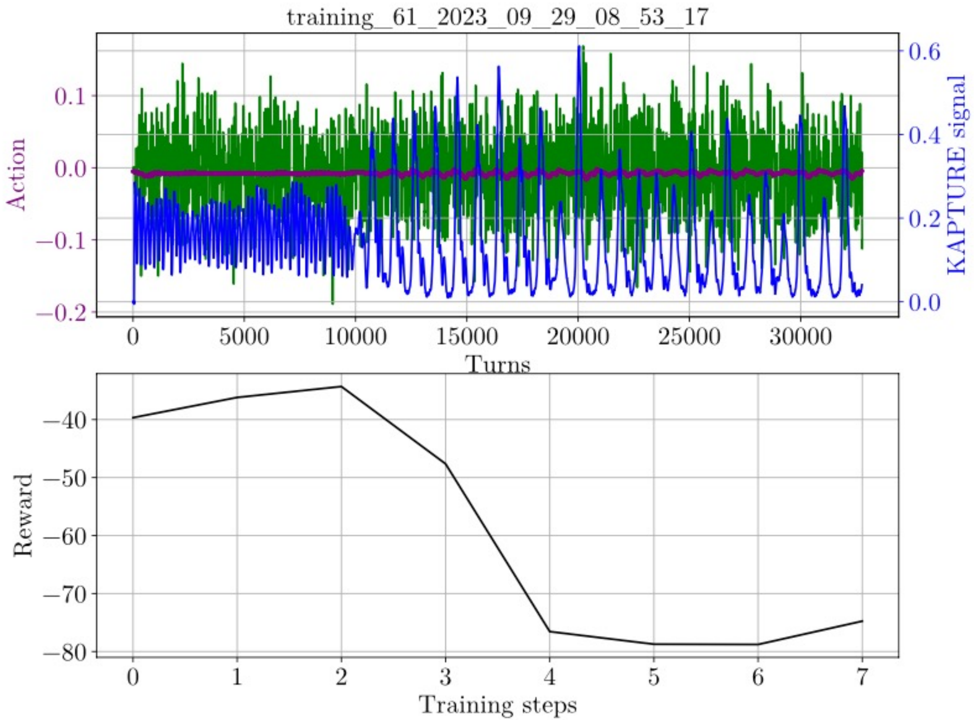
Green: action taken from distribution

Real action that reaches the cavity: ?

Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC



Energy = 1.3
Current = 0.3947
Alpha = 0.00836
RF voltage = 767.26
fsync = 8.859
Decimation = 16
Ramping steps = 240

-((rollout_buffer.T[2]-0.3)**2)

Purple: mean of action distribution
(policy output, non cumulated)

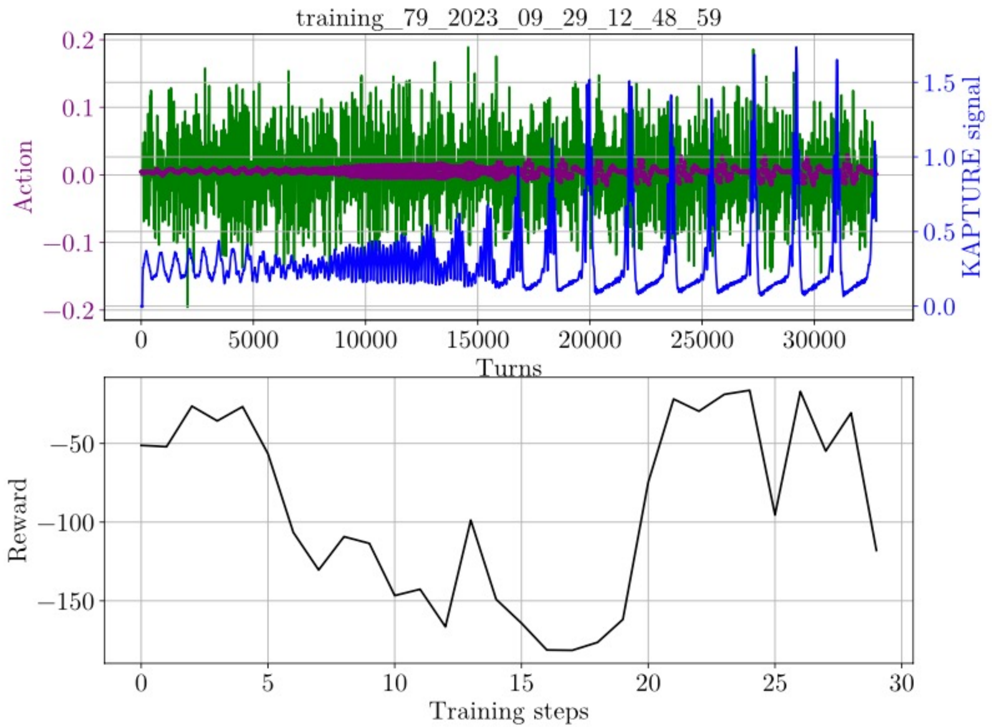
Green: action taken from distribution

Real action that reaches the cavity: ?

Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC



Energy = 1.3
Current = 0.2752
Alpha = 0.00836
RF voltage = 767.16
fsync = 8.096
Decimation = 16
Ramping steps = 245

-((rollout_buffer.T[2]-0.3)**2)

Purple: mean of action distribution
(policy output, non cumulated)

Green: action taken from distribution

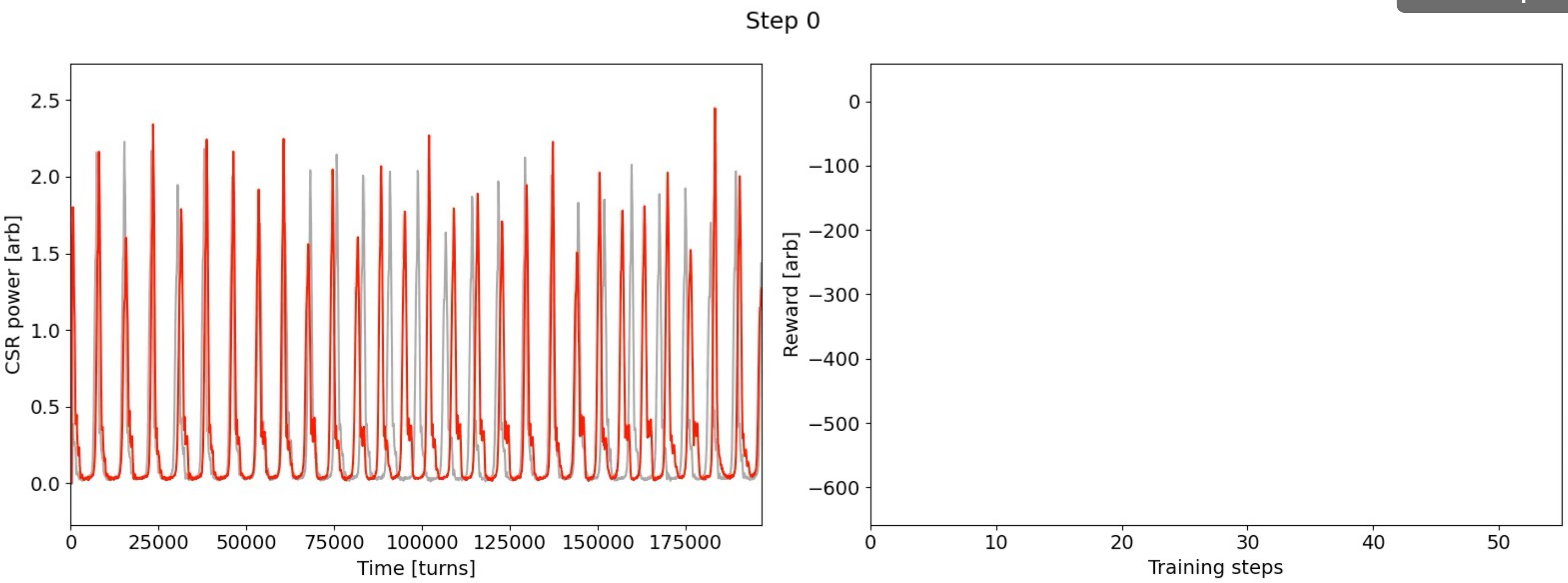
Real action that reaches the cavity: ?

Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC

L. Scomparin

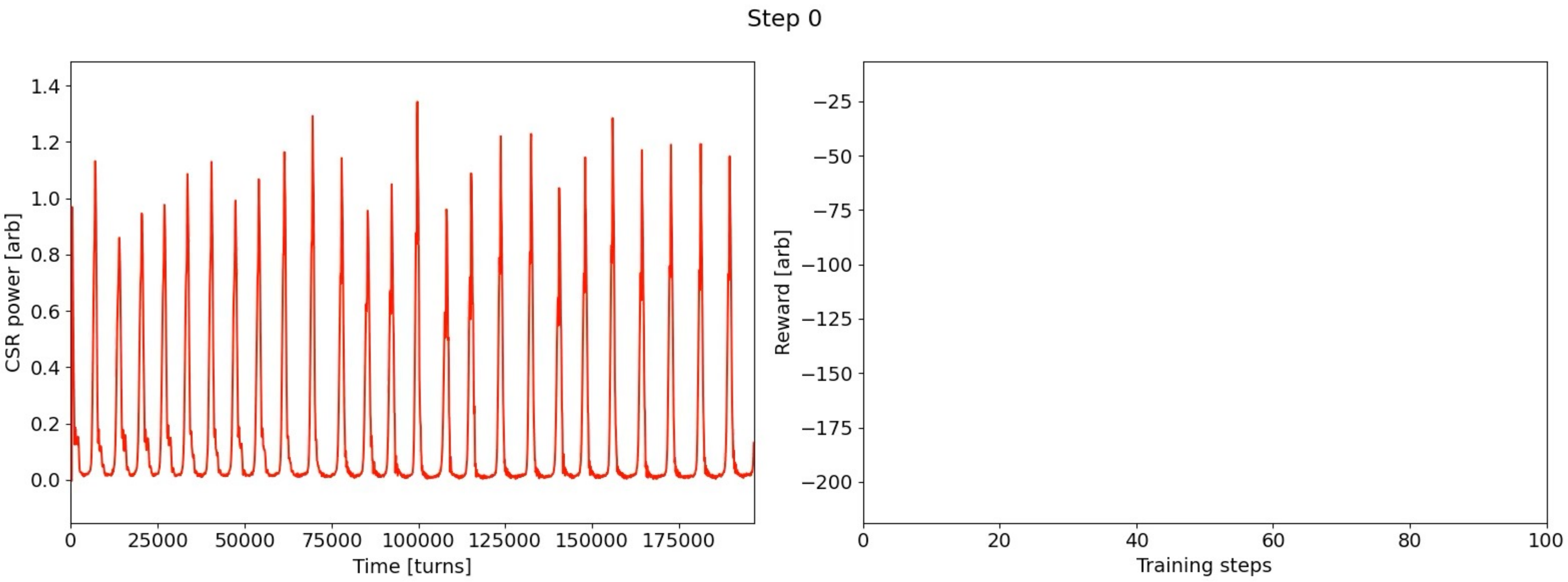


Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC

L. Scomparin



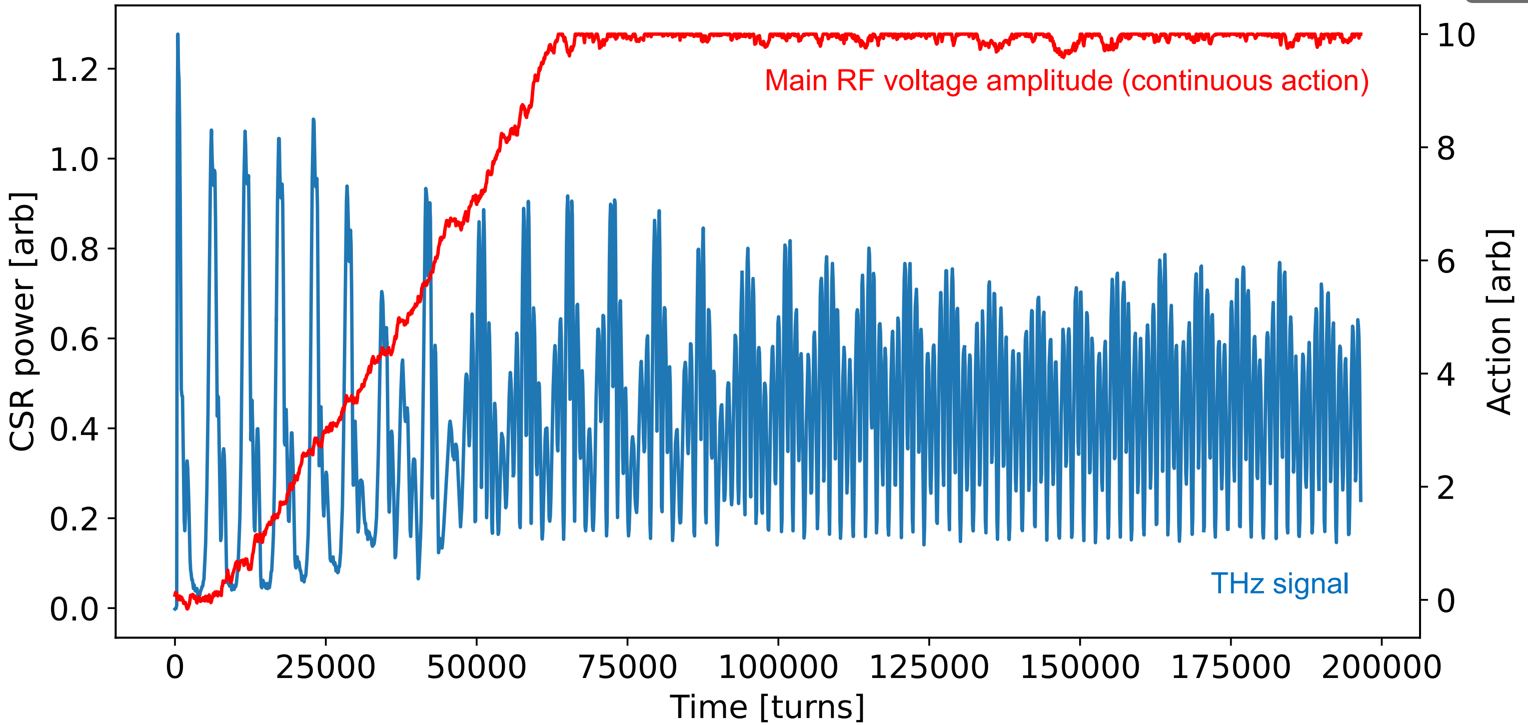
Control of the microbunching instablity with RL

Preliminary results

Single bunch, 1 ADC

Karlsruhe Institute of Technology

L. Scomparin



Thank you for your attention!

What questions do you have for me?

Karlsruhe Institute of Technology

Institute for Beam Physics and Technology

Dr. Andrea Santamaria Garcia
AI4Accelerators team leader

andrea.santamaria@kit.edu
https://twitter.com/ansantam
https://www.linkedin.com/in/ansantam/
https://github.com/ansantam
https://instagram.com/ansantam