First Steps Towards Mitigation of Harmonic Orbit Perturbations with Reinforcement Learning at BESSY II

Luis Vera Ramírez, Pierre Schnizer,
Tom Mertens, Markus Ries

Helmholtz-Zentrum Berlin

luis.vera_ramirez@helmholtz-berlin.de

EPICS Collaboration Fall Meeting, 22/10/2020

Table of Contents

Problem Description

Control via Bluesky/Naus

Reinforcement Learning

Experiments

- Simulations
- First Tests @ BESSY II

Summary and Outlook

References

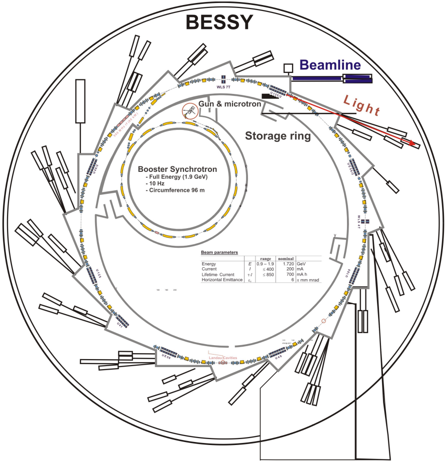


Table of Contents

Problem Description

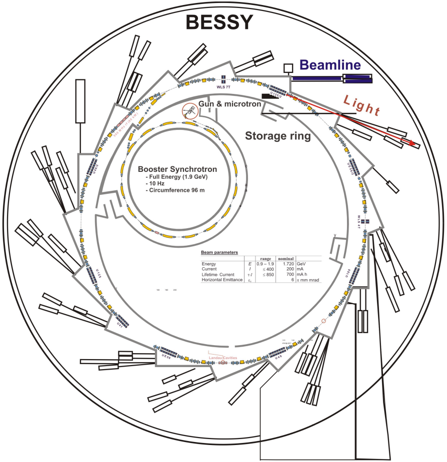
Control via Bluesky/Naus

Reinforcement Learning

Experiments
- Simulations
- First Tests @ BESSY II

Summary and Outlook

References



HZB Helmholtz Zentrum Berlin

Follow-up of Olivier Churlaud's work ([Chu16])

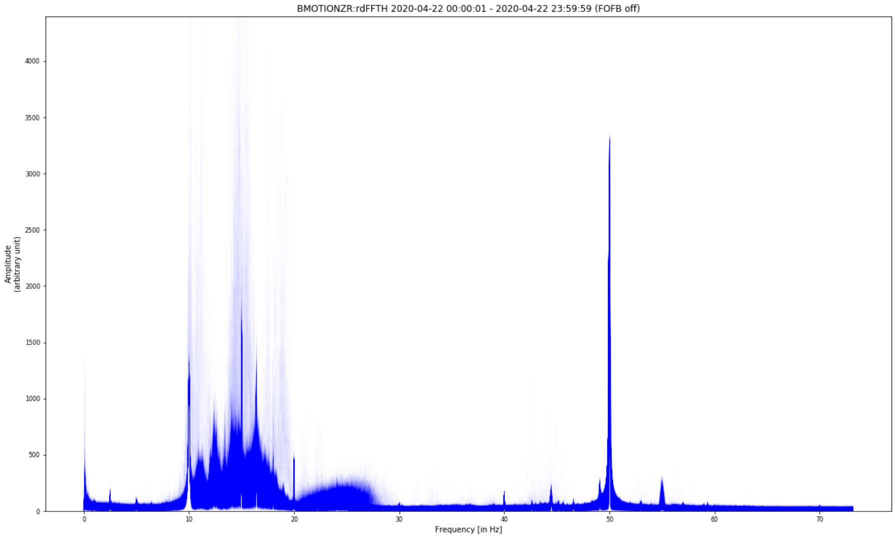
- To achieve light radiation with high quality brilliance and brightness over time, the light source itself must be very stable and the electron beam very small.
- The stability of the orbit (the ideal trajectory of the particles) must be **below the transverse beam dimensions** - in BESSY II, $100 \times 20\mu$m.
- As the precision of the positioning of the magnets is limited, some errors may destabilize the orbit.

$\rightarrow$ Mostly reduced with *traditional* correction methods - **Fast Orbit Feedback**, based on the response matrix inverse calculation with SVD

- But the environment also produces **perturbations**, e.g. the 50 Hz of the main power or some imperfectly isolated magnetic sources (like the booster at 10 Hz), among others.

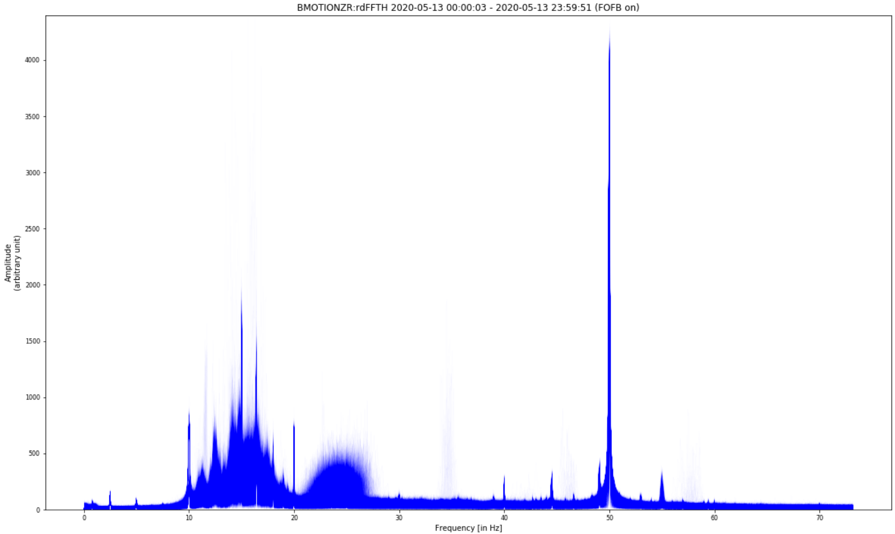
$\rightarrow$ **To be corrected with Reinforcement Learning?**

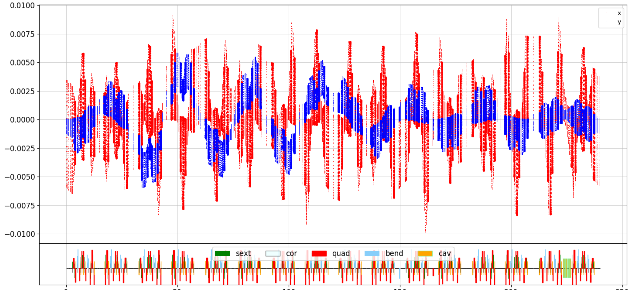
Horizontal beam motion spectrum without fast orbit correction:



Problem Description: Orbit FFT

Horizontal beam motion spectrum with fast orbit correction:





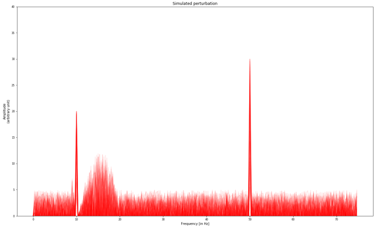

```
from ocelot.cpbd.elements import *

DG9L2D1R = Drift(l=0.6155, eid='DG9L2D1R')
M_FOMZ2D1R = Marker(eid='M_FOMZ2D1R')
DF9L2D1R = Drift(l=0.521, eid='DF9L2D1R')
DE9L2D1R = Drift(l=0.6485, eid='DE9L2D1R')
BPMZ43D1R = Marker(eid='BPMZ43D1R')
DD9L2D1R = Drift(l=0.4025, eid='DD9L2D1R')
BPMZ44D1R = Marker(eid='BPMZ44D1R')
DB9L2D1R = Drift(l=0.5485, eid='DB9L2D1R')
BPMZ5D1R = Marker(eid='BPMZ5D1R')
DA9L2D1R = Drift(l=0.07, eid='DA9L2D1R')
S4M2D1RL= Sextupole(l=0.08, k2=27.0435,
    tilt=0.0, eid='S4M2D1RL')
HS4M2D1R= Hcor(l=0.0, angle=0.0, eid='HS4M2D1R')
S4M2D1RR= Sextupole(l=0.08, k2=27.0435,
    tilt=0.0, eid='S4M2D1RR')
D08L2D1R = Drift(l=0.153, eid='D08L2D1R')
( . . . )
```

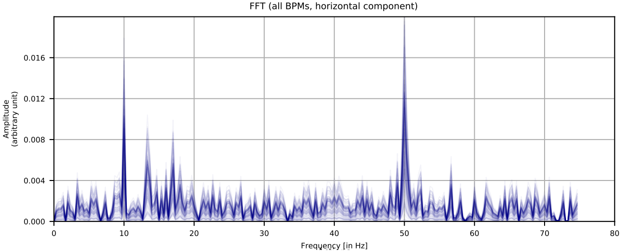
- The software framework OCELOT ([AGTZ14]) allows more interactive and modularized control of the BESSY representation.
- It allowed us to extend the first experiments with code based on [Chu16].
- The simulation performance could be optimized and accelerated more easily.

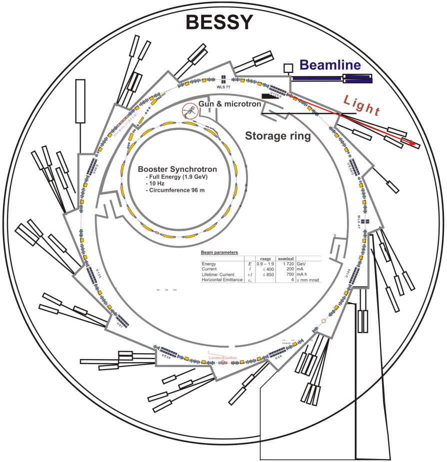
Simulation of perturbed beam motion with OCELOT:

- Synthetic, randomized, harmonic perturbations defined in [Chu16]:



- Spectrum of the horizontal beam motion with the synthetic perturbation applied to the horizontal offset of the quadrupole Q4M2D1R at 150Hz:

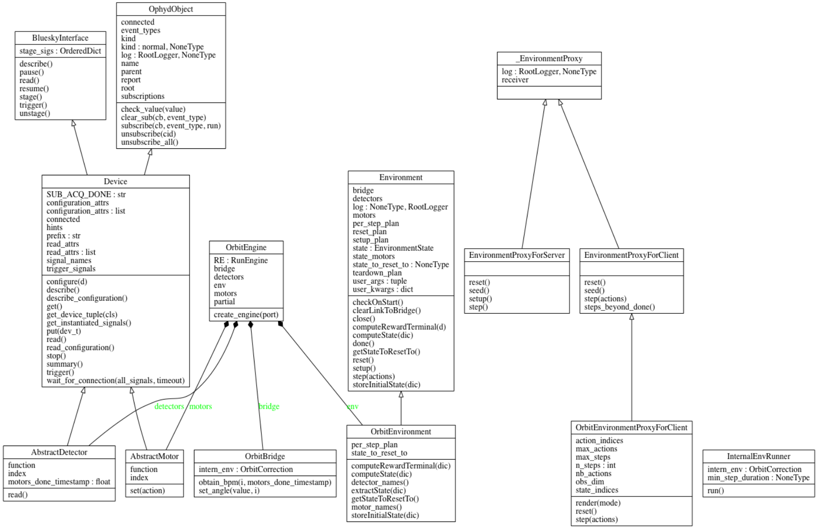



BESSY
Beamline
Light
Gun & microtron
Storage ring
Booster Synchrotron
- Full Energy (1.9 GeV)
- 10 Hz
- Circumference 96 m
WLS 7T
Landau Cavities
Beam parameters
range
nominal
Energy
E
0.9 – 1.9
1.720
GeV
Current
I
≤ 400
200
mA
Lifetime· Current
τ·I
≤ 850
700
mA h
Horizontal Emittance
ε_x
6
π mm mrad

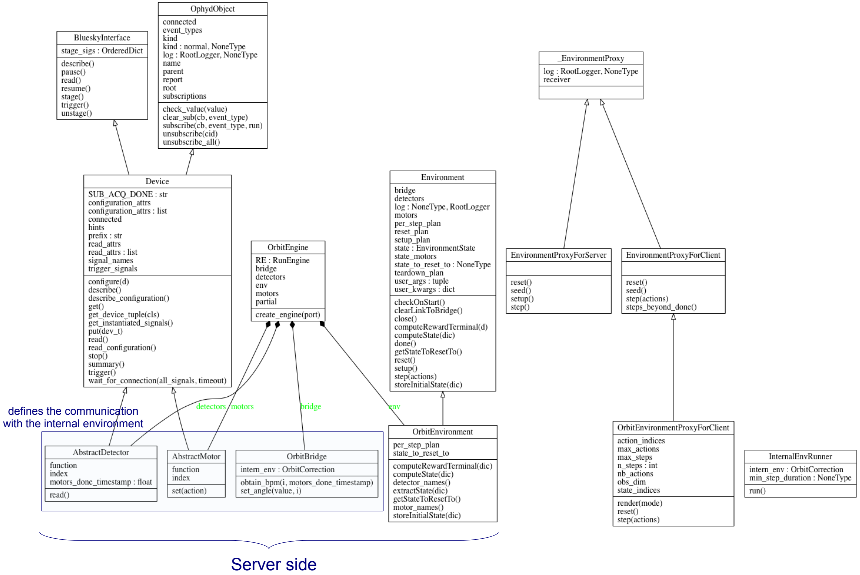
Control via Bluesky/Naus: Concept

- We aimed to create an interaction framework whose interfaces remained completely unchanged when training a ML-model with simulations or at the real machine.
- It corresponds to the philosophy of the *Digital Twin*.
- For this we used **Naus**, based on **bluesky** and **ophyd** (already presented by Pierre Schnizer - *Bluesky at BESSY II: A measurement script metamorphosis*)

Control via Bluesky/Naus: Environment Diagram

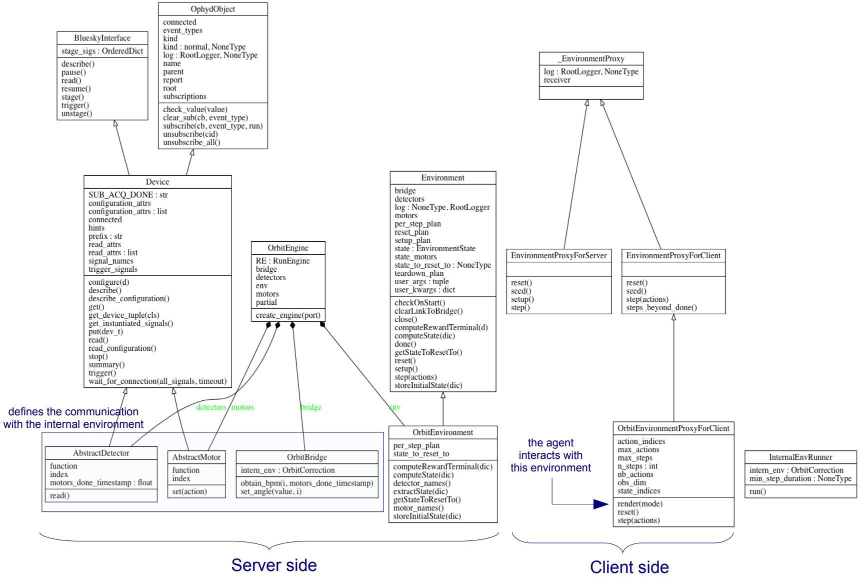


Control via Bluesky/Naus: Environment Diagram

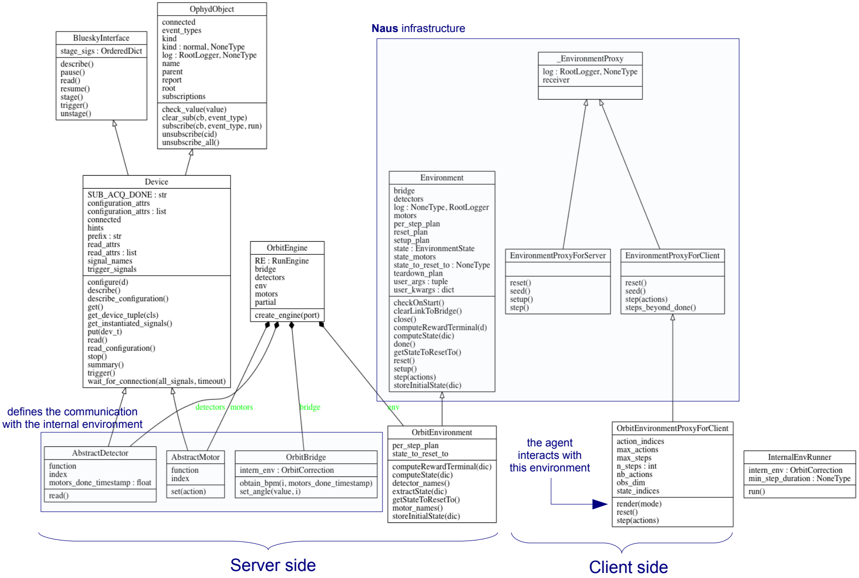


Control via Bluesky/Naus: Environment Diagram

HZB Helmholtz Zentrum Berlin



Control via Bluesky/Naus: Environment Diagram



Control via Bluesky/Naus: Environment Diagram

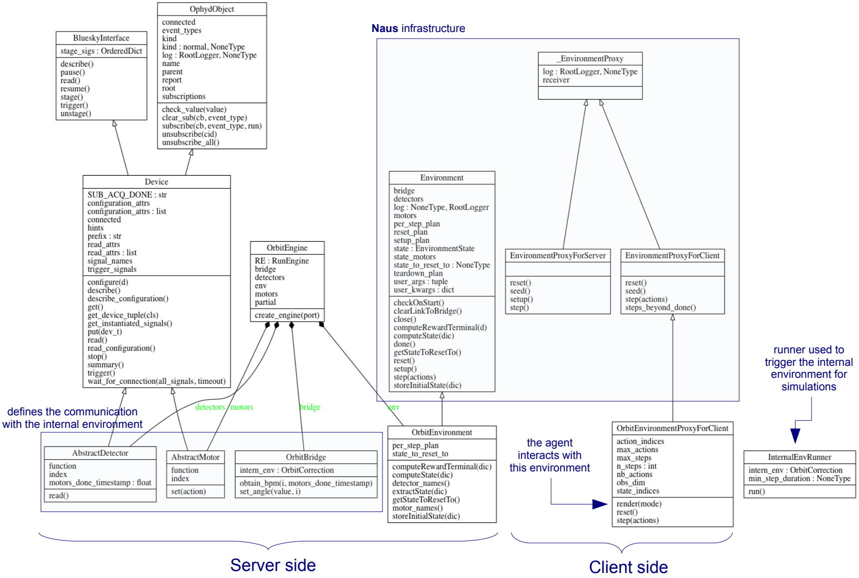
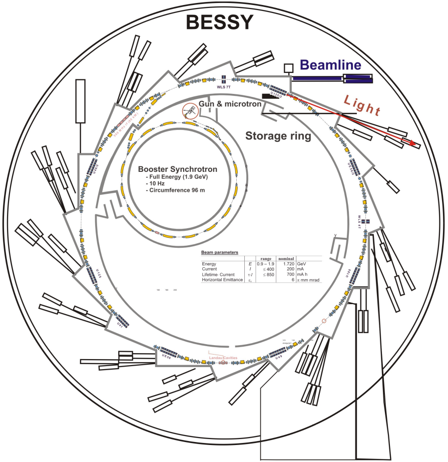


Table of Contents



Reinforcement Learning: Machine Learning Landscape

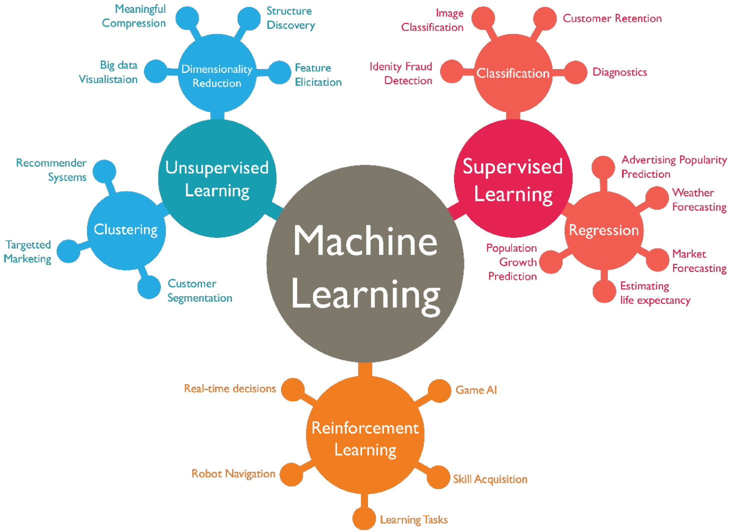
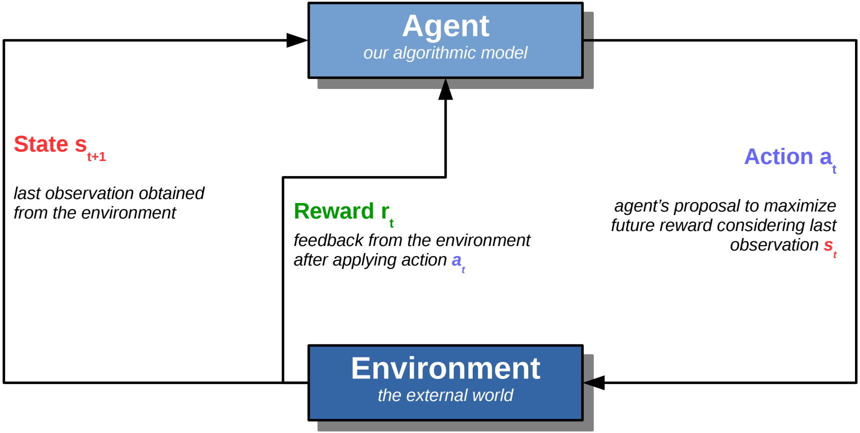
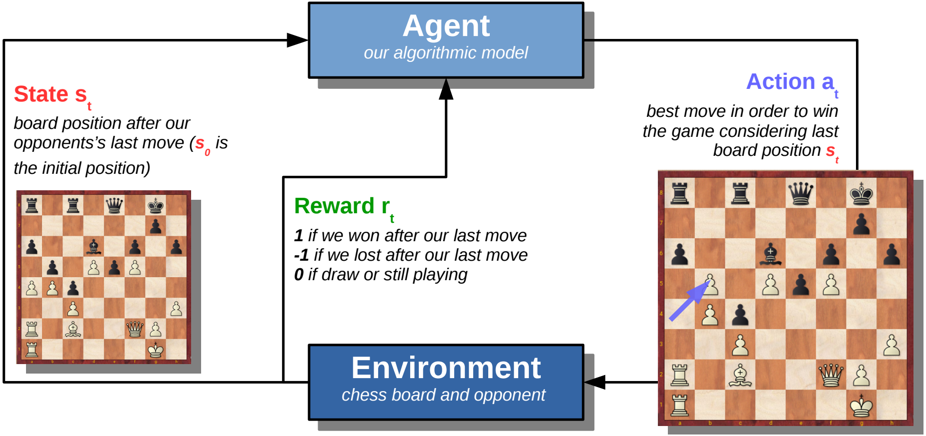


Figure from http://www.isaziconsulting.co.za/machinelearning.html



- **Step**: a single state-action-reward-state interaction loop.
- **Episode**: a succession of steps until a *terminal state* is reached.

HZB Helmholtz Zentrum Berlin



- **Step**: one white's and black's movement.
- **Episode**: a complete game.

- **Deep Deterministic Policy Gradient** [LHP+16]: Actor-Critic Reinforcement Learning algorithm for continuous environments.
- The **Q-function** (estimation of the future reward for a given state-action pair) and the **policy** (map between states and actions) are *approximated* with Neural Networks.

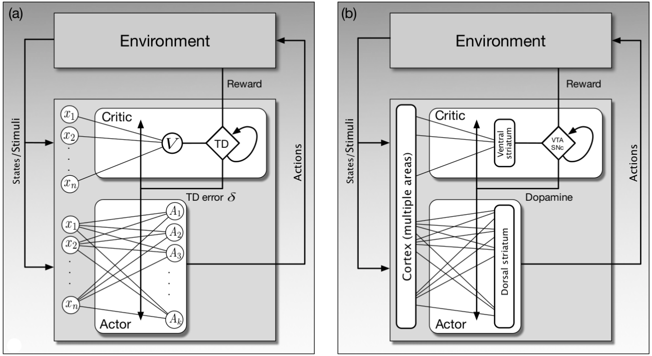
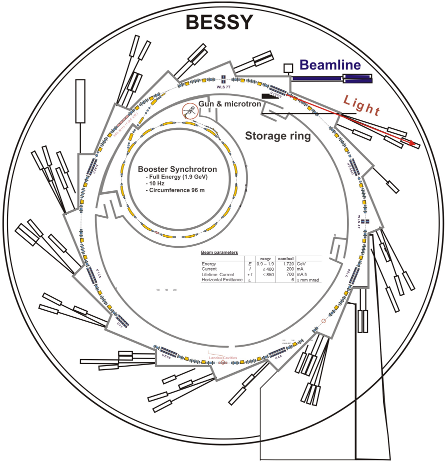


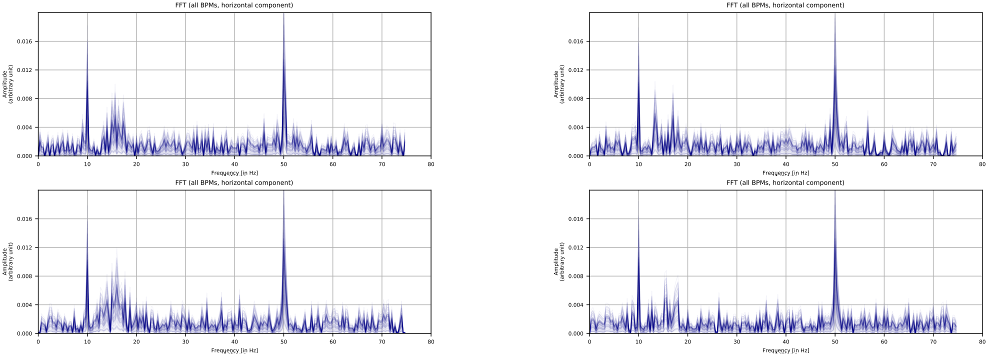
Figure: From [SB18]



BESSY
Beamline
Light
Gun & microtron
Storage ring
Booster Synchrotron
- Full Energy (1.9 GeV)
- 10 Hz
- Circumference 96 m
WLS 7T
Beam parameters
		range	nominal	
Energy	E	0.9 – 1.9	1.720	GeV
Current	I	≤ 400	200	mA
Lifetime- Current	τ·I	≤ 850	700	mA h
Horizontal Emittance	εx		6	π mm mrad
Landau Cavities

Simulations: Some tests with DDPG and OCELOT

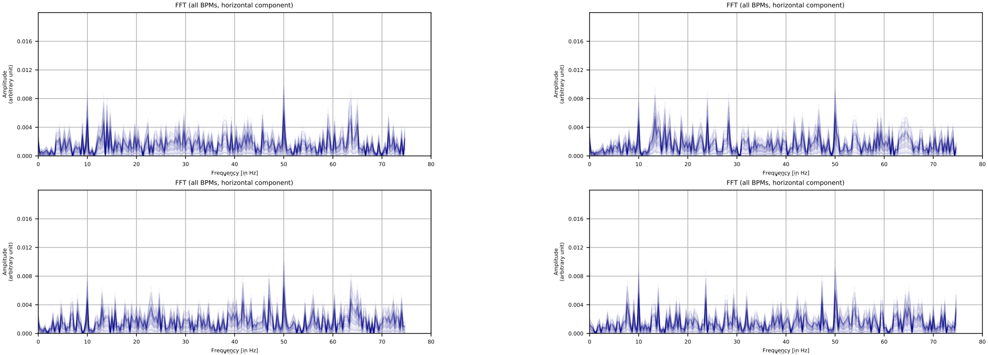
- Simulated perturbation applied to **Q4M2D1R.dx** with resolution = 150Hz
- Horizontal steerer **HS4M2D1R** modified ($\rightarrow$ **action**)
- x component of the BPM (beam position monitor) **BPMZ6D1R** read with **windows size = 30** ($\rightarrow$ **state**) - *the Deep RL Agent stays in the time domain!*



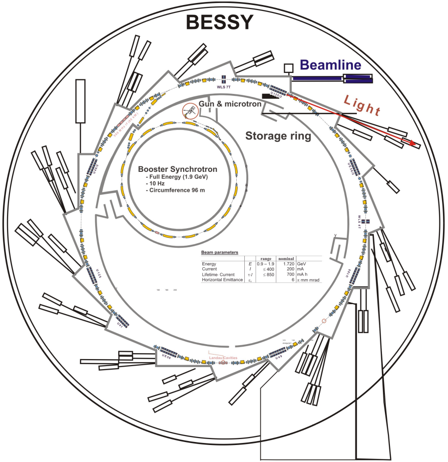
Simulation of perturbed beam motion spectrum

Simulations: Some tests with DDPG and OCELOT

- Simulated perturbation applied to **Q4M2D1R.dx** with resolution = 150Hz
- Horizontal steerer **HS4M2D1R** modified ($\rightarrow$ **action**)
- x component of the BPM (beam position monitor) **BPMZ6D1R** read with **windows size = 30** ($\rightarrow$ **state**) - *the Deep RL Agent stays in the time domain!*



Simulation of perturbed beam motion spectrum **corrected with RL-Agent**


BESSY
Beamline
Light
Gun & microtron
Storage ring
Booster Synchrotron
- Full Energy (1.9 GeV)
- 10 Hz
- Circumference 96 m
WLS 7T
Beam parameters
range
nominal
Energy E 0.9 – 1.9 1.720 GeV
Current I ≤ 400 200 mA
Lifetime- Current τ·I ≤ 850 700 mA h
Horizontal Emittance εx 6 π mm mrad

- In July 2020 we managed to set up the infrastructure for RL-based correction of harmonic perturbations during machine commissioning.

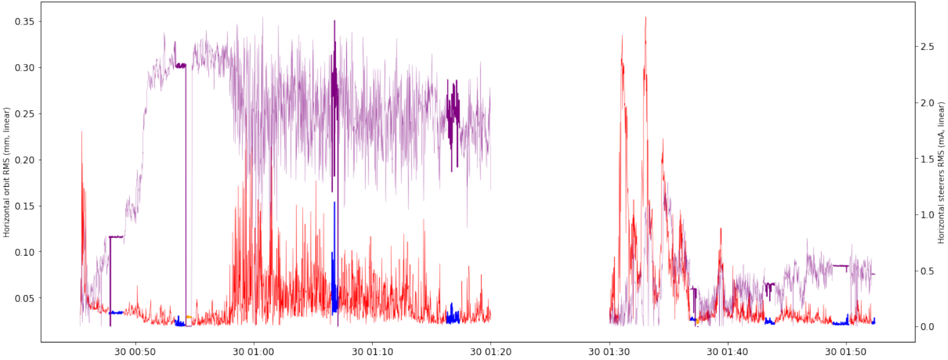
$\rightarrow$ First *plausibility tests* of the Naus-based framework **up to 20Hz** were carried out succesfully.

- In September 2020 we carried out new tests, focussing on the code performance in order to accelerate the interaction loop and so get first meaningful learning results.

$\rightarrow$ A direct zmq-communication with the mBox (fast orbit correction infrastructure) was established, allowing an acceleration of the RL-interaction loop **up to 100Hz**.

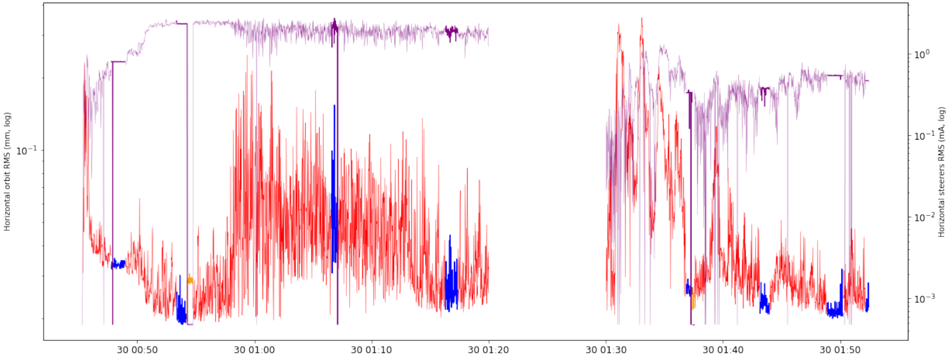
Experiments: Last test runs (30/09/20)

- **State**: all active BPMs (102) with window size 10.
- **Action**: all horizontal steerers (48) modified up to $\pm$ 6 mA.
- **Reward**: exponential transformation of the (horizontal) BPM norm.



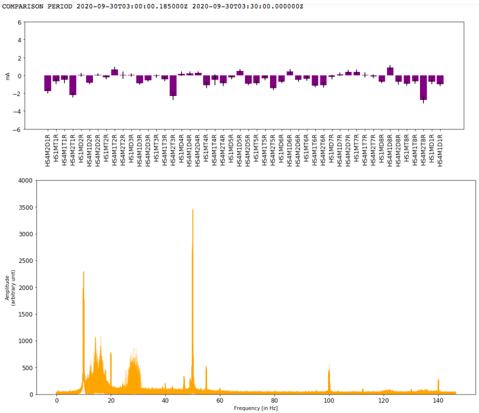
BPM RMS norm: red = exploration, blue = exploitation, orange = agent off (comparison)
Learning rate: learn at every step at $\sim$ 26.6 Hz (left); learn after 50 steps at $\sim$ 100 Hz (right)

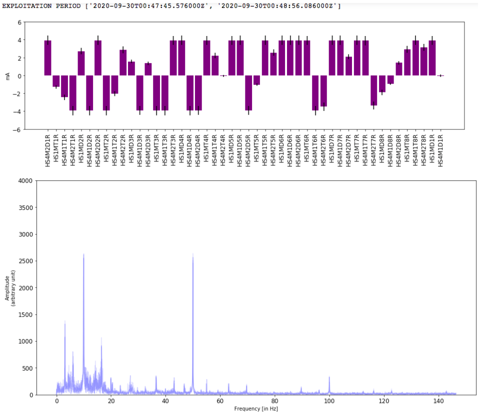
- **State**: all active BPMs (102) with window size 10.
- **Action**: all horizontal steerers (48) modified up to $\pm$ 6 mA.
- **Reward**: exponential transformation of the (horizontal) BPM norm.



BPM RMS norm: red = exploration, blue = exploitation, orange = agent off (comparison)
Learning rate: learn at every step at $\sim$ 26.6 Hz (left); learn after 50 steps at $\sim$ 100 Hz (right)

- **State**: all active BPMs (102) with window size 10.
- **Action**: all horizontal steerers (48) modified up to $\pm$ 6 mA.
- **Reward**: exponential transformation of the (horizontal) BPM norm.

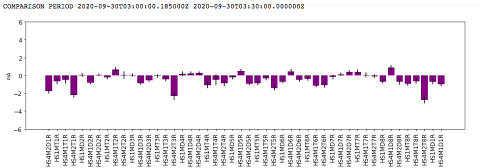


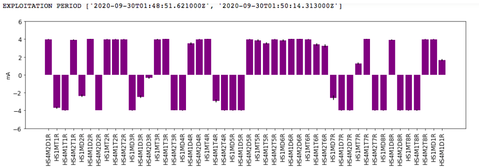


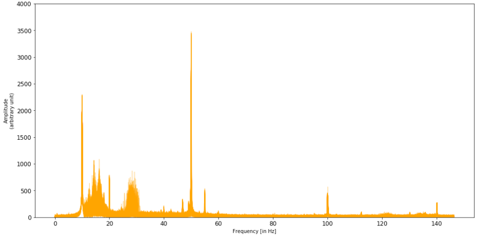
Beam Motion Spectrum: blue = exploitation, orange = agent off (comparison)
Learning rate: learn at every step at $\sim$ 26.6 Hz

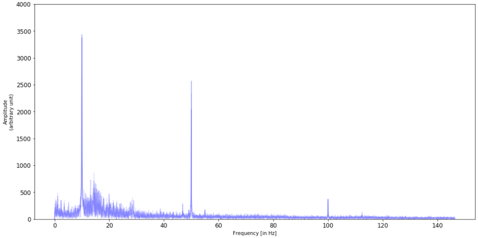
Experiments: Last test runs (30/09/20)

- **State**: all active BPMs (102) with window size 10.
- **Action**: all horizontal steerers (48) modified up to $\pm$ 6 mA.
- **Reward**: exponential transformation of the (horizontal) BPM norm.









Beam Motion Spectrum: blue = exploitation, orange = agent off (comparison)
Learning rate: learn after 50 steps at $\sim$ 100 Hz

Table of Contents

Summary and Outlook

- The problem of the mitigation of harmonic orbit perturbations at BESSY II is being faced with the help of **Reinforcement Learning Agents**.
- An *universal interaction environment* with shared interfaces for simulations and experiments at the machine was developed and tested up to 20Hz.
- Further tests up to 100Hz were carried out with a simplified version of the environment, giving first meaningful results.

Summary and Outlook

- Next steps:
  - Improve performance of the Naus-environment.
  - Improve synchronization with the mBox.
  - Accelerate RL-algorithms.
- Further open projects with ML @ BESSY II:
  - Optimization of **booster current** and **injection efficiency** with RL.
  - Prediction and optimization of **vertical beam size**
  - **Anomaly detection systems** (e.g. orbit position) with Isolation Forests

Table of Contents

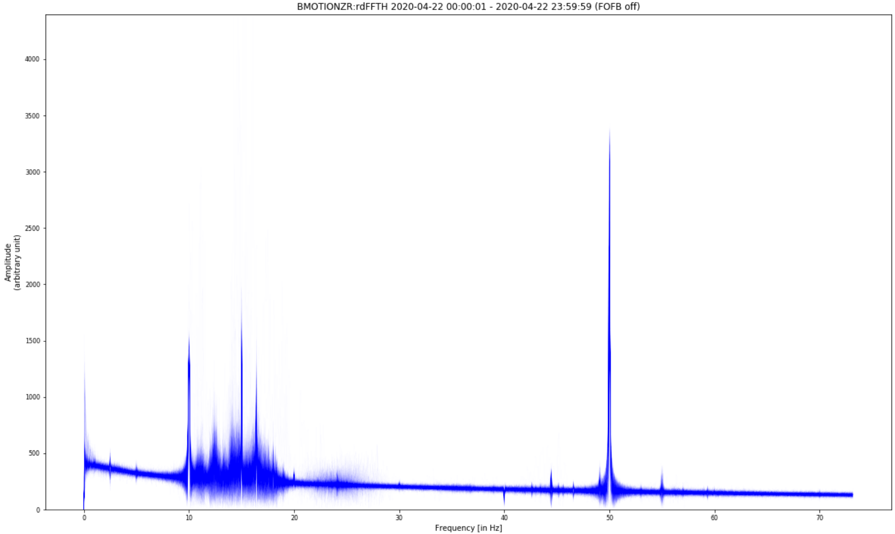
[AGTZ14] Ilya Agapov, Gianluca Geloni, Sergey Tomin, and Igor Zagorodnov. OCELOT: A software framework for synchrotron light source and FEL studies. *Nuclear instruments & methods in physics research / A*, 768:151 – 156, 2014. (c) Elsevier B.V.

[Chu16] Olivier Churlaud. Localization and correction of orbit perturbations in bessy ii storage ring. Master's thesis, TU Berlin, HZB, 8 2016. An optional note.

[LHP+16] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. In *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, 2016.

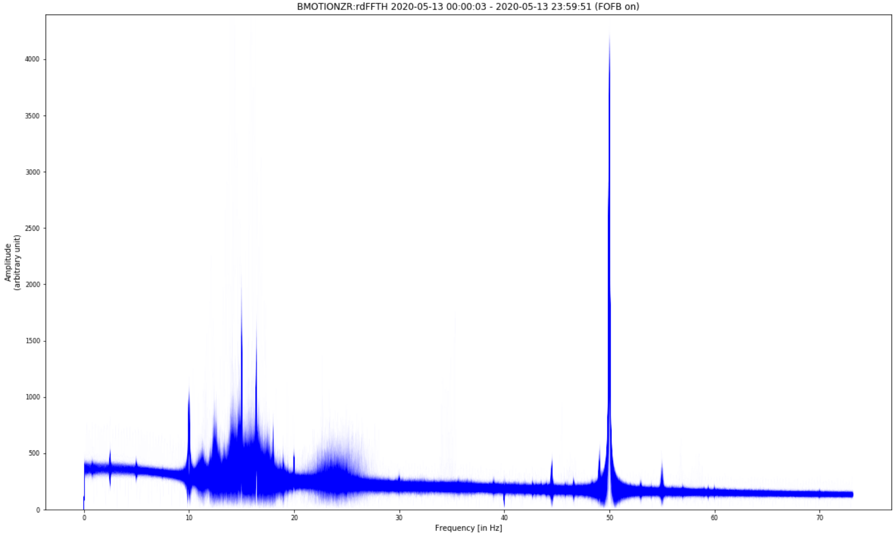
[SB18] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

Horizontal beam motion spectrum during injections without fast orbit correction:



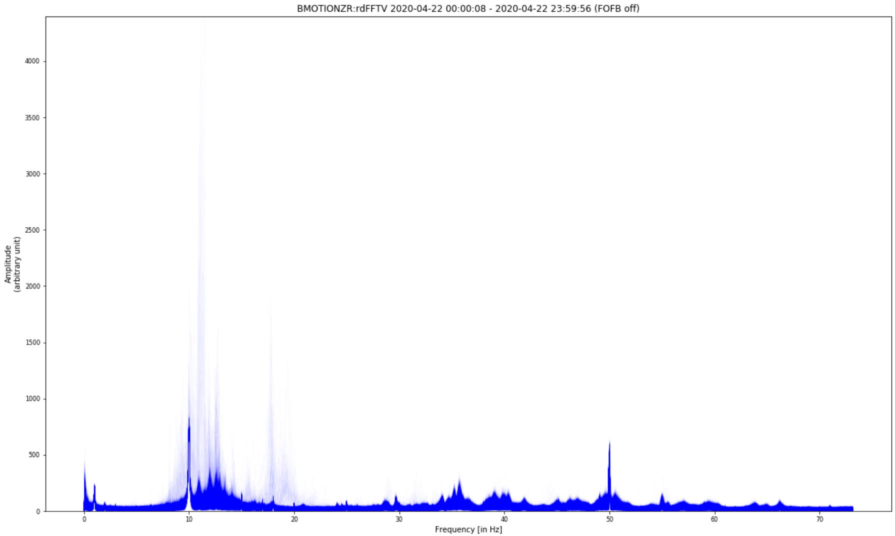
Backup: Orbit FFT

Horizontal beam motion spectrum during injections with fast orbit correction:

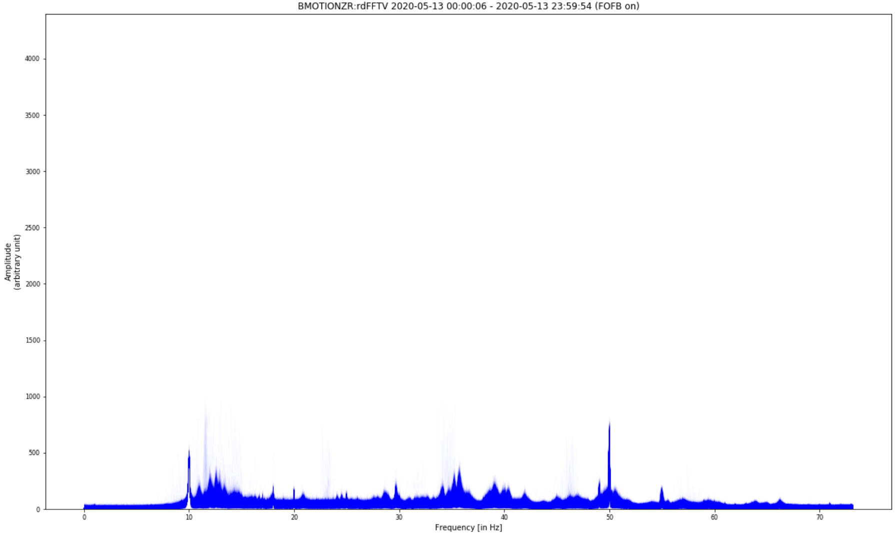


Backup: Orbit FFT

Vertical beam motion spectrum without fast orbit correction:



Vertical beam motion spectrum with fast orbit correction:



Fast Orbit Correction Schema at BESSY - extracted from [Chu16]:

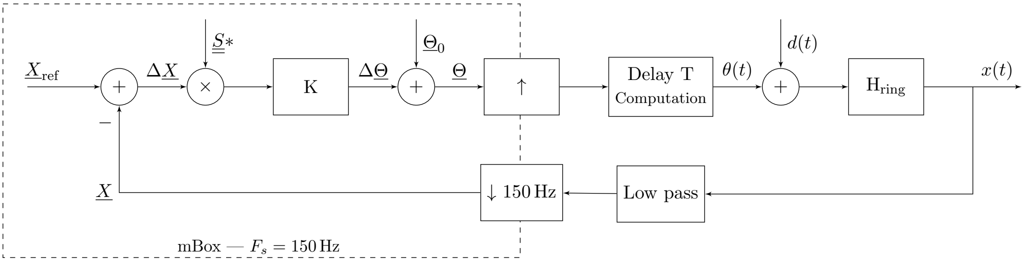


Figure 4.4: The full model, K being the corrector to define

HZB Helmholtz Zentrum Berlin

Synthetic, randomized, harmonic perturbations defined in [Chu16]:

```
def real_perturbation(t_max, fs):
    t = np.arange(int(fs*t_max))/fs
    N = t.size
    Fs = 1/(t[1]-t[0])
    freqs = np.fft.fftfreq(N, 1/Fs)
    freqs_half = freqs[:N//2+1]
    cm_fft = 5*np.random.random(N//2+1)*np.exp(1j*2*np.pi*np.random.random(N//2+1))

    idxmin = np.argmin(abs(freqs_half - 9))
    idx20 = np.argmin(abs(freqs_half - 20))
    for k in range(idxmin, idx20):
        cm_fft[k] = 0.1*cm_fft[k]*(5 - (freqs_half[k] - 11)*(freqs_half[k] - 20))

    nprand = np.random.random
    cmph10 = 2*np.pi*nprand()
    cm_fft[np.argmin(abs(freqs_half - 0))] = 0
    cm_fft[np.argmin(abs(freqs_half - 10))] = 20*np.exp(1j*cmph10)
    cm_fft[np.argmin(abs(freqs_half - 50))] = 30*np.exp(1j*2*np.pi*nprand())
    cm_fft[-1] = 0

    cm_fft = np.concatenate((cm_fft[:-1], np.flipud(cm_fft.conjugate())[:-1]))
    cm_fft *= N/2/np.max(np.abs(cm_fft))
    cm = np.fft.ifft(cm_fft).real

    return cm
```