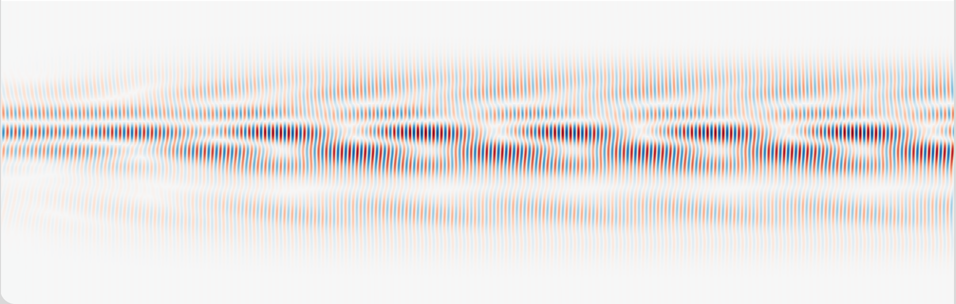
Karlsruhe Institute of Technology

Feedback Design for Control of the Micro-Bunching Instability based on Reinforcement Learning

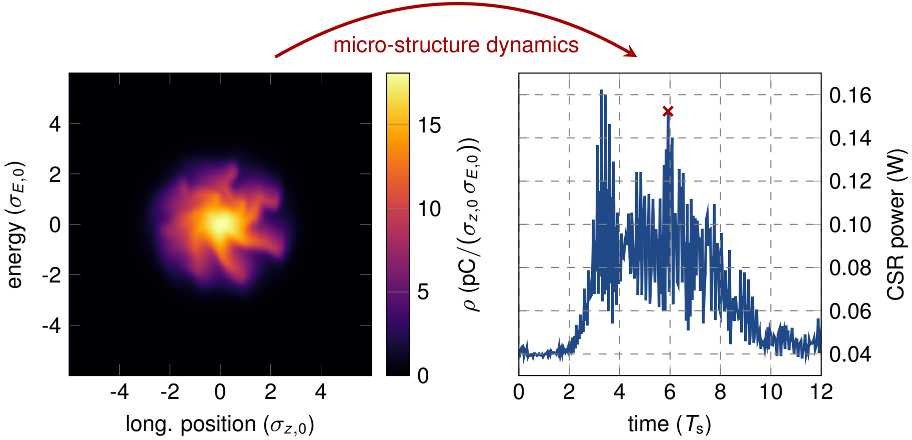
Tobias Boltz, Miriam Brosi, Erik Bründermann, Bastian Härer, Peter Kaiser, Christoph Pohl, Patrick Schreiber, Minjie Yan, Tamim Asfour and Anke-Susanne Müller | June 24, 2019

Laboratory for Applications of Synchrotron Radiation (LAS)



Micro-Bunching Instability

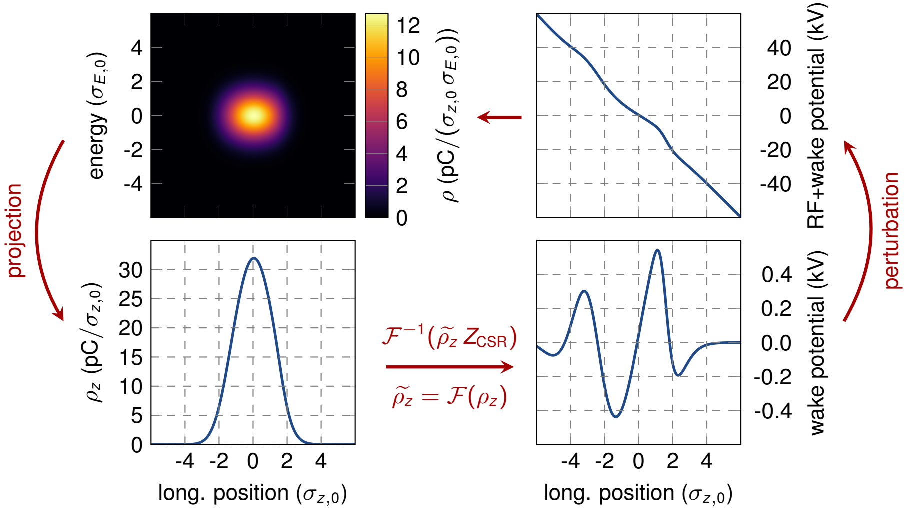
Micro-Bunching and CSR Power Fluctuations



⇒ continuous variation of charge distribution leads to fluctuating CSR

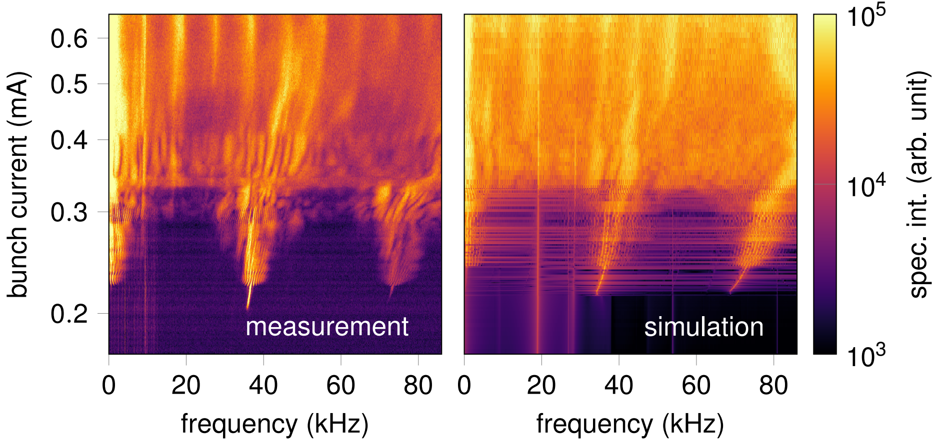
Micro-Bunching Instability

CSR self-interaction



Micro-Bunching Instability

CSR Power Spectrogram: Dependency on Bunch Current



Simulation code: Parallelized VFP solver ***Inovesa*** (`https://github.com/Inovesa/Inovesa`)
Schönfeldt, P. *et al.*, Phys. Rev. Accel. Beams 20 (2017)

Reinforcement Learning

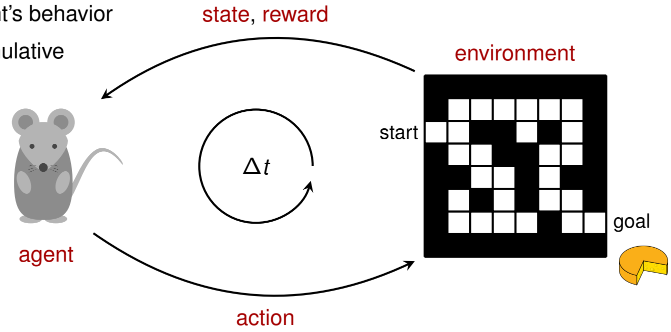
…in a Nutshell: Learning from Interaction

- goal-directed learning from interaction (trial-and-error search)
- mathematical foundation: Markov decision process (MDP)

  "The future is independent of the past given the present."

policy: defines agent's behavior

goal: maximize cumulative reward



Reinforcement Learning

. . . in a Nutshell: Finding better Policies

- value function q_π is the expected cumulative reward following policy π
- general policy iteration (GPI)
  - policy evaluation: learning the value function
  - policy improvement: exploiting the gained knowledge

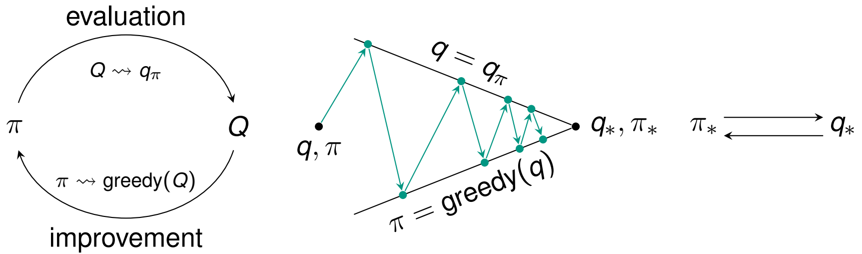
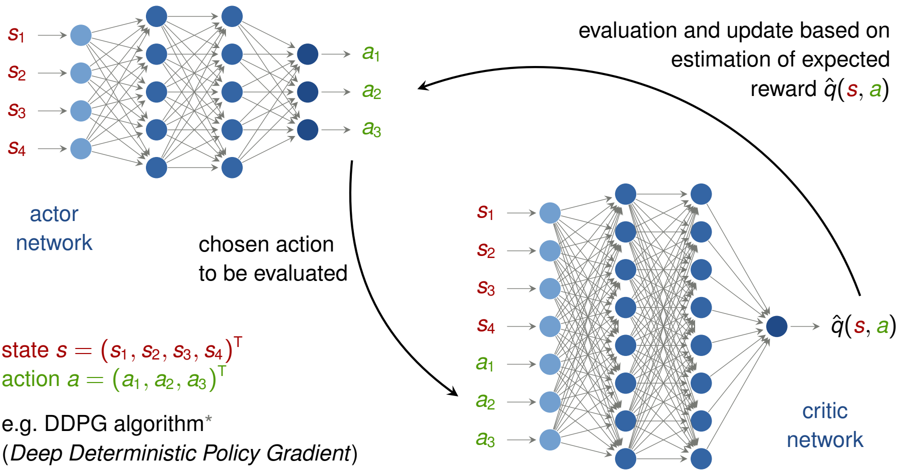


Figure: adapted from *Reinforcement Learning*, Sutton, R.S. and Barto, A.G., 2nd edition, MIT Press (November 2018)

Reinforcement Learning

…in a Nutshell: Actor-Critic System using NNs



state $s = (s_1, s_2, s_3, s_4)^\mathsf{T}$
action $a = (a_1, a_2, a_3)^\mathsf{T}$

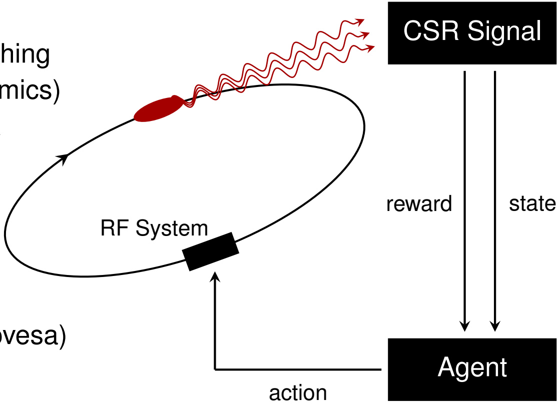
e.g. DDPG algorithm*
(*Deep Deterministic Policy Gradient*)

**Continuous control with deep reinforcement learning*, Lillicrap, T.P. *et al.* (2015), `https://arxiv.org/abs/1509.02971`

RL based Control of Micro-Bunching

Definition of Partially Observable MDP

goal: control micro-bunching (longitudinal beam dynamics) to optimize emitted CSR



proof of principle:
control in simulation (Inovesa)

implementation:
THz diagnostics (KAPTURE) and RF system at KARA

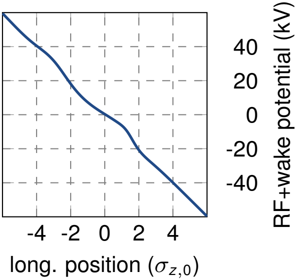
RL based Control of Micro-Bunching

Observation, Reward and Action

- observation: *hidden state* of electron bunch
  $\mathbf{o} = (\mu, \sigma, m, f_{\mathrm{max}}, A_{\mathrm{max}}, \varphi_{\mathrm{max}})^{\mathrm{T}} \Rightarrow$ full CSR signal or phase space?

- reward function: optimization of emitted CSR signal
  $R = w_1\mu - w_2\sigma \quad$ with $\quad w_1, w_2 > 0 \quad \Rightarrow$ best ratio w_1/w_2?

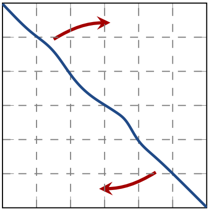
- action: modifications to the RF system
  1) scale RF amplitude and phase:
     $\mathbf{a} = (V_{\mathrm{RF}}, \varphi_{\mathrm{RF}})^{\mathrm{T}}$
  2) restrict to modulations of V_{RF} and φ_{RF}:
     $\mathbf{a} = (A_V, f_V, A_\varphi, f_\varphi)^{\mathrm{T}}$



RL based Control of Micro-Bunching

Observation, Reward and Action

- observation: *hidden state* of electron bunch
  $\mathbf{o} = (\mu, \sigma, m, f_{\max}, A_{\max}, \varphi_{\max})^{\mathrm{T}} \Rightarrow$ full CSR signal or phase space?
- reward function: optimization of emitted CSR signal
  $R = w_1\mu - w_2\sigma \quad \text{with} \quad w_1, w_2 > 0 \quad \Rightarrow$ best ratio w_1/w_2?
- action: modifications to the RF system
  1) scale RF amplitude and phase:
     $\mathbf{a} = (V_{\mathrm{RF}}, \varphi_{\mathrm{RF}})^{\mathrm{T}}$
  2) restrict to modulations of V_{RF} and φ_{RF}:
     $\mathbf{a} = (A_V, f_V, A_\varphi, f_\varphi)^{\mathrm{T}}$



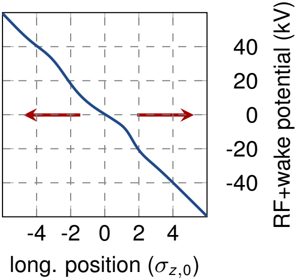
RL based Control of Micro-Bunching

Observation, Reward and Action

- observation: *hidden state* of electron bunch
  $\mathbf{o} = (\mu, \sigma, m, f_{\text{max}}, A_{\text{max}}, \varphi_{\text{max}})^{\text{T}} \Rightarrow$ full CSR signal or phase space?

- reward function: optimization of emitted CSR signal
  $R = w_1\mu - w_2\sigma$ with $w_1, w_2 > 0$ $\Rightarrow$ best ratio w_1/w_2?

- action: modifications to the RF system
  1) scale RF amplitude and phase:
     $\mathbf{a} = (V_{\text{RF}}, \varphi_{\text{RF}})^{\text{T}}$
  2) restrict to modulations of V_{RF} and φ_{RF}:
     $\mathbf{a} = (A_V, f_V, A_\varphi, f_\varphi)^{\text{T}}$



RL based Control of Micro-Bunching

Karlsruhe Institute of Technology

Tools and Packages

- simulation code Inovesa is written in C++, everything else in Python
- (mainly) used Python packages:

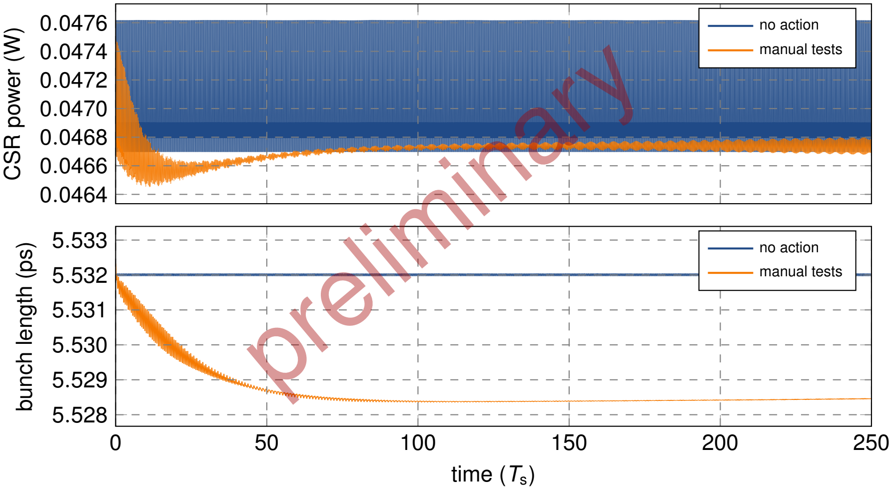
Gym Keras TensorFlow

NumPy PyTorch

- (so far) testing and prototyping is done on standard Desktop PC
- ideas towards implementation include FPGA and GPU systems (cooperation with IPE)

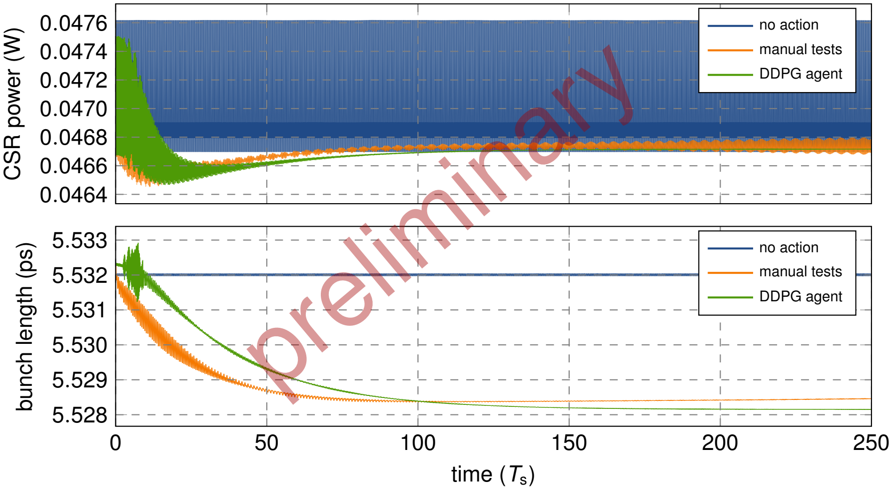
RL based Control of Micro-Bunching

First Results using the DDPG Algorithm



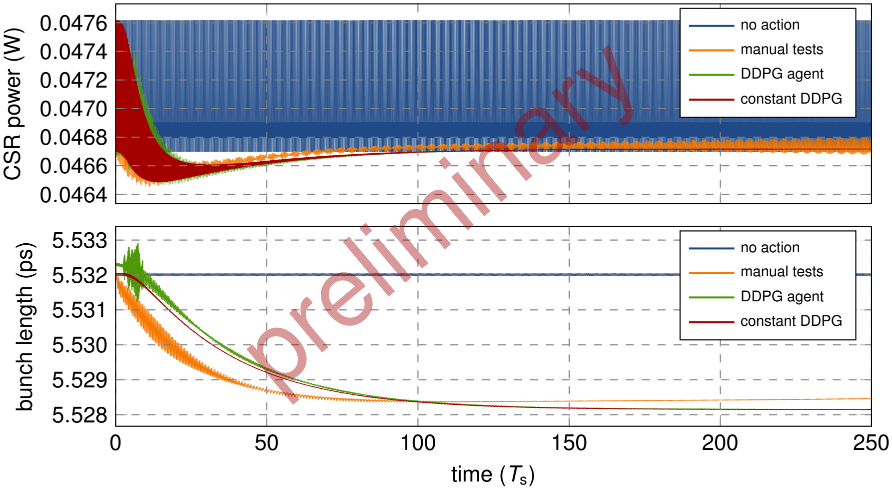
RL based Control of Micro-Bunching

First Results using the DDPG Algorithm



RL based Control of Micro-Bunching

First Results using the DDPG Algorithm



Summary and Outlook

Open Questions

- early results indicate stationary optimization might be sufficient, i.e. finding and repeating the best action (*multi-armed bandit problem*)

however, ...

- action is expected to depend on the state (CSR self-interaction)
- different bunch currents, machine settings and reward functions need to be explored $\Rightarrow$ is control possible with a singular agent?
- what happens in a noisy environment, i.e. the real accelerator?

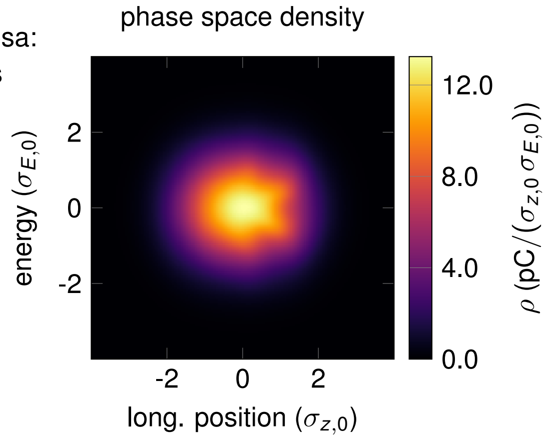
- definition of observation / state (retaining the Markov property)
- choice of action space $\Rightarrow$ what influences the micro-bunching?
- transferability to different control tasks / instabilities?

Thank you for your attention!

Backup

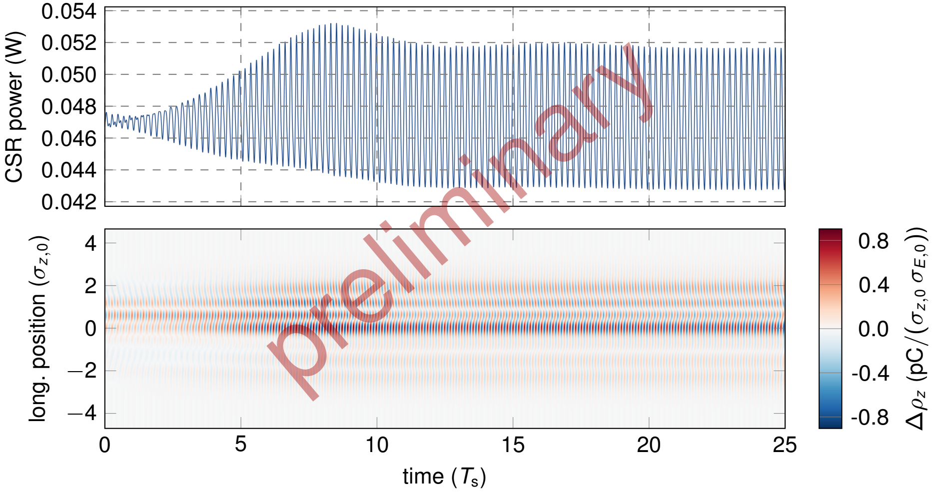
Markov Property and Phase Space

- Markov property in Inovesa: configuration parameters and initial phase space density determine results (*true state*)
- not just a feature of VFP solvers, but something that's rooted in the definition of phase spaces



Backup

Effects of RF Amplitude Modulation ($A_V = 0.2\ V_0$, $f_V = 4.78\ f_s$)



Backup

Effects of RF Amplitude Modulation ($A_V = 0.2\ V_0$, $f_V = 5\ f_s$)