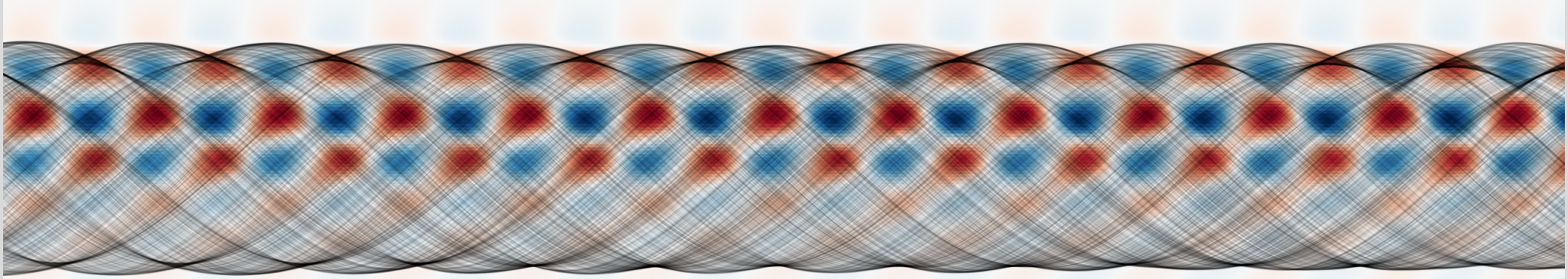
Karlsruhe Institute of Technology

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

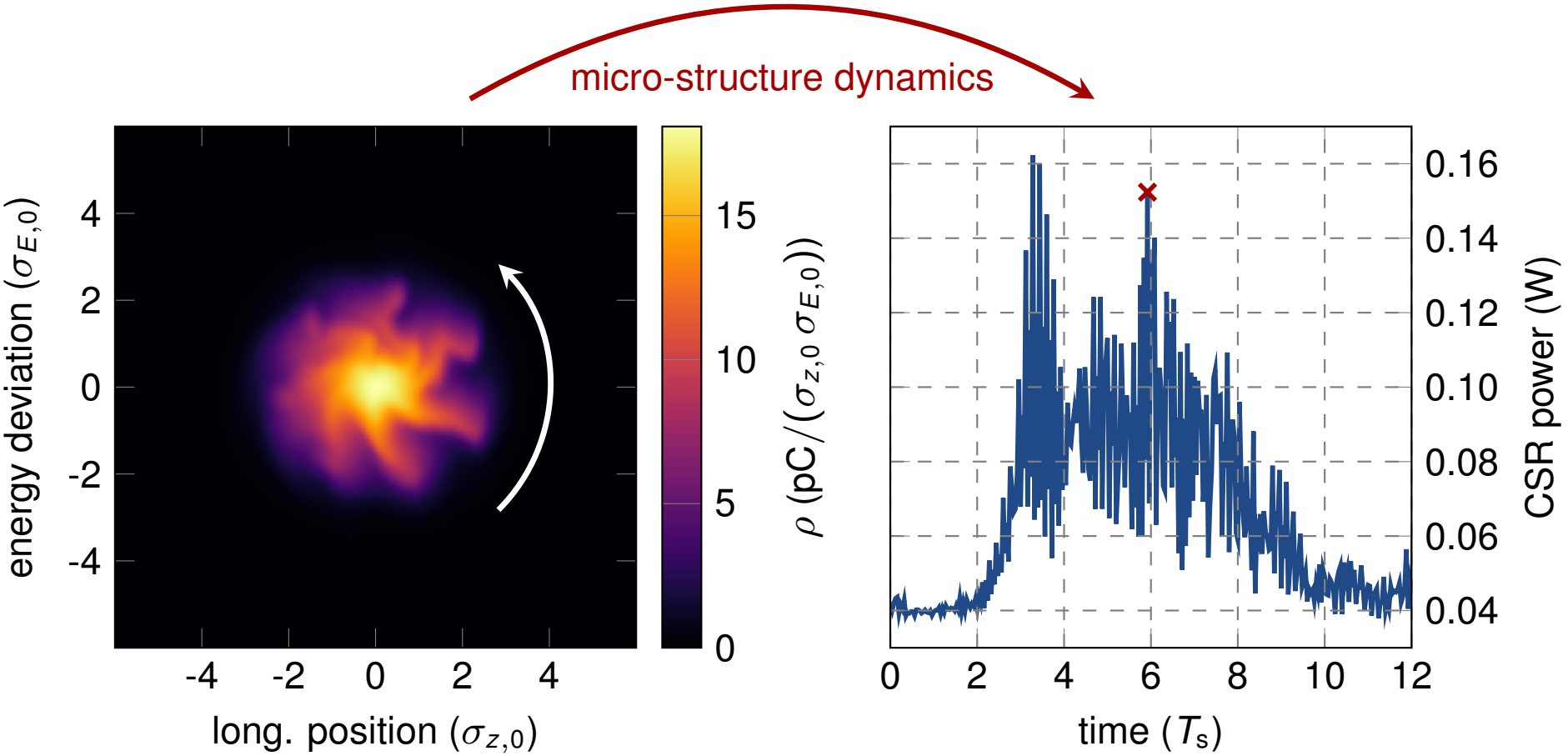
Tobias Boltz, Miriam Brosi, Erik Bründermann, Bastian Härer, Peter Kaiser, Christoph Pohl, Patrick Schreiber, Minjie Yan, Tamim Asfour and Anke-Susanne Müller | September 26, 2019

Laboratory for Applications of Synchrotron Radiation (LAS)



Micro-Bunching Instability

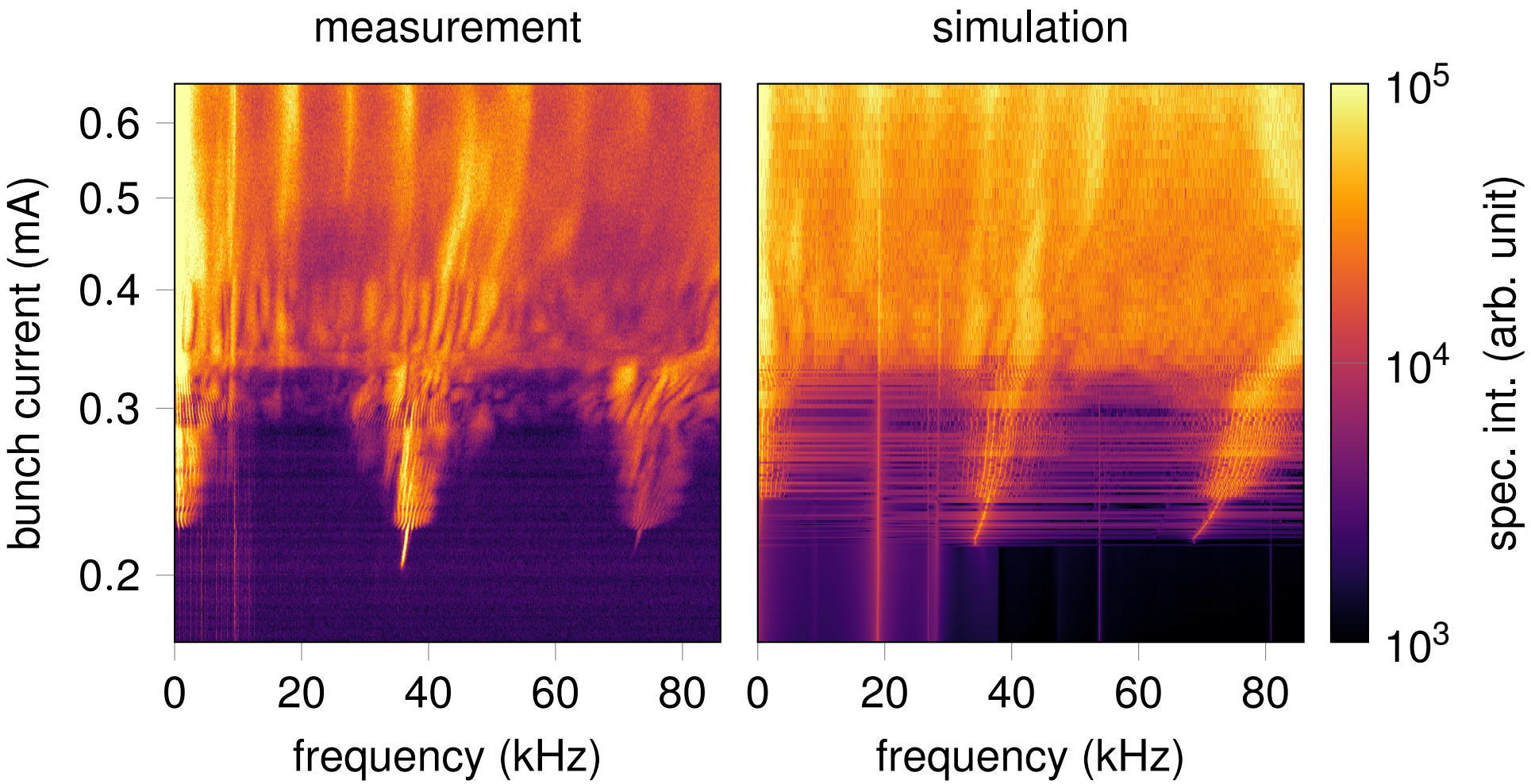
Micro-Bunching and CSR Power Fluctuations



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

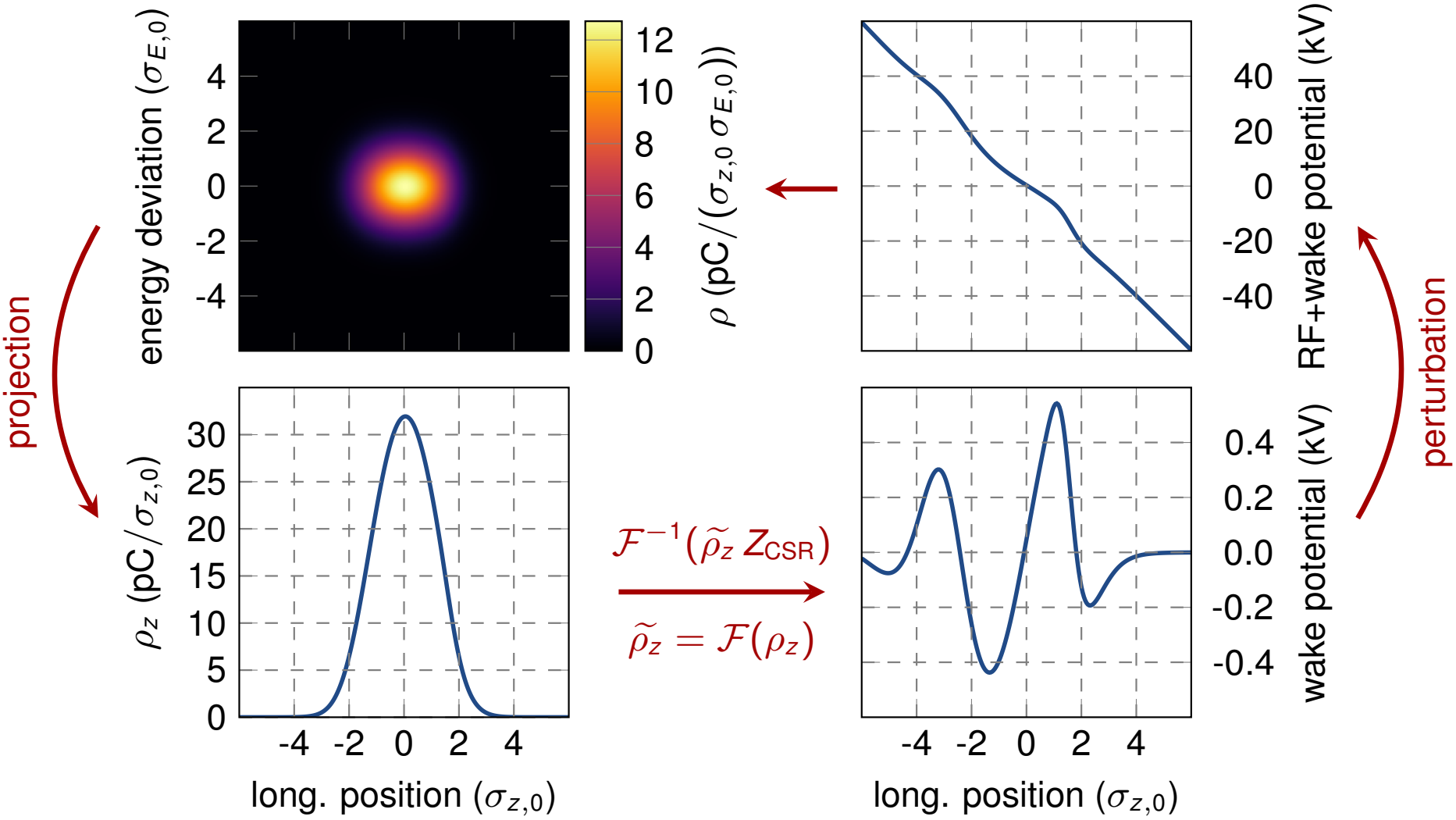
Micro-Bunching Instability

CSR Power Spectrogram: Dependency on Bunch Current



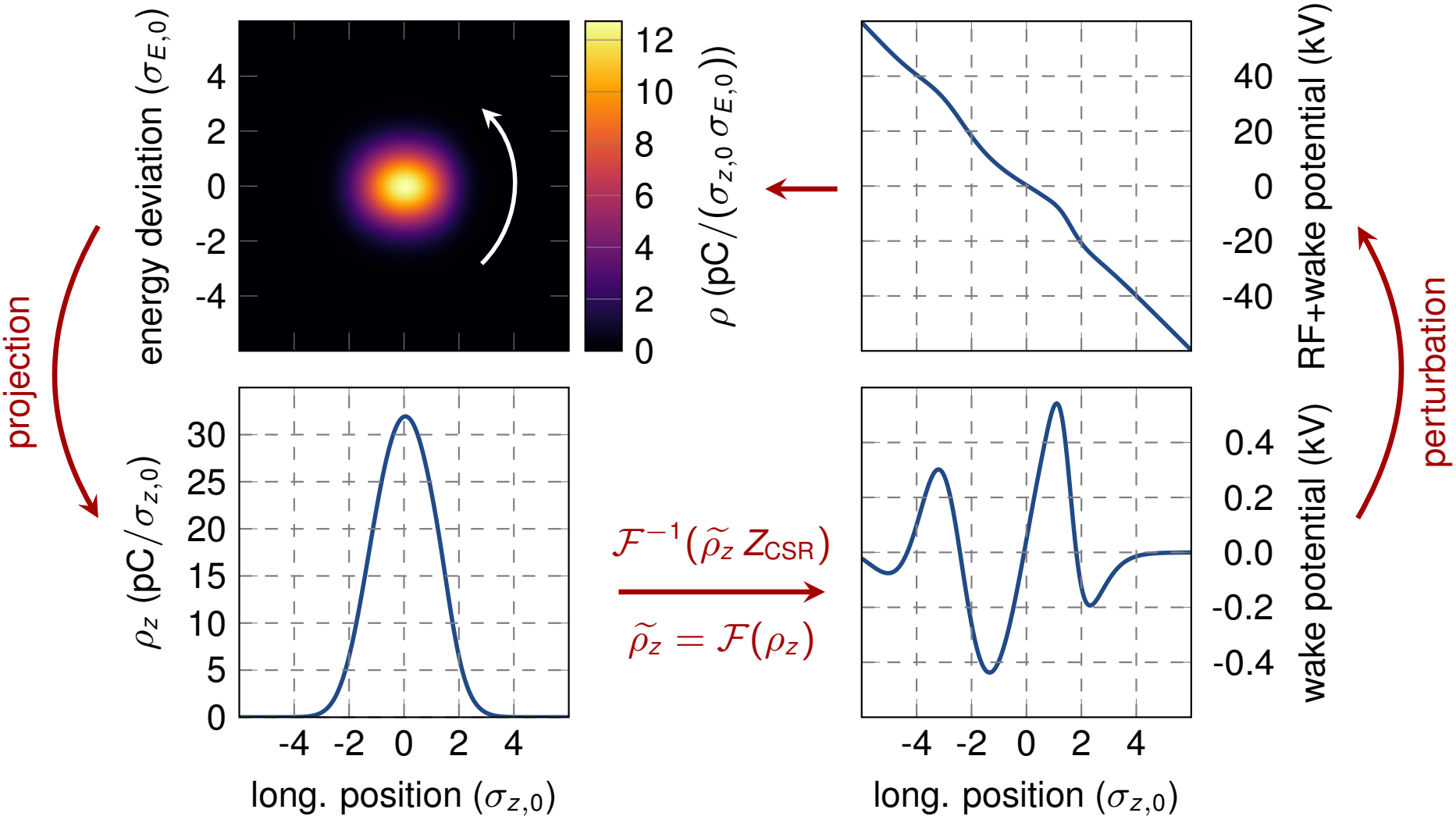
Micro-Bunching Instability

CSR self-interaction



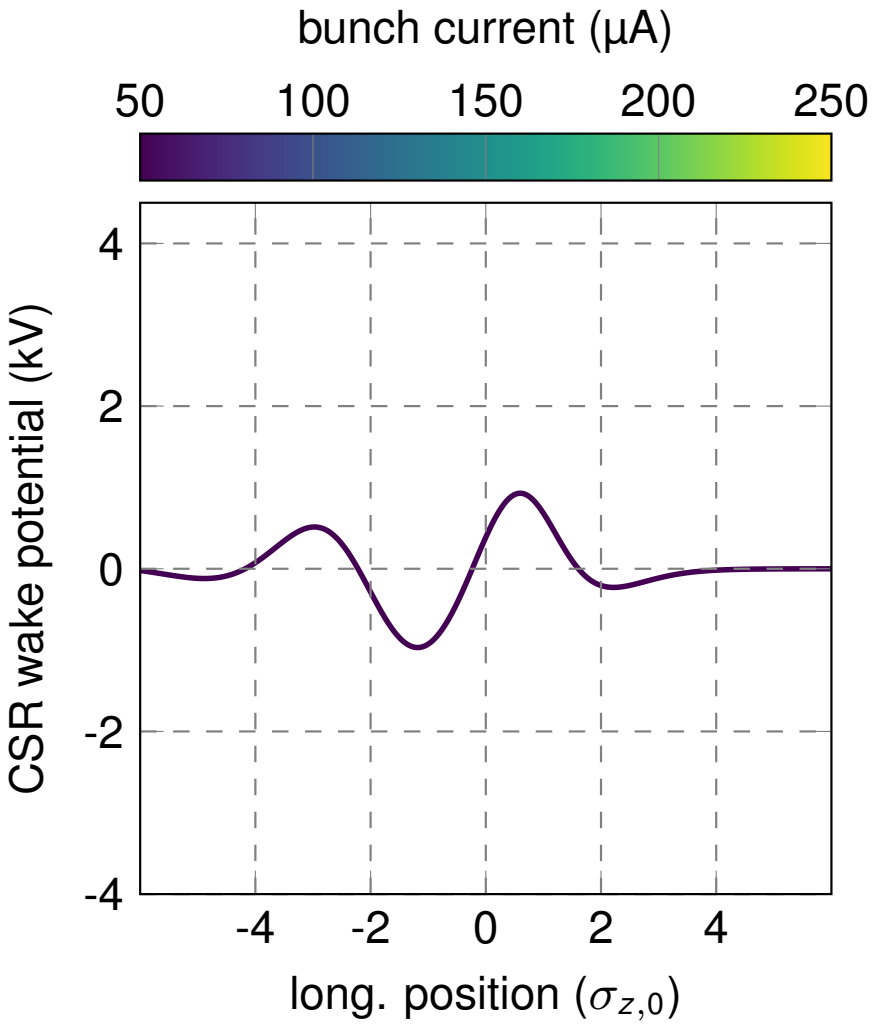
Micro-Bunching Instability

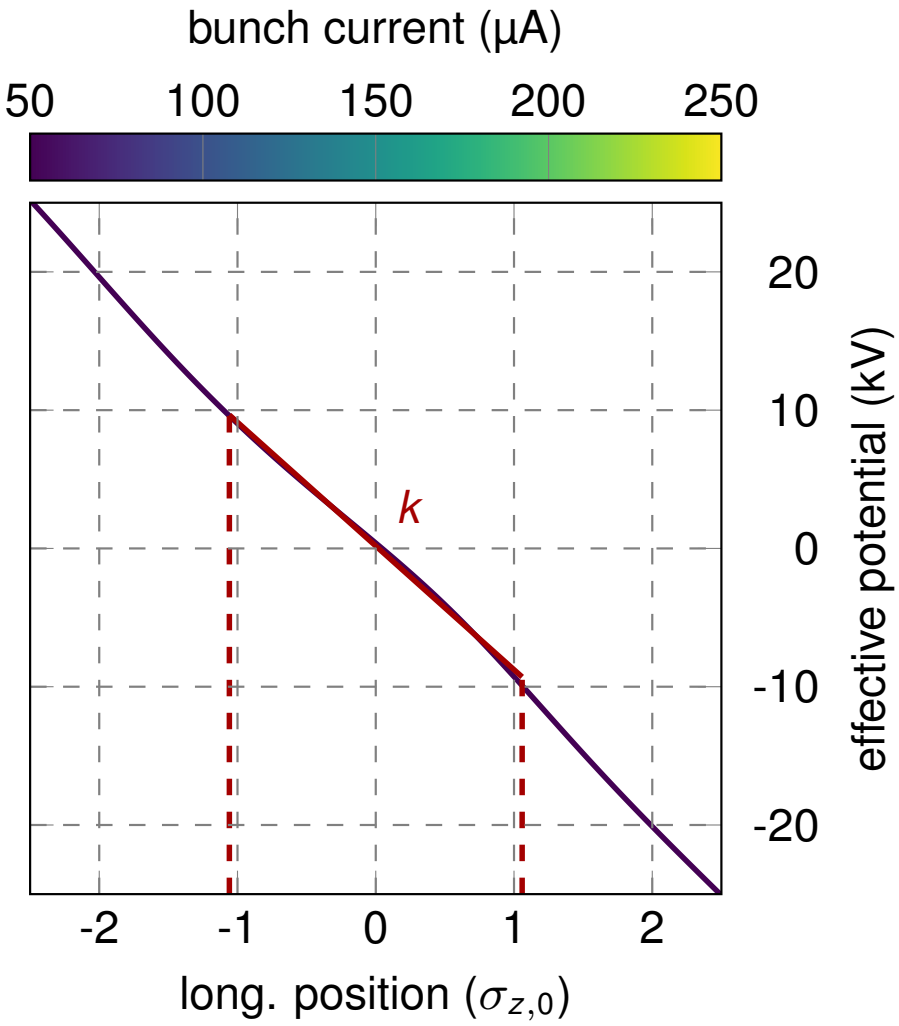
CSR self-interaction



Single Particle Synchrotron Motion

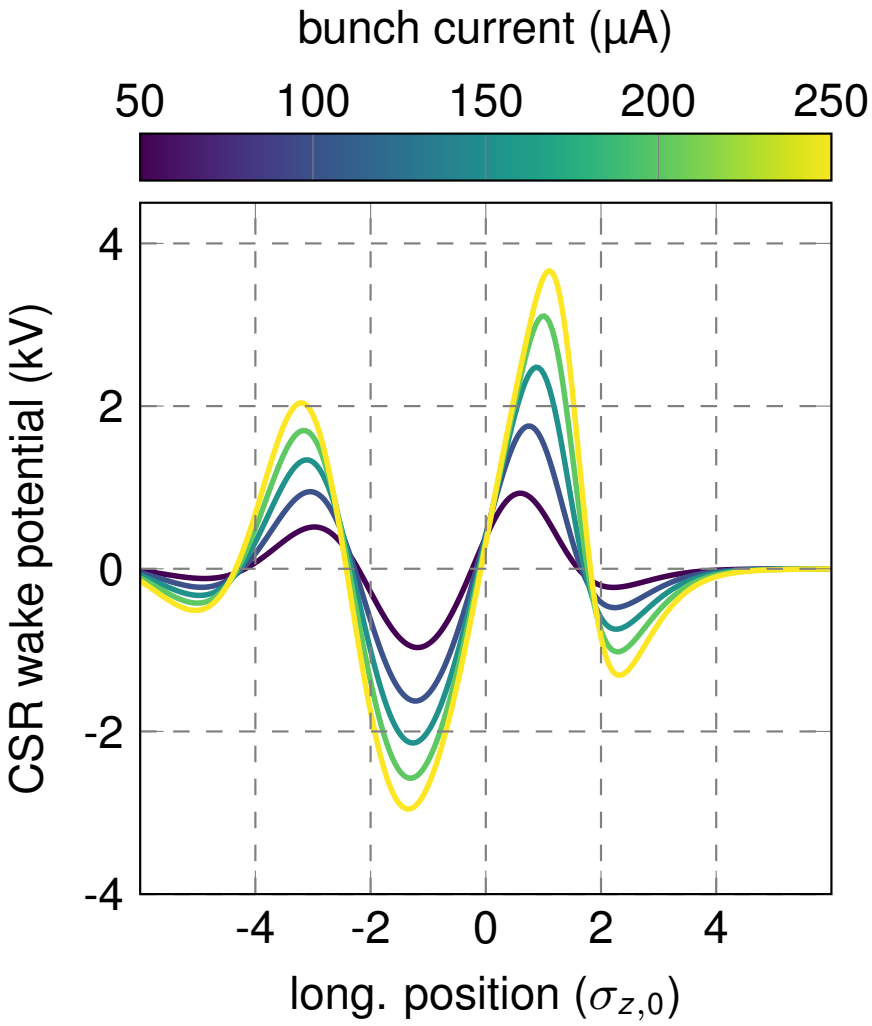
Perturbation of the Restoring Force

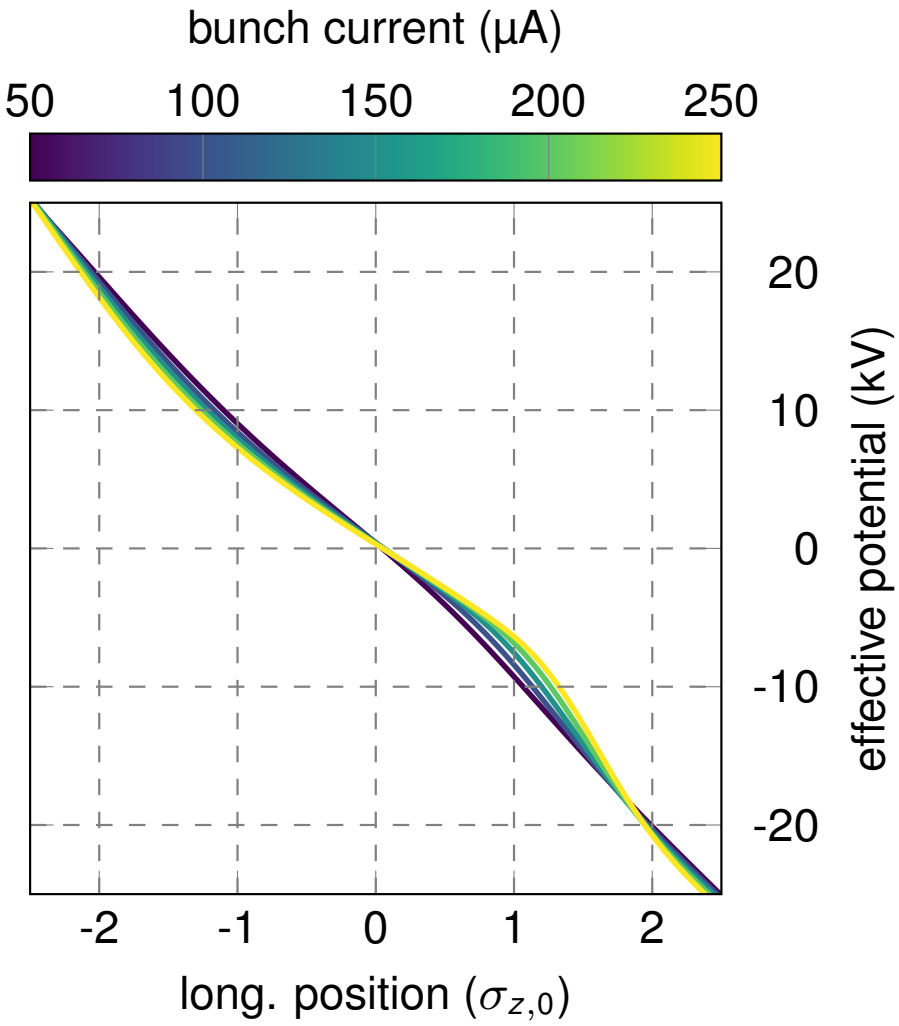




Single Particle Synchrotron Motion

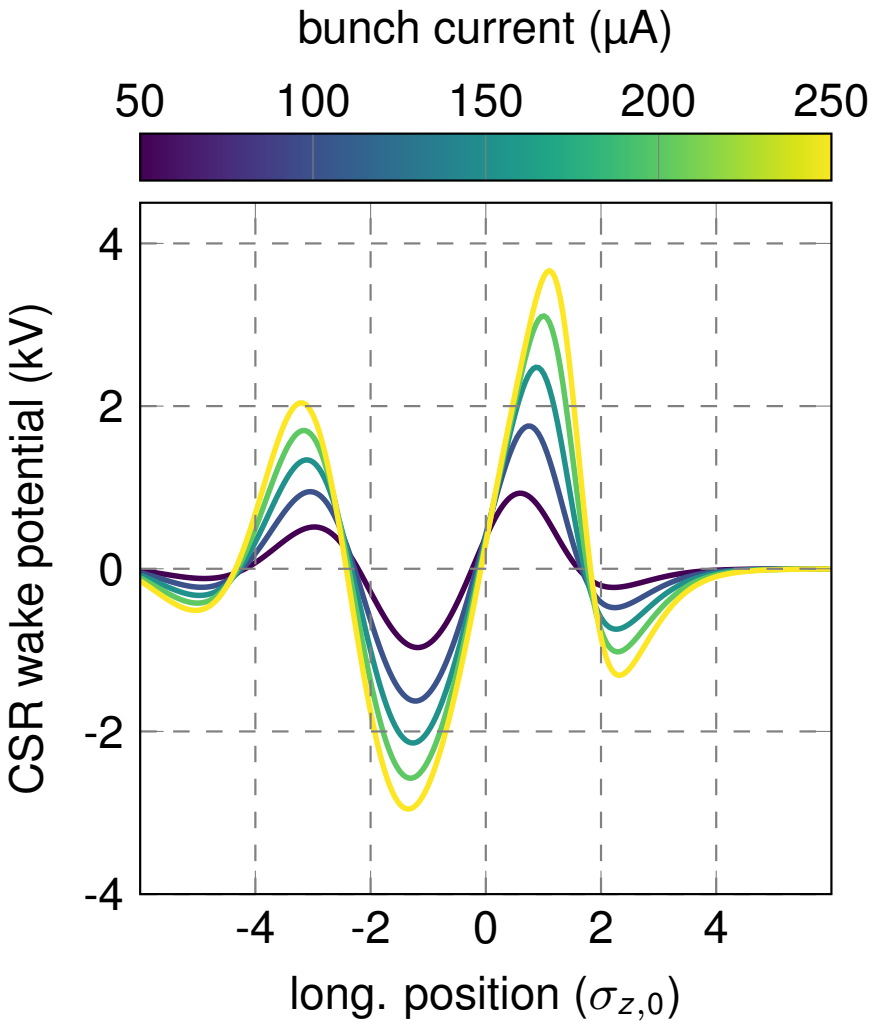
Perturbation of the Restoring Force

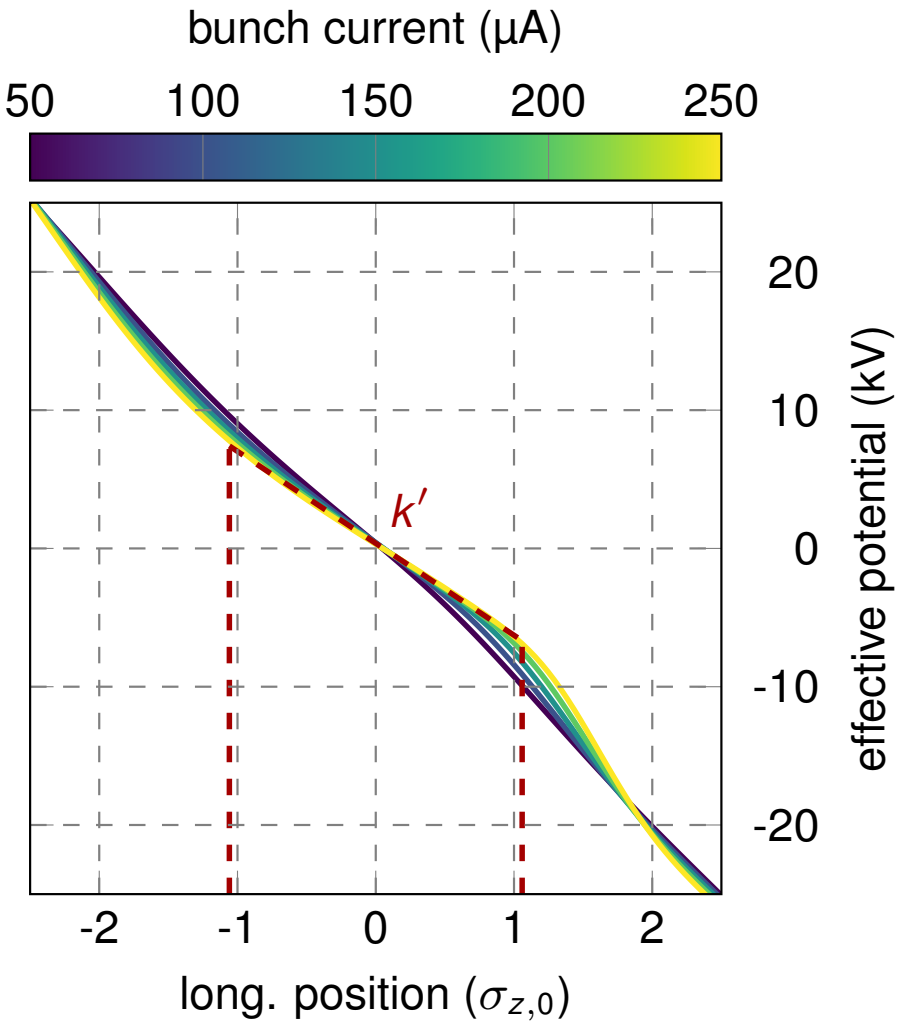




Single Particle Synchrotron Motion

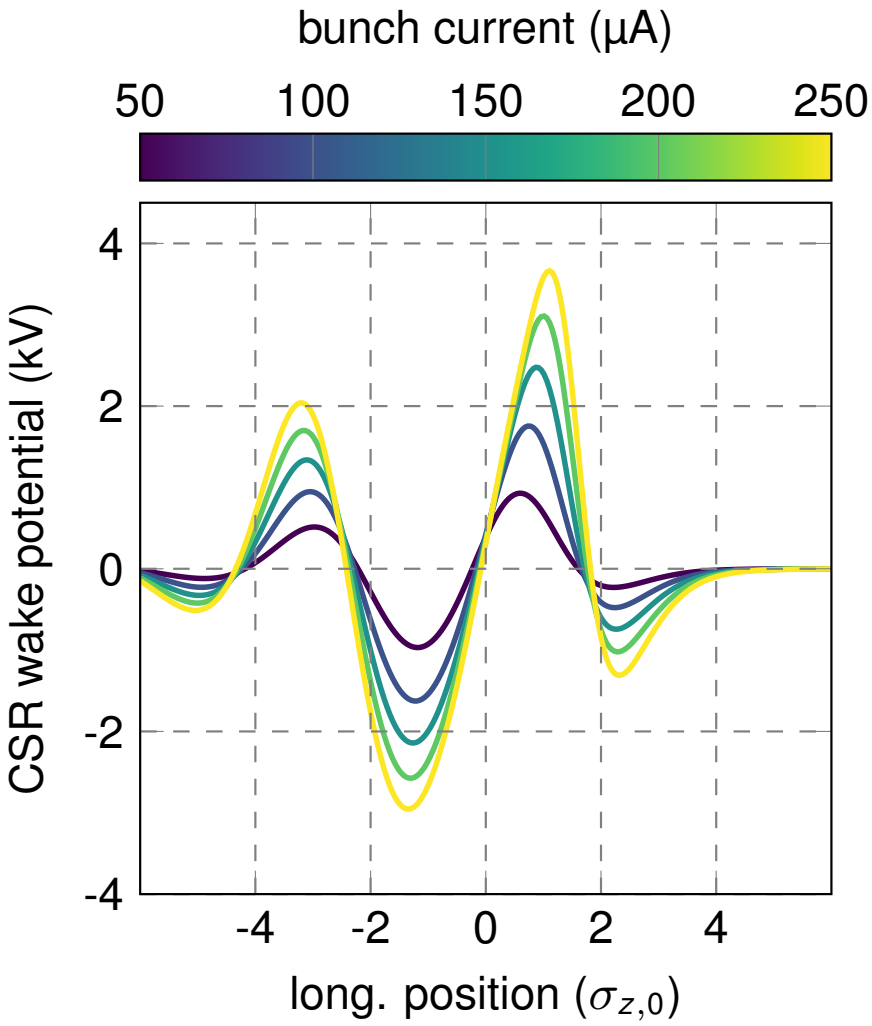
Perturbation of the Restoring Force

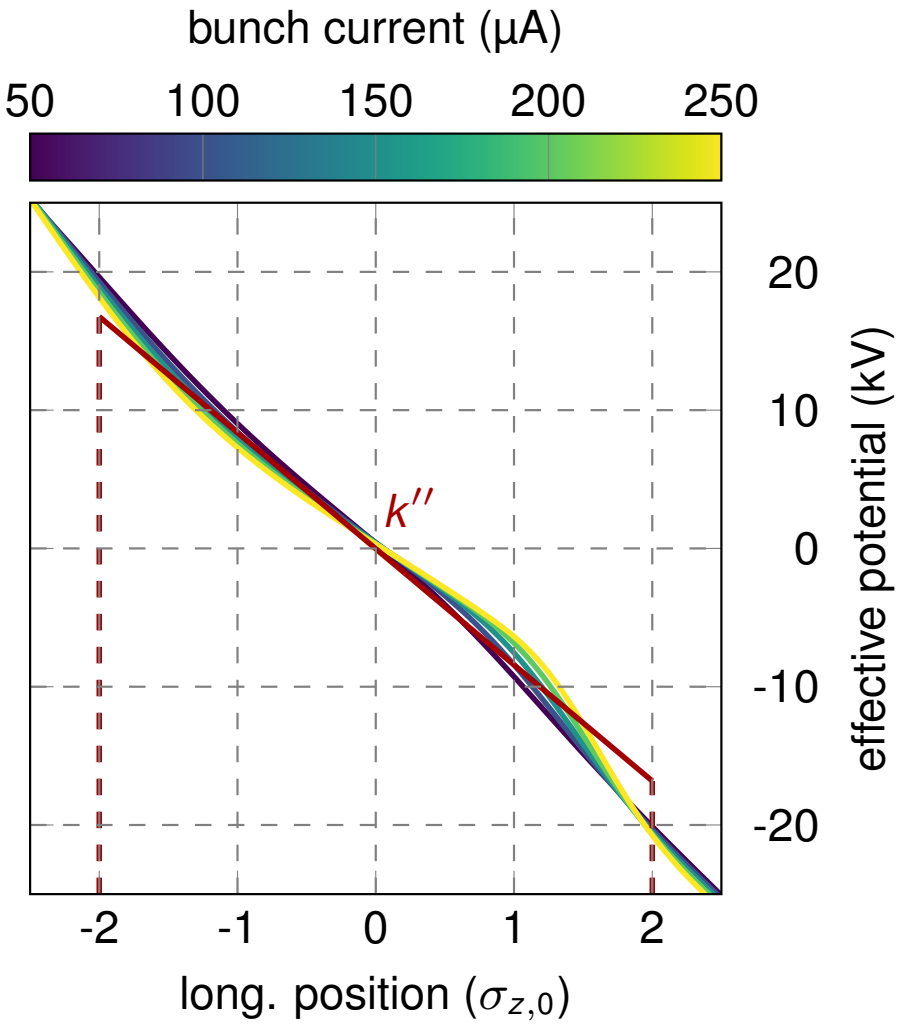




Single Particle Synchrotron Motion

Perturbation of the Restoring Force



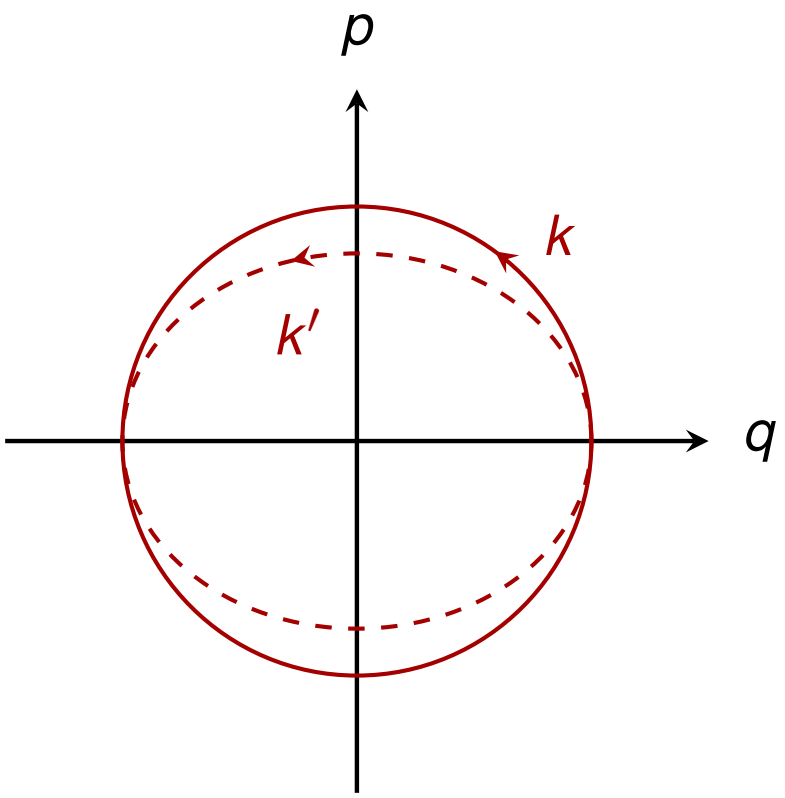


Single Particle Synchrotron Motion

1D-Harmonic Oscillator with perturbed Restoring Force

- RF potential acts as a linear restoring force

  $\Rightarrow$ 1D-harmonic oscillator
  $\Rightarrow$ particle motion in generalized coordinates is perfectly circular
- CSR wake potential introduces position-dependent perturbation that affects the strength of the restoring force $k' < k$
- particle motion in generalized coordinates becomes elliptical



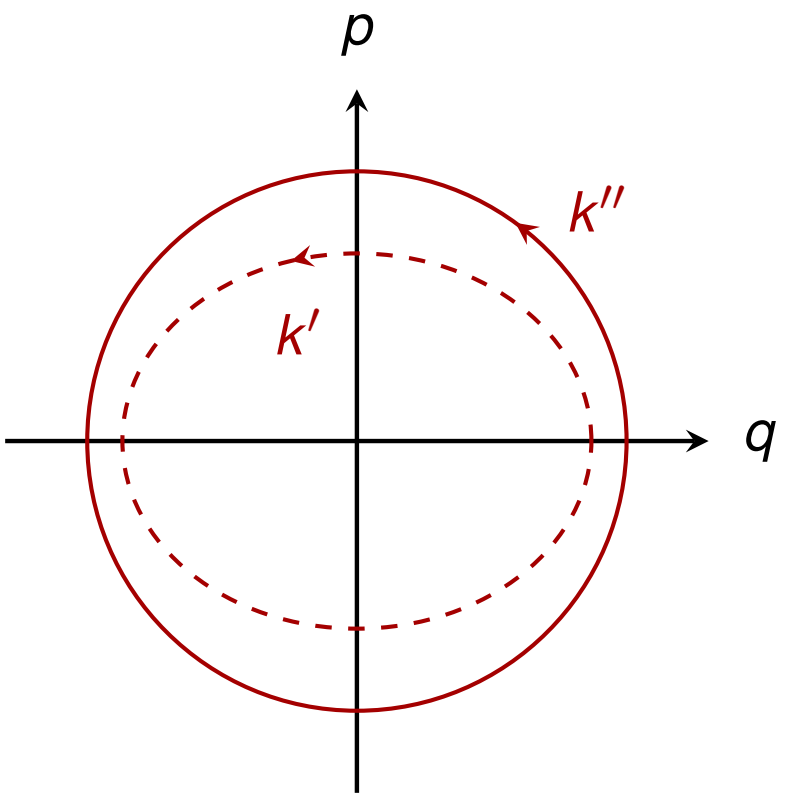
Single Particle Synchrotron Motion

1D-Harmonic Oscillator with perturbed Restoring Force

- RF potential acts as a linear restoring force

  $\Rightarrow$ 1D-harmonic oscillator

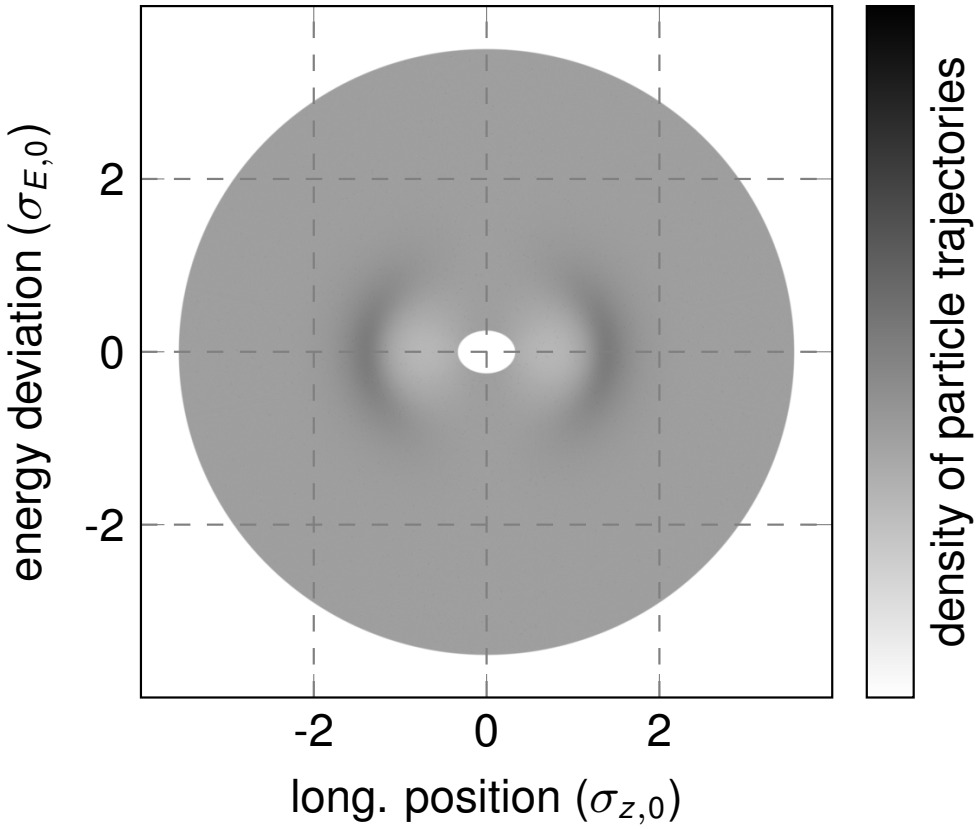
  $\Rightarrow$ particle motion in generalized coordinates is perfectly circular
- CSR wake potential introduces position-dependent perturbation that affects the strength of the restoring force $k' < k$
- particle motion in generalized coordinates becomes elliptical



Single Particle Synchrotron Motion

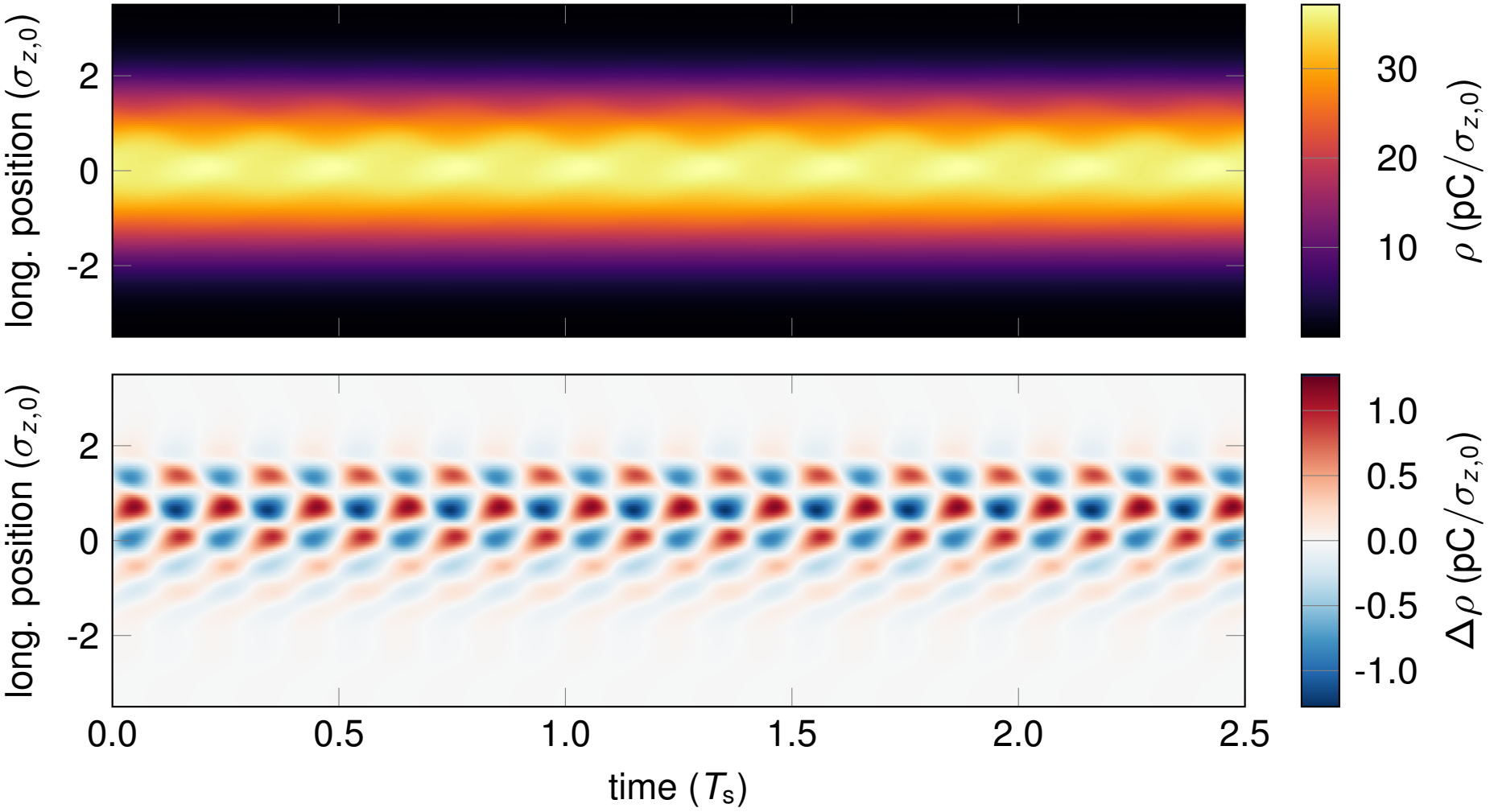
Elliptical Deformation of Trajectories below Threshold

- visualization of particle trajectories with uniformly distributed energies
- position-dependent elliptical deformation of particle trajectories
  $\Rightarrow$ local charge densities
- quadrupole-like mode below the instability threshold



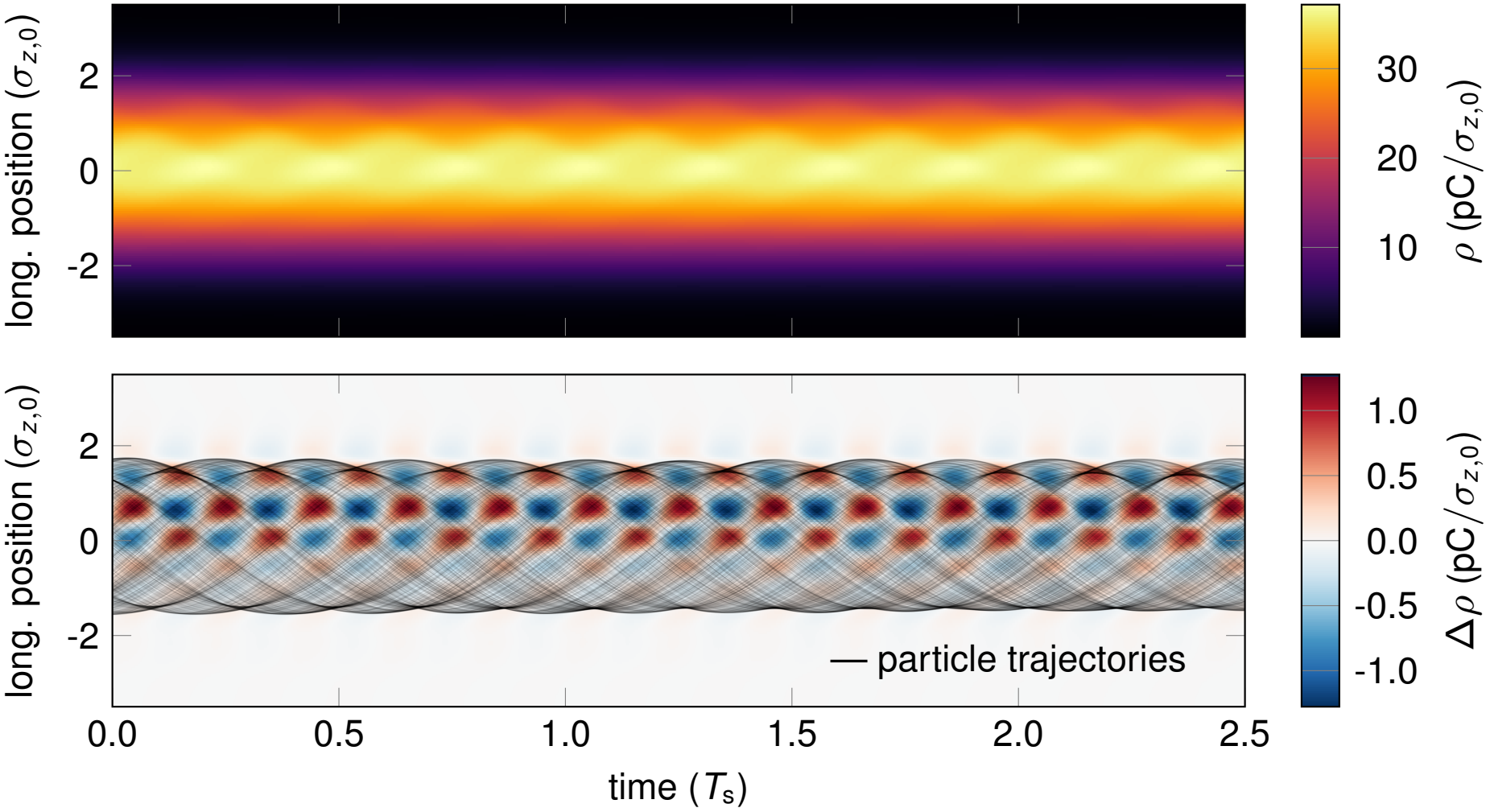
Single Particle Synchrotron Motion

Elliptical Deformation of Trajectories above Threshold



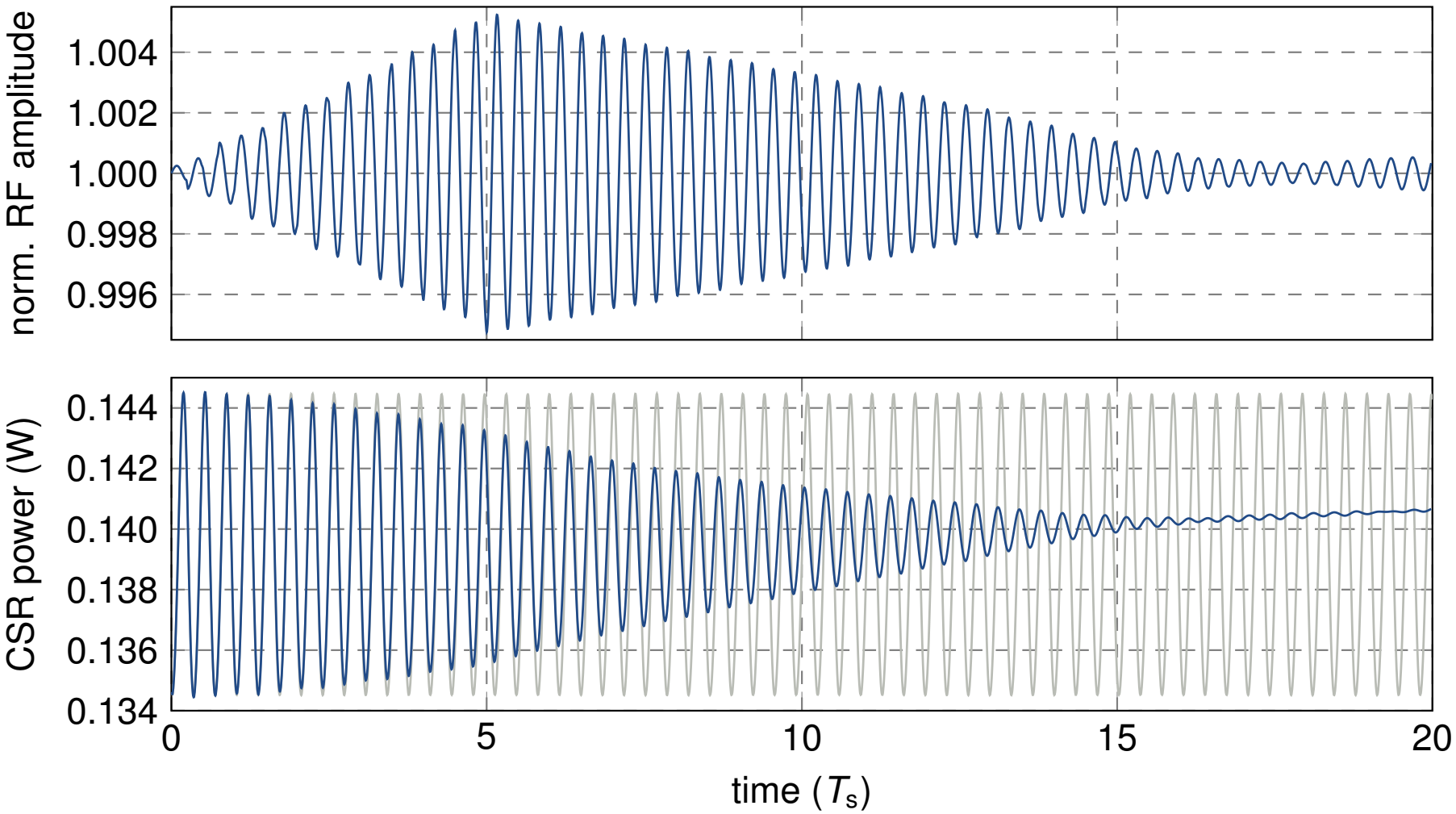
Single Particle Synchrotron Motion

Elliptical Deformation of Trajectories above Threshold



Micro-Bunching Control

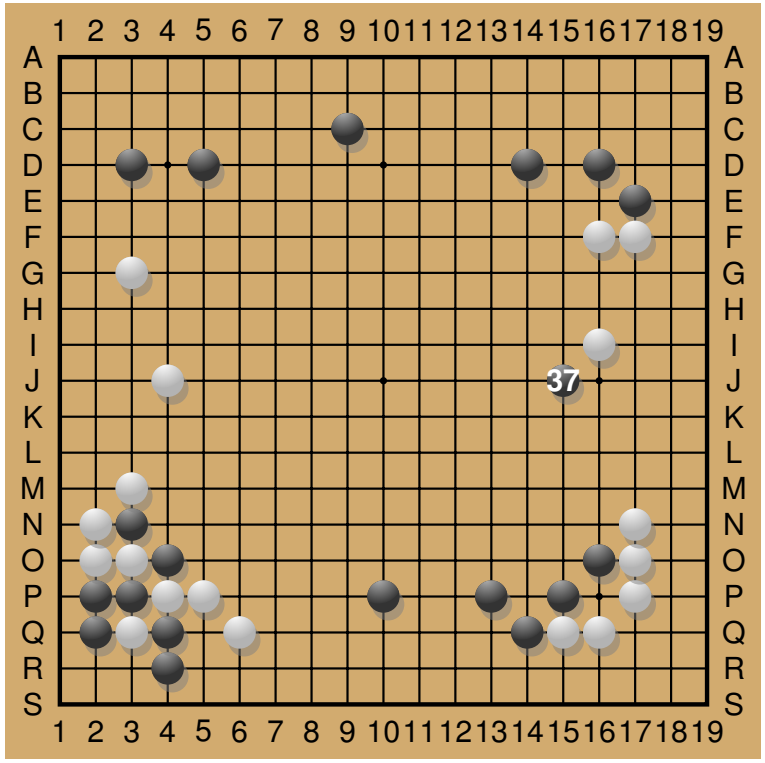
. . . via Dynamic RF Amplitude Modulation



Reinforcement Learning

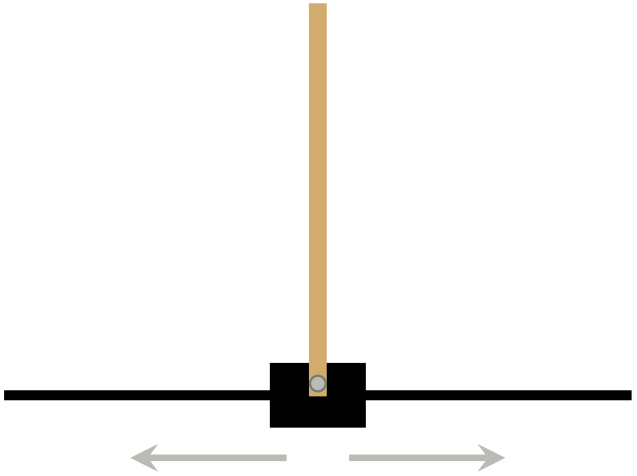
Developing some Intuition . . .

- goal-directed learning from interaction (trial-and-error search)
- finding a sequence of actions (e.g. moves in a game)
- taken actions affect the following states (e.g. board positions)



cartpole balancing

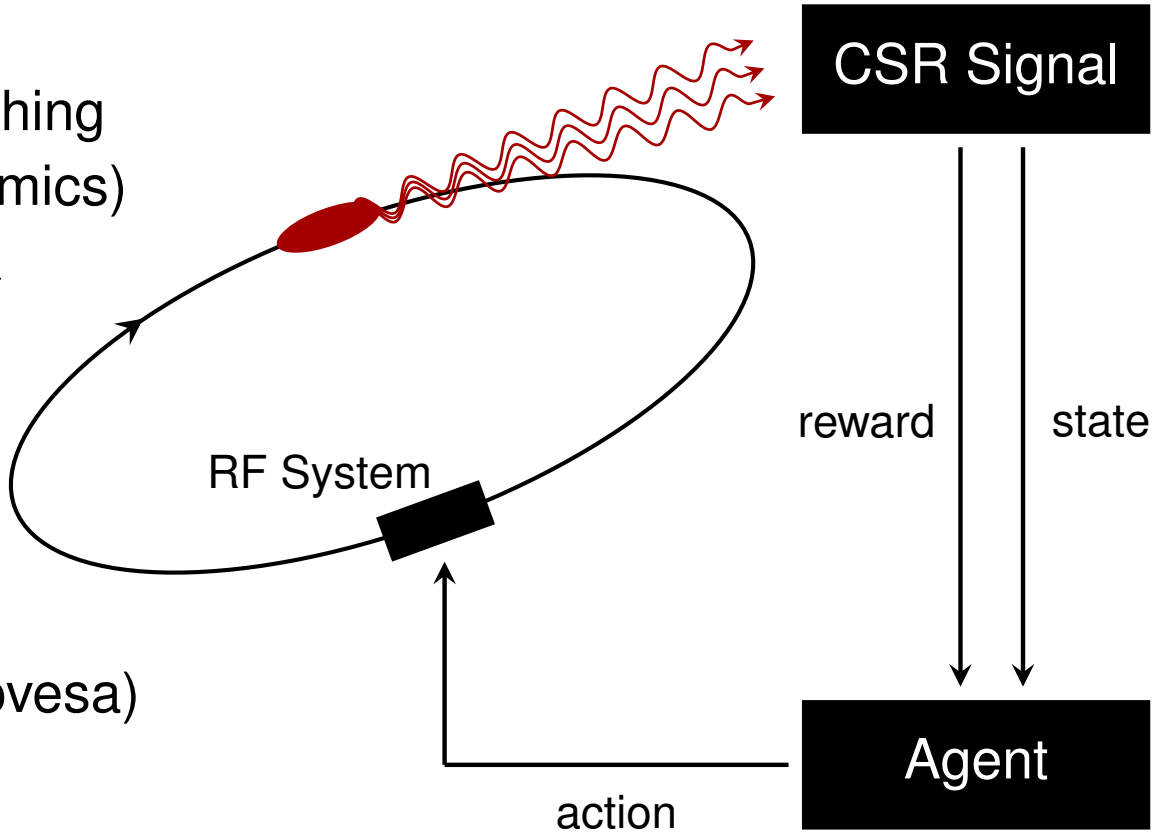
(a textbook RL problem)



RL based Control of Micro-Bunching

Feedback Scheme

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



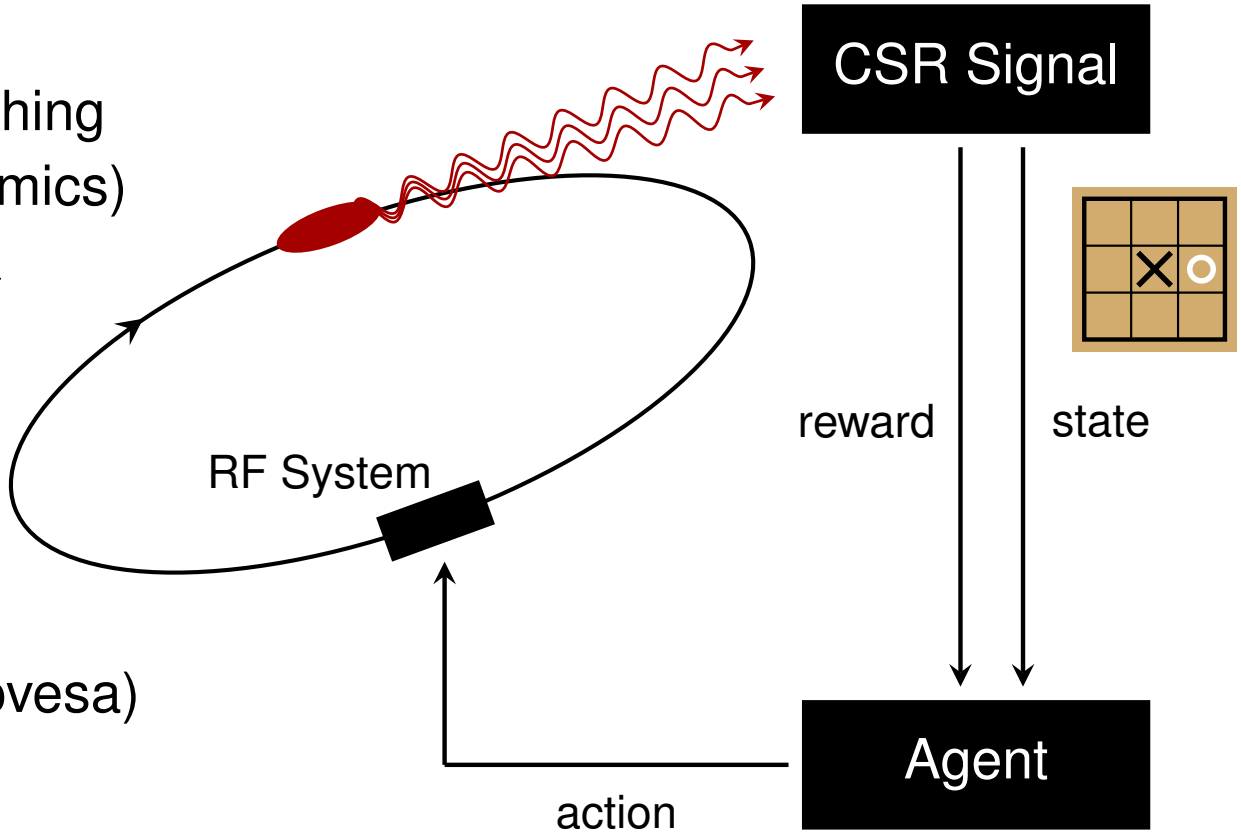
proof of principle:
control in simulation (Inovesa)

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

RL based Control of Micro-Bunching

Feedback Scheme

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



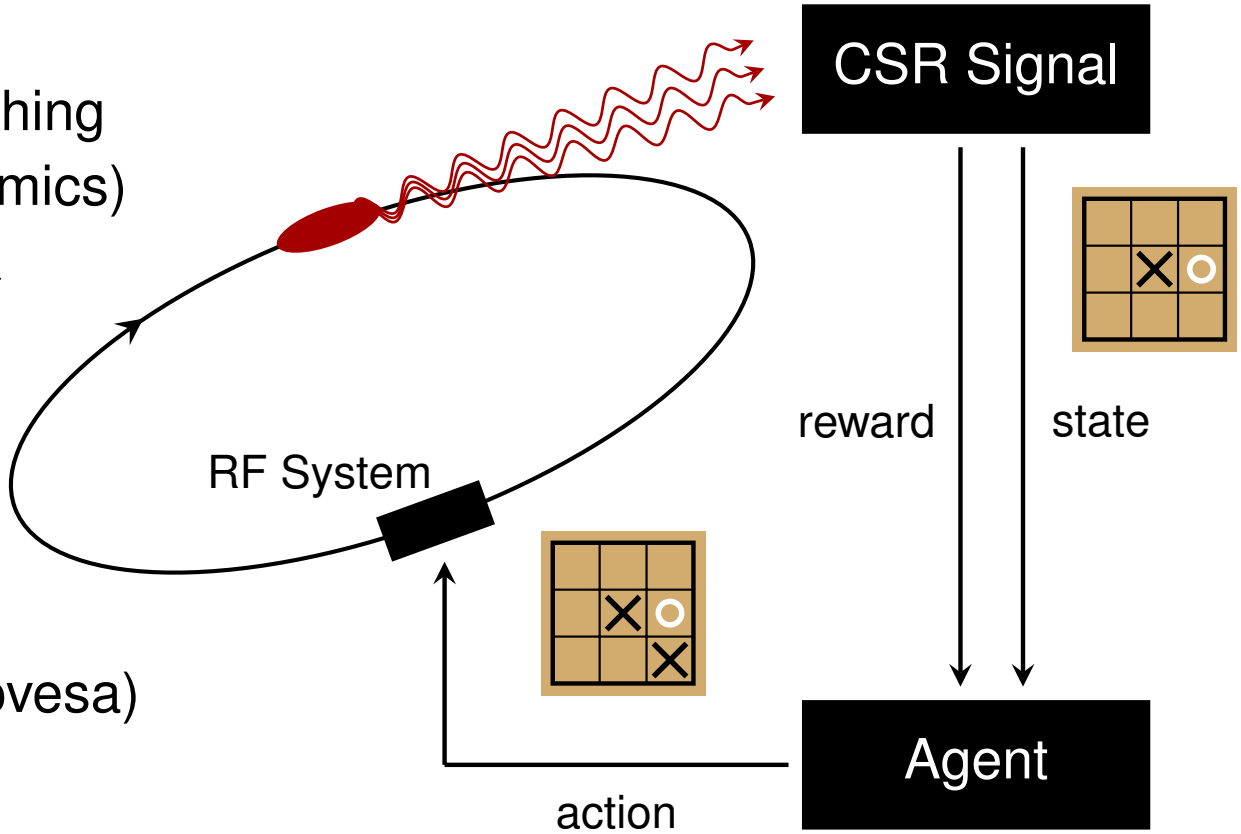
proof of principle:
control in simulation (Inovesa)

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

RL based Control of Micro-Bunching

Feedback Scheme

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: state of electron bunch / micro-bunching dynamics

  1) in theory/simulations: longitudinal charge distribution

  2) measurable: CSR power signal $\Rightarrow$ feature vector encoding the state

- reward function: optimization of emitted CSR signal

  $R = w_1 \mu_{\mathrm{CSR}} - w_2 \sigma_{\mathrm{CSR}}$ with $w_1, w_2 > 0$

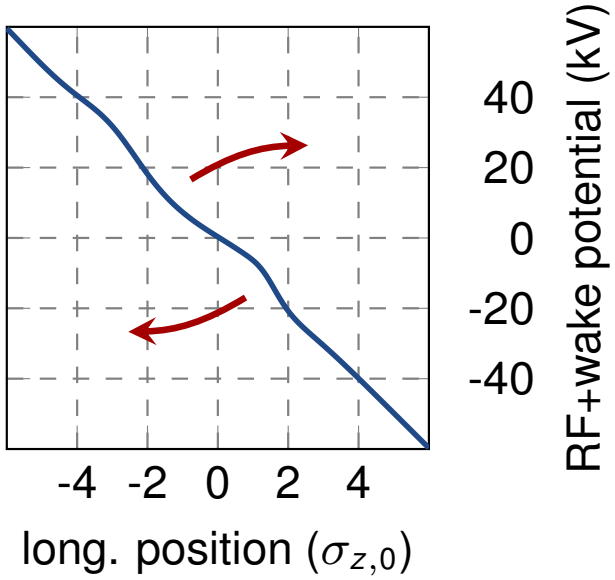
  $\Rightarrow$ maximize μ_{CSR}, minimize σ_{CSR}!

- action: RF amplitude modulation

  dynamically adjust amplitude & frequency

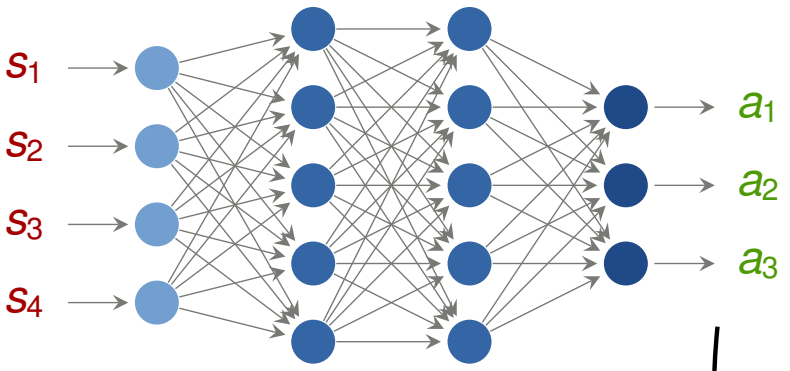
  $\Rightarrow$ counteract perturbation by CSR

  $\Rightarrow$ include RF phase?



RL based Control of Micro-Bunching

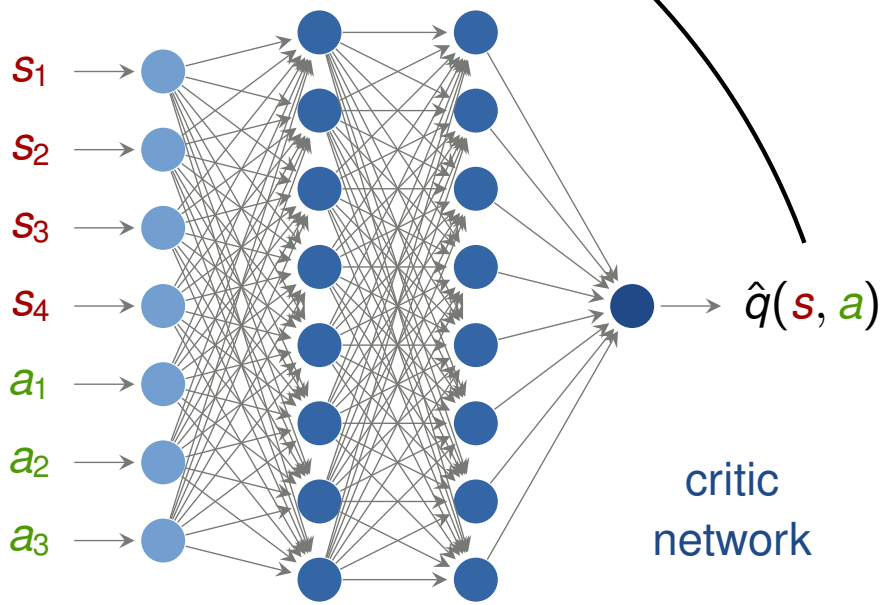
. . . behind the curtain: Actor-Critic System using NNs



actor network

chosen action to be evaluated

evaluation and update based on estimation of expected reward $\hat{q}(s, a)$



critic network

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

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`

Summary

- goal: control micro-bunching $\Rightarrow$ optimize CSR emission
- CSR self-interaction: perturbation is dependent on the state of the micro-bunching dynamics
  $\Rightarrow$ countermeasure should be state-dependent as well!
- CSR wake potential causes perturbation of effective restoring force
  $\Rightarrow$ compensate via dynamic RF amplitude modulation scheme
- interaction with the bunch changes the micro-bunching dynamics
  $\Rightarrow$ sequence of actions is required *(deal with consequences)*

Further Questions

- dependency on machine settings / current $\Rightarrow$ control in one system?
- general-purpose approach $\Rightarrow$ transferability to other control tasks?

Thank you for your attention!

Backup

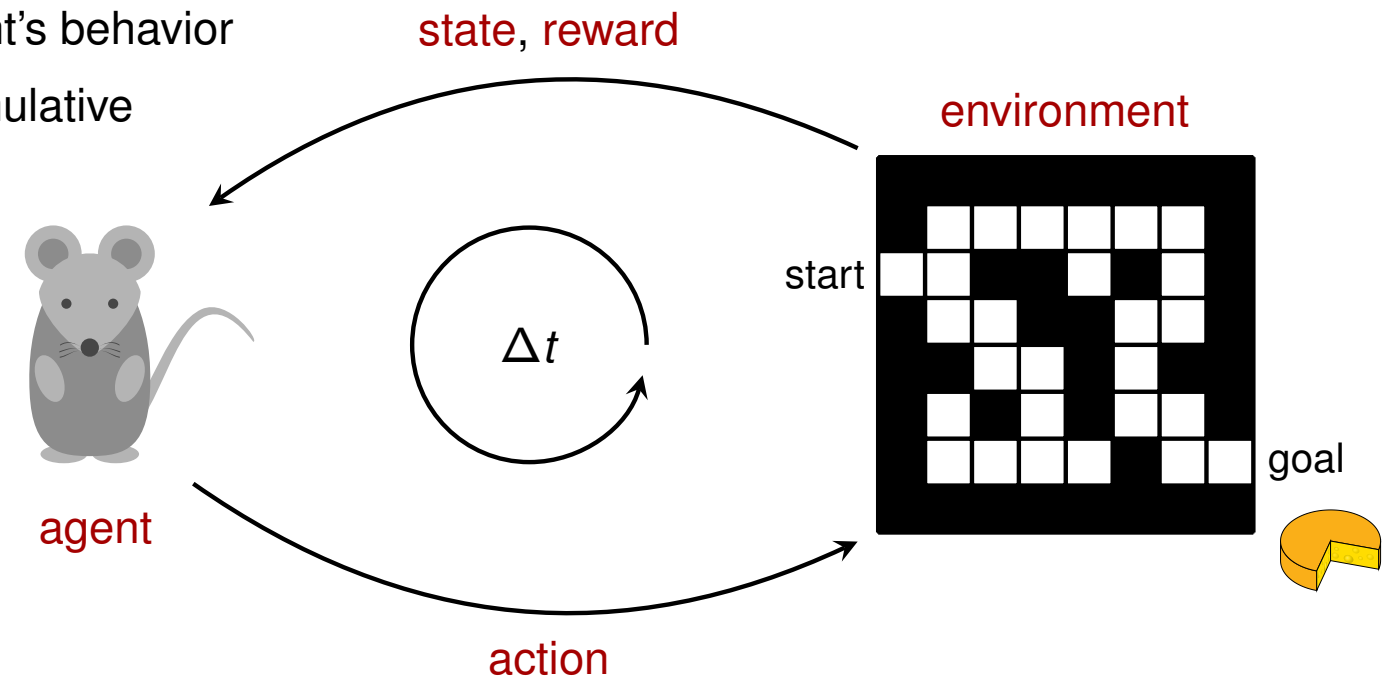
RL in a Nutshell: Learning from Interaction

- mathematical foundation: Markov decision process (MDP)

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

policy: defines agent's behavior

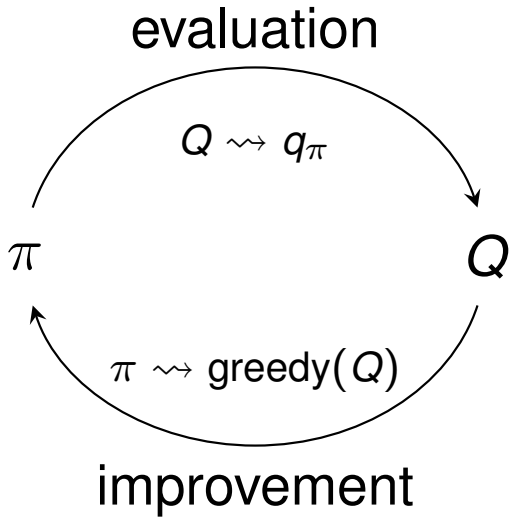
goal: maximize cumulative reward



Backup

RL 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



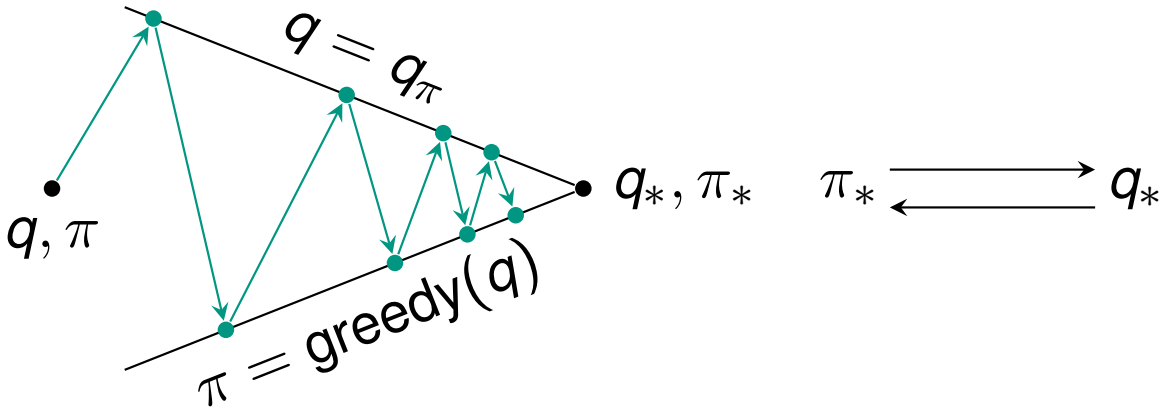


Figure: adapted from *Reinforcement Learning*, Sutton, R.S. and Barto, A.G., 2nd edition, MIT Press (November 2018)