ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPGW017

FEEDBACK DESIGN FOR CONTROL OF THE MICRO-BUNCHING INSTABILITY BASED ON REINFORCEMENT LEARNING

T. Boltz*, M. Brosi, E. Bründermann, B. Haerer, P. Kaiser, C. Pohl, P. Schreiber, M. Yan, T. Asfour, A.-S. Müller
Karlsruhe Institute of Technology, Karlsruhe, Germany

Abstract

The operation of ring-based synchrotron light sources with short electron bunches increases the emission of coherent synchrotron radiation (CSR) in the THz frequency range. However, the micro-bunching instability resulting from self-interaction of the bunch with its own radiation field limits stable operation with constant intensity of CSR emission to a particular threshold current. Above this threshold, the longitudinal charge distribution and thus the emitted radiation vary rapidly and continuously. Therefore, a fast and adaptive feedback system is the appropriate approach to stabilize the dynamics and to overcome the limitations given by the instability. In this contribution, we discuss first efforts towards a longitudinal feedback design that acts on the RF system of the KIT storage ring KARA (Karlsruhe Research Accelerator) and aims for stabilization of the emitted THz radiation. Our approach is based on methods of adaptive control that were developed in the field of reinforcement learning and have seen great success in other fields of research over the past decade. We motivate this particular approach and comment on different aspects of its implementation.

MICRO-BUNCHING INSTABILITY

Self-interaction of short electron bunches with their own radiation field can have a significant impact on the longitudinal beam dynamics in a storage ring. Above a given threshold current, this leads to dynamically changing micro-structures in the longitudinal charge distribution and thus to fluctuating CSR emission (illustrated in Fig. 1). This phenomenon is commonly referred to as micro-bunching or micro-wave instability. The CSR self-interaction, as its driving force, is conveniently described by the resulting wake potential

$$V_{\mathrm{CSR}}(q) = \int_{-\infty}^{\infty} \widetilde{\rho}(\omega) Z_{\mathrm{CSR}}(\omega) e^{i\omega q} \mathrm{d}\omega \,, \qquad (1)$$

where $q = (z - z_s)/\sigma_{z,0}$ denotes the generalized longitudinal position, $\widetilde{\rho}(\omega)$ the Fourier-transformed longitudinal bunch profile and $Z_{\mathrm{CSR}}(\omega)$ the CSR-induced impedance of the storage ring. This additional potential acts as a perturbation to the accelerating RF potential and as such, influences the temporal evolution of the longitudinal phase space density. As the charge distribution in phase space varies, so may the longitudinal bunch profile $\rho(q)$ as its projection, which in turn causes changes in the CSR wake potential $V_{\mathrm{CSR}}(q)$. This dynamic process can be simulated numerically using the Vlasov-Fokker-Planck (VFP) solver Inovesa [1], which has shown great qualitative agreement with measurements at the KIT storage ring KARA [2].

Previous efforts towards the control of the micro-bunching instability have mainly been focused on suppression of the CSR self-interaction by making adjustments to the impedance budget of the storage ring, e.g, [3, 4]. More recently, a linear RF feedback has been used to influence the CSR bursting pattern at SOLEIL [5]. Depending on the application, the formation of micro-structures on the longitudinal charge distribution can also be desirable as it leads to the emission of CSR at higher frequencies, reaching up to the low THz range. Extensive control over the longitudinal beam dynamics would thus provide the opportunity of optimizing the emitted CSR for each application individually. Given the nature of the instability, such efforts require a fast and adaptive feedback in order to deal with the dynamic variation of the CSR wake potential on time scales comparable to the synchrotron period.

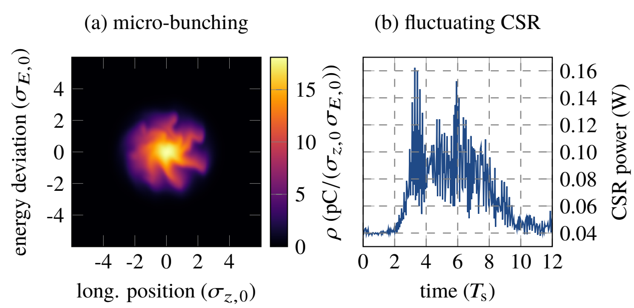

Figure 1: (a) The CSR self-interaction of the bunch causes the formation of micro-structures in the longitudinal phase space density. (b) Their continuous variation leads to fluctuations in the emitted CSR power. The illustrated dynamics are simulated with the VFP solver Inovesa.

REINFORCEMENT LEARNING

A detailed introduction to the subject can be found in [6], on which the following brief description is based.

Reinforcement learning is the computational approach to goal-directed learning from interaction with an environment. It is different to other sub-fields in machine learning as its learning paradigm does not require a pre-existing data set. Instead, learning takes place in an iterative process based on the general concept of trial-and-error search.

The learner and decision maker, usually called the *agent*, continuously interacts with the *environment* while seeking to

* tobias.boltz@kit.edu

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPGW017

improve its behavior. At every time step, the agent perceives the current situation, the *state* S_t of the environment, and performs an *action* A_t. Based on the chosen action, the agent receives a scalar *reward* R_t and finds itself in a new state. Eventually, the agent's goal is defined as to maximize the reward received over time.

Formally, the reinforcement learning problem is described as a Markov decision process (MDP). In its most rigorous form, the MDP demands a perfect fulfillment of the Markov property, which puts a specific restriction on the sequence of states

$$p(S_{t+1}|S_t) = p(S_{t+1}|S_1, \dots, S_t)\,, \tag{2}$$

where $p(S_{t+1}|S_t)$ denotes the conditional probability of transitioning to state S_{t+1} given the previous state S_t. Thus, the Markov property constrains the probability to reach state S_{t+1} to be only dependent on the previous state S_t and not on any other state in the past. The state S_t is therefore required to capture all relevant information regarding the transition dynamics of the environment. While this rigorous formalism is very useful for modeling a wide range of problems and allows precise theoretical statements, the Markov property can sometimes be difficult to fulfill precisely in practical applications. Nevertheless, recent efforts in reinforcement learning research have proven quite successful (e.g. [7, 8]), and lead to a new wave of attention for the field.

Overall, reinforcement learning represents a very powerful, adaptable and scalable framework that, due to the generality of its approach, can be applied to a large variety of control problems.

FEEDBACK DESIGN

In case of the micro-bunching instability and its simulation via VFP solvers, the formulation of a Markovian process is straightforward. The starting conditions for the numerical solution of the VFP equation are given by an initial charge distribution in the longitudinal phase space and a set of constant parameters. Subsequently, the VFP equation can be solved iteratively to simulate the temporal evolution of the charge distribution under these pre-defined conditions [9]. At any point in time, computation of the next time step is entirely based on the charge distribution at the preceding time step (neglecting constant parameters). Defining the temporal sequence of the longitudinal charge distributions in phase space as the state signal

$$S_t \doteq \psi_t(z, E) \tag{3}$$

thus yields a Markov process, fully satisfying Eq. (2).

To obtain an MDP, the Markov process still needs to be augmented by proper definitions of a reward function and an action space. As the primary interest lies in the CSR emitted by the electron bunch, we define the reward function based on the CSR power time series

$$R_t \doteq R_t(P_{t,\mathrm{CSR}})\,. \tag{4}$$

It is worth noting that choosing the reward function is a very crucial point of any reinforcement learning problem, as it alone defines the goal the agent is trying to achieve. In this context, aiming to stabilize the emitted CSR, the choice can be as simple as

$$R_t \doteq \omega_1 \mu_{t':t} - \omega_2 \sigma_{t':t}\,, \tag{5}$$

where $\mu_{t':t}$ and $\sigma_{t':t}$ denote the mean and standard deviation of the time series $P_{t,\mathrm{CSR}}$ in the interval $[t', t]$, and $\omega_{1,2} > 0$ are simple weighting factors. Clearly, this function expresses the desire of having a CSR power signal of high intensity and low fluctuation. However, whether or not it does so in the best possible and most desirable way is less obvious and still under investigation.

Finally, what remains is the choice of action space. In principle, anything capable of influencing the micro-bunching dynamics and thus the CSR power signal can be considered. Yet, one particularly promising idea seems to be centered around the RF system of the storage ring in order to counteract the additional perturbation by the CSR wake potential. Following this train of thought, a straightforward choice of the action space is

$$A_t \in \{V_{\mathrm{RF}} \times \varphi_{\mathrm{RF}}\}\,, \tag{6}$$

where V_{RF} denotes the RF amplitude and φ_{RF} the RF phase. This choice leaves the agent with an option for a trivial solution, as the dependency of the instability threshold on the RF amplitude is well established [10–12]. The agent may therefore chose to continuously reduce the RF amplitude until the instability threshold is crossed and the dynamics are stabilized just naturally. To circumvent this issue, modifications to the RF system can be restricted to mere modulations of V_{RF} and φ_{RF}, maintaining the same effective values. One exemplary choice is

$$A_t \in \{A_V \times f_V \times A_\varphi \times f_\varphi\}\,, \tag{7}$$

where $A_{V,\varphi}$ and $f_{V,\varphi}$ denote the amplitude and frequency of sinusoidal modulations of the RF amplitude and RF phase. Preliminary studies using our VFP solver Inovesa indicate the effectiveness of influencing the micro-bunching dynamics by RF modulations and they have also been tested experimentally in the past, e.g. [13, 14]. A temporally adaptable RF modulation scheme is a promising proposition to exert influence on the longitudinal beam dynamics in the micro-bunching instability as it provides the required flexibility to respond to the varying perturbation by the CSR wake potential over continuous time.

Feasibility of the State Signal

Using the definition of the MDP discussed above, reinforcement learning solution methods can be applied to train an agent on simulation data. With this in mind, the VFP solver Inovesa has already been extended to support RF modulations and communication with other processes during runtime in order to allow interaction with such an agent. First tests are currently ongoing.

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPGW017

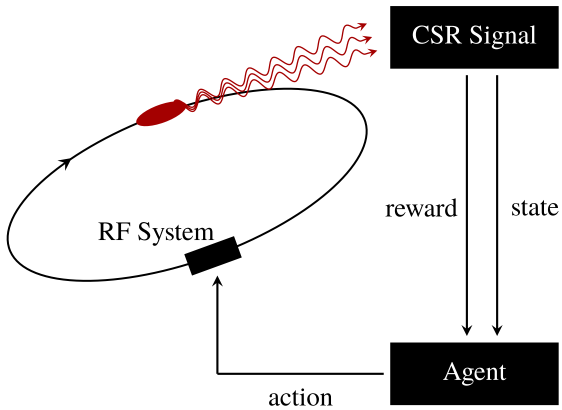

Figure 2: General feedback scheme using the CSR power signal to construct both, the state and reward signal of the Markov decision process (MDP).

However, particularly the choice of the state signal in form of the longitudinal phase space density is challenging to realize for an actual storage ring. Although first efforts towards phase space tomography have been made at KARA, this type of information is not yet accessible. Nevertheless, several diagnostic systems are in place, which can provide information about the longitudinal beam dynamics. As the projection of $\psi_t(z, E)$ on the longitudinal axis, the longitudinal bunch profile can be measured by an electro-optical near-field setup on a turn-by-turn basis [15–17]. Complementary information about the energy distribution can be obtained by measuring the horizontal bunch profile in a dispersive section of the accelerator using a fast-gated camera [18–20]. However, the simplest and most robust way of acquiring information regarding the state of the electron bunch is by using the CSR power signal $P_{t,\mathrm{CSR}}$ itself. As the emitted CSR power is strongly correlated to the micro-bunching dynamics within the bunch, a state signal can be constructed based entirely on that:

$$S_t \doteq S_t(P_{t,\mathrm{CSR}}) \,. \tag{8}$$

One way of implementing this idea is by choosing a handcrafted feature vector, which is designed to capture information about the state of the micro-bunching. Based on previous studies of the CSR power signal in the micro-bunching instability, one exemplary choice is

$$S_t \doteq (\mu_{t':t}, \sigma_{t':t}, m_{t':t}, f_{\max}, A_{\max}, \varphi_{\max})^{\mathrm{T}} \,, \tag{9}$$

where $m_{t':t}$ represents a slow trend in the amplitude of the CSR power. The variables $f_{\max}, A_{\max}, \varphi_{\max}$ denote the frequency, amplitude and phase of the main component in the Fourier transform of the time series $P_{t,\mathrm{CSR}}$ in the preceding interval $[t', t]$. The modified feedback scheme is illustrated in Fig. 2.

It should be noted, the adjusted state signal in Eq. (9) is far apart from the originally proposed definition in Eq. (3). Whether or not enough of the Markov property can be retained using this definition is unclear and has to be verified in practice. In the best case, the provided condensed information is sufficient to yield a fast learning rate of the agent and convergence to a satisfying extent of control over the CSR power signal. If these goals can not be met experimentally, the state signal should be extended to carry more information in order to satisfy the Markov property in Eq. (2) as closely as possible.

Beyond that, the agent's step width $\Delta t = t - t'$ has to be considered. As the micro-bunching dynamics typically occur at time scales of several multiples of the synchrotron period, the step width Δt might have to be chosen quite small in order to react to these fast changes. Whether or not this can be relaxed to slower interaction rates has to be tested empirically. Overall, the synchrotron period ($\Delta t \approx T_\mathrm{s}$), which is in the order of several kHz at KARA, seems to be a reasonable starting value.

SUMMARY AND OUTLOOK

A feedback scheme aiming to establish extensive control over the longitudinal beam dynamics in the micro-bunching instability needs to be capable of reacting to the fast and dynamic variation of the CSR wake potential. As the perturbation explicitly depends on the state of the electron bunch, the action or countermeasure should, in general, be expected to be state-dependent as well. Reinforcement learning represents a very potent approach to model these dynamics and to apply solution methods, which optimize for a pre-defined goal in form of a scalar reward function. The required formulation of a Markov decision process is well-motivated due to the Markov property of VFP solvers and conceptually outlined in this contribution. It should be noted, the chosen action space implies that control is not necessarily achieved via direct suppression of the CSR self-interaction, but by an additional, active and dynamic interaction with the longitudinal charge distribution.

Finally, the outlined feedback scheme is not necessarily restricted to the micro-bunching instability driven by the CSR impedance. Different collective effects can be modeled in form of Eq. (1) and simulated using a VFP solver. A successful implementation may thus be easily transferable to control tasks of different longitudinal instabilities at storage rings.

ACKNOWLEDGEMENT

T. Boltz and P. Schreiber acknowledge the support by the DFG-funded Doctoral School "Karlsruhe School of Elementary Particle and Astroparticle Physics: Science and Technology (KSETA)". This research is in part supported by the Innovationspool AMALEA (Accelerating Machine Learning for Physics) in the Helmholtz Association's Programme "Matter and Technologies".

REFERENCES

[1] P. Schönfeldt *et al.*, "Parallelized Vlasov-Fokker-Planck solver for desktop personal computers", *Phys. Rev. Accel. Beams*, vol. 20, p. 030704, 2017. https://github.com/Inovesa/Inovesa

ISBN: 978-3-95450-208-0 doi:10.18429/JACoW-IPAC2019-MOPGW017

[2] J. L. Steinmann *et al.*, "Continuous bunch-by-bunch spectroscopic investigation of the microbunching instability", *Phys. Rev. Accel. Beams*, vol. 21, p. 110705, 2018.

[3] K. L. F. Bane and K. Oide, "Simulations of the Longitudinal Instability in the New SLC Damping Rings", in *Proc. 16th Particle Accelerator Conf. (PAC'95)*, Dallas, TX, USA, May 1995, paper WAC16, pp. 3105–3108.

[4] K. Bane *et al.*, "Impedance Calculation and Verification in Storage Rings", *SLAC Technical Report*, No. SLAC-PUB-11007, 2005. doi:10.2172/839763

[5] C. Evain *et al.*, "Stable coherent terahertz synchrotron radiation from controlled relativistic electron bunches", *Nature Physics*, 2019. doi:10.1038/s41567-019-0488-6

[6] R. S. Sutton and A. G. Barto, *Reinforcement Learning*. Cambridge, MA, USA: MIT Press, 2018.

[7] V. Mnih *et al.*, "Human-level control through deep reinforcement learning", *Nature*, vol. 518, pp. 529–533, 2015.

[8] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search", *Nature*, vol. 529, pp. 484–489, 2016.

[9] R. L. Warnock and J. A. Ellison, "A General Method for Propagation pf the Phase Space Distribution, with Application to the Sawtooth Instability", *SLAC Technical Report*, No. SLAC-PUB-8404, 2000.

[10] K. L. F. Bane, Y. Cai, and G. Stupakov, "Threshold studies of the microwave instability in electron storage rings", *Phys. Rev. ST Accel. Beams*, vol. 13, p. 104402, 2010.

[11] P. Kuske, "CSR-driven Longitudinal Single Bunch Instability Thresholds", in *Proc. 4th Int. Particle Accelerator Conf. (IPAC'13)*, Shanghai, China, May 2013, paper WEOAB102, pp. 2041–2043.

[12] M. Brosi *et al.*, "Fast mapping of terahertz bursting thresholds and characteristics at synchrotron light sources", *Phys. Rev. Accel. Beams*, vol. 19, p. 110701, 2016.

[13] Y. Shoji and T. Takahashi, "Coherent Synchrotron Radiation Burst from Electron Storage Ring under External RF Modulation", in *Proc. 11th European Particle Accelerator Conf. (EPAC'08)*, Genoa, Italy, Jun. 2008, paper MOPC048, pp. 178–180.

[14] J. L. Steinmann, "Diagnostics of Short Electron Bunches with THz Detectors in Particle Accelerators", Ph.D. thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany, 2019.

[15] N. Hiller *et al.*, "A Setup for Single Shot Electro Optical Bunch Length Measurements at the ANKA Storage Ring", in *Proc. 2nd Int. Particle Accelerator Conf. (IPAC'11)*, San Sebastian, Spain, Sep. 2011, paper TUPC086, pp. 1206–1208.

[16] L. Rota *et al.*, "KALYPSO: A Mfps Linear Array Detector for Visible to NIR Radiation", in *Proc. 5th Int. Beam Instrumentation Conf. (IBIC'16)*, Barcelona, Spain, Sep. 2016, pp. 740–743. doi:10.18429/JACoW-IBIC2016-WEPG46

[17] S. Funkner *et al.*, "High throughput data streaming of individual longitudinal electron bunch profiles", *Phys. Rev. Accel. Beams*, vol. 22, p. 022801, 2019.

[18] P. Schütze *et al.*, "A Fast Gated Intensified Camera Setup for Transversal Beam Diagnostics at the ANKA Storage Ring", in *Proc. 6th Int. Particle Accelerator Conf. (IPAC'15)*, Richmond, VA, USA, May 2015, pp. 872–875. doi:10.18429/JACoW-IPAC2015-MOPHA039

[19] B. Kehrer *et al.*, "Synchronous detection of longitudinal and transverse bunch signals at a storage ring", *Phys. Rev. Accel. Beams*, vol. 21, p. 102803, 2018.

[20] B. Kehrer *et al.*, "Turn-by-Turn Horizontal Bunch Size and Energy Spread Studies at KARA", presented at the 10th Int. Particle Accelerator Conf. (IPAC'19), Melbourne, Australia, May 2019, paper WEPGW016, this conference.