Extremum Seeking Control for the Optimization of Heavy Ion Beam

重イオンビーム輸送系へのExtremum Seeking制御の適用

Accelerator Control session
Particle Accelerator Society, Japan (PASJ) 2021

08/11/2021

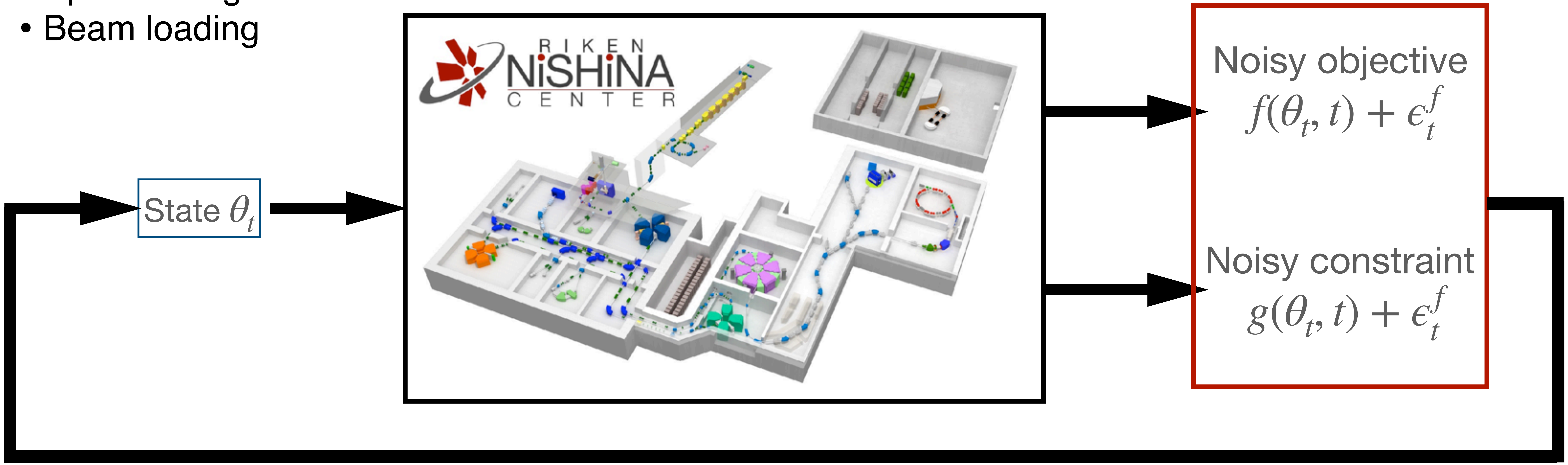
Hiroki Fujii*, Akito Uchiyama, Osamu Kamigaito, Nobuhisa Fukunishi (Riken Nishina Center)
Alexander Scheinker (LANL)
Gessner Spencer (SLAC)

*hiroki.fujii@riken.jp

Background: Safe autonomy of accelerator complex

The accelerator complex is a large-scale MIMO system in nature. (Optics, RF phase and amplitude, etc…)

- **Dynamics of intense charged particle bunches affected by:**
  - **Components drift unpredictably with time, misalignments**
    - Uncertain and time varying particle distribution
  - **Collective effects:**
    - Space charge
    - Beam loading

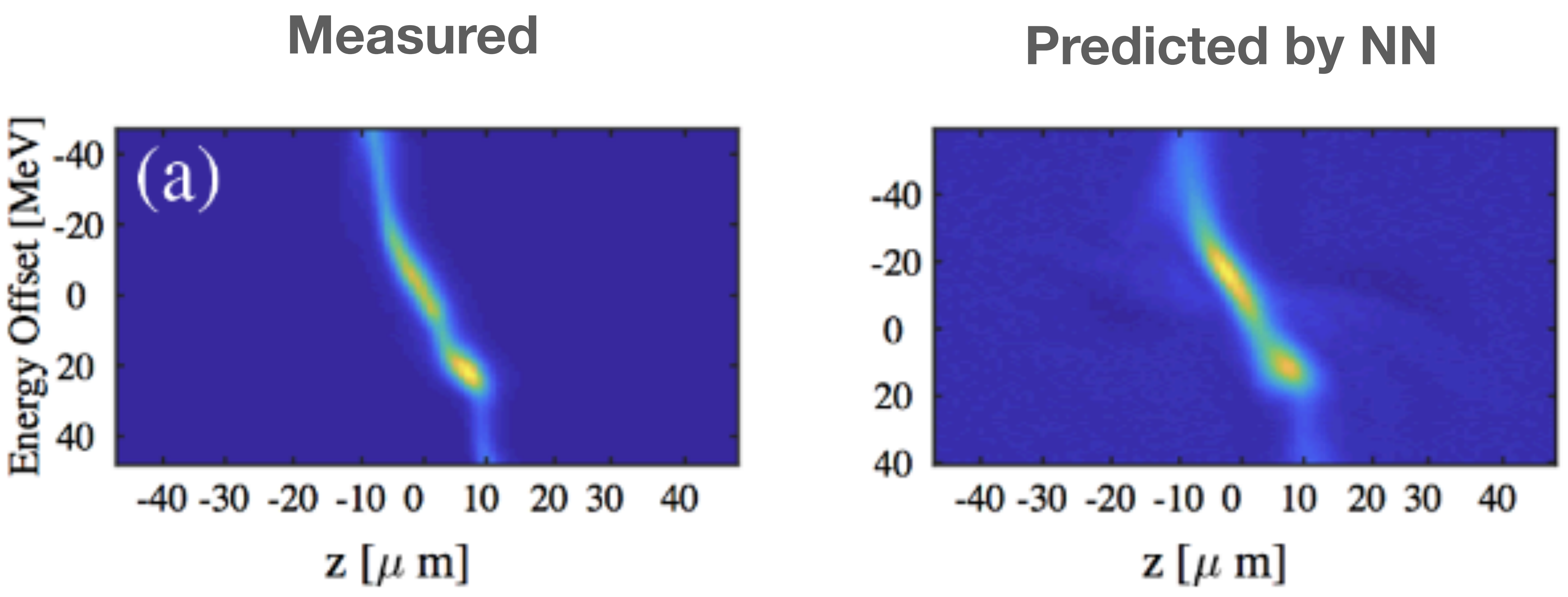


We take measurement based approach for $f(\theta), g(\theta)$ in the real operation.

Goal : $\min_{\theta} f(\theta) \quad \mathrm{s.t.}\ g(\theta) \geq 0$ Safety constraints : $g(\theta_t) \geq 0$ for all t

Recent Researches using machine learning (ML) methods for accelerators

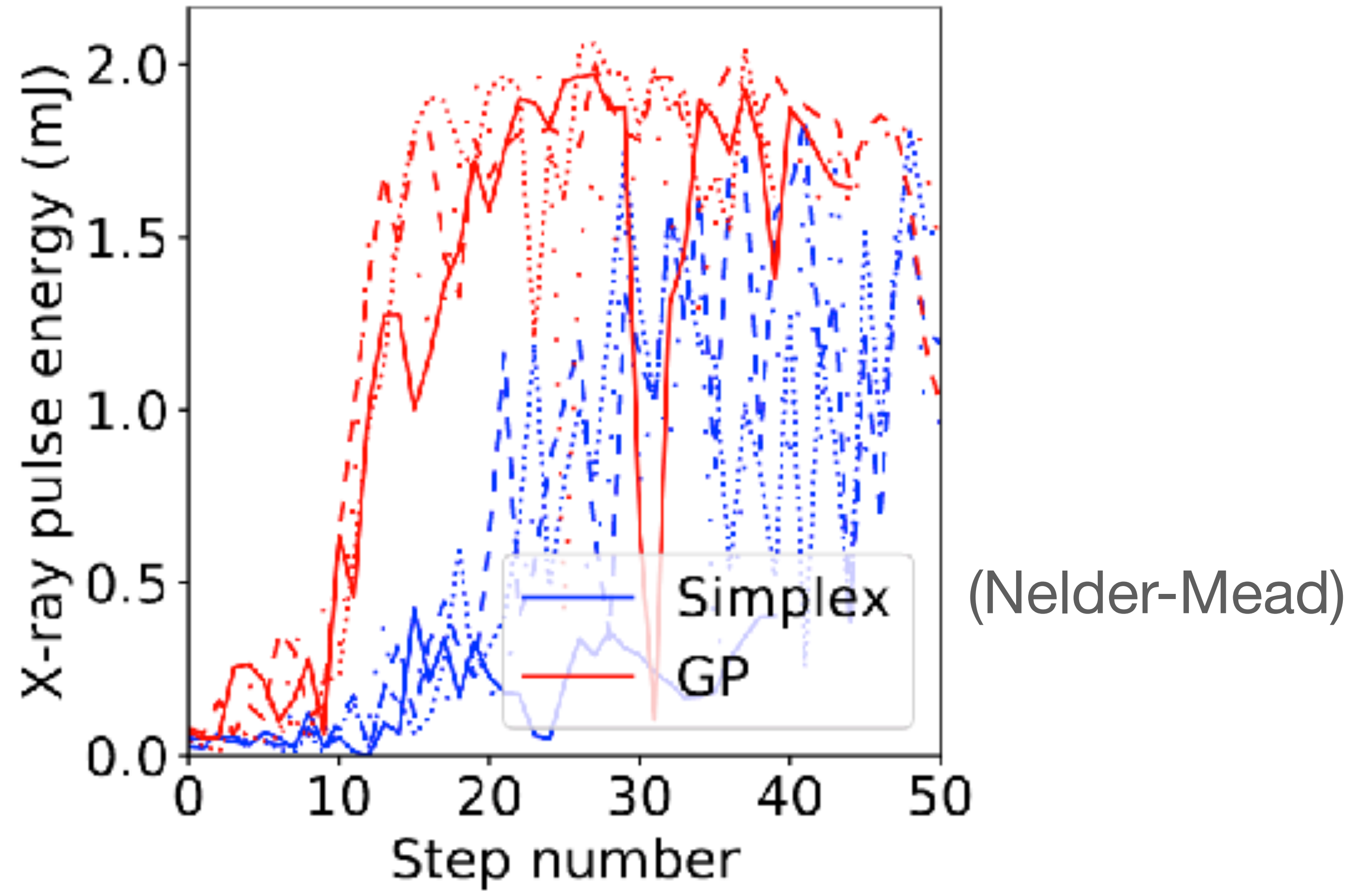
NN based approaches, Modeling of “distribution”



- Experimentally mapped the relation between accelerator input and longitudinal phase space (LPS)
- Possible to predict longitudinal phase space even when they are not measured

C. Emma et al., Physical Review Accelerators and Beams 21.11, 112802, 2018.

Gaussian process (Bayesian optimization)



(c) Live 12 quadrupole optimization

J. Duris, et. al., arXiv:1909.05963v1, 2019

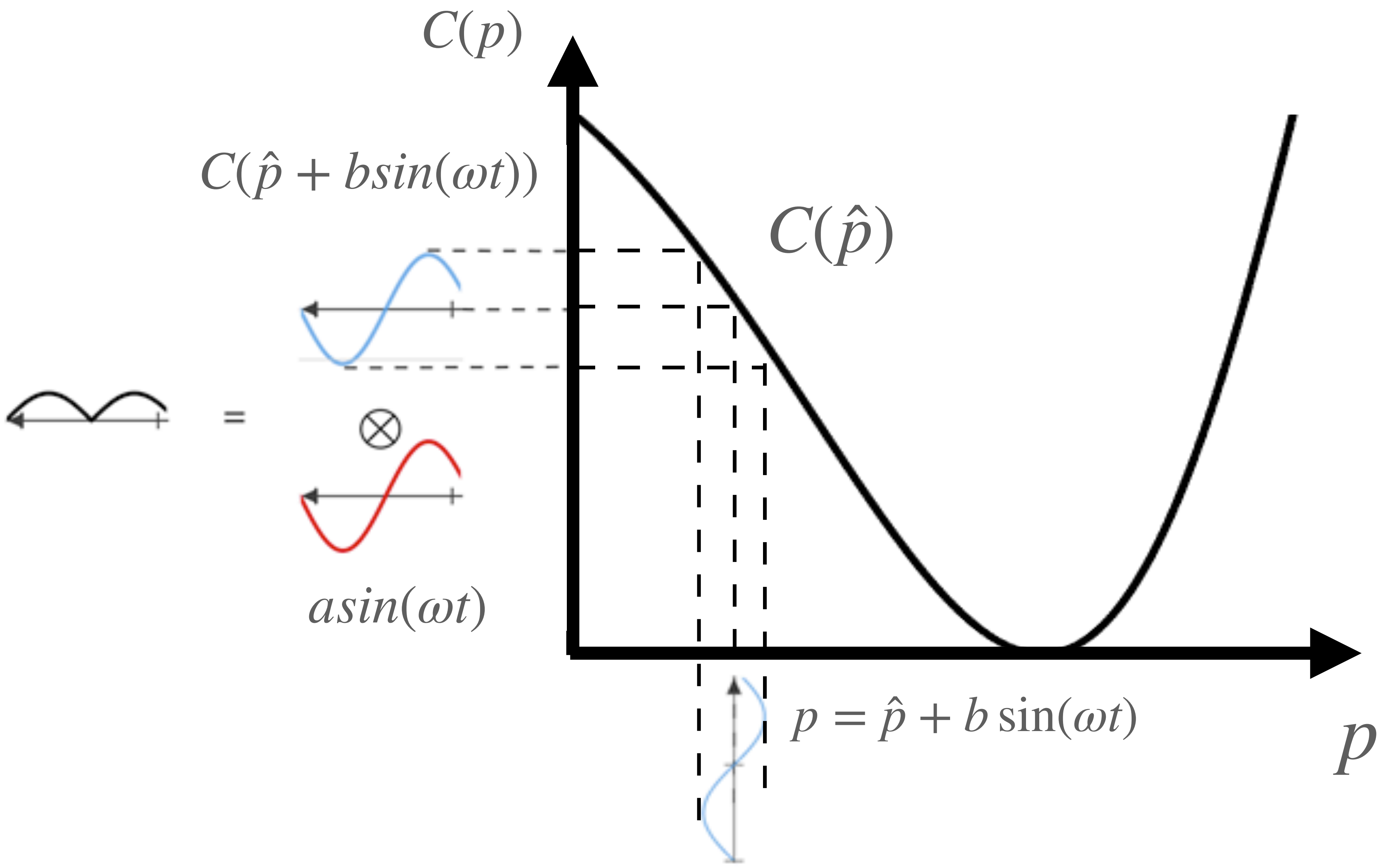
Ongoing : Nishi et.al, Applying GP Optimier developed by Spring-8 team @RIBF (Ref to Iwai et. al., WEOB02)

Limitations

- **Global parameter search -> Is the optimization process damage-free when searching broad parameters?**
- **Vulnerable to the machine drift and fluctuations -> Adaptive control method preferred in some tuning cases**

Model-free adaptive control (Extremum seeking control, ESC)

Dither based extremum seeking



1. Bounded adaptive feedback technique developed for **Multi-input system,**

$$\frac{dp_j}{dt} = \sqrt{\alpha\omega_j}\cos[\omega_j t + k_j C(\mathbf{p}, t)], \quad C(p, t) = \bar{C}(p, t) + noise,$$

$$\omega_j \neq \omega_i \text{ (Orthogonality)}$$

result in the **average** parameter dynamics, which is a gradient descent of cost $C(\mathbf{p}, \mathbf{t})$

$$\frac{d\bar{p}_j}{dt} = -\frac{k_j\alpha}{2}\frac{\partial\bar{C}}{\partial\bar{p}_j}$$

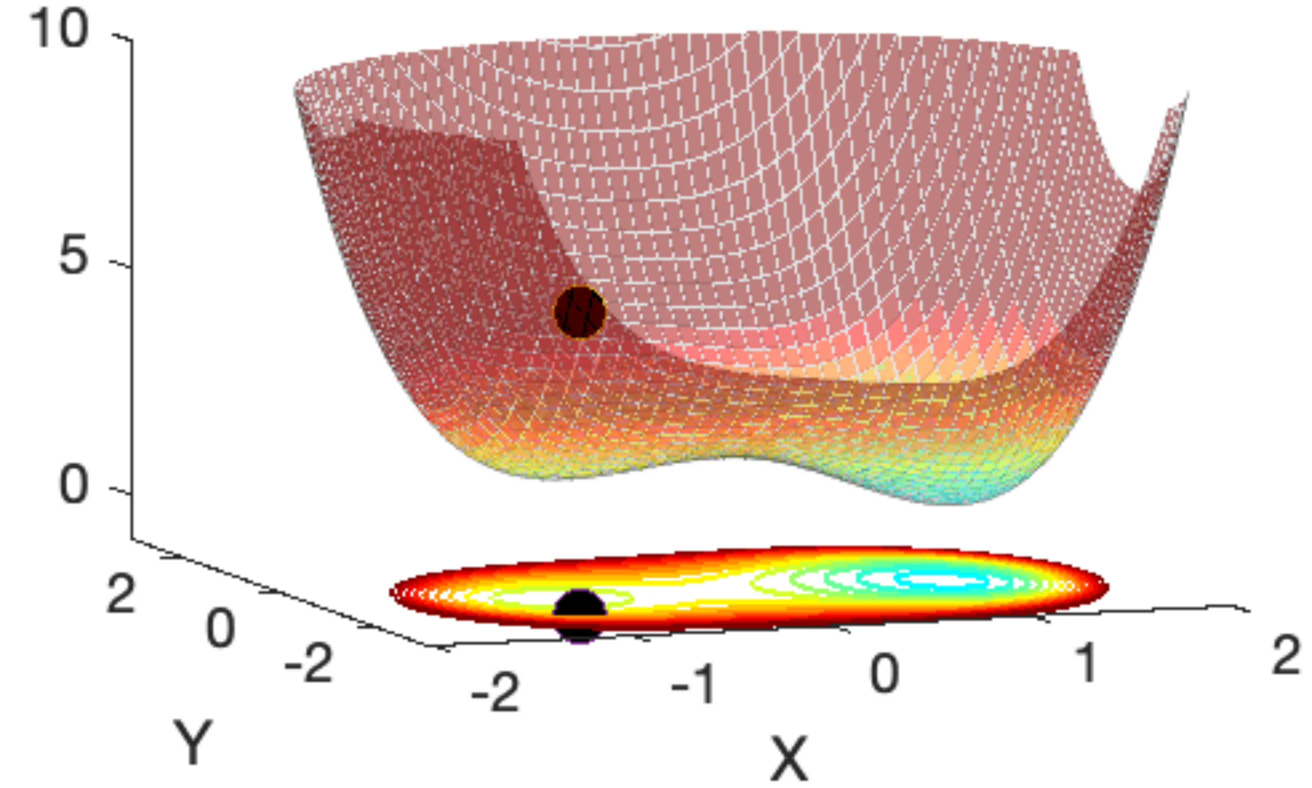
2. Parameter search **within small parameter range** without pre-training

$$\left|\frac{dp}{dt}\right| = \left|\sqrt{\alpha\omega_j}\cos(\omega_j t + kC)\right| \leq \sqrt{\alpha\omega_j}$$

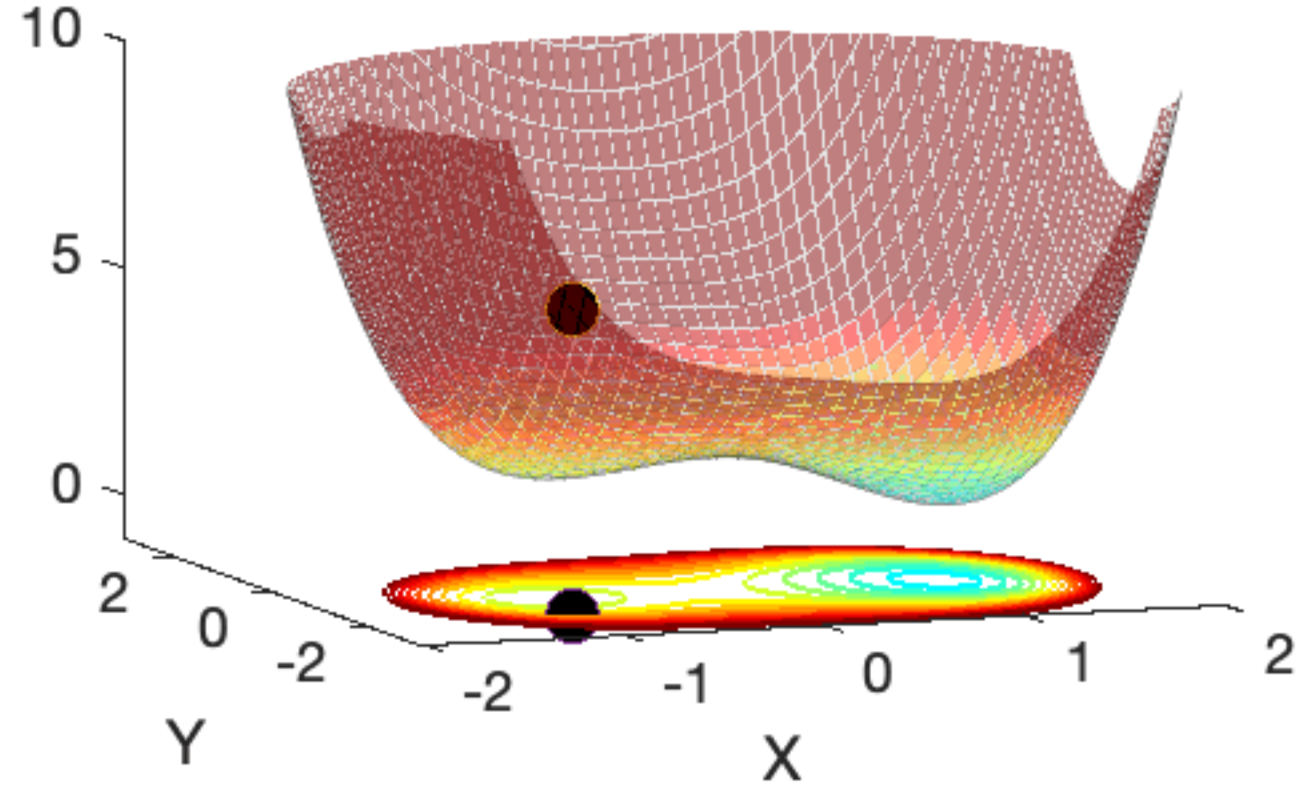
- **Can apply the algorithm to a time-varying system**
- **Rapid response to the increase of cost function is advantageous for damage protection**

Animation of ESC for adaptive state space

$r_0 = \sqrt{\alpha_i \omega_i} = 0.8$



$r_0 = \sqrt{\alpha_i \omega_i} = 1.2$



Red dot : temporal position $(x, y, C(x, y))$

Black dot : 50 steps averaged position $\overline{(x, y, C(x, y))}_{[t-50, t]}$

2D visualization on time-varying function with 2 local minima (Cassini oval)

$$C(x, y) = (x^2 + y'^2)^2 - 2x^2 + 2y'^2 + (-0.5x)$$

$$y' = y - t$$

Adding slope

Adding time varying shift in +y

Red line's trajectory

$$x_t = \gamma^t \sqrt{\alpha_x \omega_x} \cos(\omega_x t + kC(x, y)), \quad \omega_x = 1.0$$

$$y_t = \gamma^t \sqrt{\alpha_x \omega_y} \cos(\omega_y t + kC(x, y)), \quad \omega_y = 1.3$$

$$\gamma = 0.999 : \text{amplitude decay factor}$$

- **Work for time-varying state map**
- **Potentially overcome stuck in local minimum depending on the dither-amplitude**

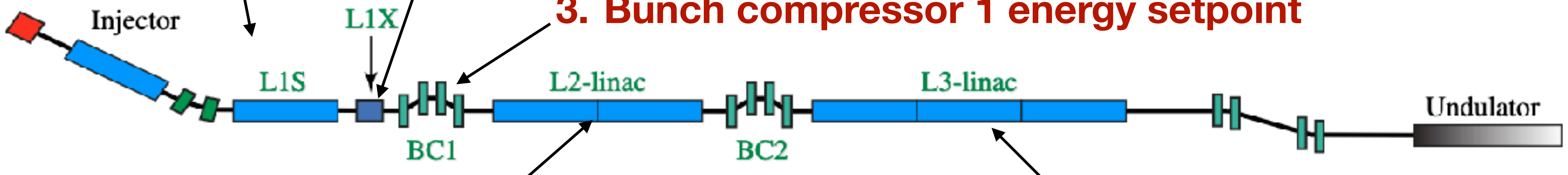
Application examples of ESC: Noisy case

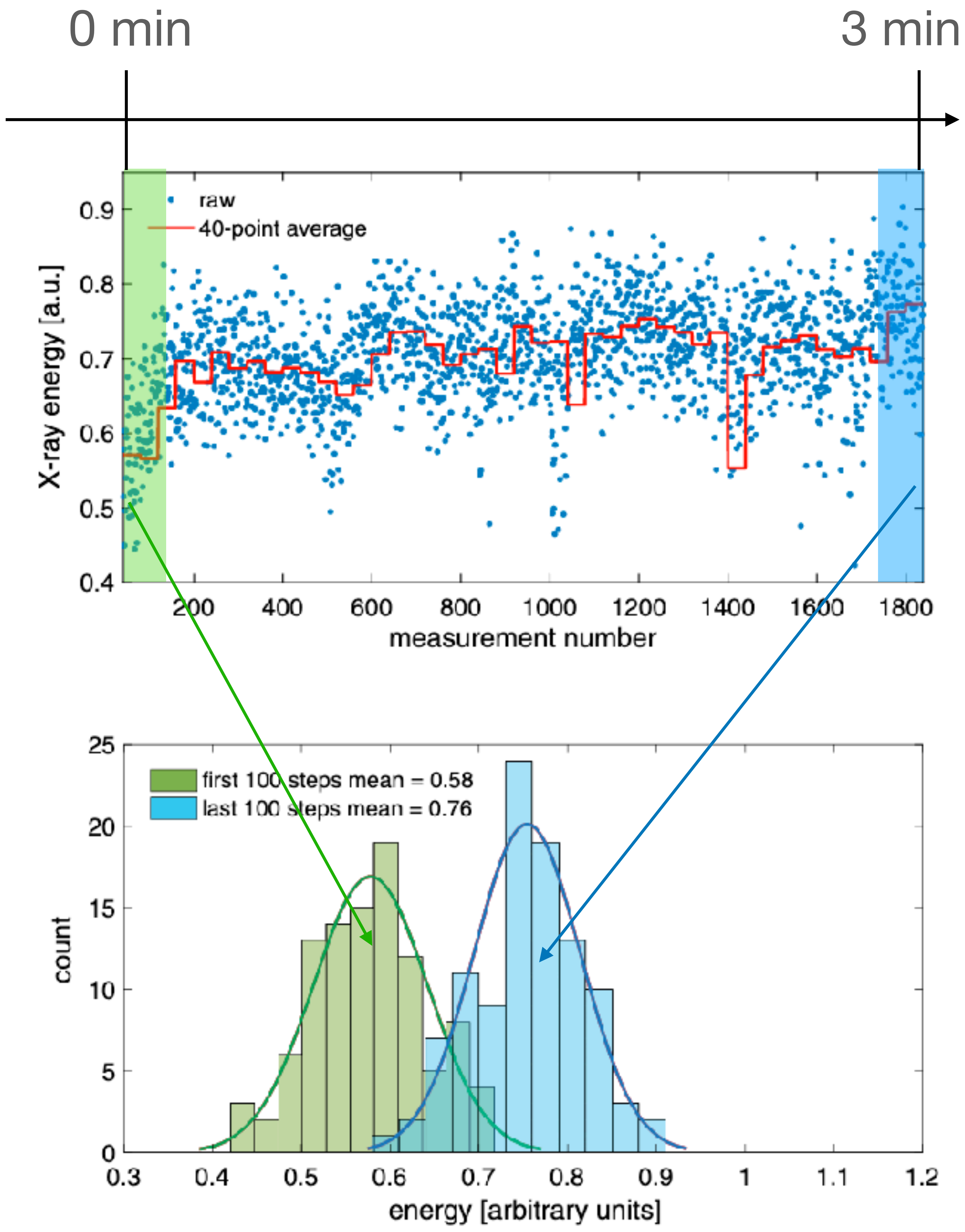
A. Scheinker, et. al., Phys. Rev. Accelerators and Beams 22, 082802 (2019)

LCLS accelerator's RF phase and bunch compressor settings

6 parameters simultaneously tuned

1. Linac1 phase setpoint
2. X-band cavity compensating for energy offsets introduced by L1S
3. Bunch compressor 1 energy setpoint
4. Linac 2 phase setpoint
5. Bunch compressor 2 (BC2) energy setpoint
6. Linac 3 (L3) phase setpoint



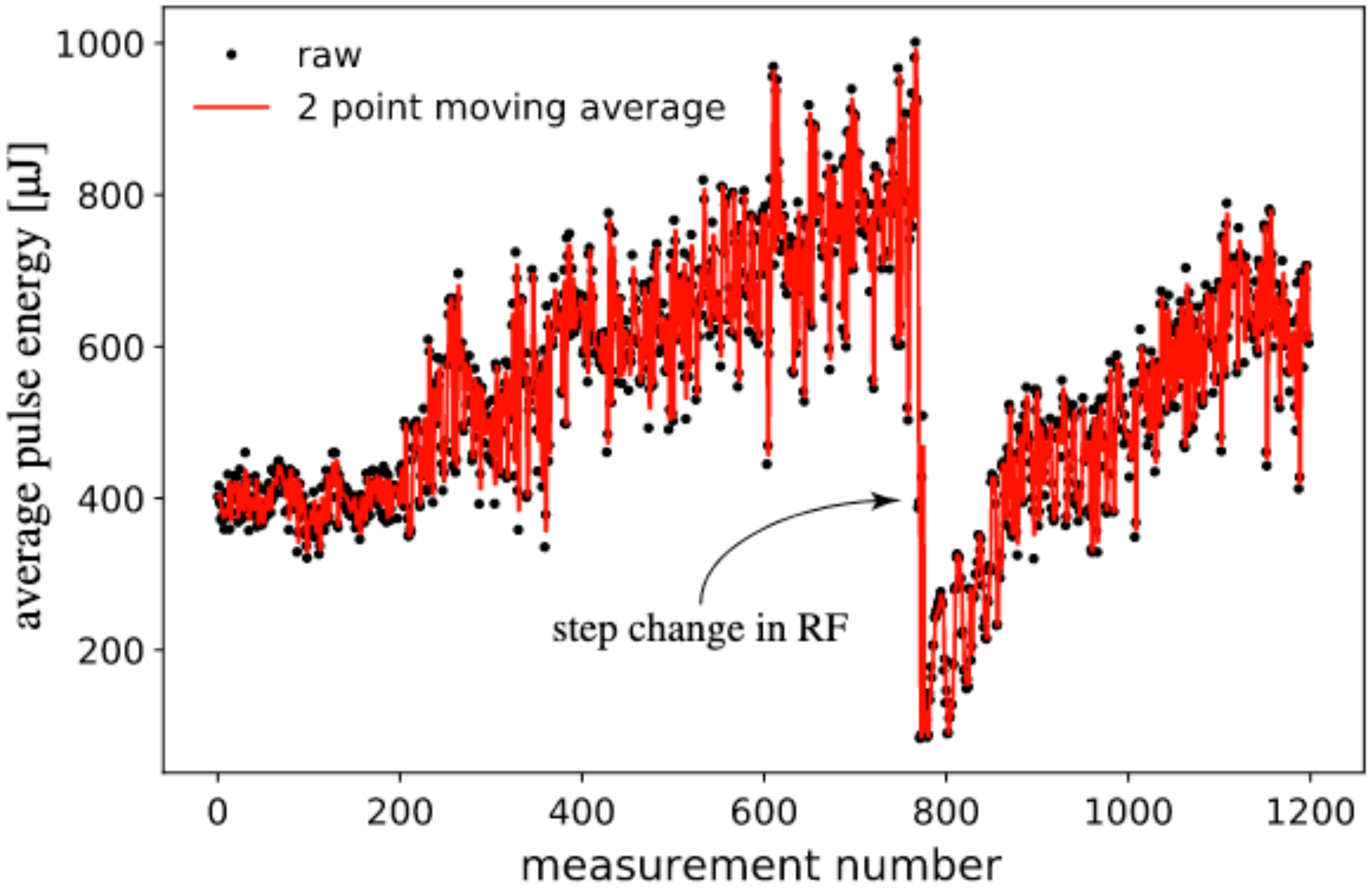


- 40-point average for 10-Hz sampled data
- FEL power increased by 30%

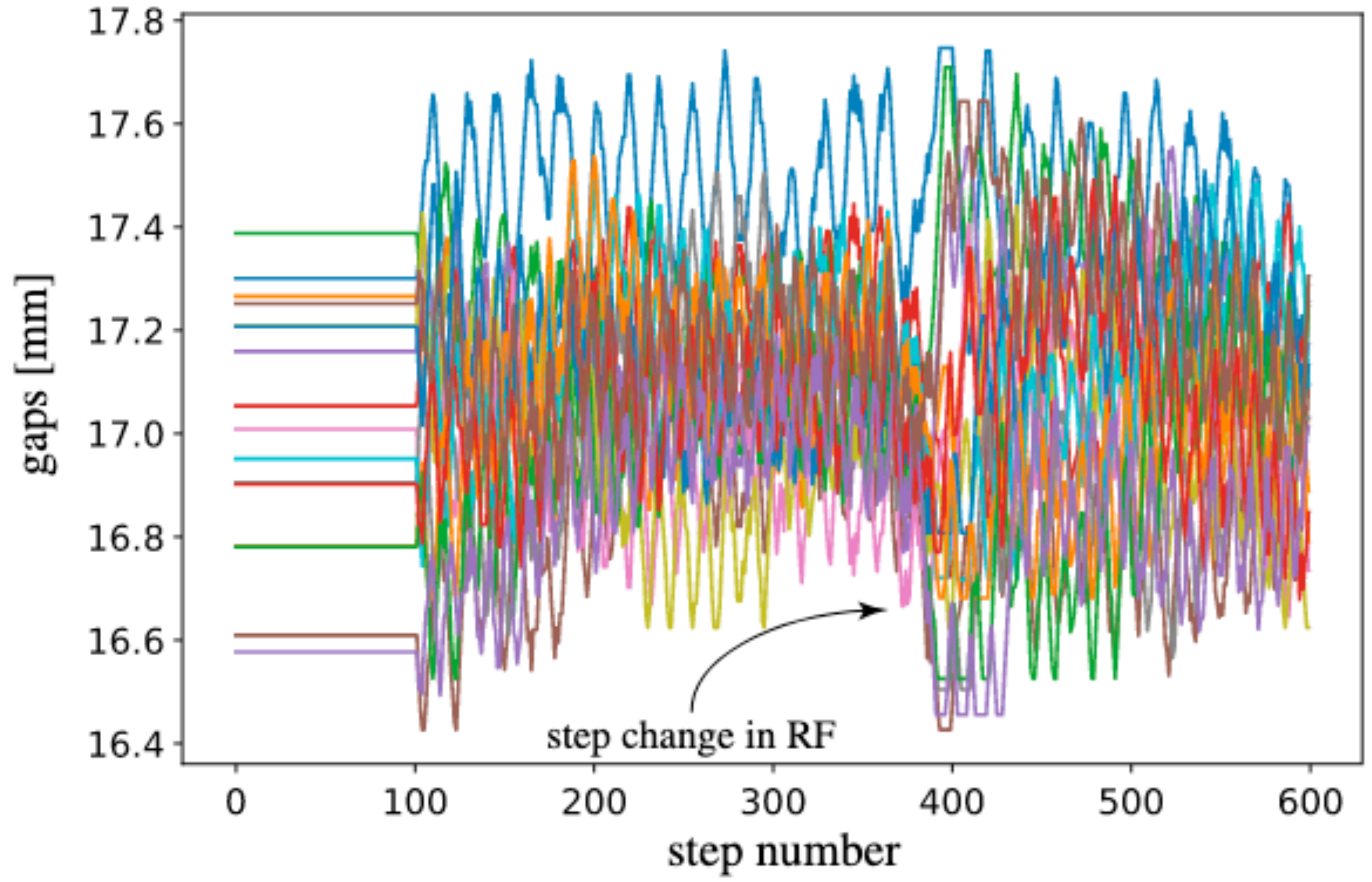
Application examples of ESC: Robust to time-varying environment

A. Scheinker, et. al., Phys. Rev. Accelerators and Beams 22, 082802 (2019)

EuXFEL sudden phase shift example,
robust to noise and sudden setting changes



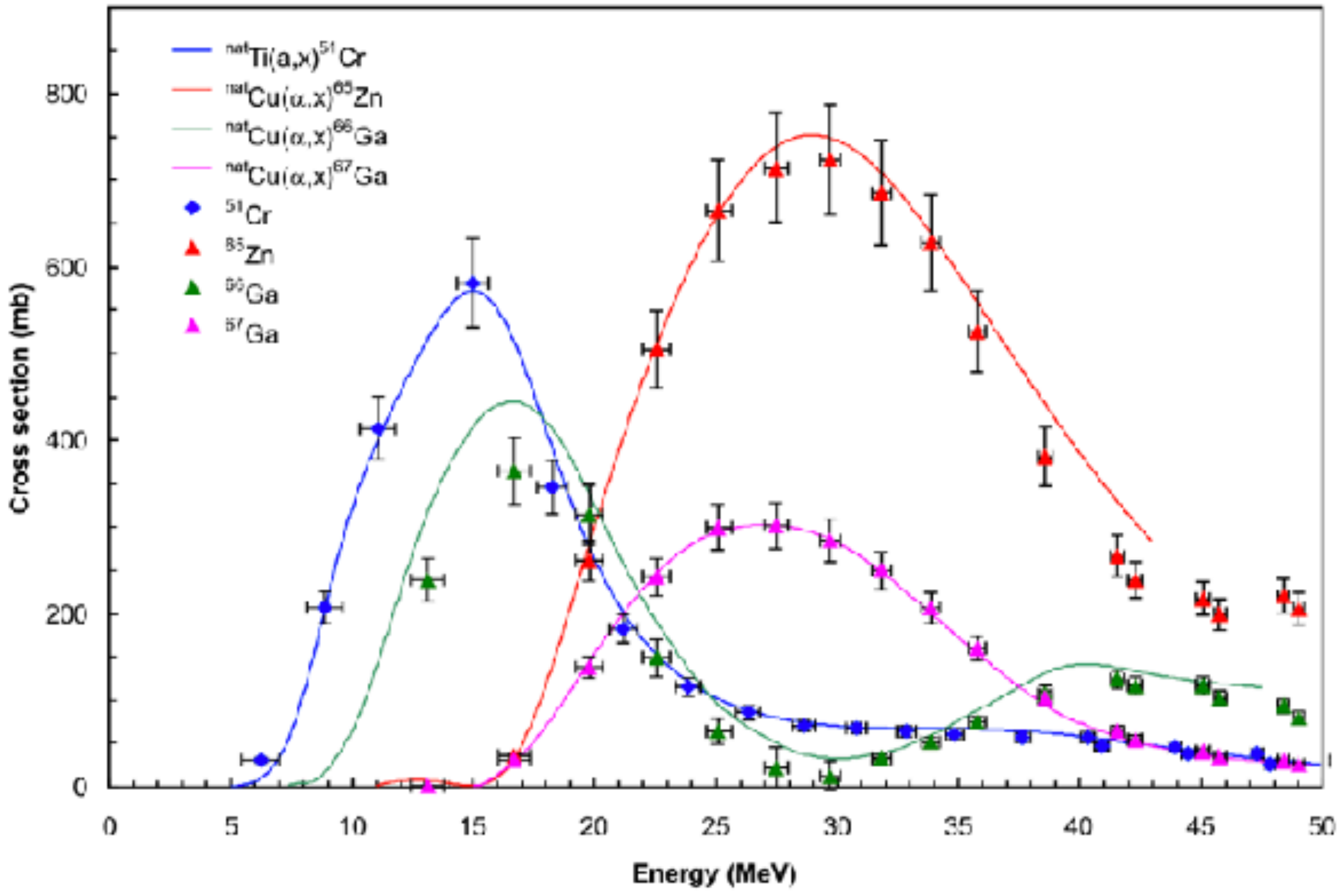
Tuning 15 phase-shifter gaps (chicanes) between each undulator section



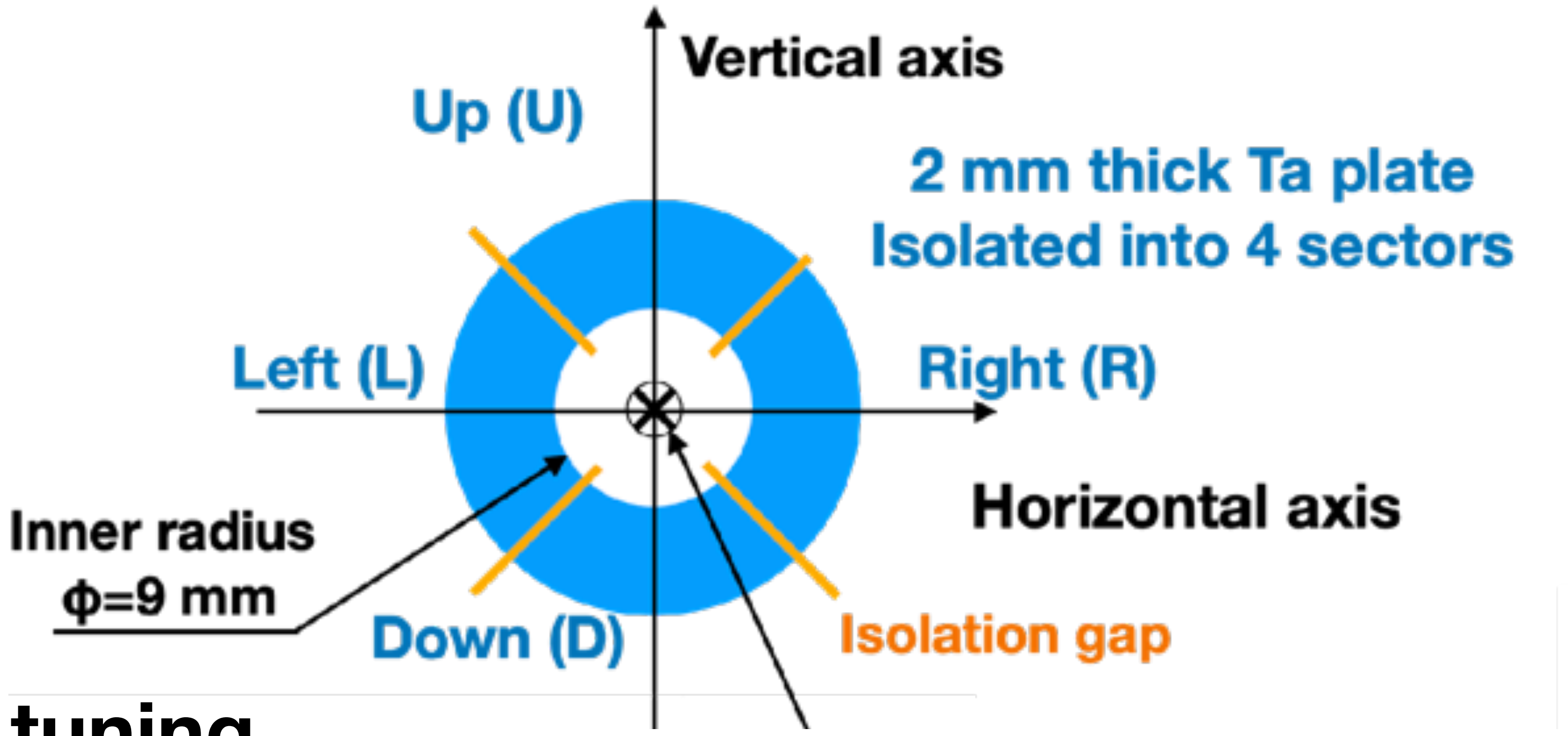
$$\left|\frac{dp_j}{dt}\right| = \left|\sqrt{\alpha\omega_j}\cos(\omega_j t + kC)\right| \leq \sqrt{\alpha\omega_j}$$

Potential ESC application at RIKEN : RI production

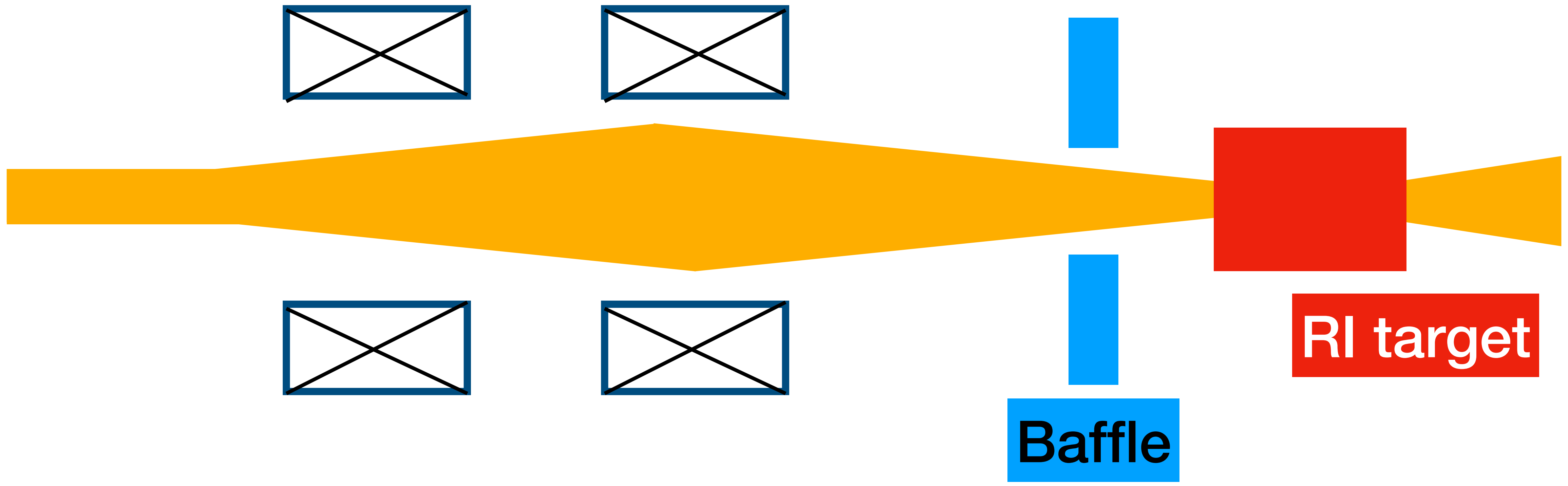
RI production cross-section for ^{4}He beams



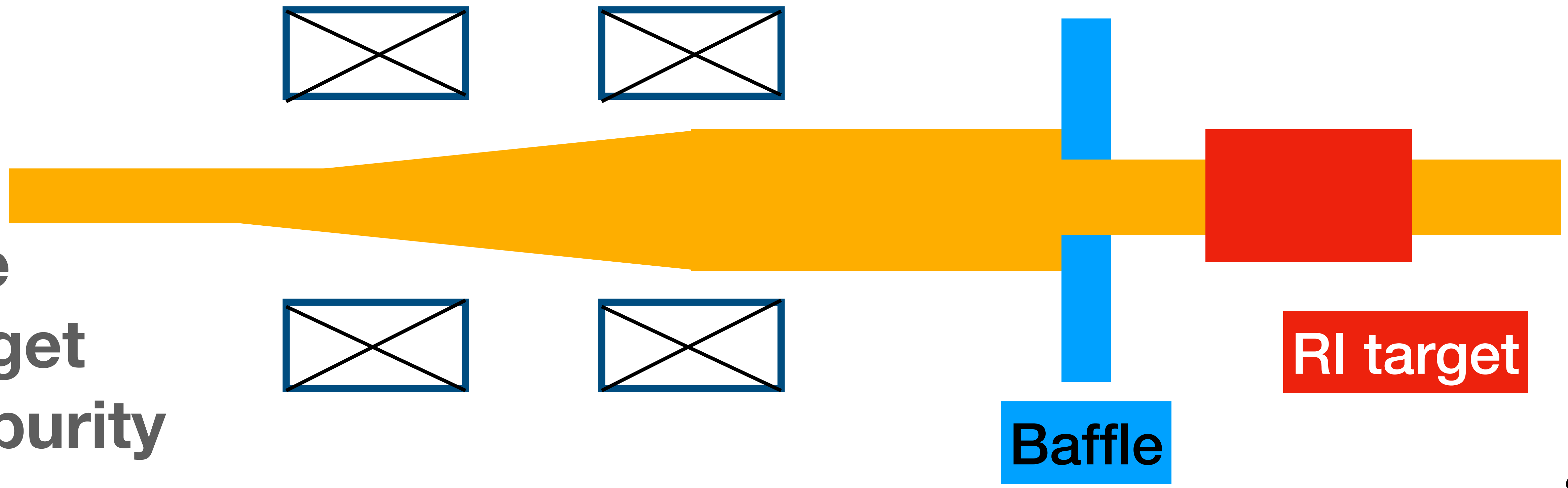
S. Takacs, et. al., Nuclear Instruments and Methods in Physics Research B 397 (2017) 33–38



(a) Beam intensity oriented tuning

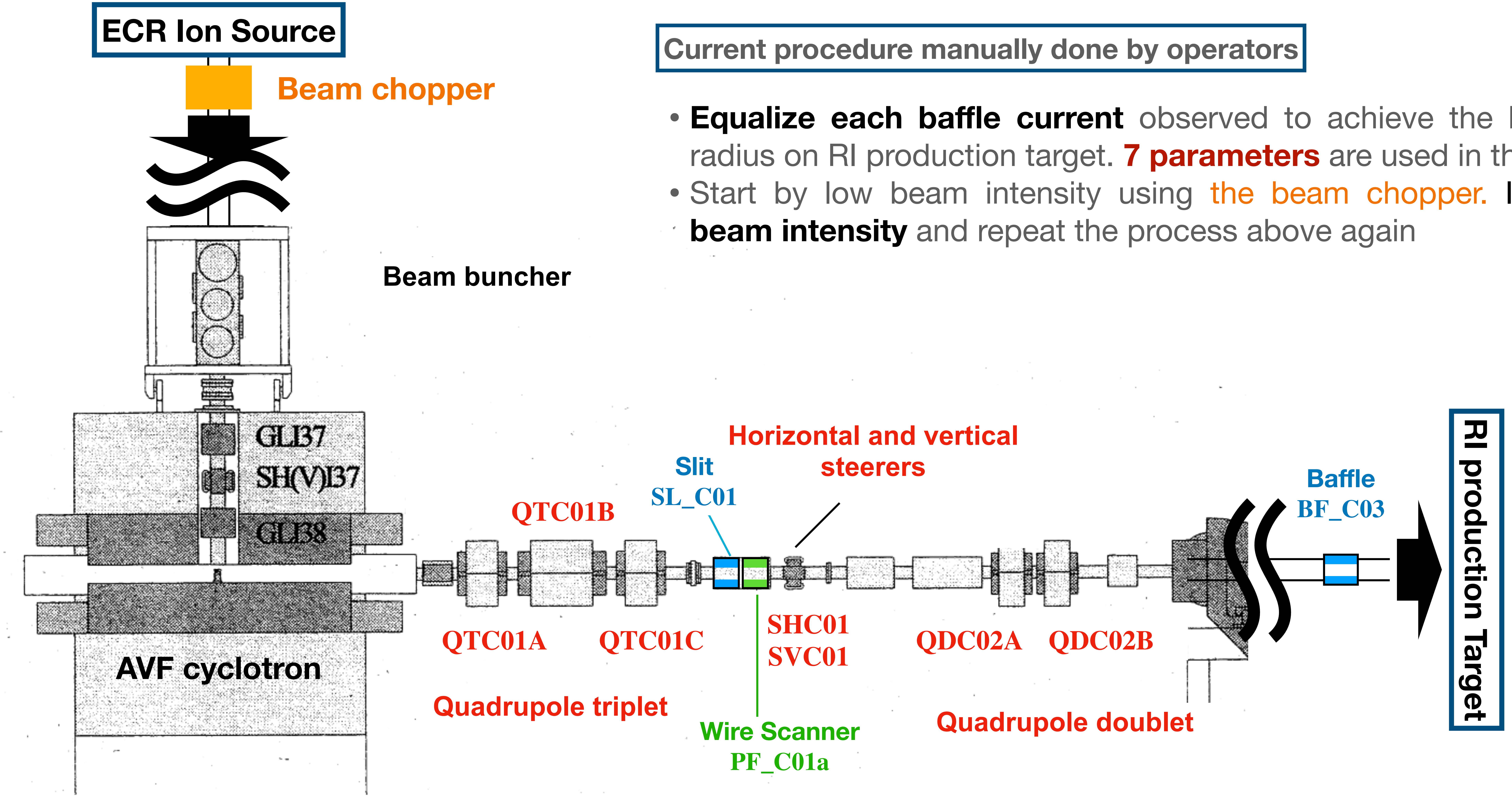


(b) Baffle current oriented tuning



- Large enough beam intensity for higher production rate
  - But should be less than the damage threshold of target
- Smooth beam profile on target, less energy spread for purity

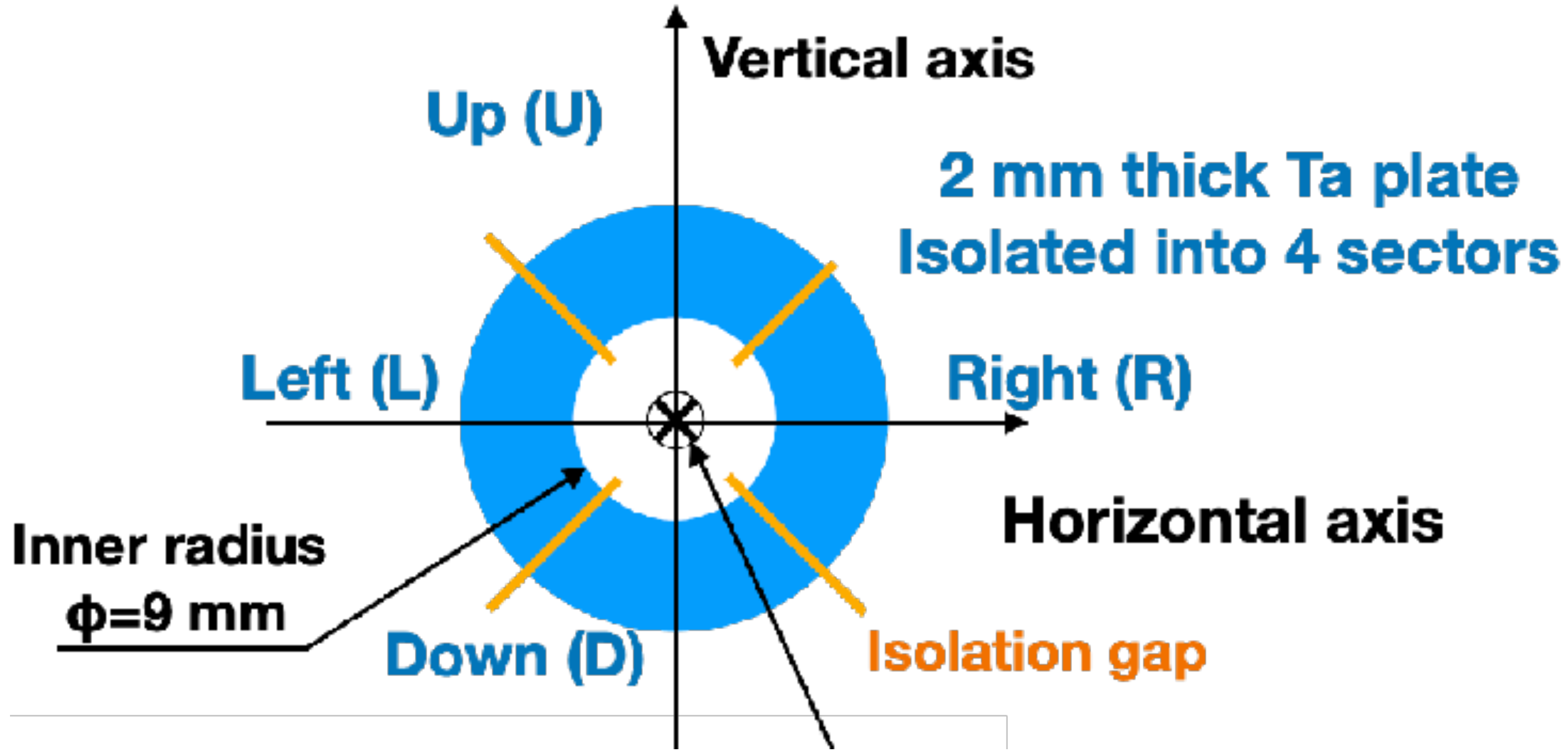
Test at AVF accelerator system, goal of this application



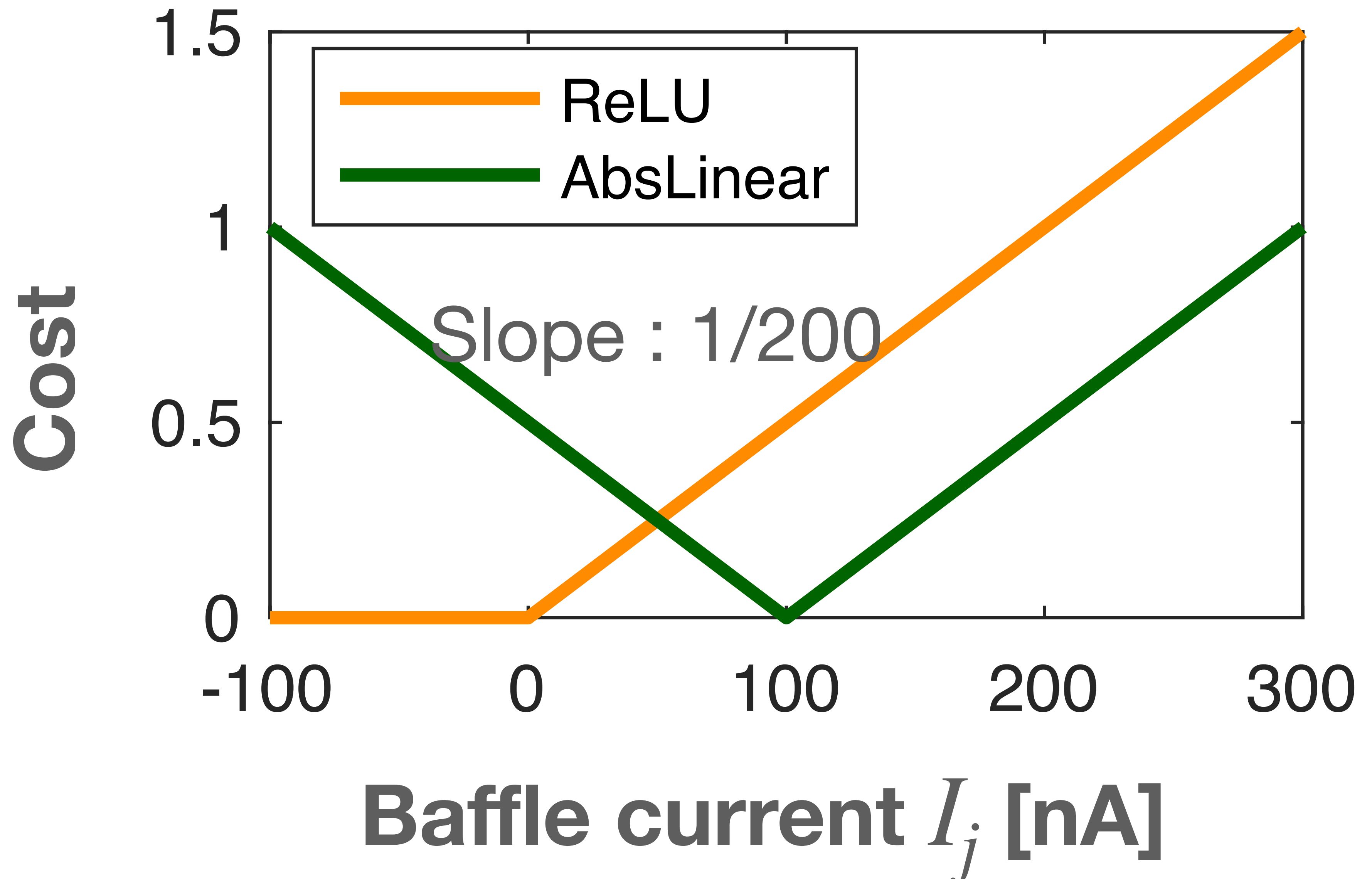
Current procedure manually done by operators

- **Equalize each baffle current** observed to achieve the large enough radius on RI production target. **7 parameters** are used in this study.
- Start by low beam intensity using the beam chopper. **Increase the beam intensity** and repeat the process above again

Experimental result : Cost function and baffle current response

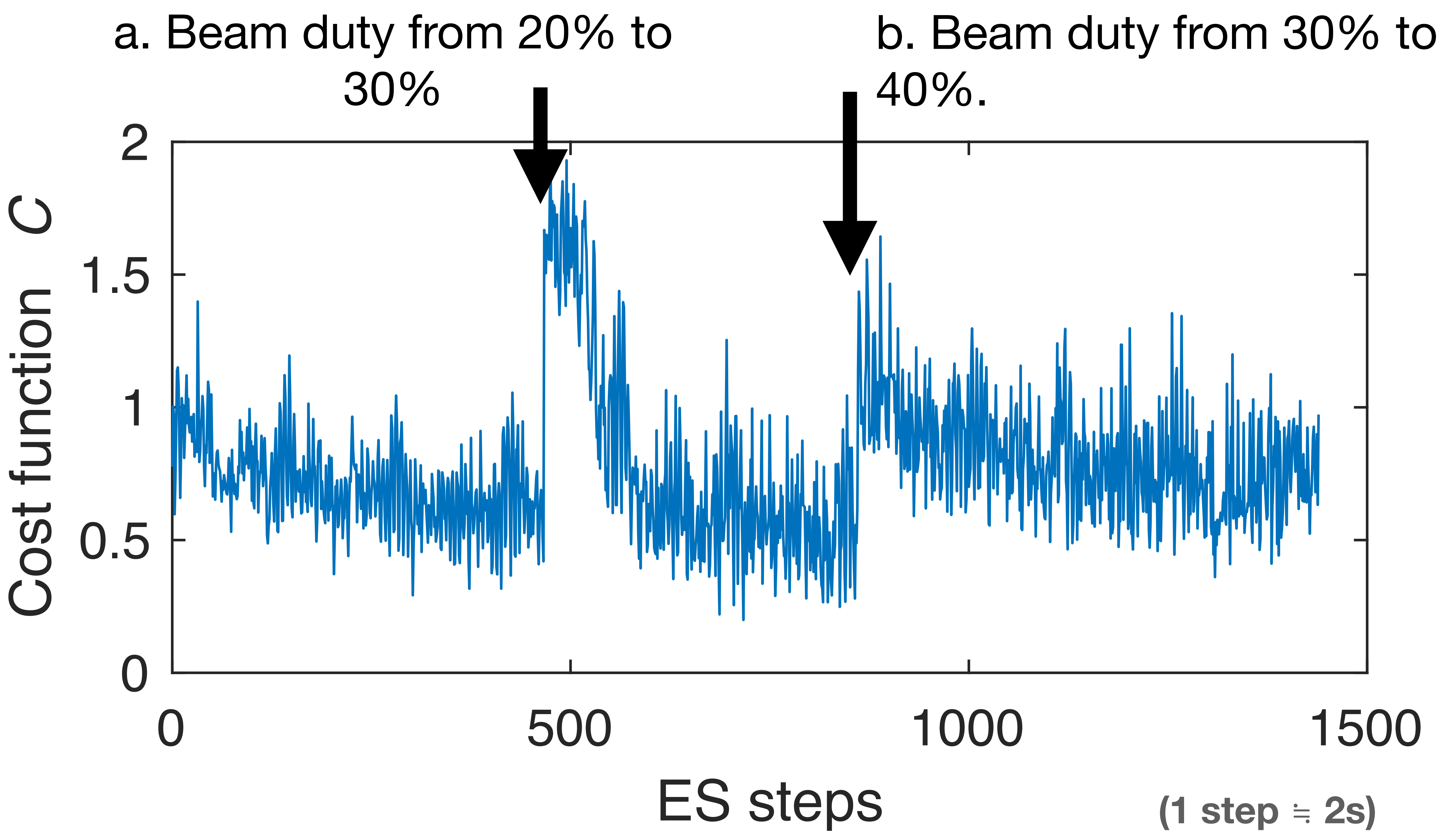


$$kC(\boldsymbol{p},t) = \sum_{j=R,L,U,D} S_j(J_j(\boldsymbol{p},t)) + w_j B_j(I_j(\boldsymbol{p},t))$$



MT : 5/5, 3 hours of Machine time

Overall cost function

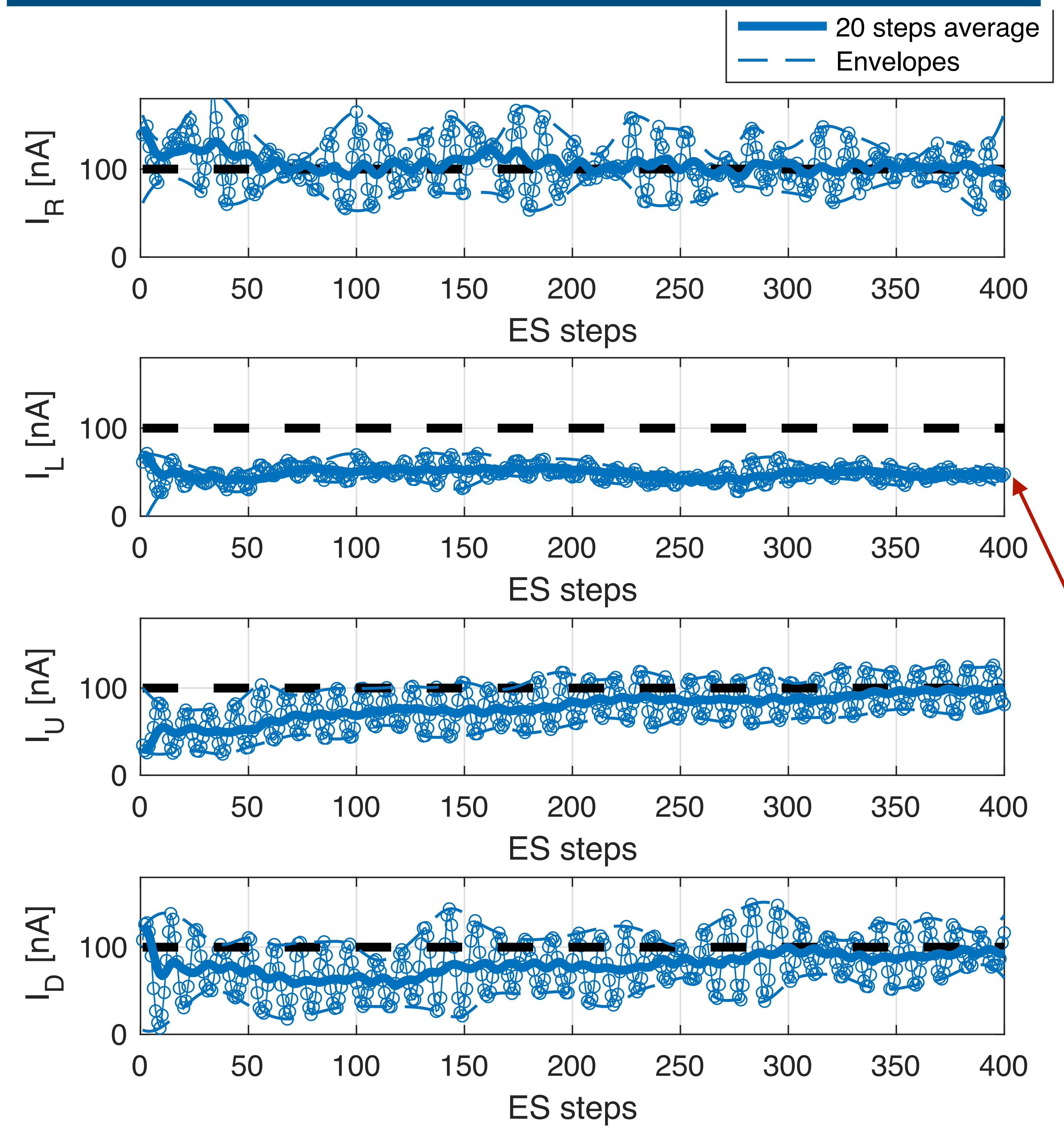


- Beam duty changed occasionally to show the robust response of the method
- Thermal effects on the AVF trim coils are compensated while this operation.

Baffle current result when $w_j B_j(I_j)$ equally weighted ($w_j = 1$)

$$kC(\mathbf{p}, t) = \sum_{j=R,L,U,D} S_j(J_j(\mathbf{p}, t)) + w_j B_j(I_j(\mathbf{p}, t))$$

Baffle current to be controlled

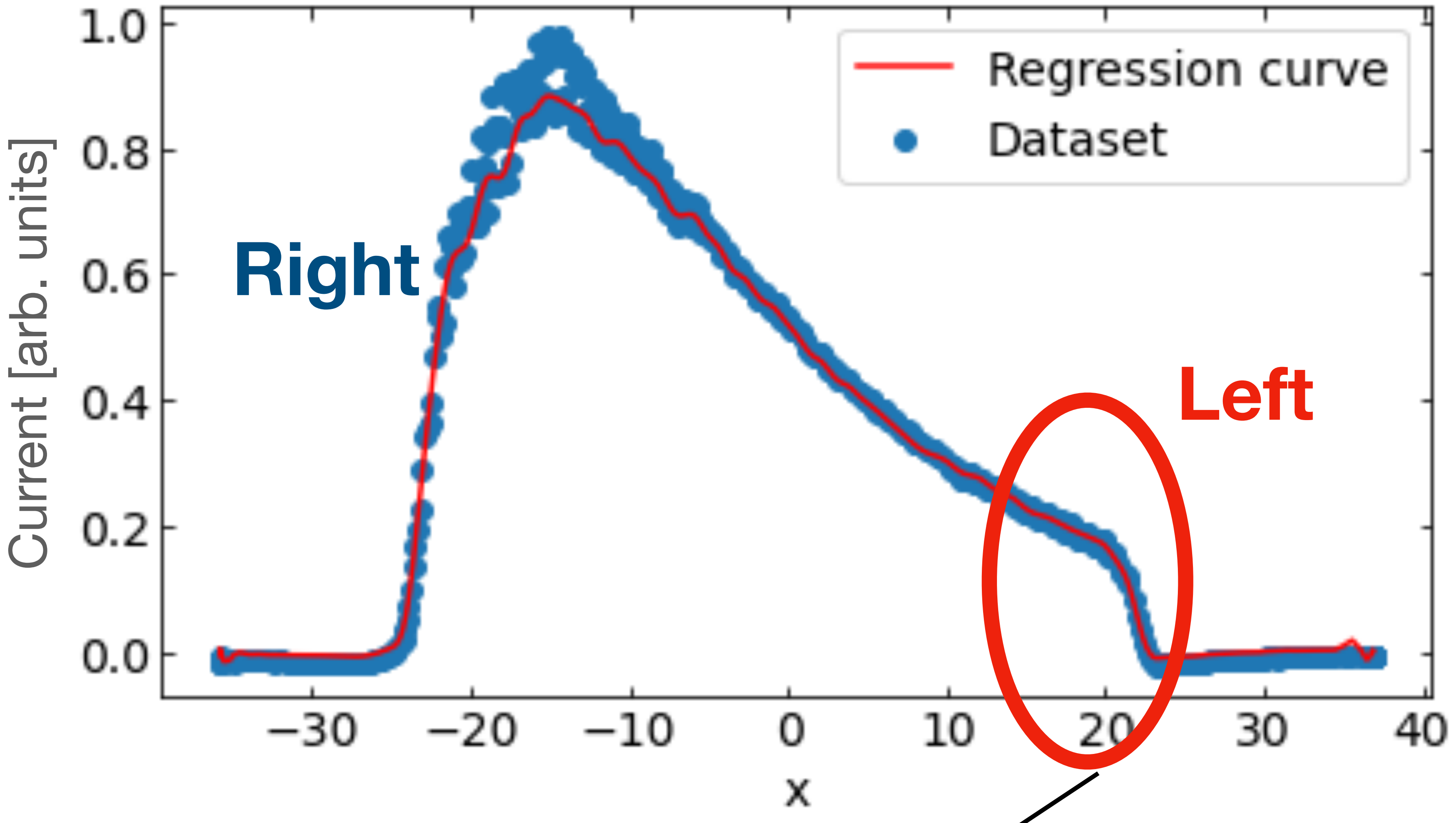


Right, Up, Down signal approaching to the target set point

Left current signal is not controlled to the setpoint

Left baffle current has less amplitude compared to the other components

Transverse beam profile

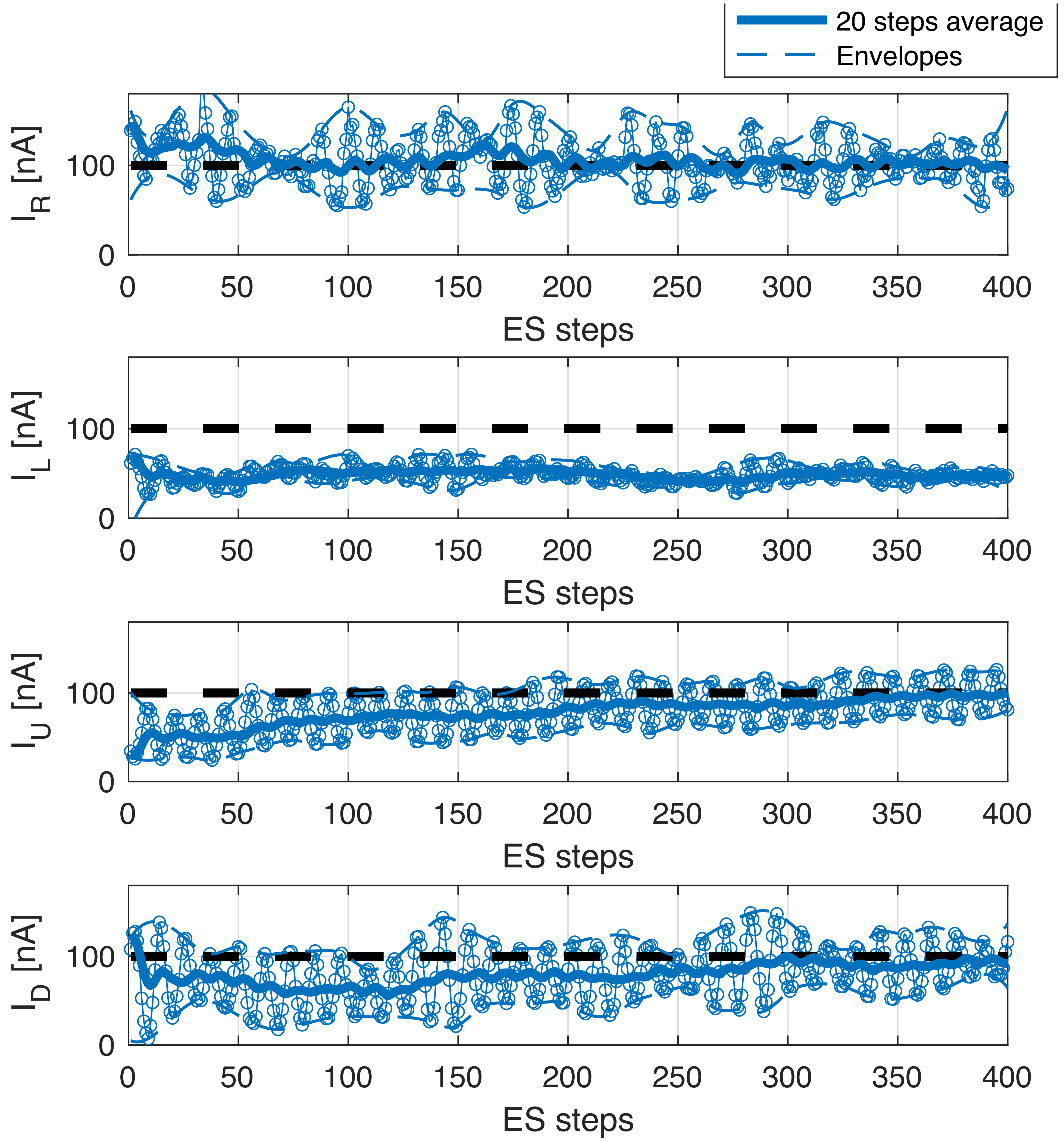


Corresponds to **Left** baffle current
Characterize smaller sensitivity

Motivates the adaptive weighting of cost function w_j

Adaptive weighting of objective function sensitivity

$$kC(\mathbf{p},t) = \sum_{j=R,L,U,D} S_j\left(J_j(\mathbf{p},t)\right) + w_j B_j(I_j(\mathbf{p},t))$$



Update w_j depending on the variance of the past objective function's fluctuation of N-data

Time-series observations

$$\mathbf{B}(I_j) = [B(0), B(T), B(2T), \ldots, B(NT)],$$

Simple moving average

$$\bar{\mathbf{B}}(I_j) = SMA(\mathbf{B}(\mathbf{I_j}))$$

Dither history on objectives

$$B_{fluc} = \mathbf{B_j}(\mathbf{I_j}) - \bar{\mathbf{B}}_\mathbf{j}(\mathbf{I_j})$$

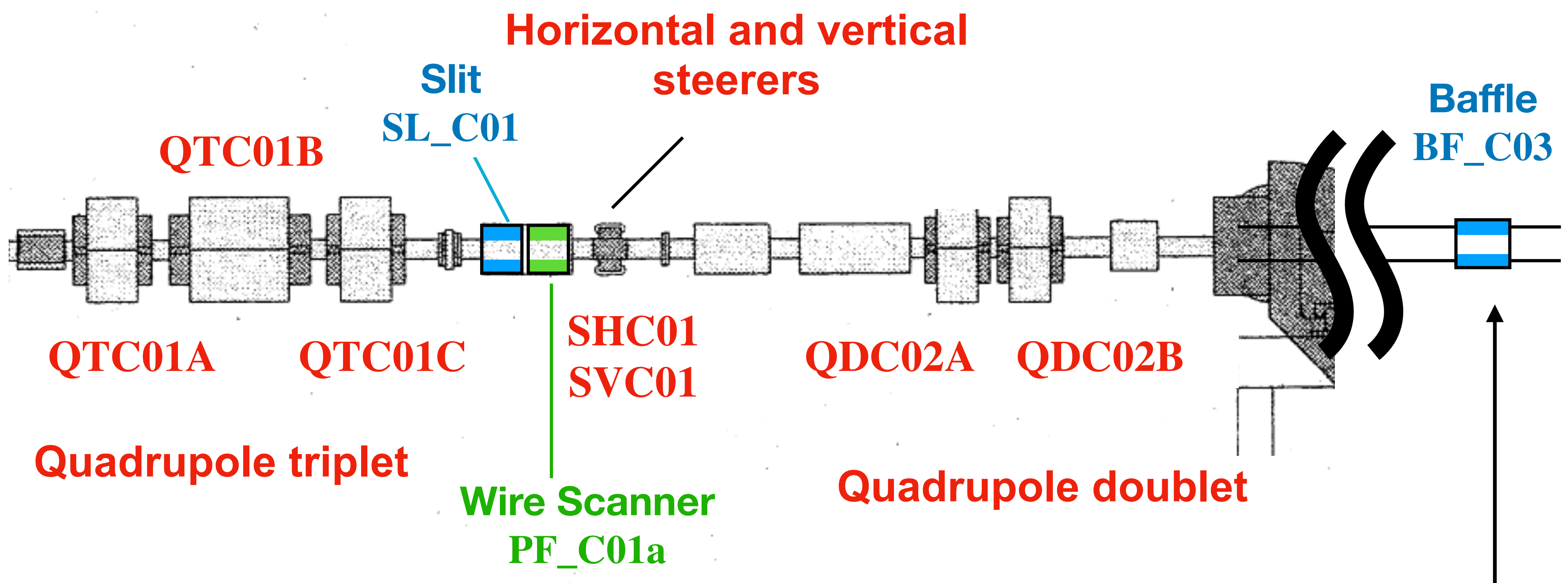
Determine each weight by variance of B_{fluc}

$$w_{new} = \sqrt{\frac{0.5^2}{\text{Variance of } B_{fluc}}}$$

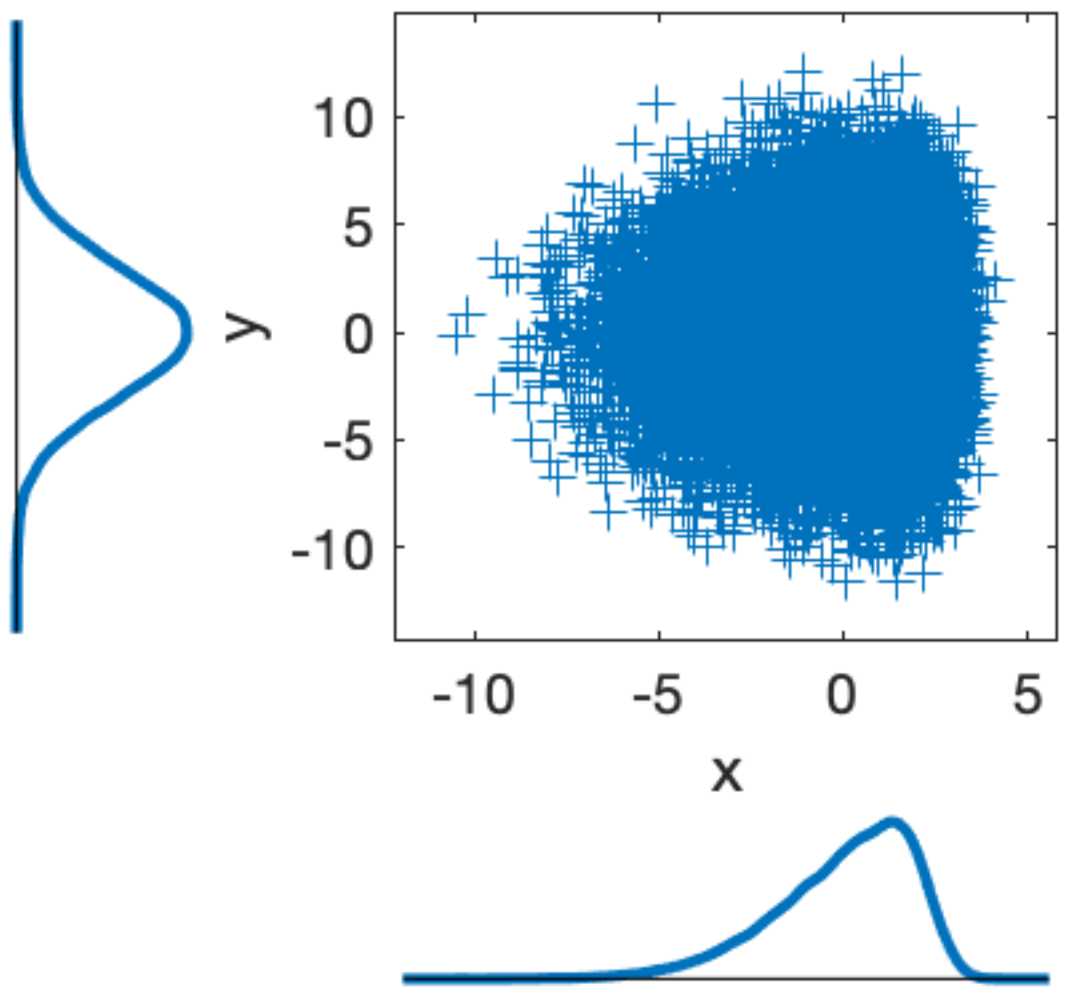
Simulation results for the setup similar to the experiment

$$kC(\mathbf{p},t) = \sum_{j=R,L,U,D} S_j(J_j(\mathbf{p},t)) + w_j B_j(I_j(\mathbf{p},t))$$

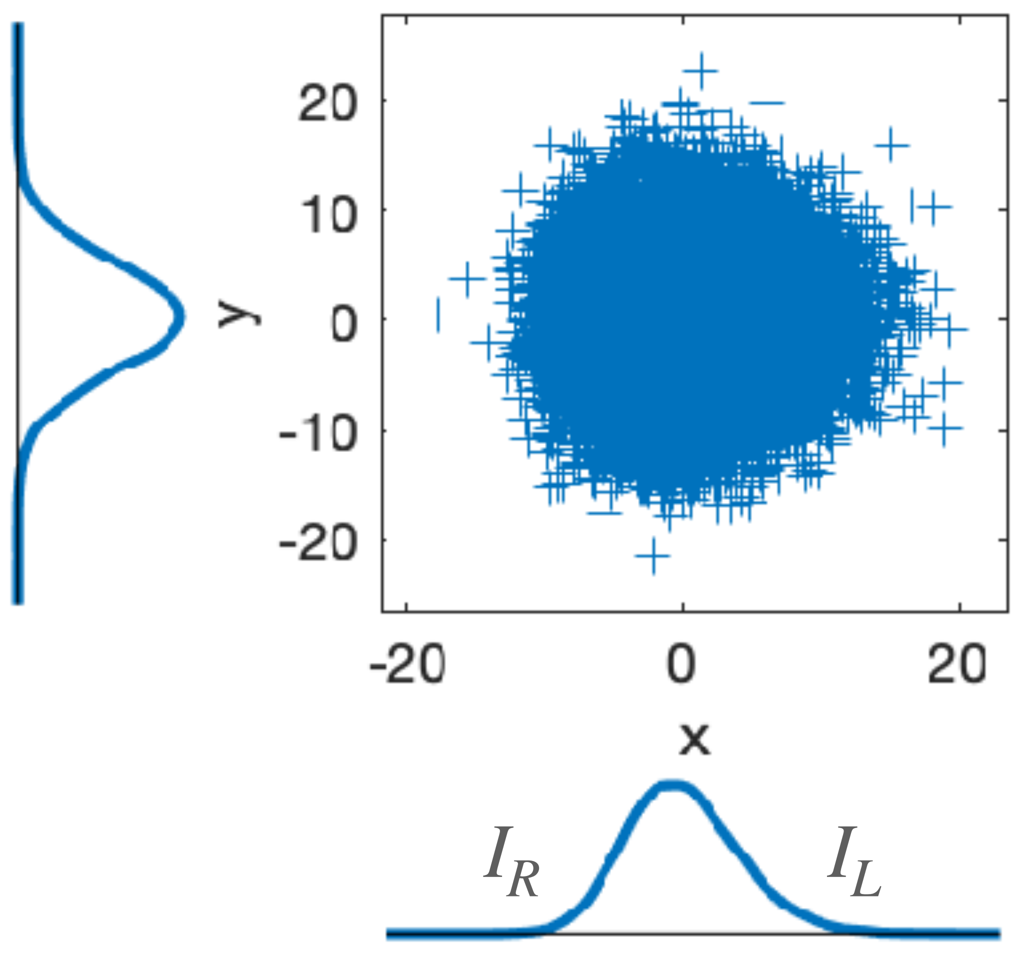
Generated horizontally skewed-gaussian profile



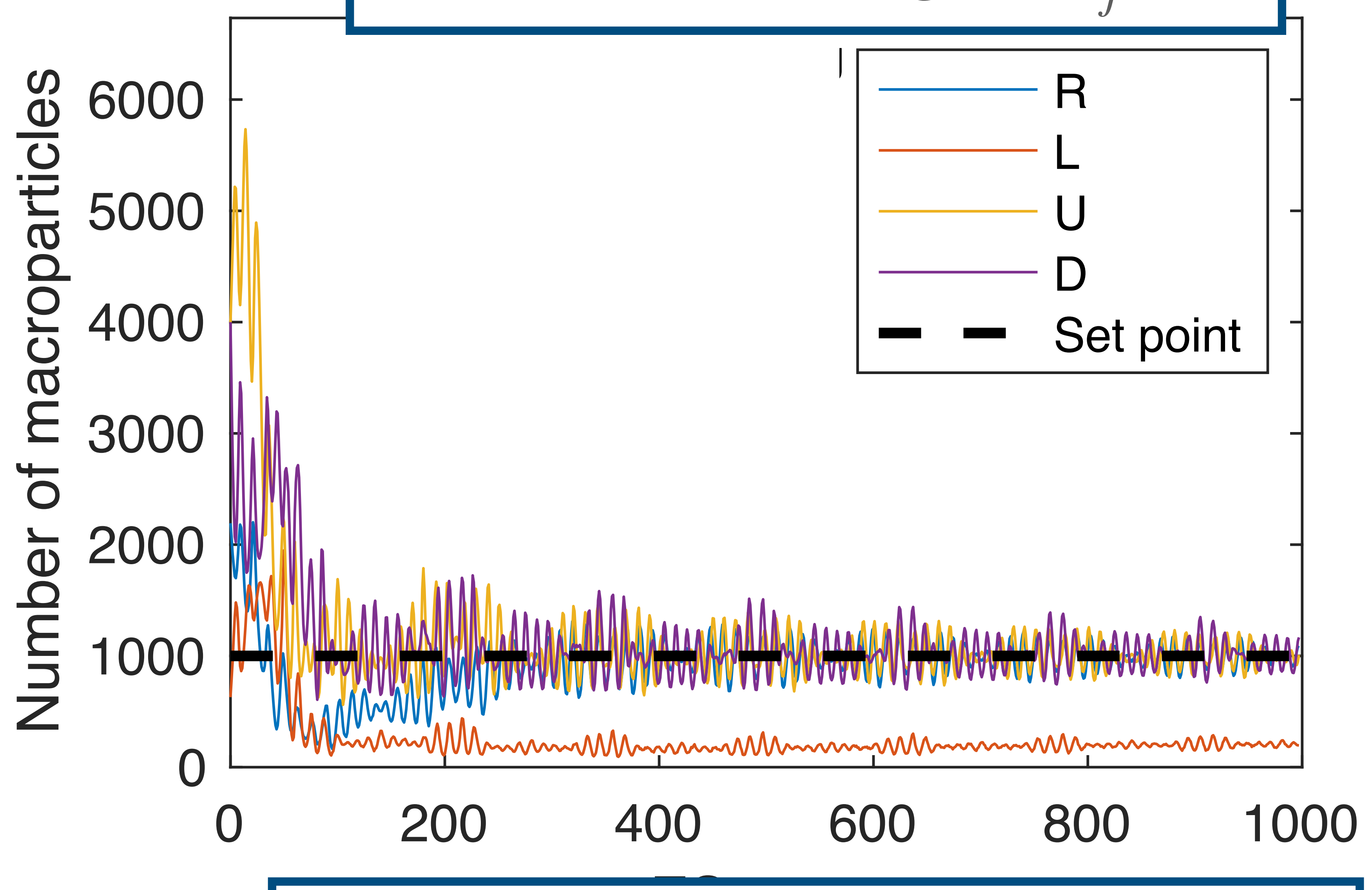
Entrance



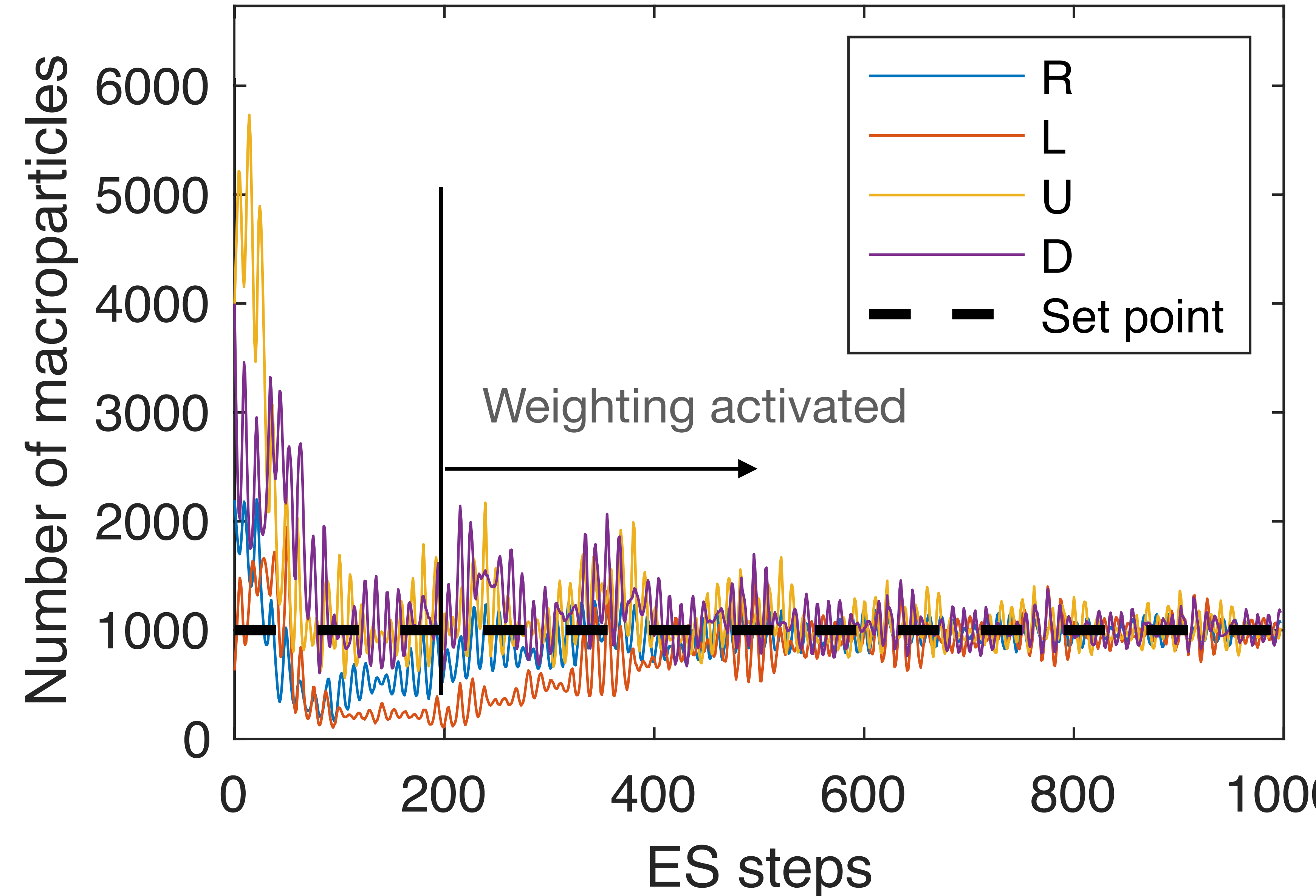
Baffle position



No update in weights $w_j = 1$



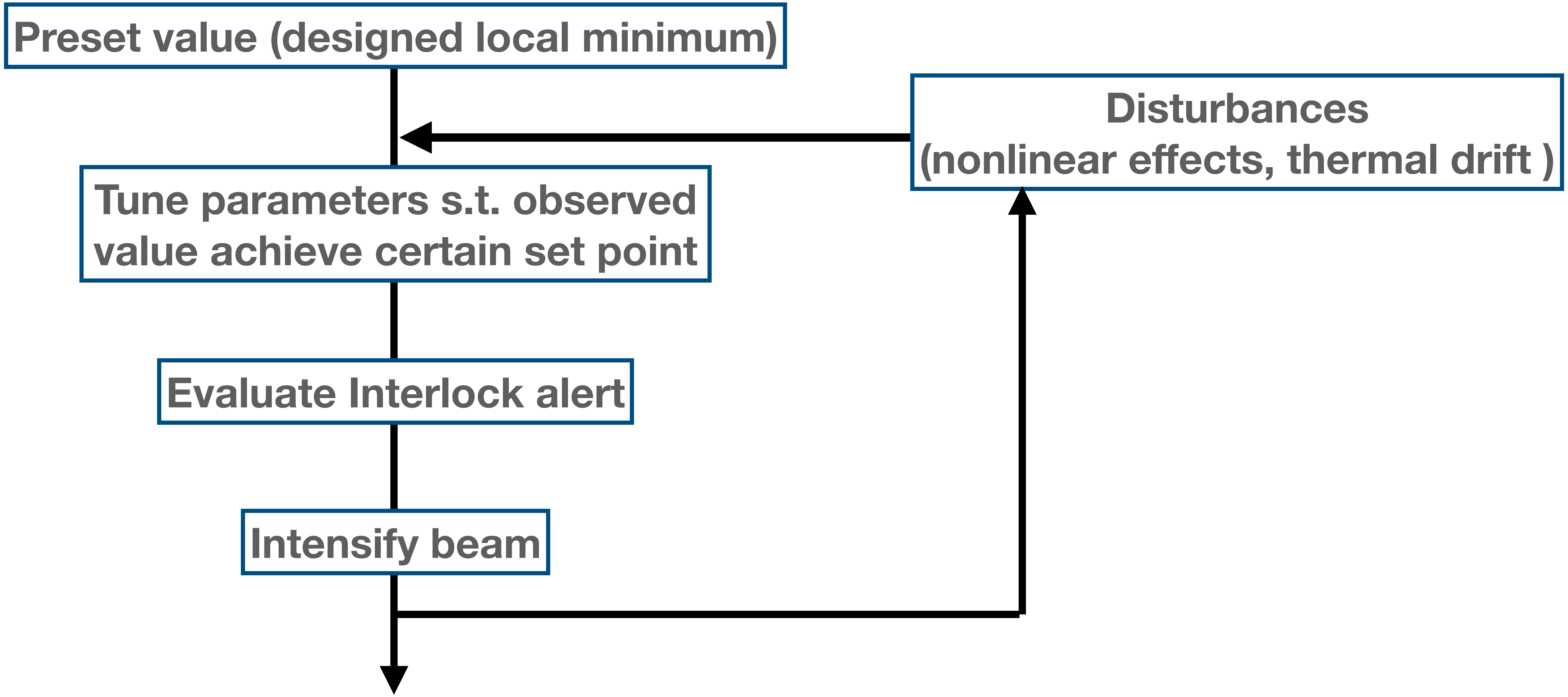
Adaptive weighting in $w_j = \sqrt{0.5^2/\sigma^2_{B_{fluc}}}$



Summary

- Introduced adaptive control scheme for the accelerator tuning
- In the experiment, cost function being minimized towards the local minimum even under sudden state change
- New adaptive weighting scheme for the gains in multi-objective function case proposed and tested via simulation.

Scenario in intensifying HIB tuning

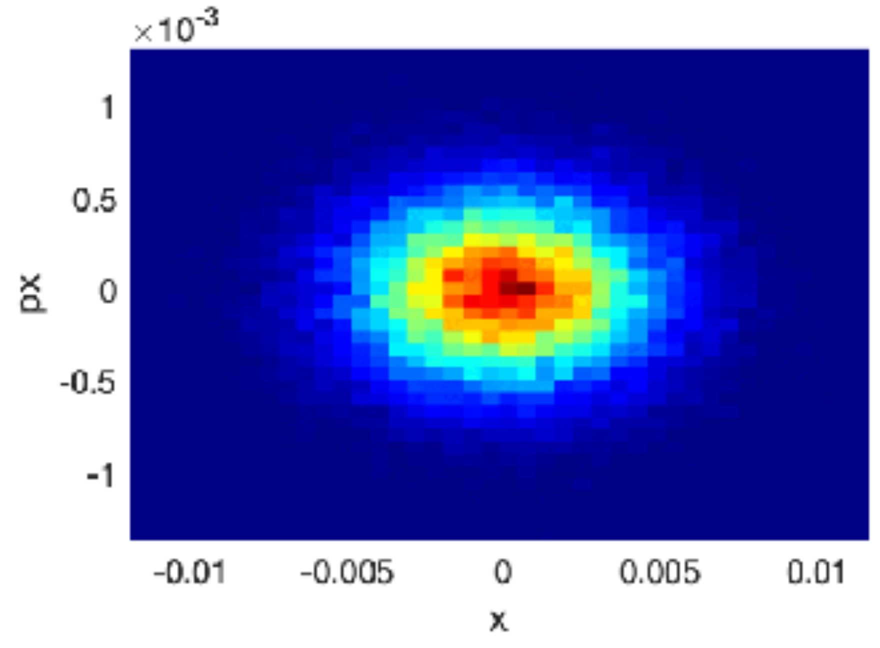


- Intensifying the beam current induces various effects
- Machine state change over time due to hardware instabilities and thermal effects

Compensation of time-varying distribution with ESC

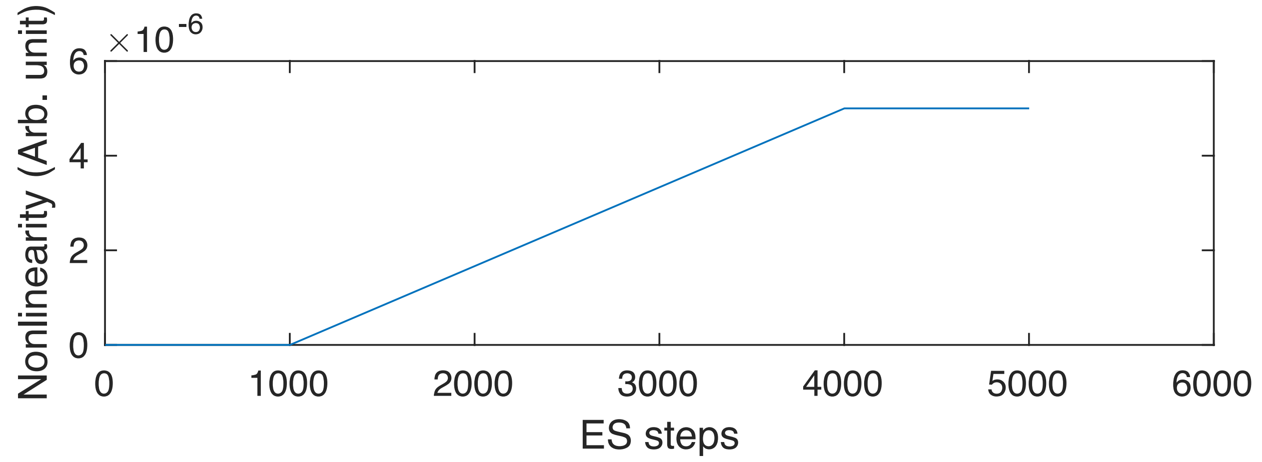
Time varying response to the nonlinear effect such as space-charge when intensifying beam

Beam evolution of x-px due to space charge effect in time

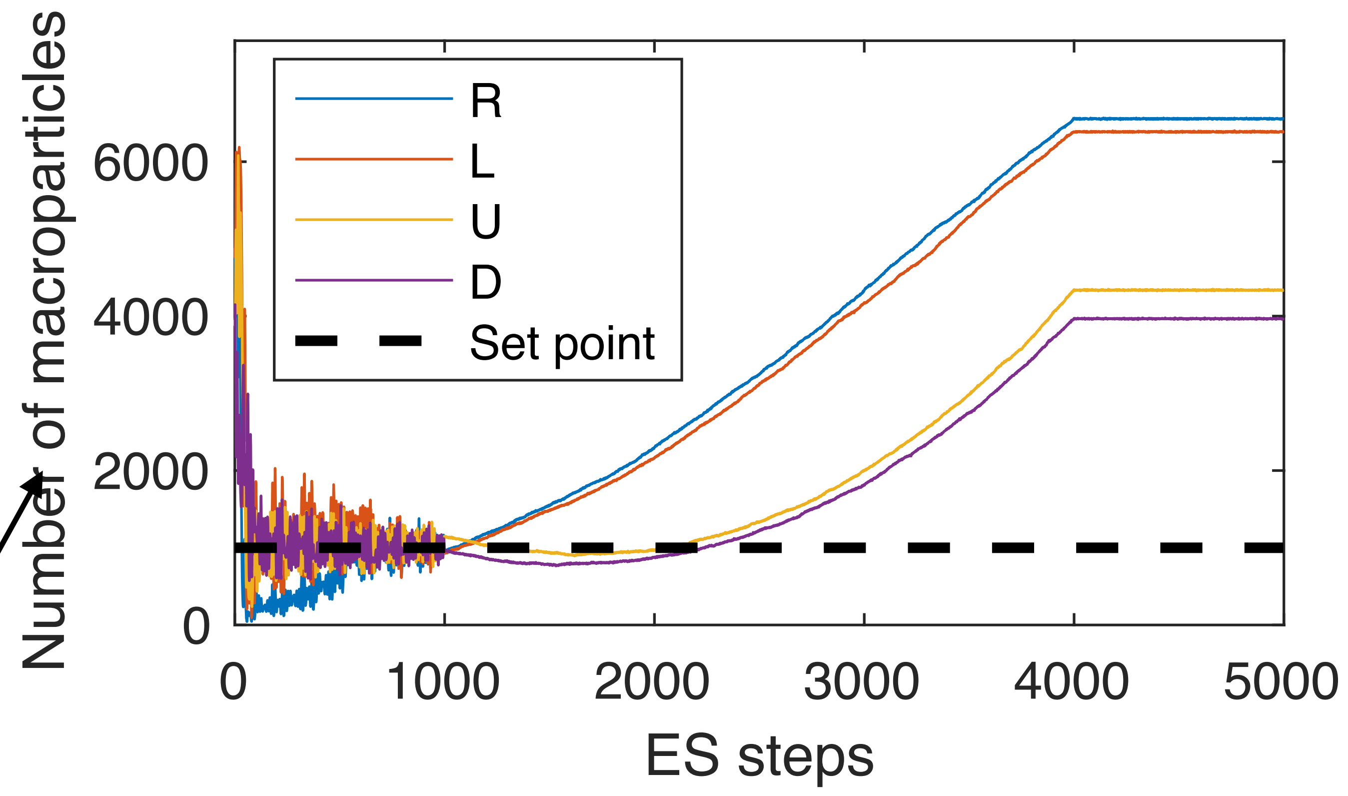


$$F_r = F_{r0} \times \frac{1 - e^{-r^2/2\sigma_r^2}}{r} = \frac{q^2 n_e}{2\pi\epsilon_0\gamma^2}\frac{1 - e^{-r^2/2\sigma_r^2}}{r}$$

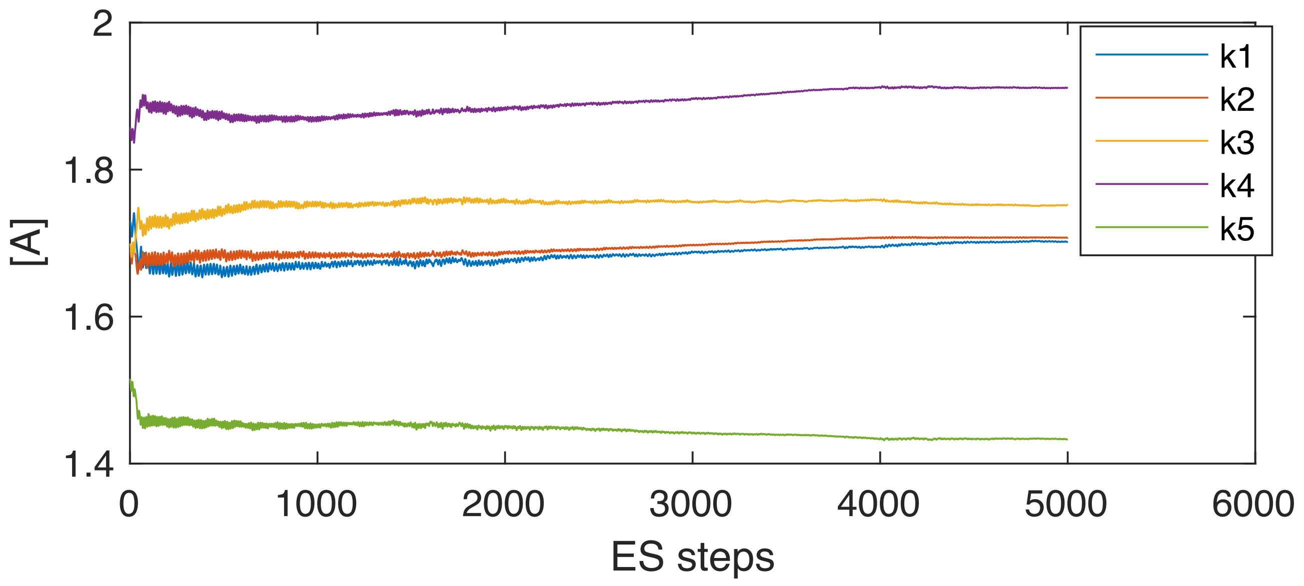
Space charge factor F_{r0}

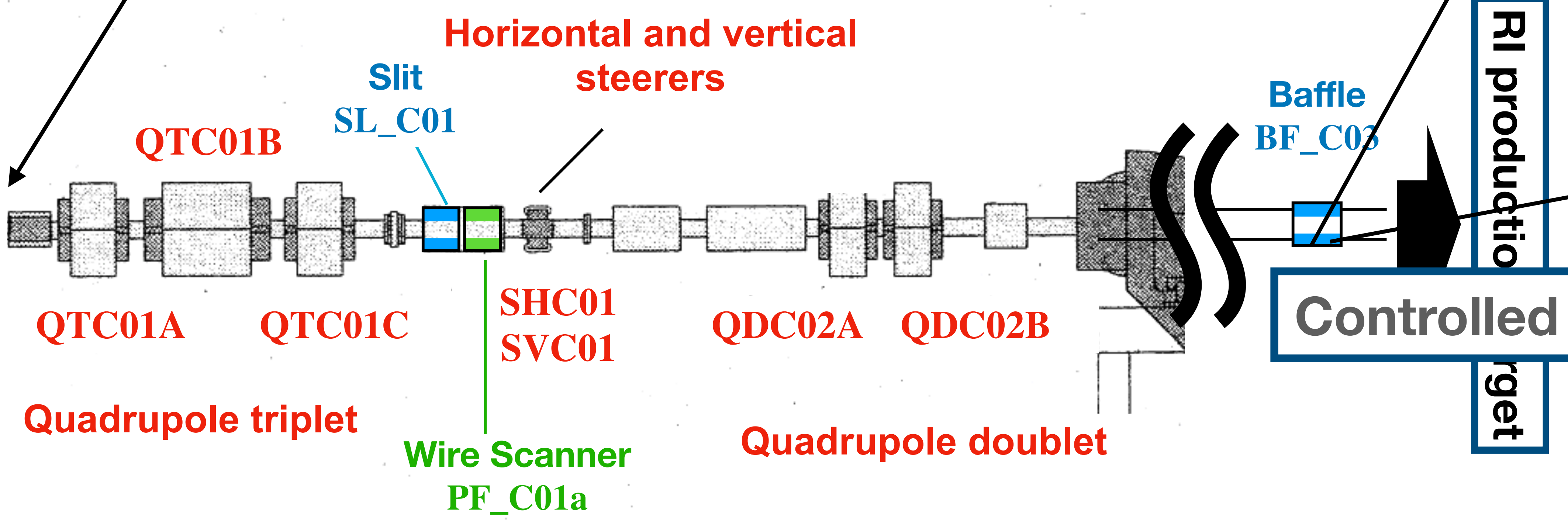


Unontrolled

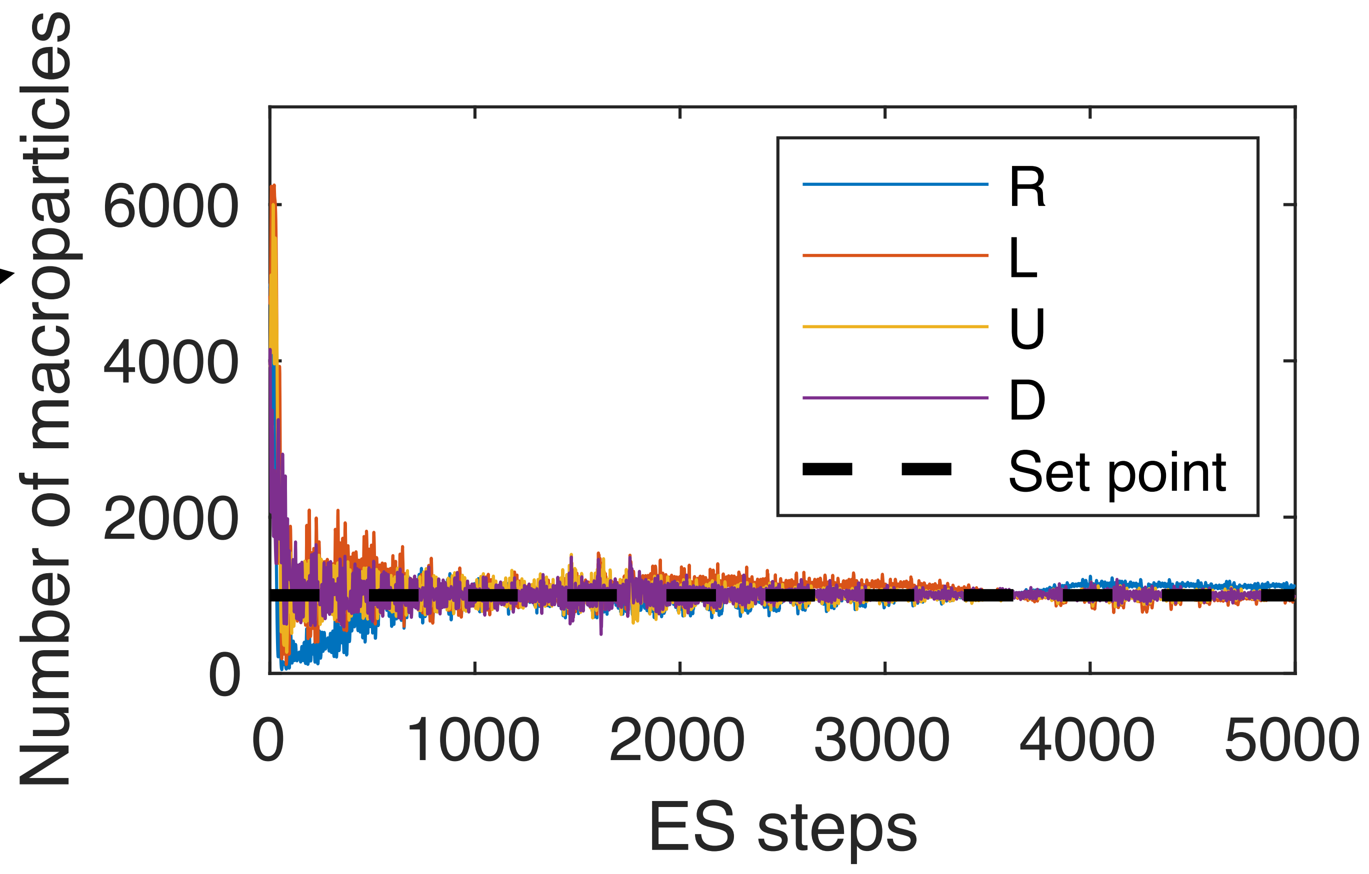


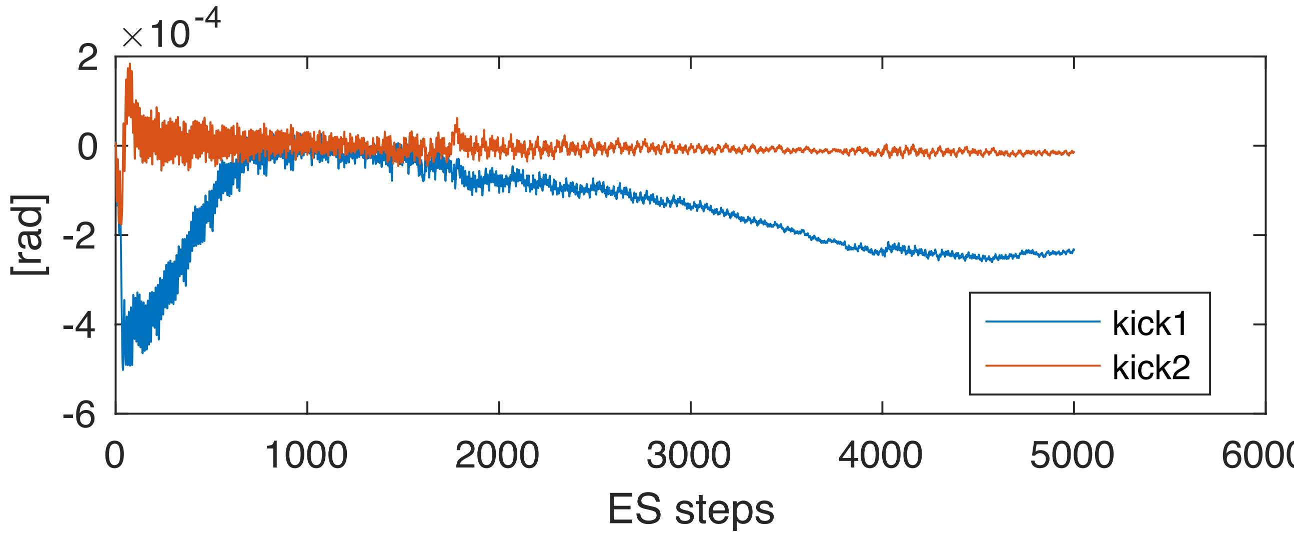
7 input params being tuned





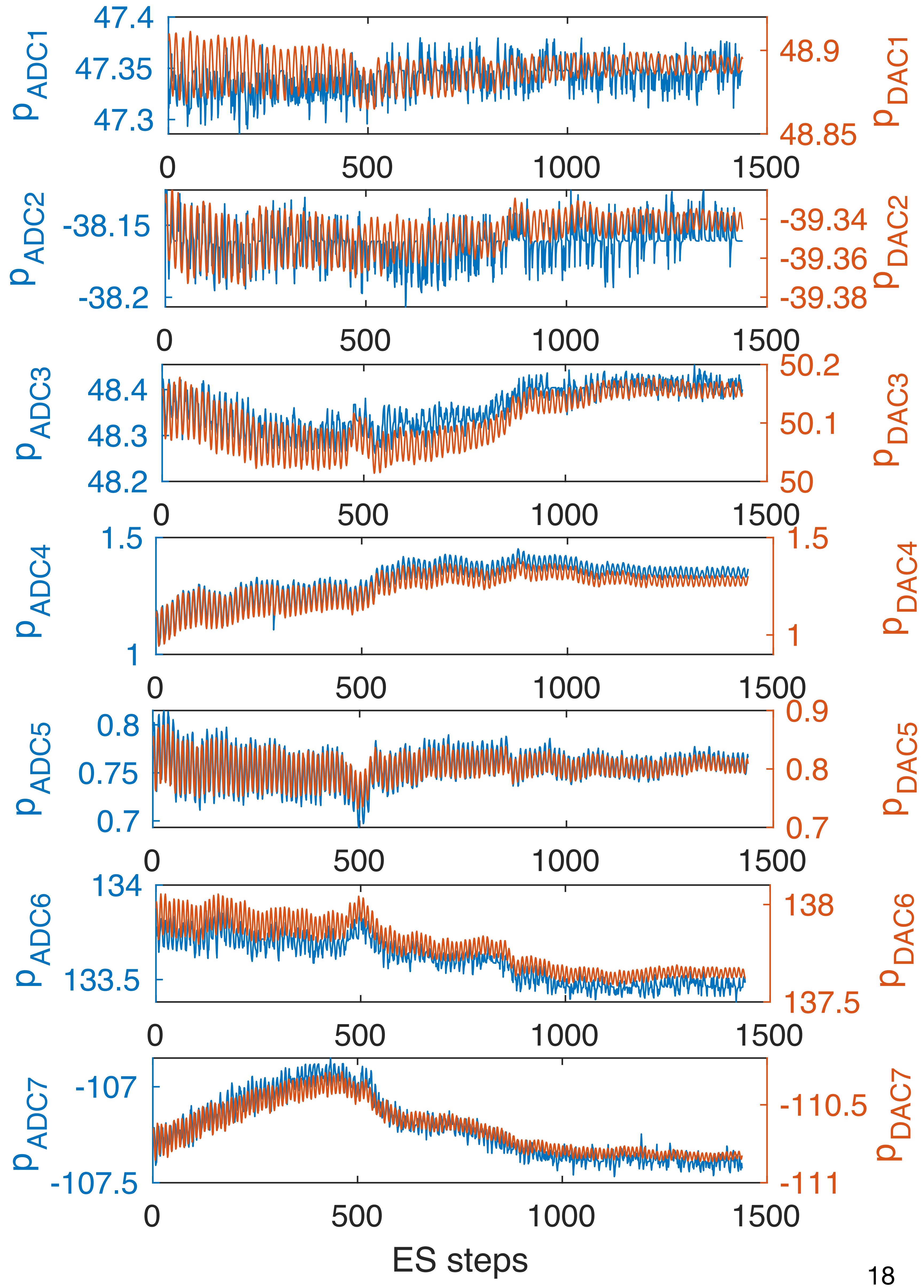
Controlled





DAC Communication error problem @ Riken

Experimental observation of tuning parameters in [A]



ESC performance would have been better if there's no communication error…

Parameters and Cosine similarity index as a result of experiment

Index of how similar two waveforms are?

$$sim_i(\boldsymbol{p}_{ADC_i}, \boldsymbol{p}_{DAC_i}) = \frac{\boldsymbol{p}_{ADC_i} \cdot \boldsymbol{p}_{DAC_i}}{\|\boldsymbol{p}_{ADC_i}\| \|\boldsymbol{p}_{DAC_i}\|}$$

Table 1: Cosine similarities between the ADC and the DAC current readouts of the power supplies.

Index i	Magnet name	Function	Similarity sim_i
1	QTC01A	QT cell	0.16
2	QTC01B		0.57
3	QTC01C		0.86
4	SHC01	Horizontal deflection	0.91
5	SVC01	Vertical deflection	0.76
6	QDC02A	QD cell	0.59
7	QDC02B		0.88

→ Not effective for the full performance of ESC operation

Some controllers cannot handle multiple DAC signals from EPICS' IOC at the same time.

Hyper parameter tuning procedure

$\frac{dp_j}{dt} = \sqrt{\alpha\omega_j}\cos[\omega_j t + kC(\mathbf{p}, t)]$, Define ω_j to equally span between 1 and 1.75

Choice of oscillation amplitude $\sqrt{\alpha\omega_j}$

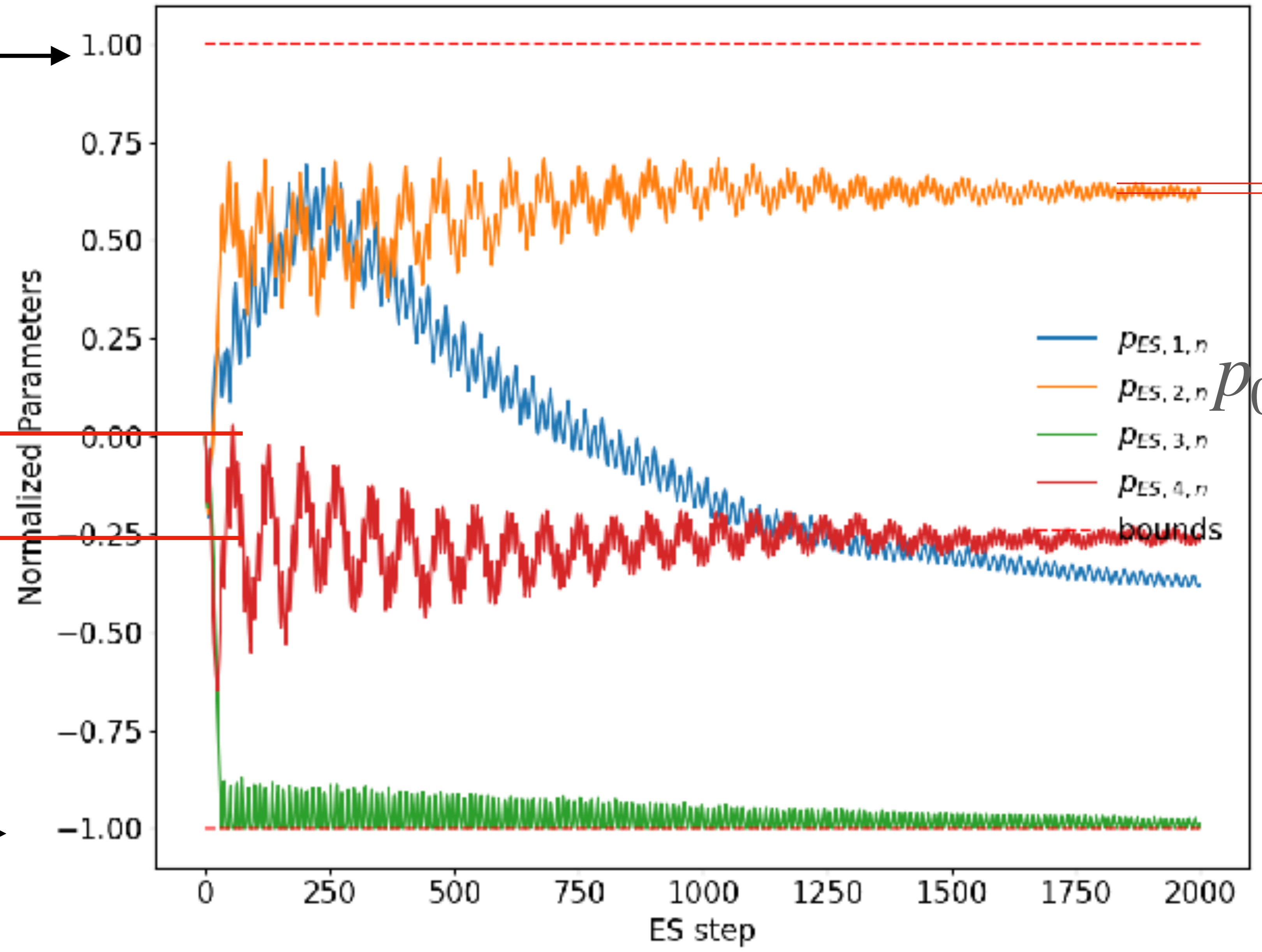
Decide p_{max} and p_{min} to limit the applicable input parameter within safe set.
These should be large enough effect on cost functions

Normalized to p_{max}

Initial oscillation amplitude

$$p_0 = 0.1 \times \frac{p_{max} + p_{min}}{2}$$

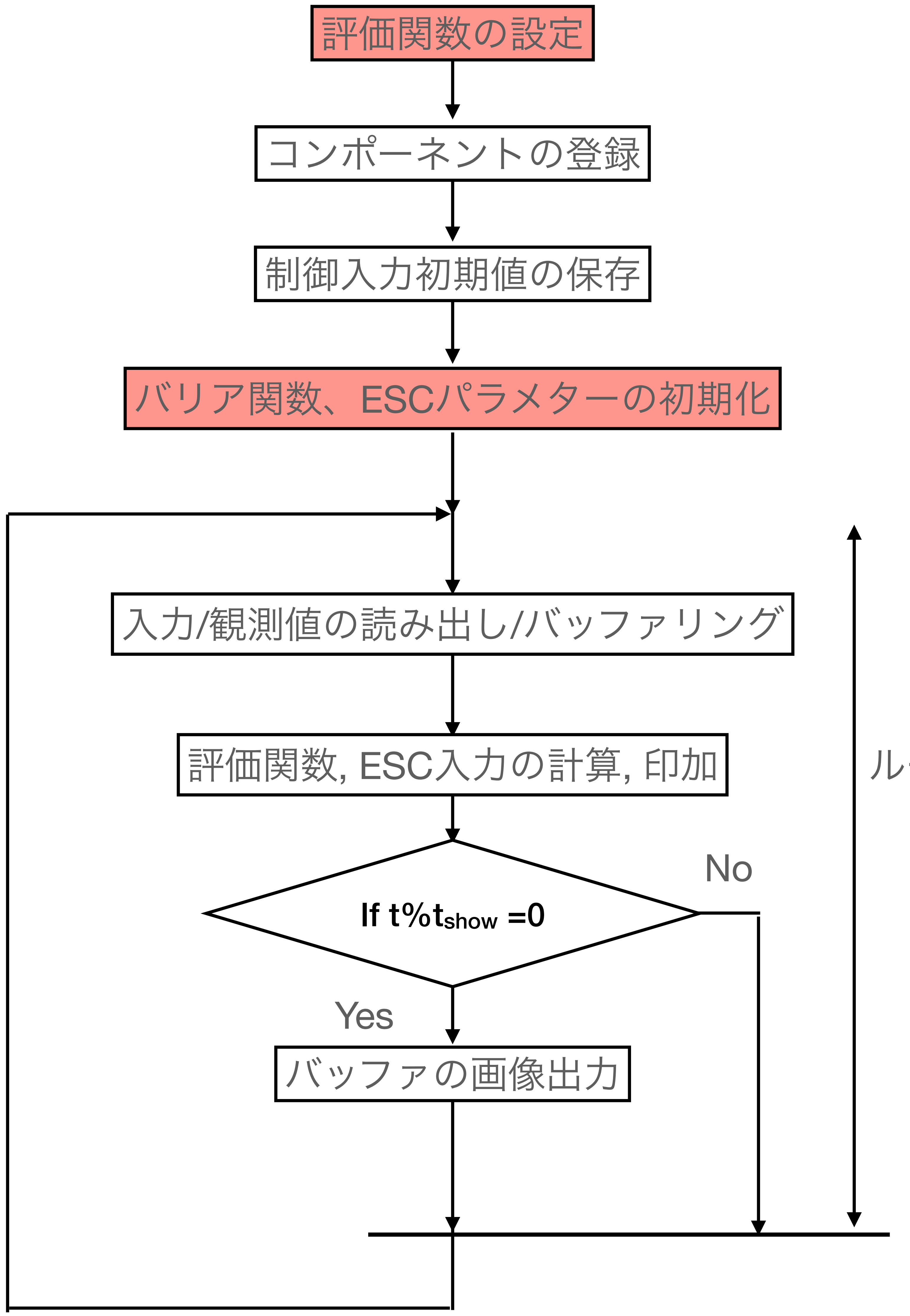
Normalized to p_{min}



Time decaying oscillation amplitude

$p_0 \exp(-\gamma t_{ES}), \quad \gamma = 0.999$

メインループの構成



評価関数の設定

ユーザーが観測出力に基づく評価関数Cを事前に定義する

```
########################################################
# Define ES cost function and use this class
########################################################
class ES_Algorithm_User(ES_Algorithm):

    def f_ES_minimize(self,p,q):
      f_val = 0
#####  Describe  evaluation function here
      return f_val
```

評価関数へのゲインkおよび揺動信号初期値の定義

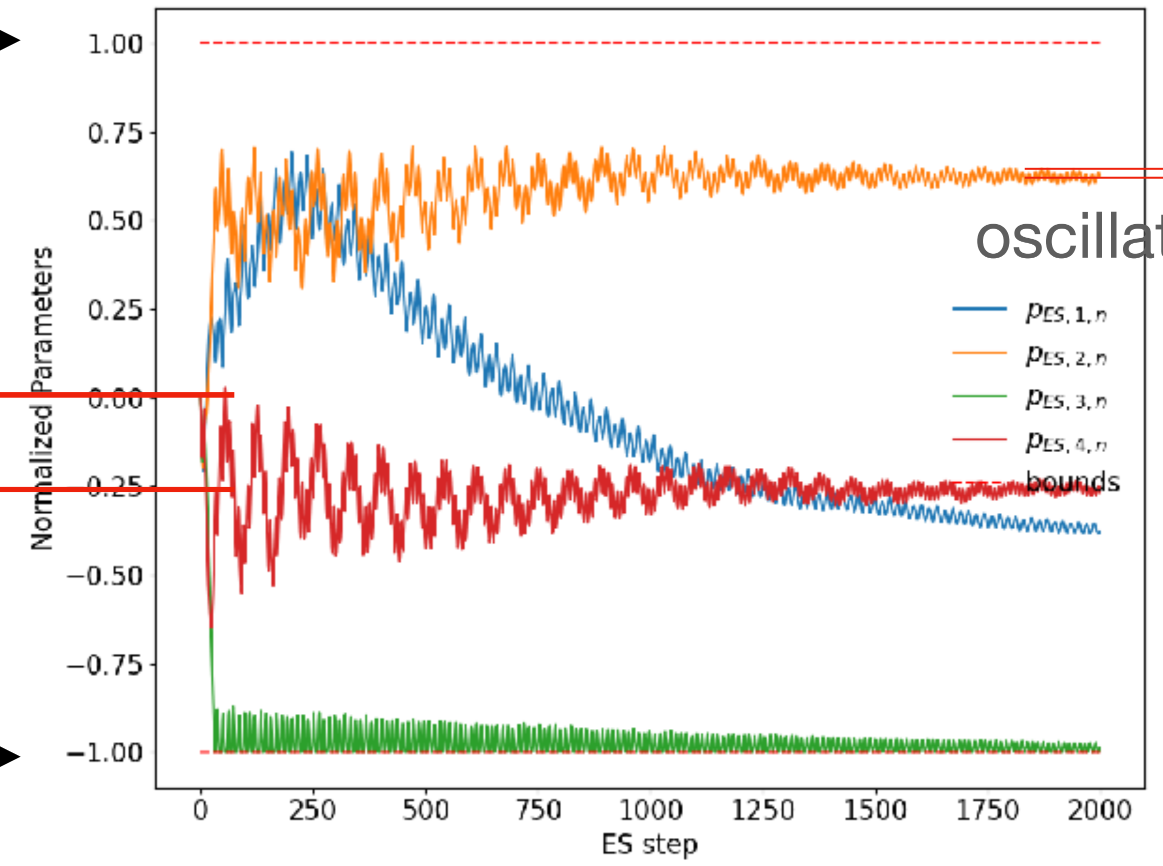
```
# in main function

#################################################################
#### Initialize ESC control class
#################################################################
  ES_steps = 110
  # Upper bounds on tuned parameters
  p_max = np.array(list(esc_input_dict.values())) + ps_allowable_diff_list
  # Lower bounds on tuned parameters
  p_min = np.array(list(esc_input_dict.values())) - ps_allowable_diff_list
  # kES needs some trial and error,  kES times B should be around the order of 100
  kES = 1e-1
  # the parameters to have normalized osciallation sizes you choose the aES as:
  oscillation_size = 0.1
  # decay_rate of input. signal decays as exp( -at)
  decay_rate = 0.999
  # initiate ES algorithm class instance
  es_class = ES_Algorithm_User(ES_steps,p_max,p_min,kES,oscillation_size)
```

p_maxに規格化

Oscillation_size=0.1
初期の揺動振幅

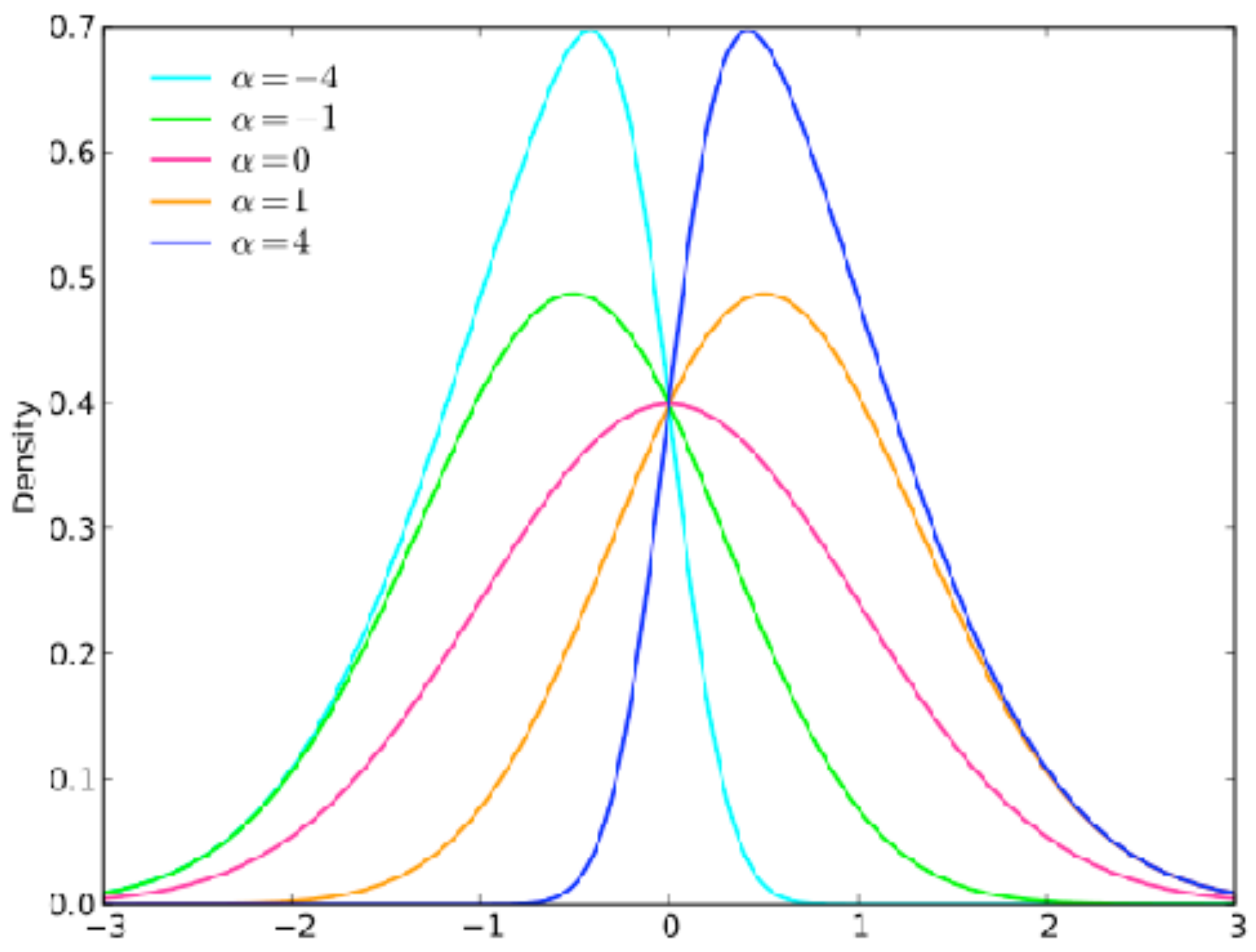
p_minに規格化

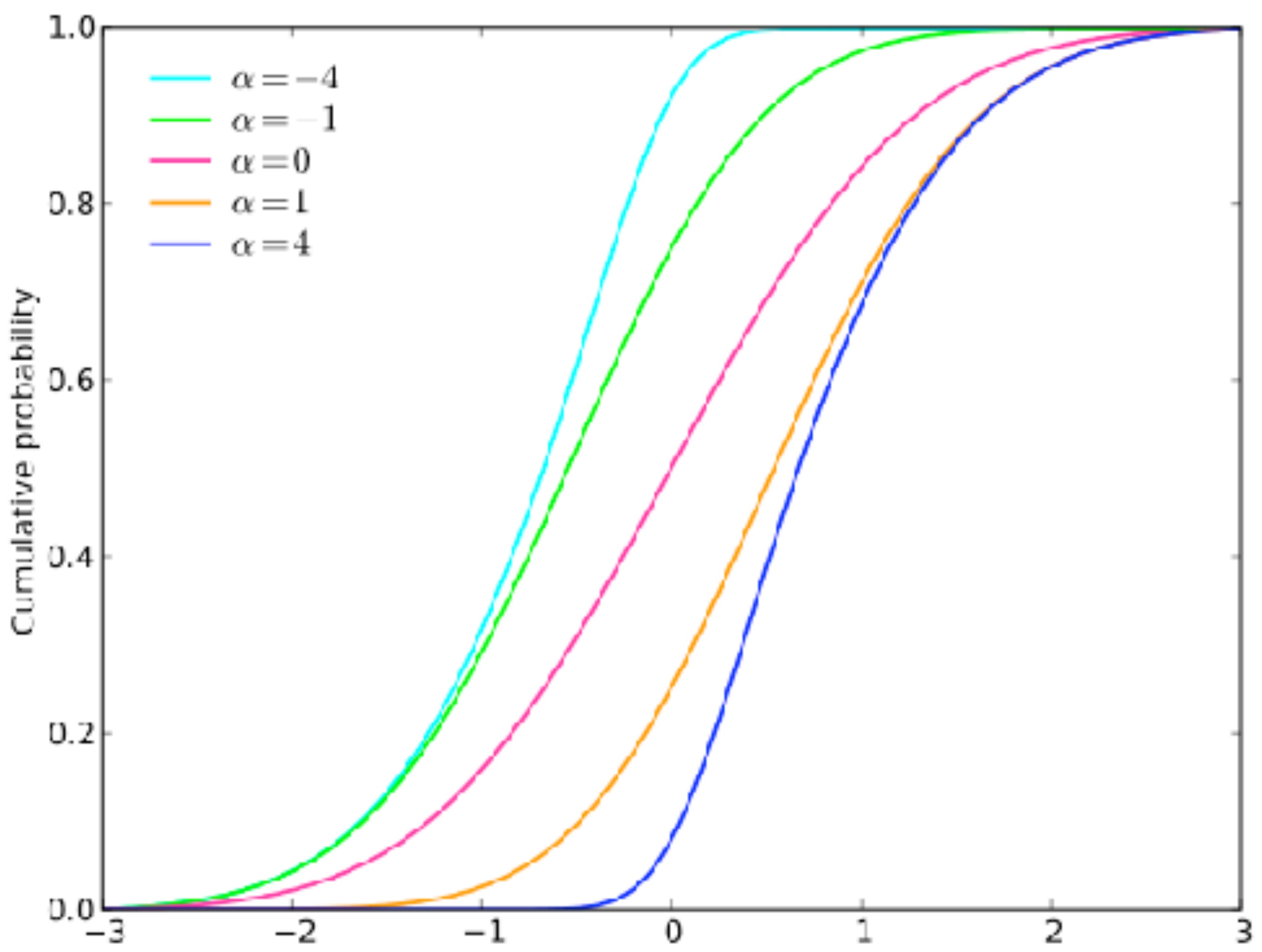


終端の揺動振幅

oscillation_size*exp(-decay_rate*ES_step)
= 0.1×exp(-0.999×2000)

Skewed gaussian profile used in the simulation





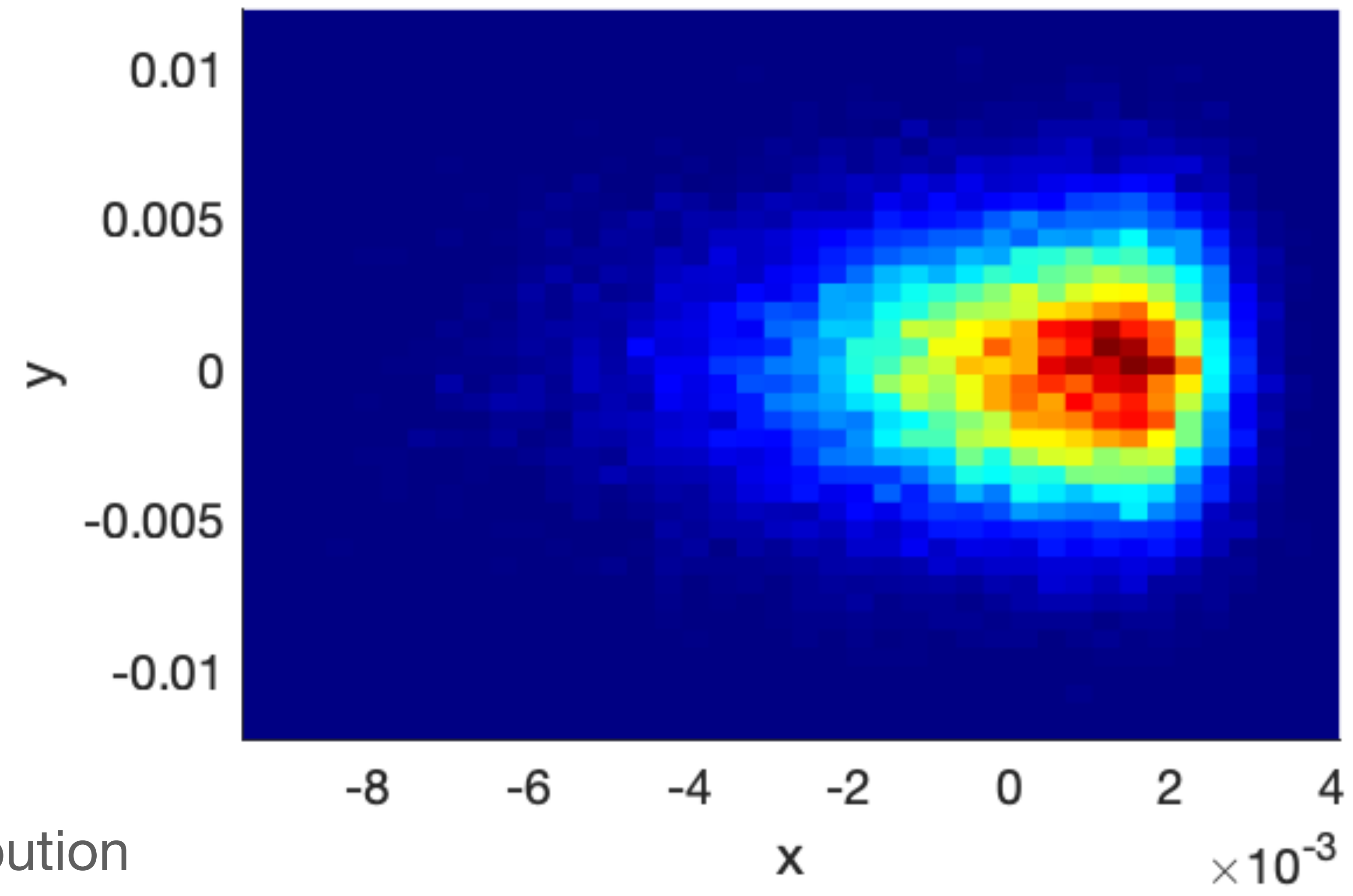
Gaussian dist. $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Skewing $\Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)\right]$

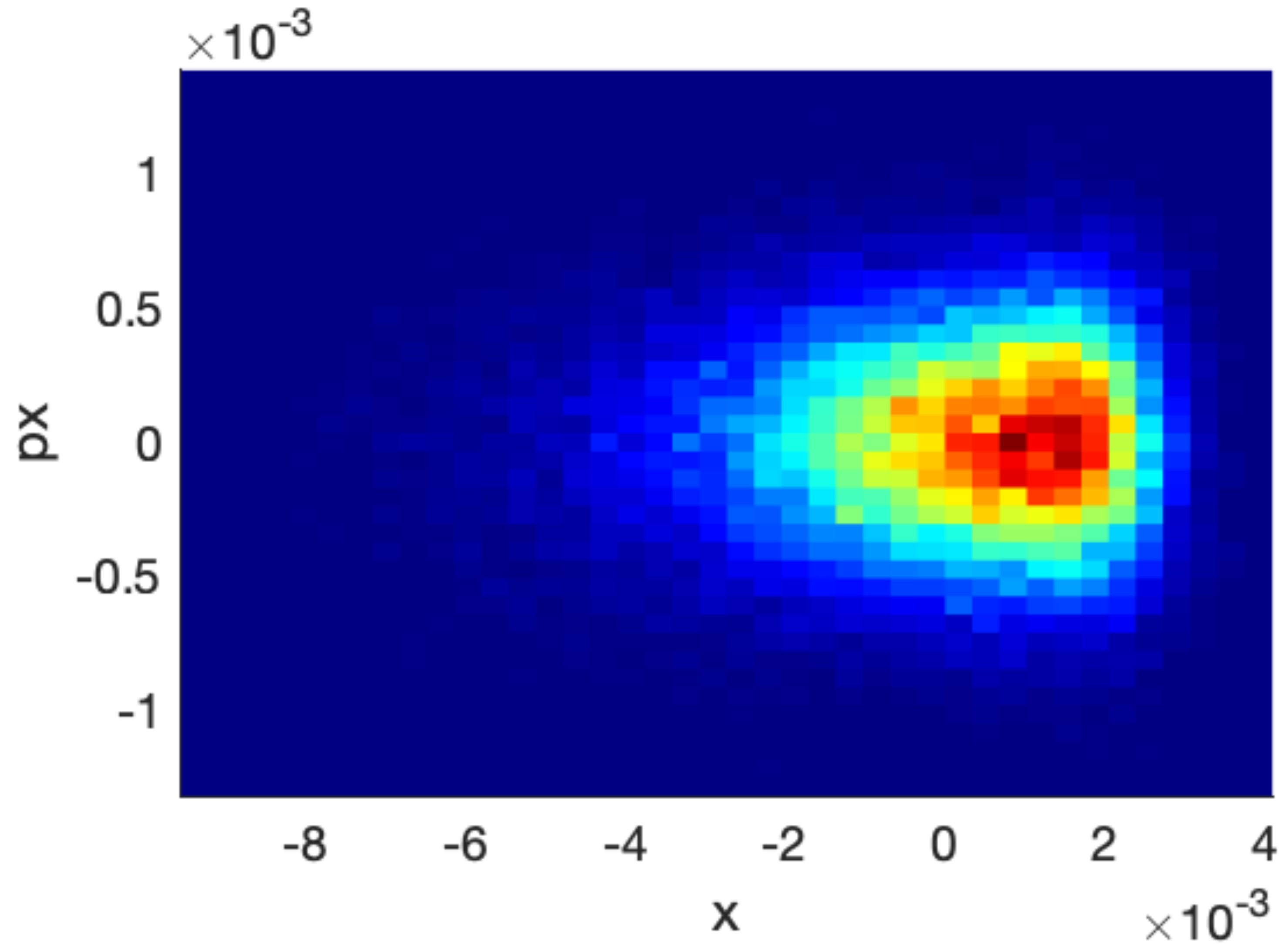
PDF $f(x) = 2\phi(x)\Phi(\alpha x).$

α=-6, applied to the profile in x. Assumed px is gaussian.

x-y projection



Initial X-px



The figures fetched from Wikipedia : Skew normal distribution
https://en.wikipedia.org/wiki/Skew_normal_distribution