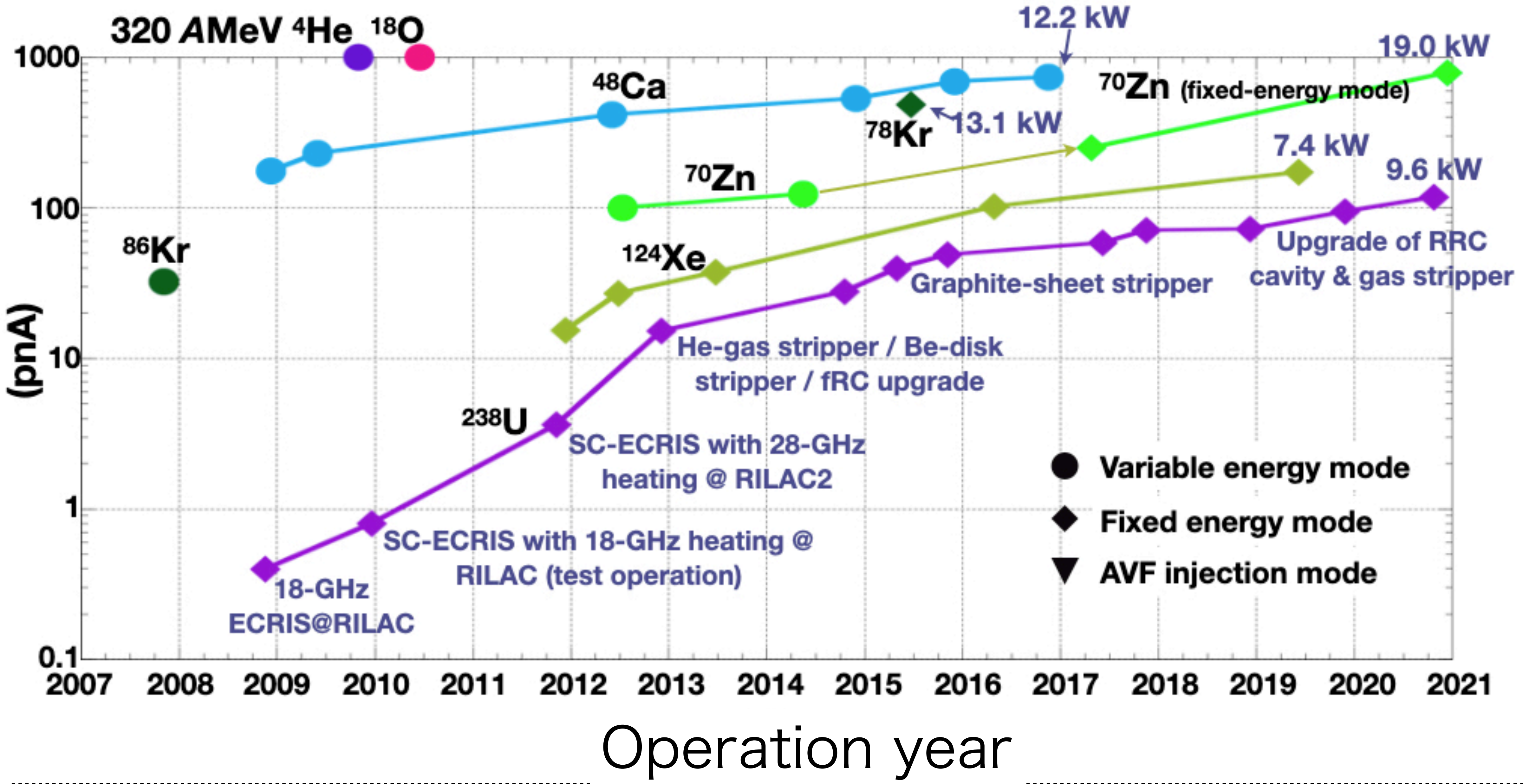
ML based accelerator operation at RIKEN

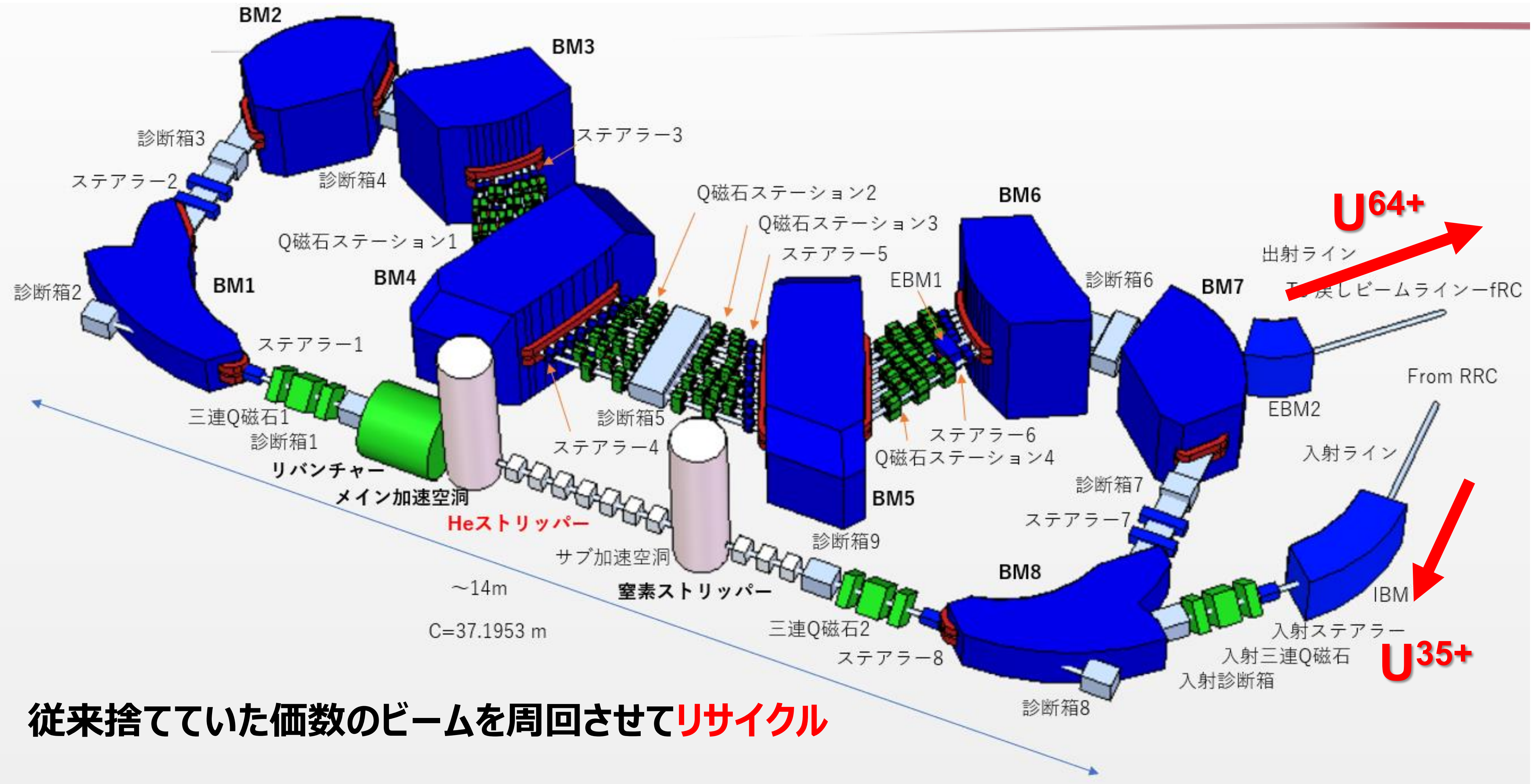
RIKEN Nishina Center for Accelerator-Based Science
Takahiro Nishi

Current status and upgrade plan of RIBF facility

Upgrade of Uranium beam intensity



荷電変換リングのデザイン図



Key concept: **Charge Stripper Ring**

Beam current is limited by low efficiency of charge stripper (~20%).

→ Beams in other charged states are recycled and re-injected into the gas stripper.

2008 → 2014: ×70

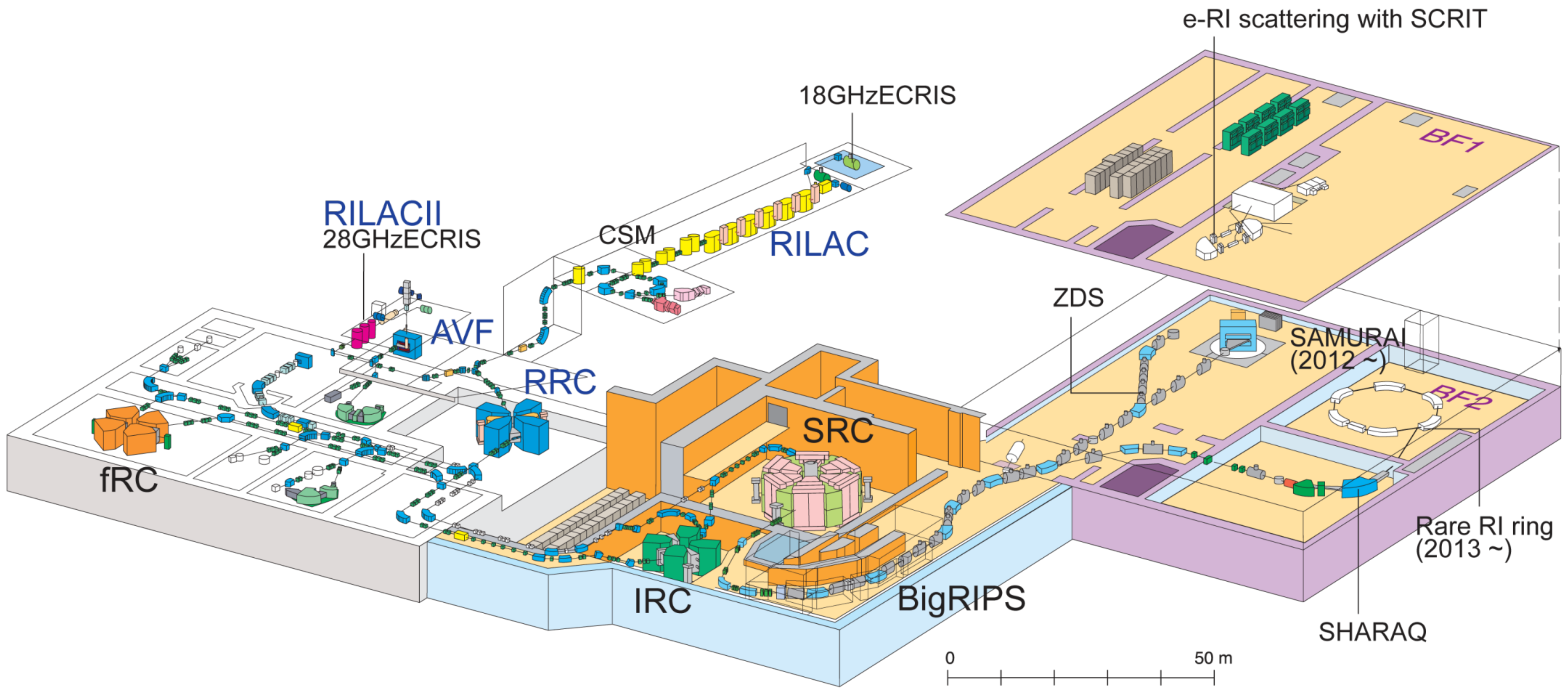
2014 → 2020: ×4

Limitation of human optimization

2020 → 2030: **×10 by ML + hardware upgrade**

Machine Learning Application for RIBF (Optimization)

□ Transport line (beam optics optimization)



Future goal : **> 1 pμA** $^{86+}$U beam (×10 times upgrade)

⇔ local beam loss ~ **0.2%** can destroy the facility

To realize high intense $^{86+}$U beam
(> 1puA / 100 kW), we should
- **suppress beam loss to 0.1 ~ 1%,**
- **continuous adjustment**

Constant and automatic optimization of beam transport parameters

- more than 600 control params.
- other environmental params.

Control by ML ⇔ **Comprehensive model using Simulation**

- non-destructive monitor
- radiation monitor
- detectors at downstream

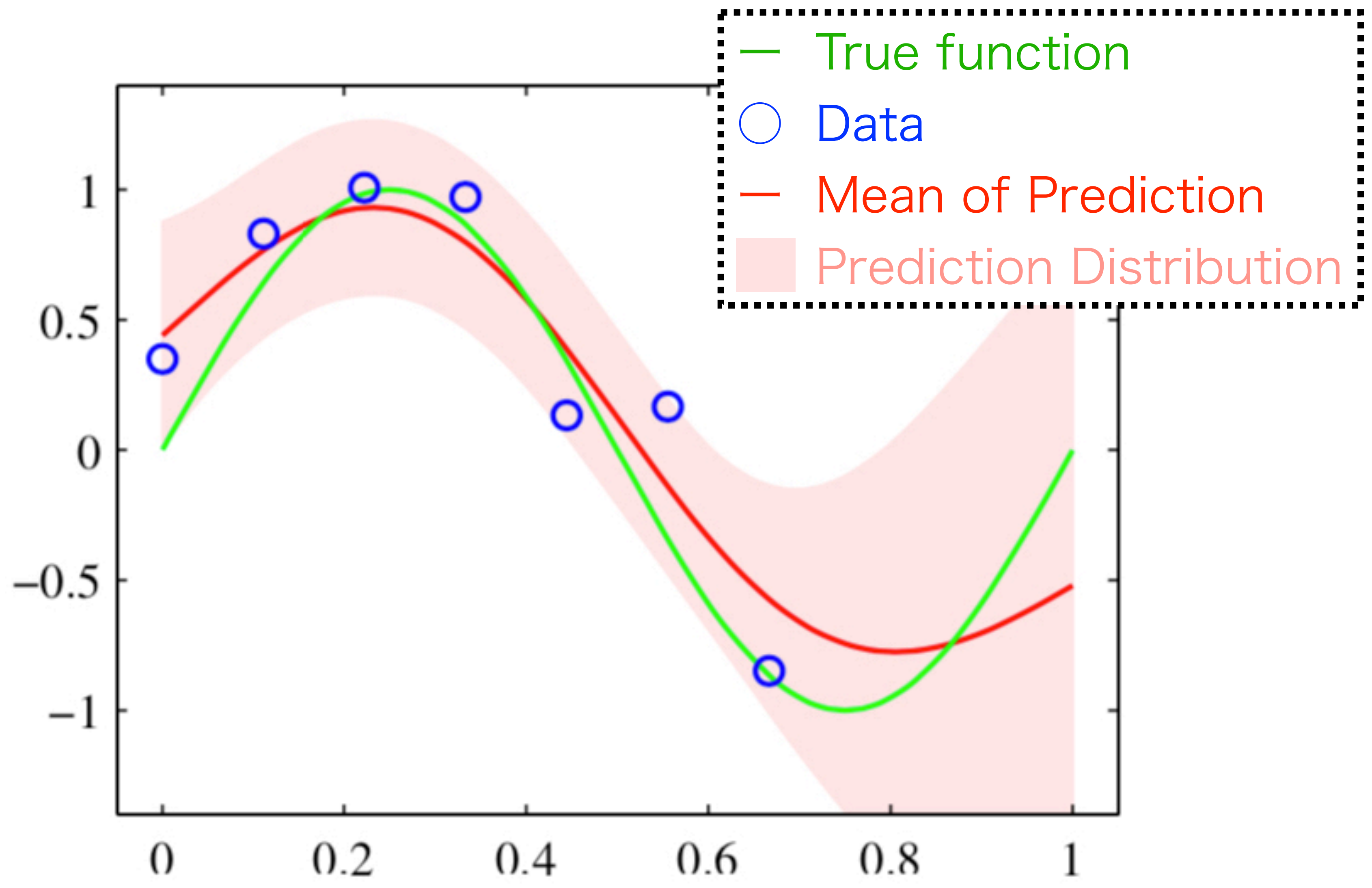
Keyword to Realize Auto Optimization of Beam Transport

Optimization method	・Manual control → auto control system via EPICS ・Optimization of multidimensional parameters ・Relatively small data set (as for DNN)
Beam monitor	・Monitor for transmission and beam spot ・Availability for high intensity beams
Safety	・Safety system to stop the beam immediately (Beam Interlock System) ・Avoiding high risk parameters during optimization

Optimization Method: Auto Optimization using Gaussian Process Regression

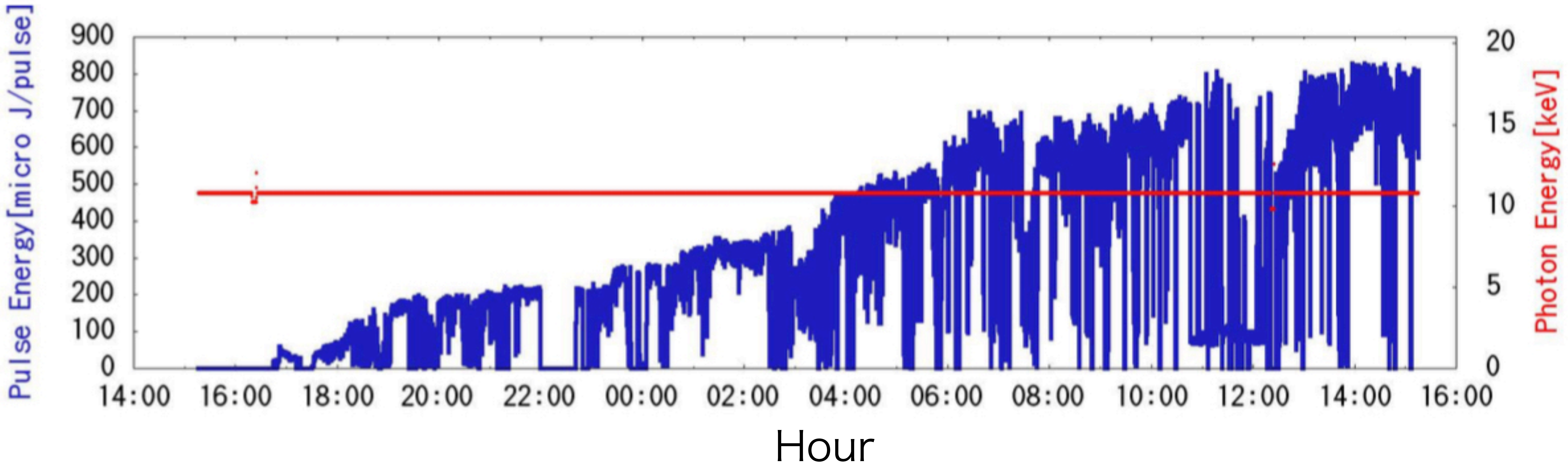
Developed / applied for usual manipulation in SACLA @ SPring-8

Gaussian Process Regression



C. M. Bishop, Pattern Recognition and Machine Learning, Springer (2006)

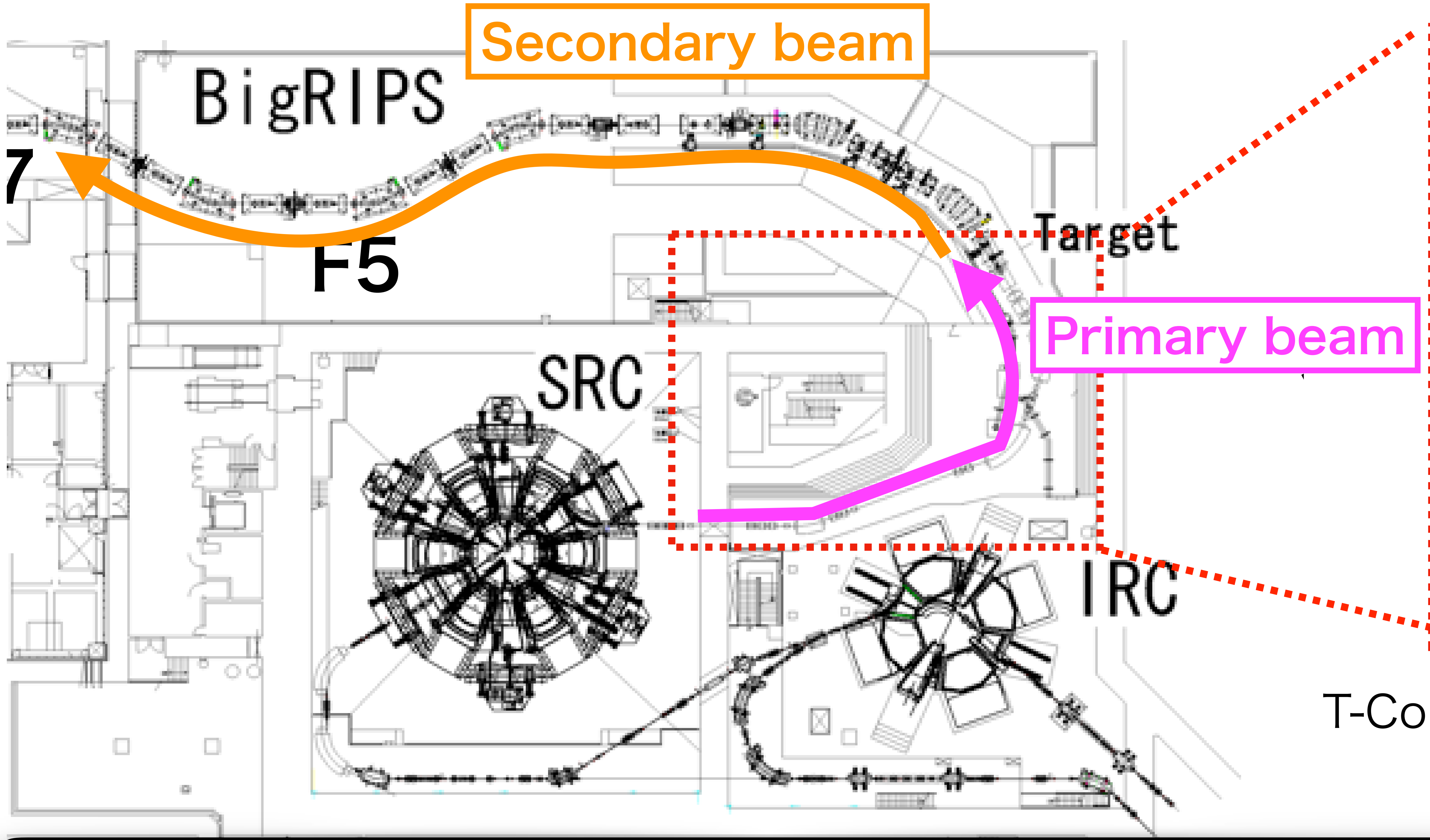
Auto tuning of XFEL@SACLA using GPR
Ex. of optimization of the leaser intensity with 15 param.

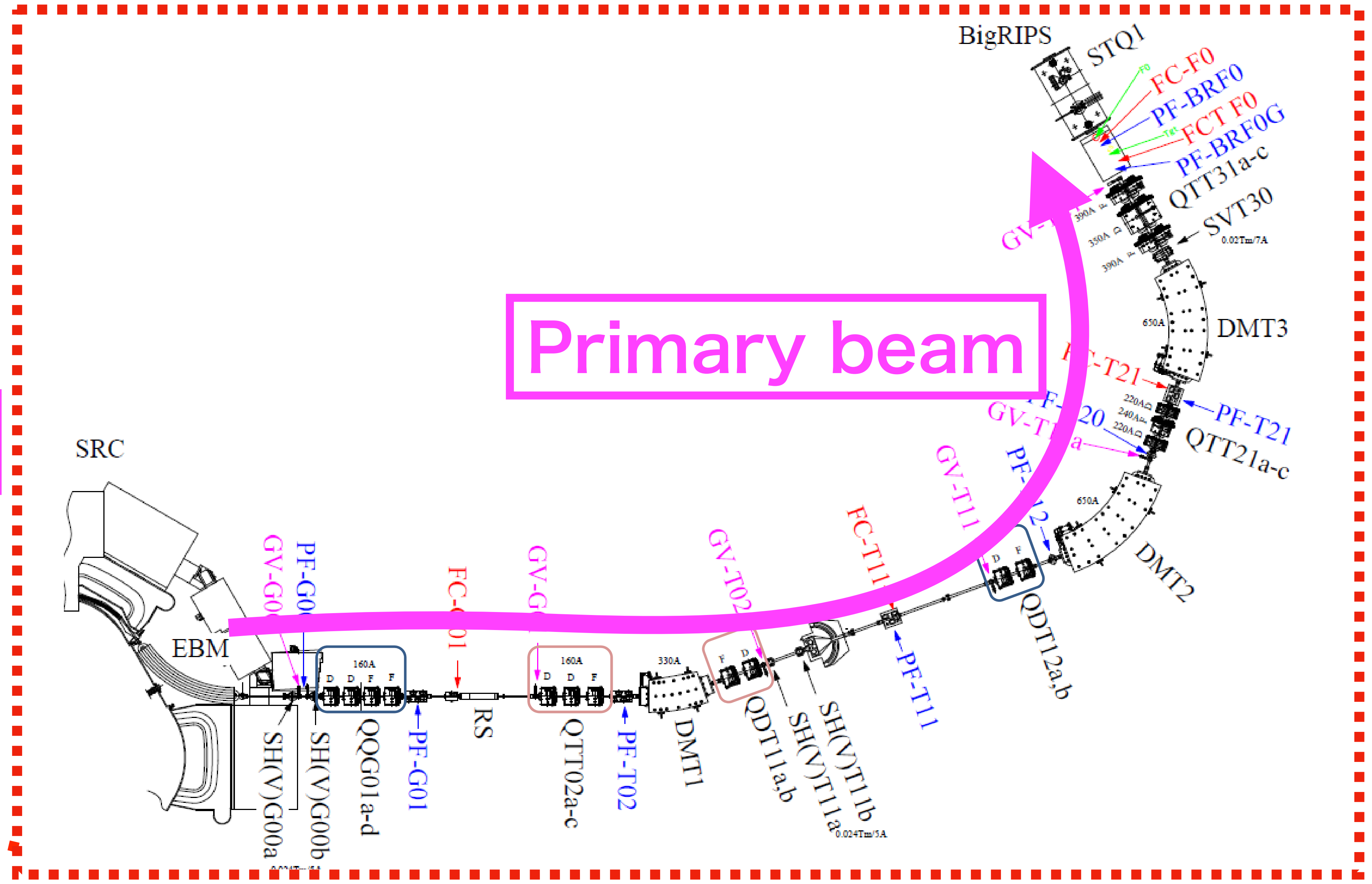


Slide from E. Iwai, “ビーム物理研究会2020”

- Handling multi-dimensional data with errors
- Avoiding local minimum
- **Relatively small step to reach optimized goal**

Attempt to Introduce ML Optimization to RIBF: “T-Course”





T-Course: Primary beam line between SRC to first target

- Fragment Separator (BigRIPS) is located at downstream → there are a lot of detectors
- Transmission ~ 90% for $^{86+}$U beam in normal operation
- Simultaneous optimization of transmission / spot shape

To realize high intense $^{86+}$U beam (> 1puA / 100 kW), we should

- **suppress beam loss to 0.1 ~ 1%,**
- **keep spot small (σ ~ 1mm).**

Detail of Algorithm: Bayesian Optimization with GPR

1. <u>Measurement of data as "initial values"</u>

   for N parameters, measurement of 2N + 1 data point : default $\pm\ \sigma_i$ (i = 1, 2, ⋯, N)

Detail of Algorithm: Bayesian Optimization with GPR

1. Measurement of data as "initial values"

   for N parameters, measurement of 2N + 1 data point : default ± σ_i (i = 1, 2, ⋯, N)

2. Create a model of the distribution of the objective function based measured data using GPR

   In Bayesian estimation, objective function has distribution for any x, i.e.,

$$p(t\,|\,x) = \frac{1}{\sqrt{2\pi}\sigma(x)} \exp\left(-\frac{(x-m(x))^2}{2\sigma(x)^2}\right)$$

m(x) and σ(x) are calculated with $\{x_i, \sigma_i\}_N$ as

$$m_N(x) = \Sigma \underline{k(x, x_i)} t_i, \qquad \sigma_N(x)^2 = \frac{1}{\beta} + \vec{\phi}(x)^T S_N \vec{\phi}(x),$$

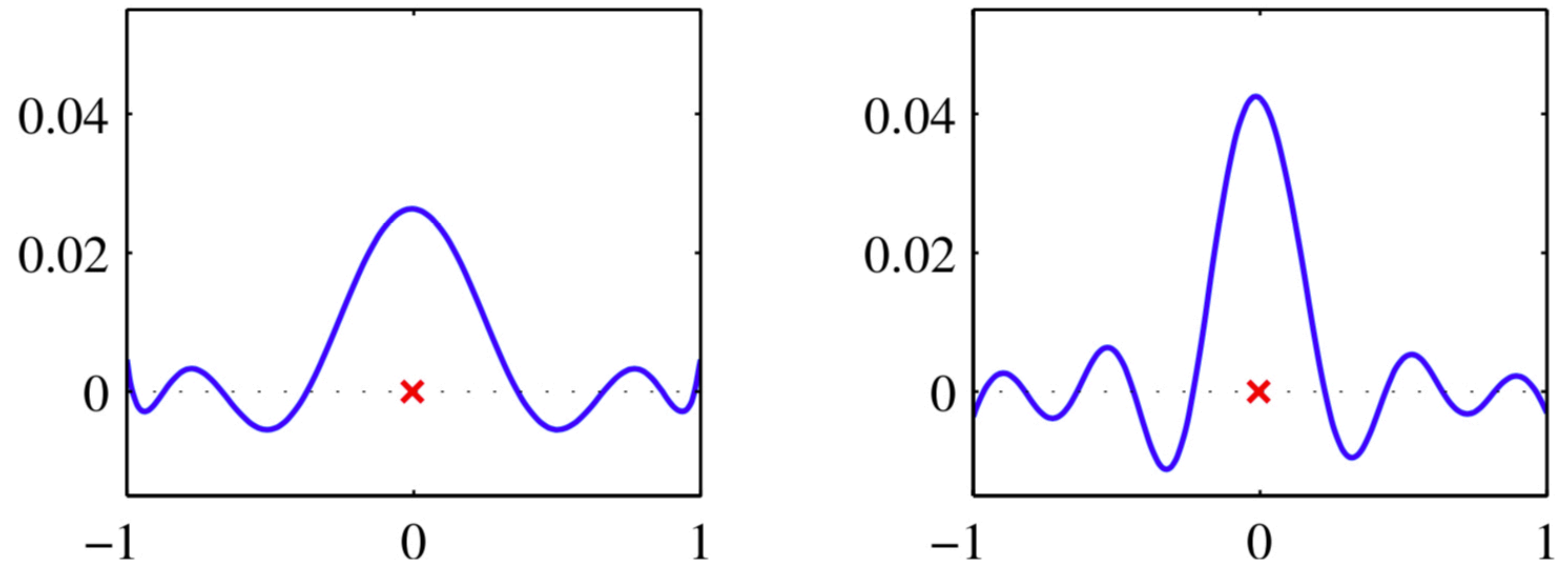
kernel function:

"Indicator of the strength

of correlation between x_i and x"

C. M. Bishop, Pattern Recognition and Machine Learning, Springer (2006)

Figure 3.11 Examples of equivalent kernels $k(x, x')$ for $x = 0$ plotted as a function of x', corresponding (left) to the polynomial basis functions and (right) to the sigmoidal basis functions shown in Figure 3.1. Note that these are localized functions of x' even though the corresponding basis functions are nonlocal.



Detail of Algorithm: Bayesian Optimization with GPR

1. Measurement of data as "initial values"
   for N parameters, measurement of 2N + 1 data point : default ± σ_i (i = 1, 2, ⋯, N)

2. Create a model of the distribution of the objective function based measured data using GPR
   In Bayesian estimation, objective function has distribution for any x, i.e.,

$$p(t|x) = \frac{1}{\sqrt{2\pi}\sigma(x)} \exp\left(-\frac{(x - m(x))^2}{2\sigma(x)^2}\right)$$

m(x) and σ(x) are calculated with $\{x_i, \sigma_i\}_N$ as

$$m_N(x) = \Sigma k(x, x_i) t_i, \qquad \sigma_N(x)^2 = \underbrace{\frac{1}{\beta}}_{\text{Intrinsic error of the measurement}} + \underbrace{\overrightarrow{\phi}(x)^T \mathrm{S_N} \overrightarrow{\phi}(x)}_{\text{Ambiguity from the estimation}},$$

Detail of Algorithm: Bayesian Optimization with GPR

1. Measurement of data as "initial values"
   for N parameters, measurement of 2N + 1 data point : default ± σ_i (i = 1, 2, ⋯, N)

2. Create a model of the distribution of the objective function based measured data using GPR

3. Calculate EI (Expected Improvement) for all x based on the created model
   Expected improvement is defined as

$$\mathbb{E}_{\alpha}(max\{f(x,\alpha) - f(\hat{x}),0\}) \equiv \int P(\alpha) max\{f(x,\alpha) - f(\hat{x}),0\} d\alpha \,.$$

   f(x, α): objective function α: error indicators P(α) : Probability function of α
   This value becomes large when the information is insufficient.

Detail of Algorithm: Bayesian Optimization with GPR

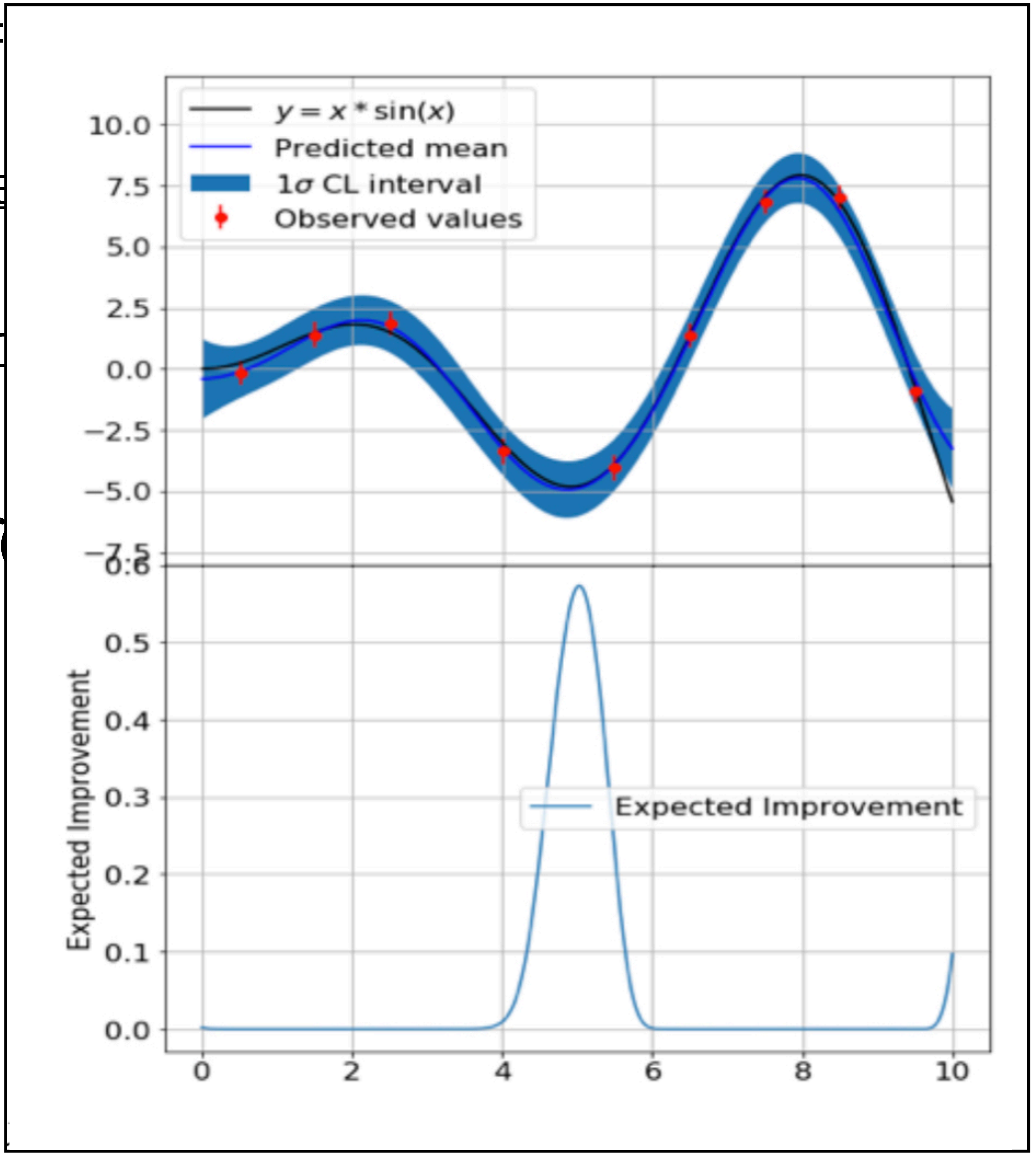
1. Measurement of data as "initial values"
   for N parameters, measurement of 2N + 1 data point : default ±

2. Create a model of the distribution of the objective function base

3. Calculate EI (Expected Improvement) for all x based on the creat
   Expected improvement is defined as

$$\mathbb{E}_\alpha(max\{f(x,\alpha) - f(\hat{x}),0\}) \equiv \int P(\alpha) max\{f(x,\alpha) - f($$

   f(x, α): objective function α: error indicators P(α) : Probability
   This value becomes large when the information is insufficient.



E. Iwai et al, Proc. of PASJ 2021 WEOB02

Detail of Algorithm: Bayesian Optimization with GPR

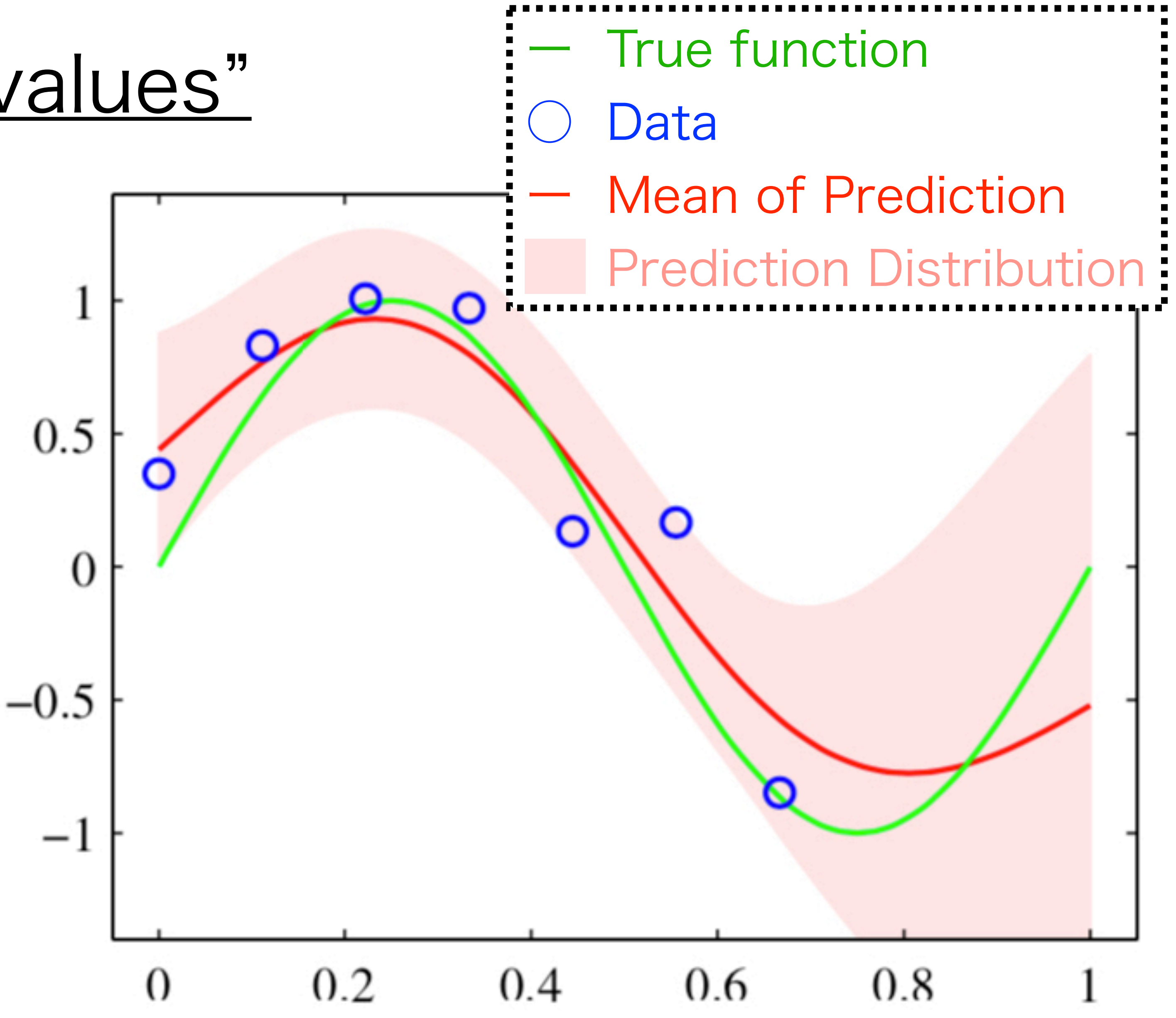
1. Measurement of data as “initial values”
   for N parameters, measurement

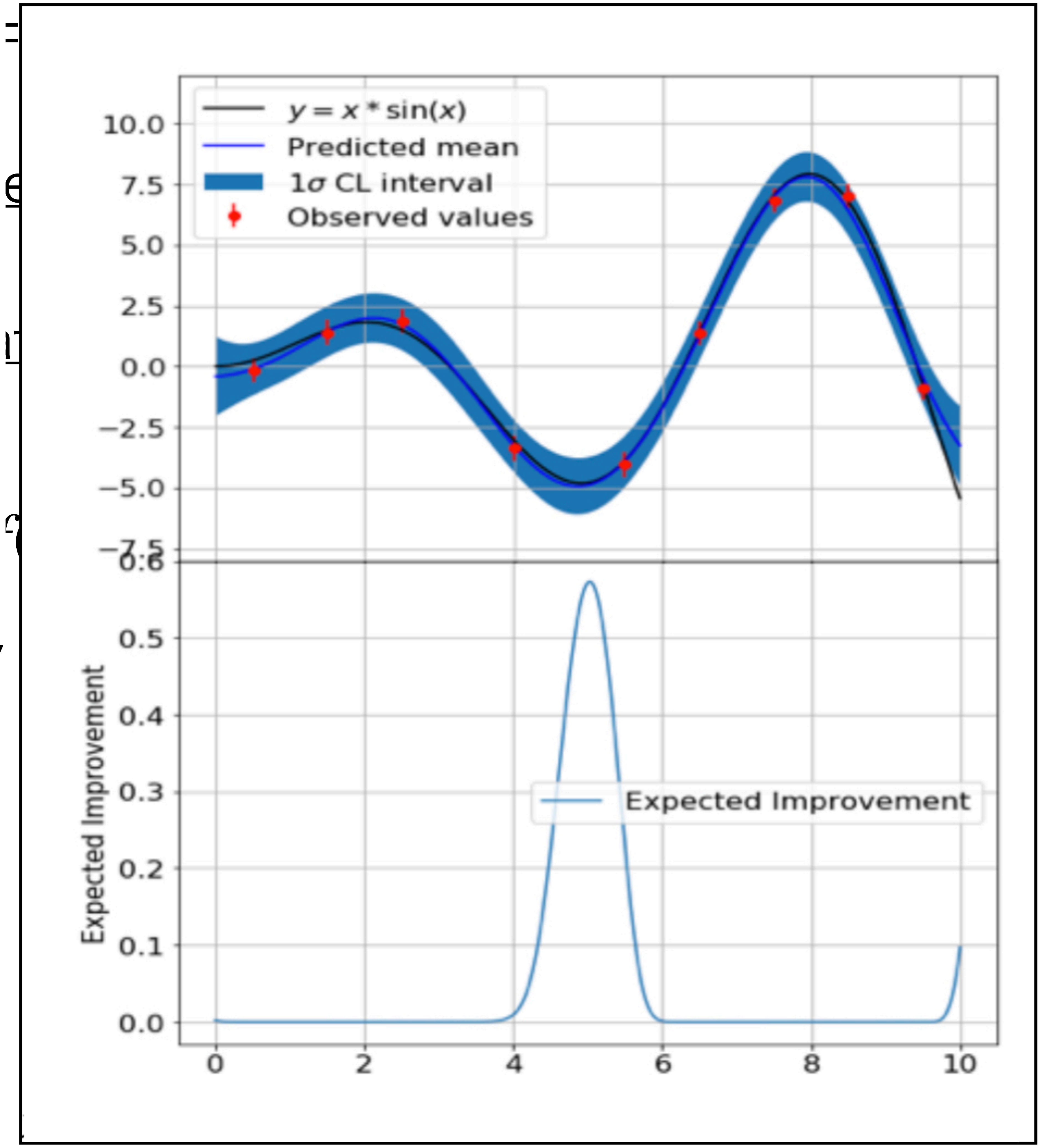
2. Create a model of the distributic

3. Calculate EI (Expected Improven
   Expected improvement is define

$$\mathbb{E}_{\alpha}(max\{f(x,\alpha) - f($$

f(x, α): objective function α: error indicators P(α) : Probability
This value becomes large when the information is insufficient.





E. Iwai et al, Proc. of PASJ 2021 WEOB02

Detail of Algorithm: Bayesian Optimization with GPR

1. Measurement of data as "initial values"
   for N parameters, measurement of 2N + 1 data point : default ± σ_i (i = 1, 2, ⋯, N)

2. Create a model of the distribution of the objective function based measured data using GPR

3. Calculate EI (Expected Improvement) for all x based on the created model
   Expected improvement is defined as

$$\mathbb{E}_{\alpha}(max\{f(x,\alpha) - f(\hat{x}),0\}) \equiv \int P(\alpha) max\{f(x,\alpha) - f(\hat{x}),0\} d\alpha \,.$$

   f(x, α): objective function α: error indicators P(α) : Probability function of α
   This value becomes large when the information is insufficient.

4. Apply a new parameter with maximum EI to the real system and measure the data
   to search the point with the maximum EI, we use the library BoTorch /GPyTorch

5. Repeat 2 ~ 4 and search the best point

Development History of GPR Optimization System@RIBF

– 1st attempt an automatic optics optimization by ML

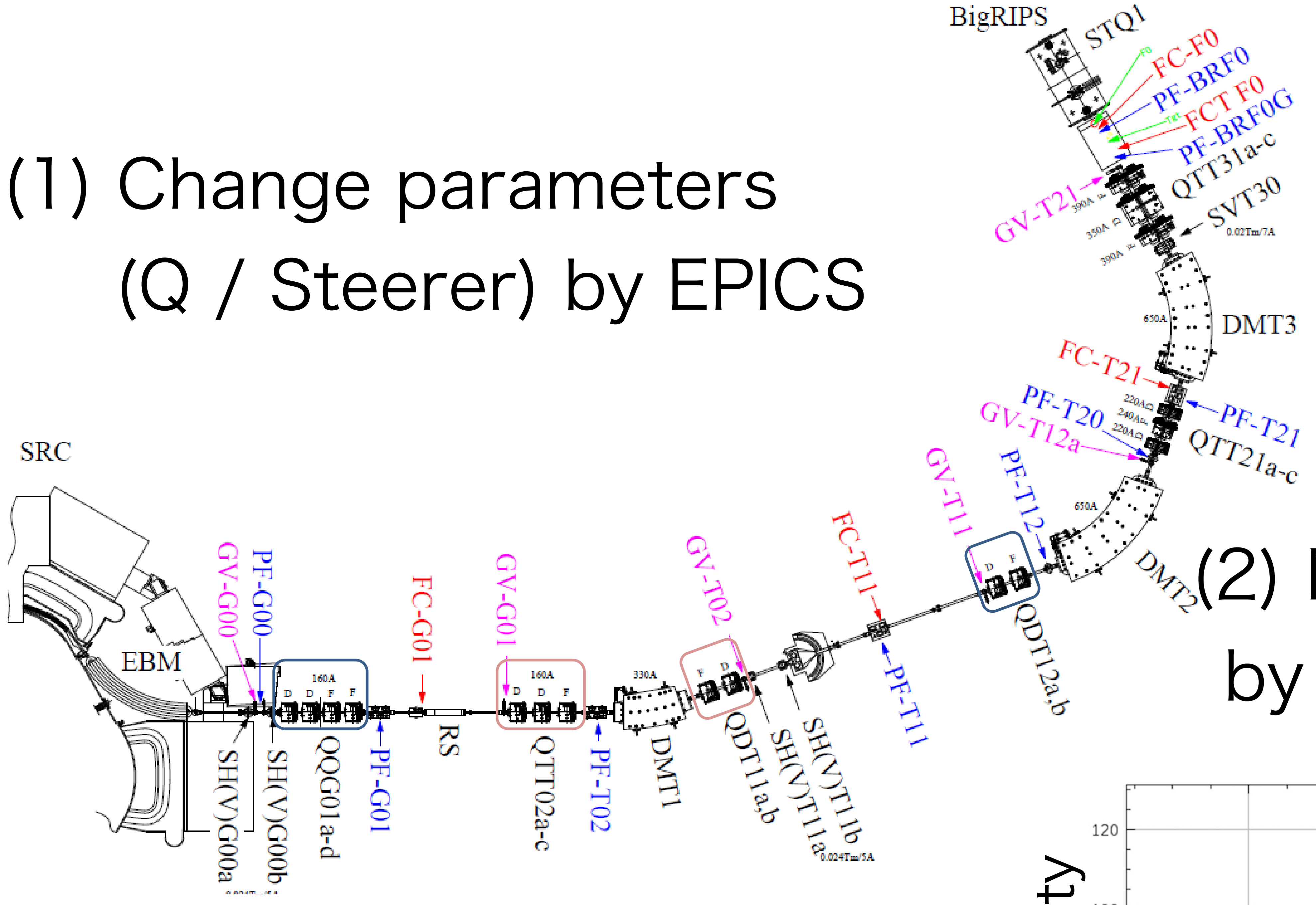
・simultaneous optimization of transmission and beam spot

– Development of indicators for high intensity beams

・secondary beam with charge conversion

Automatic Optics Optimization ①: Test with Low Intensity Beam

(1) Change parameters
(Q / Steerer) by EPICS

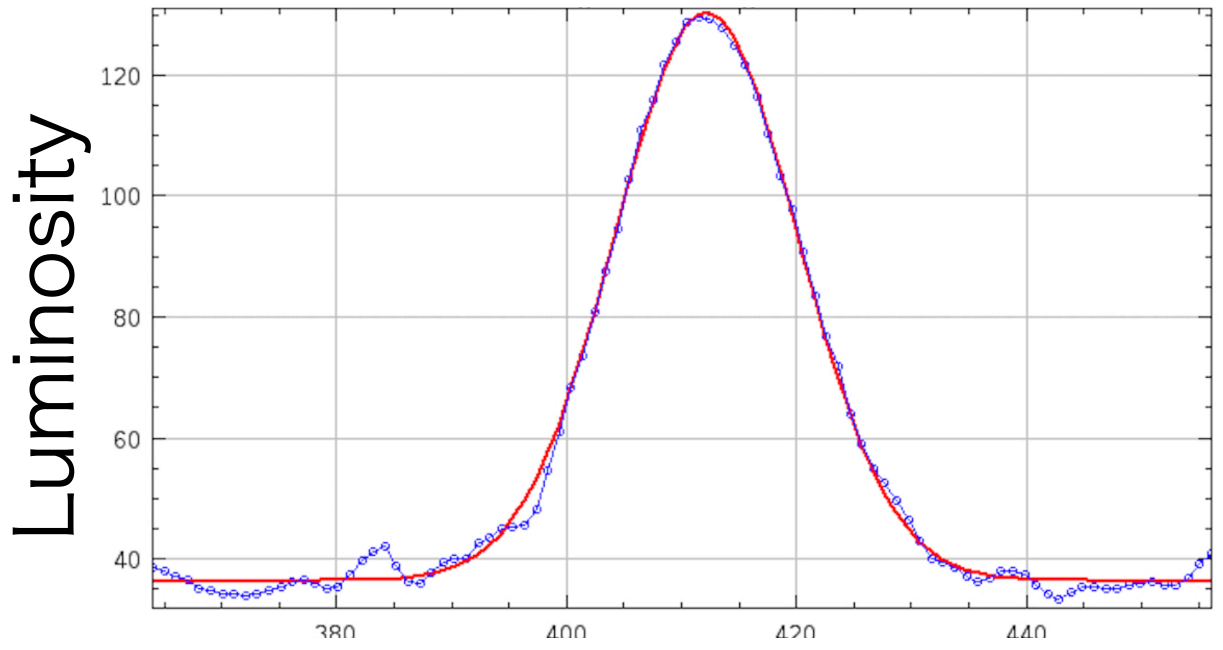


U^{86+} 345 MeV/u
~ 0.001 enA
2.0 e+7 cps

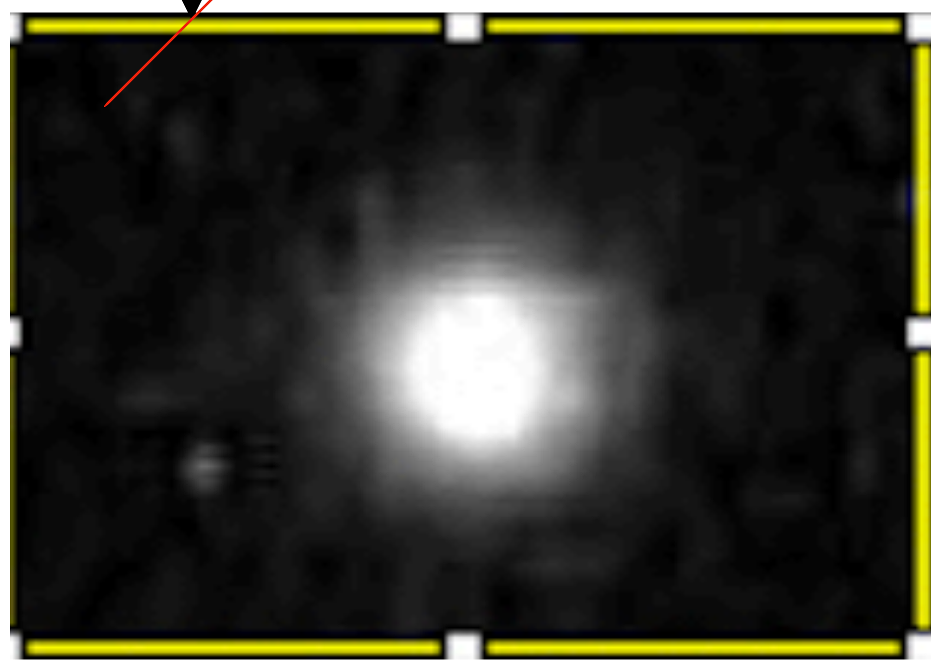
(2) Measure beam spot
by fluorescent viewer

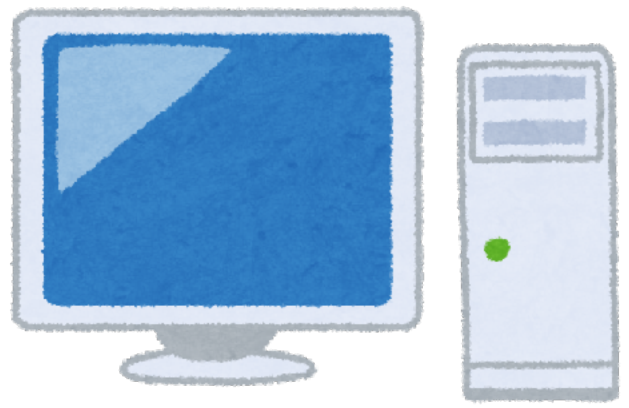
(3) Calculate objective function
Create model by GPR
Choose next param.

Gaussian Process Regression (Baysian Opimization)



horizontal position [pixels]


Viewer image

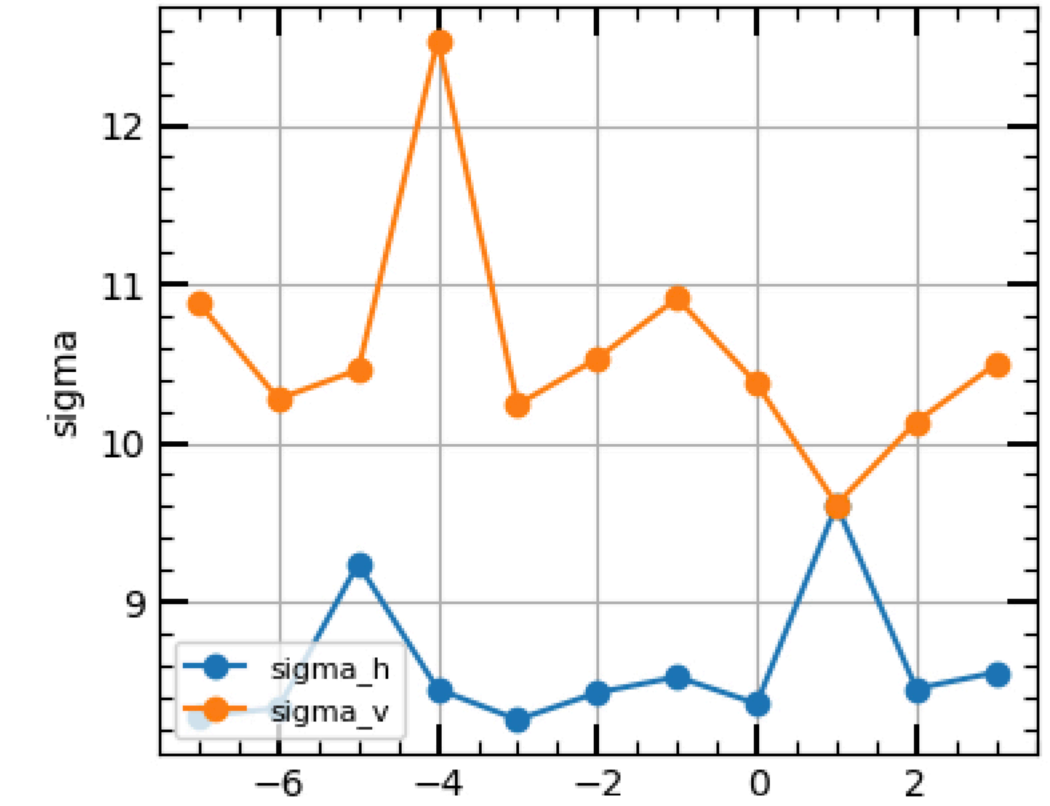


Intensity : Integral of image luminosity
Spot : Image width

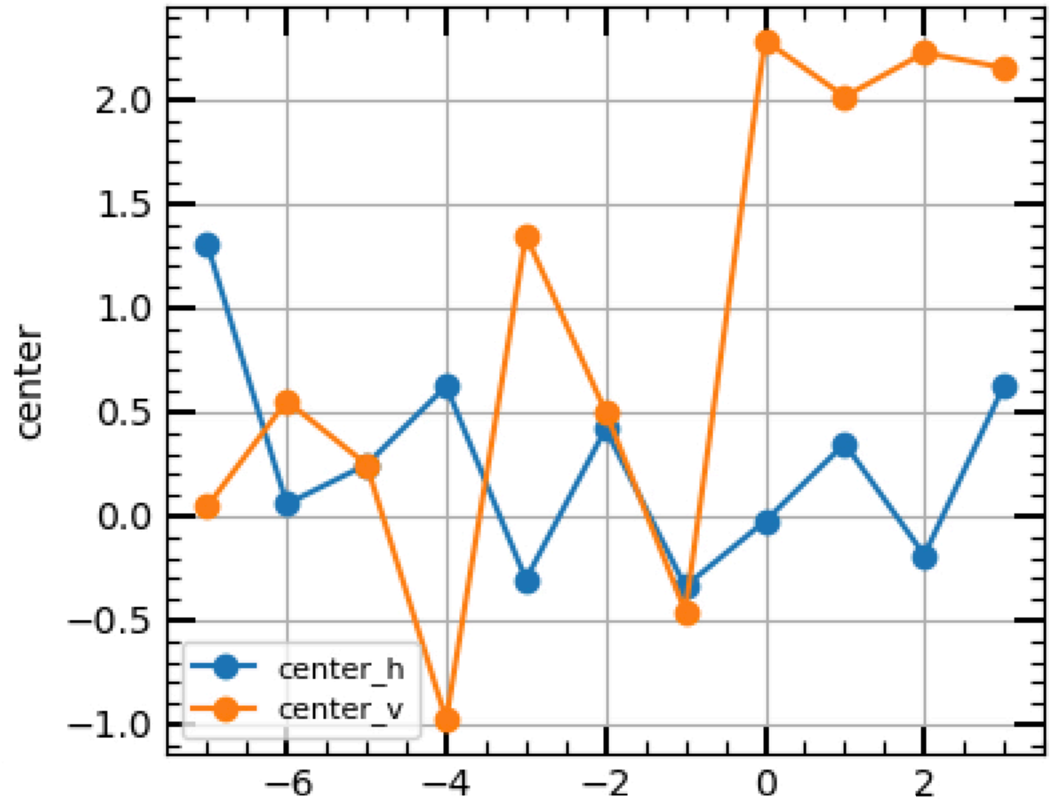
Repeat (1) — (3)
to optimize optics

1st Exp. : Auto Tuning with Low Intensity Beam

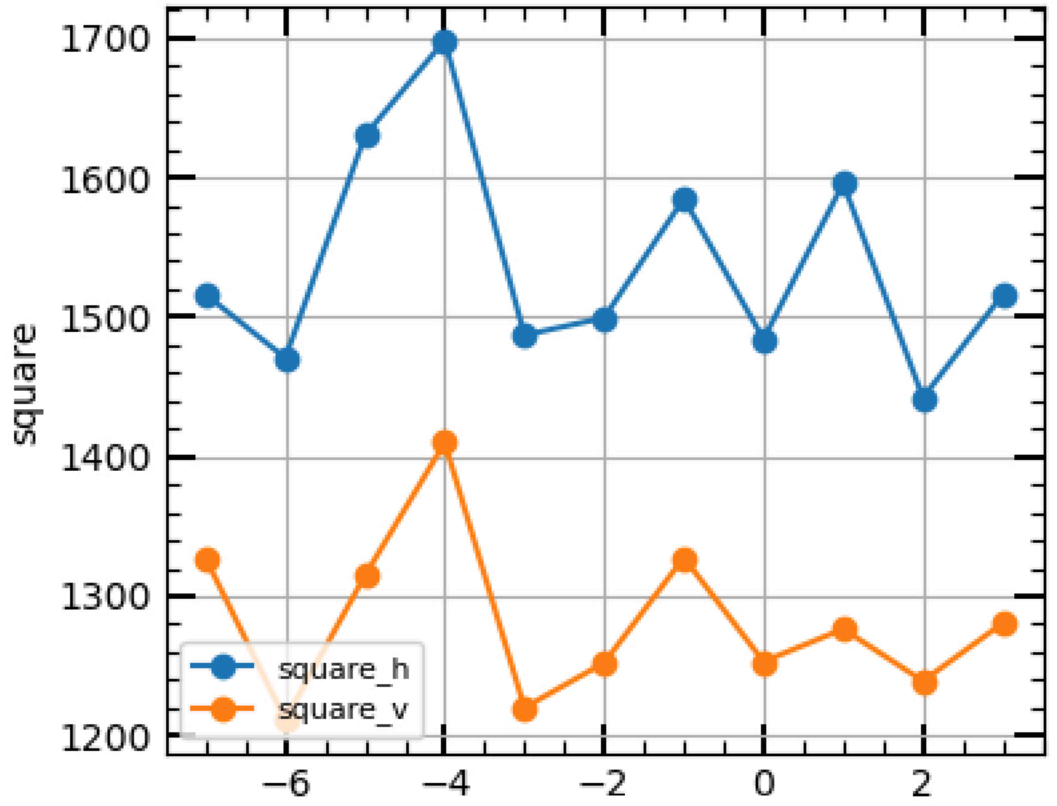
Demonstration test at 2020 Oct. 21:00 ~ 9:00



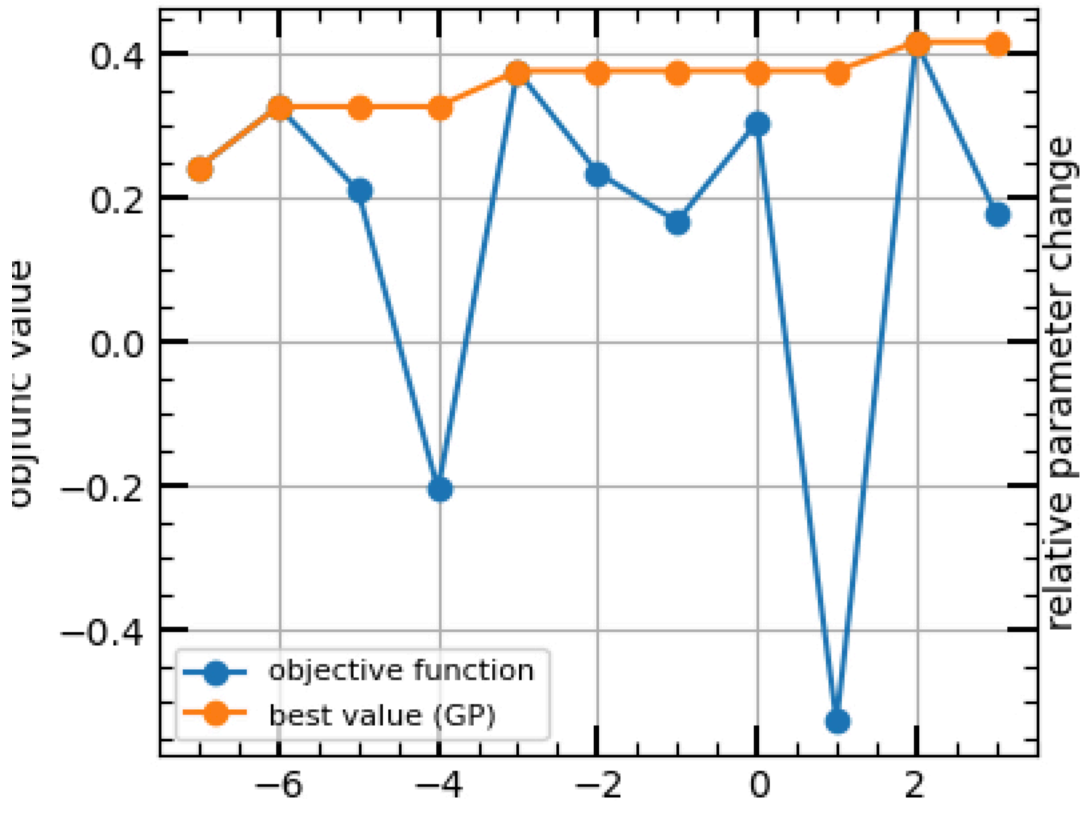
σ h, v vs Epoch



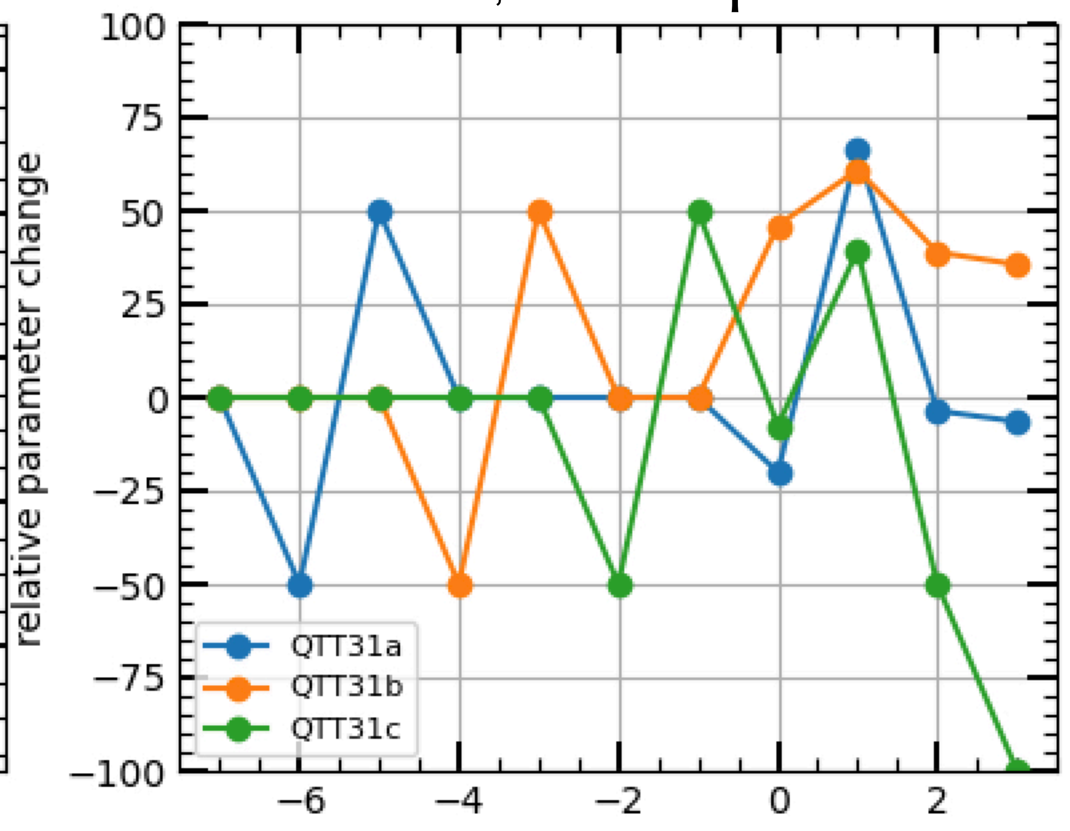
Posh, v vs Epoch



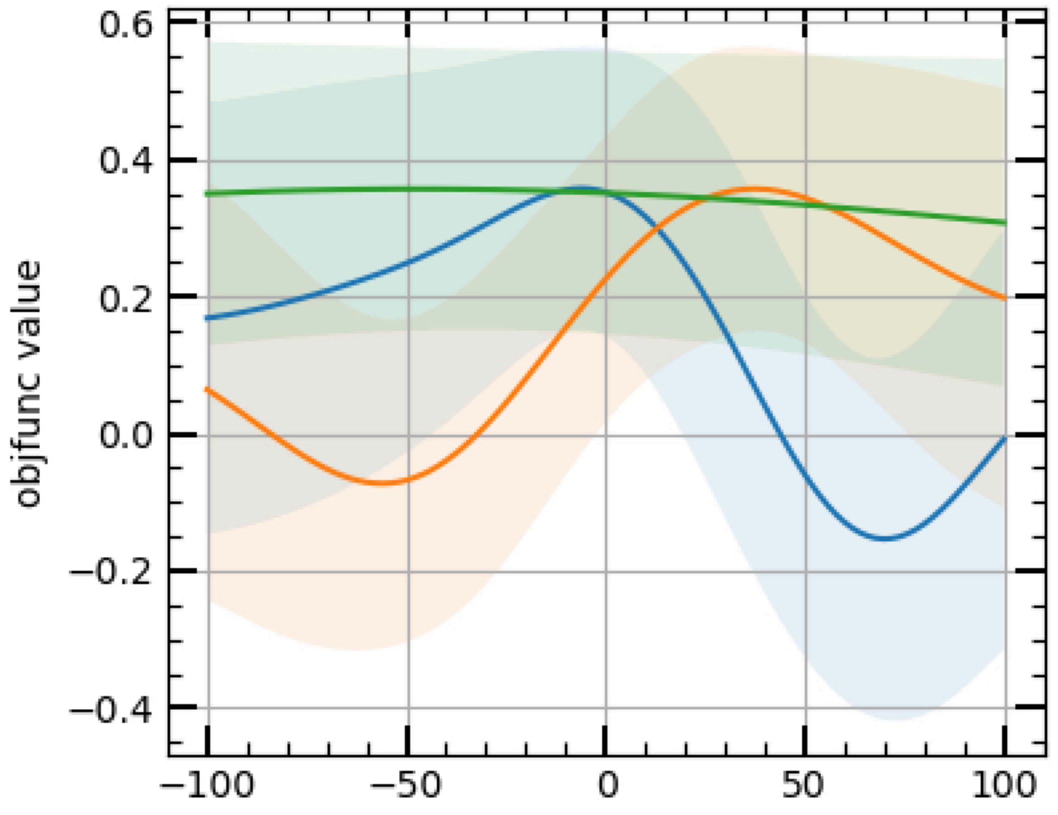
S h, v vs Epoch



Objective function vs Epoch



Current of Q vs Epoch



Objectives vs Current of Q

Initial: manually optimized params.

Goal : good transmission

small beam spot

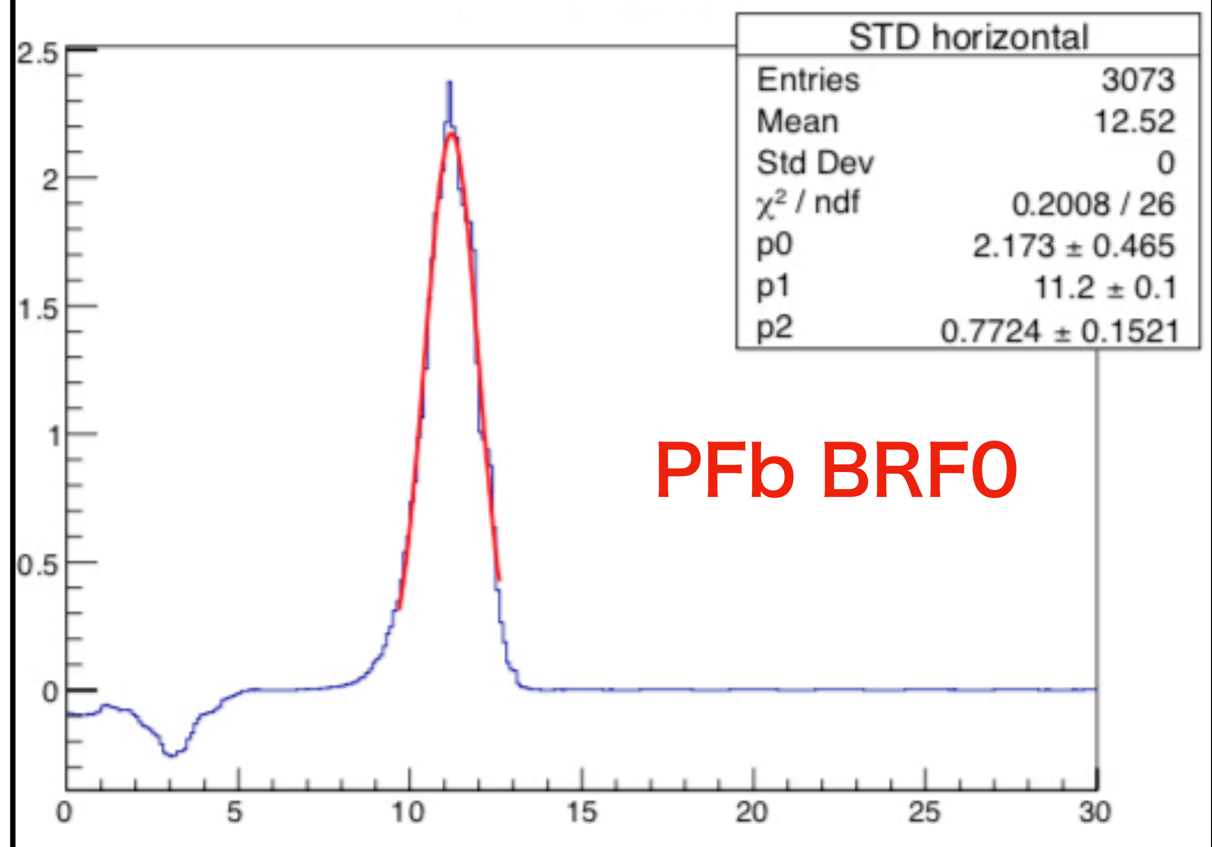
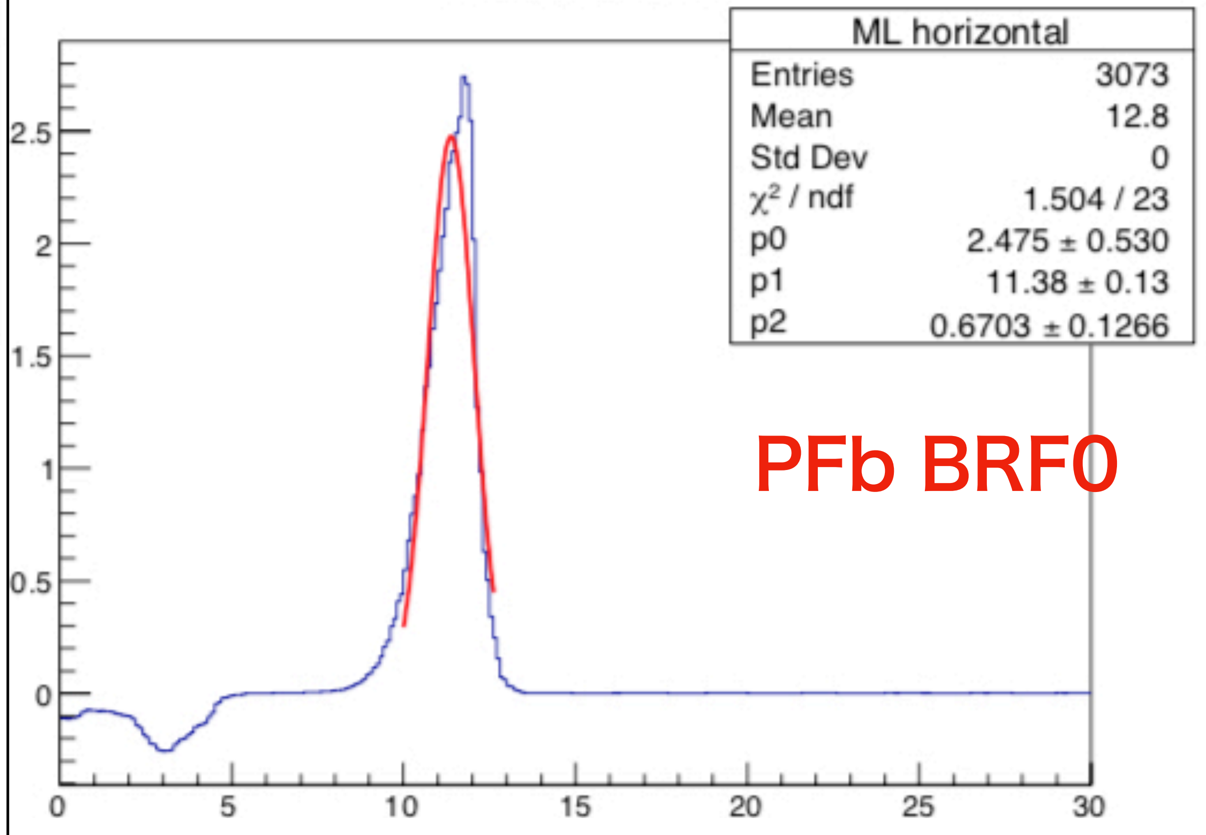
- 3 ~ 7 params. (Quadrupoles)
- 1 epoch ~ 1 s
- 1 try ~ 5 min (~300 epoch)
- ~ 0.001 enA

try optimization with several conditions

1st Exp. : Auto Tuning with Low Intensity Beam

Demonstration test at 2020 Oct. 21:00 ~ 9:00

Compare the result of manual optimization / manual + ML optimization using high intensity beams and wire scanner / Faraday cups

	Manual Optimization	Manual + ML Optimization
FCup [eμA]	7.20	7.8
Beam Dump [eμA]	7.25	8.0
Ratio (BD/FCup)※	1.01	1.03
Wire Scanner の像		

- **Transmission ↗ 2%**
- **Beam width ↘ 13%**

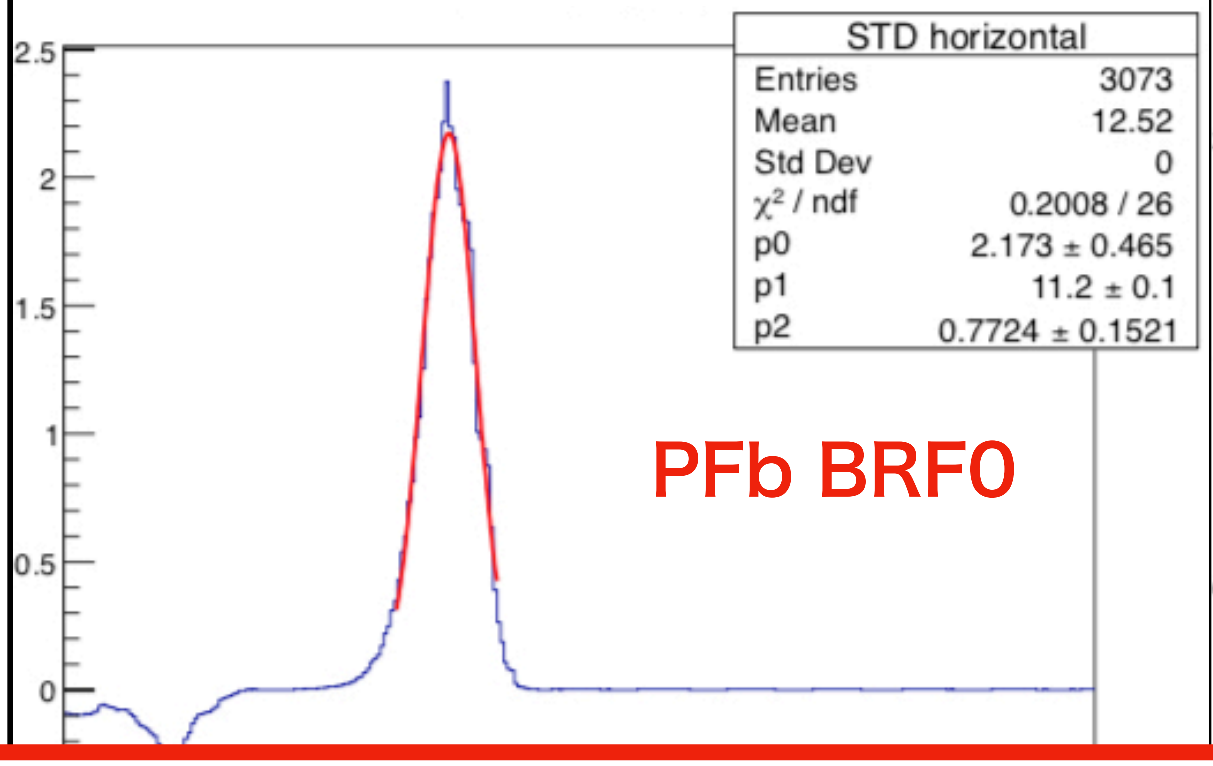
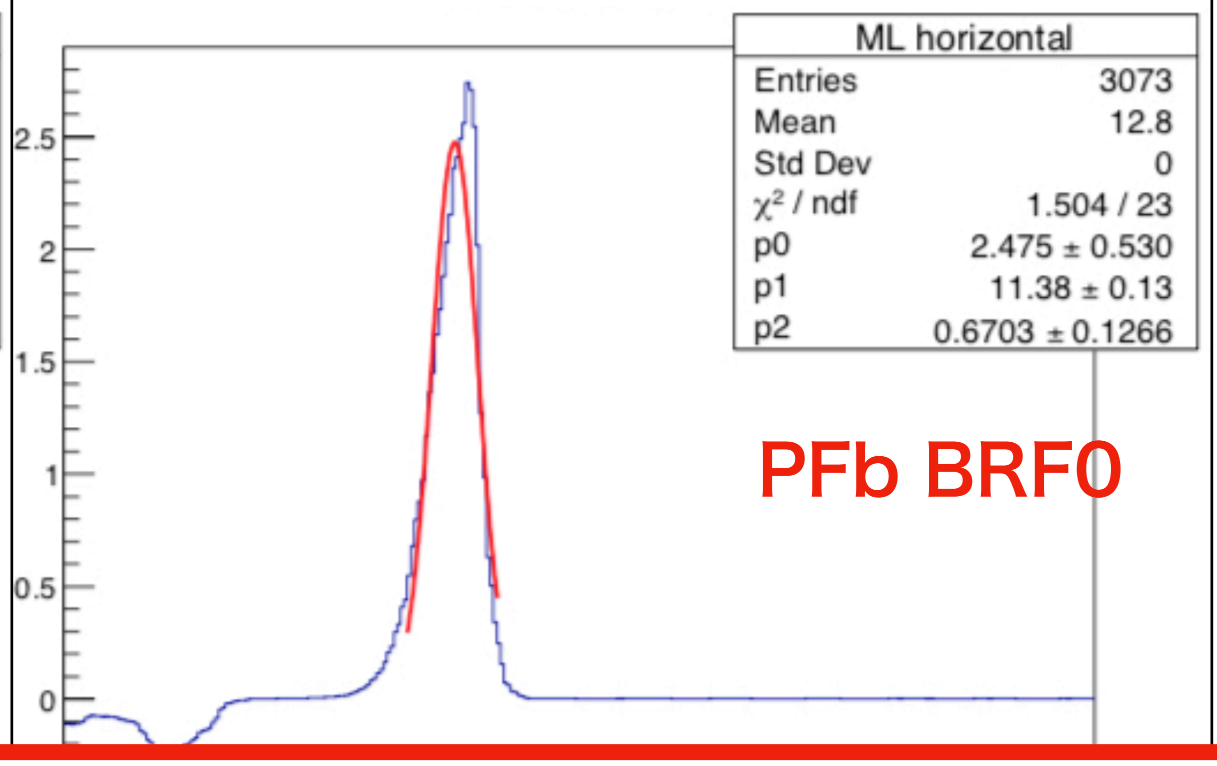
Significant improvements in the 1st test

※ Beam dump does not have good suppressor

1st Exp. : Auto Tuning with Low Intensity Beam

Demonstration test at 2020 Oct. 21:00 ~ 9:00

Compare the result of manual optimization / manual + ML optimization
using high intensity beams and wire scanner / Faraday cups

	Manual Optimization	Manual + ML Optimization
FC_{up} [eμA]	7.20	7.8
Beam Dump [eμA]	7.25	8.0
Ratio (BD/FC_{up})※	1.01	1.03
Wire Scanner の像		

- **Transmission ↗ 2%**
- **Beam width ↘ 13%**

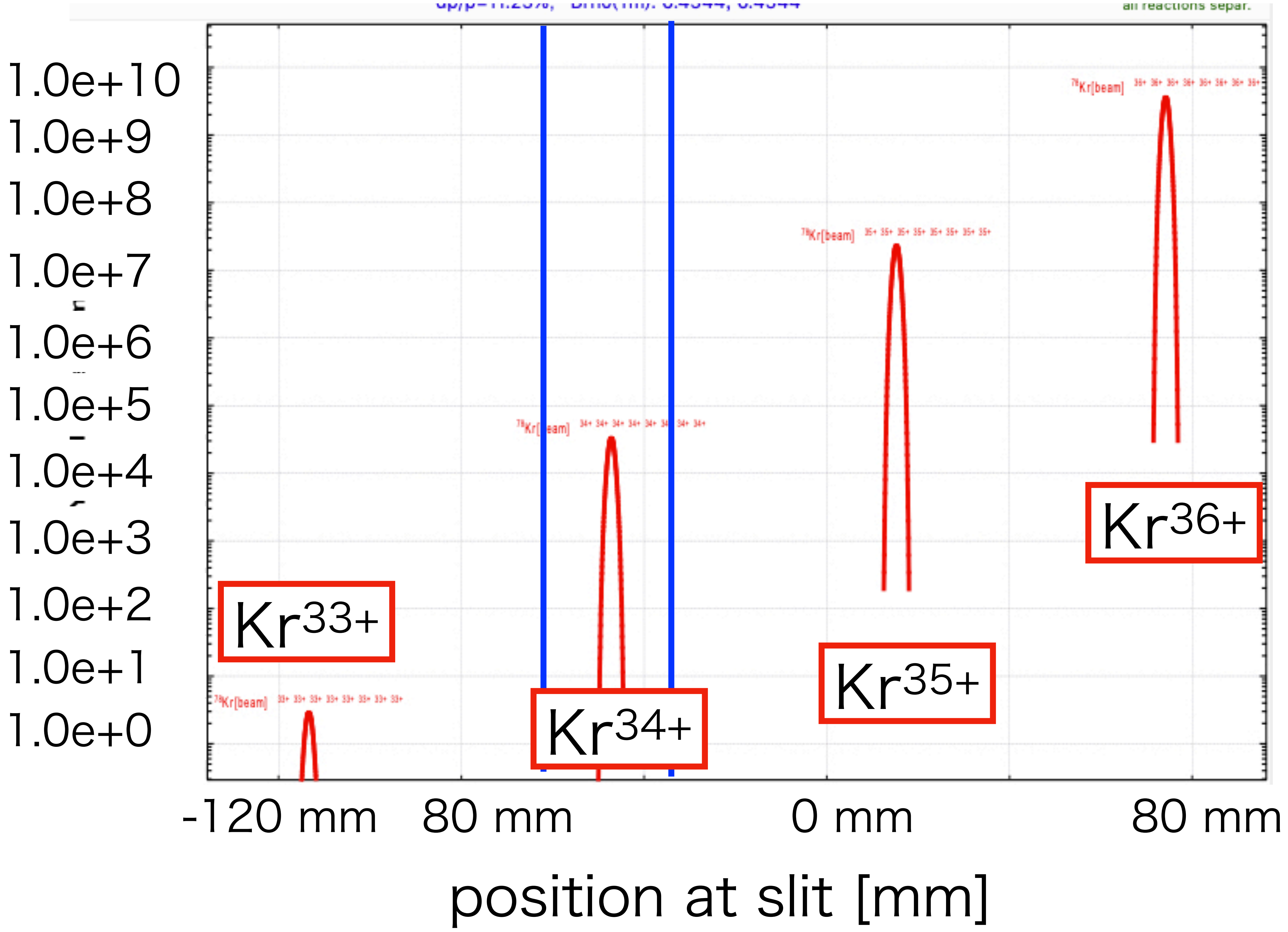
Significant improvements in the 1st test

※ Beam dump does not have good suppressor

◯ Algorithm works as expected
◯ Optics is improved with low intensity beam (limit by fluorescent viewer: 0.001 enA)
△ Effective algorithm

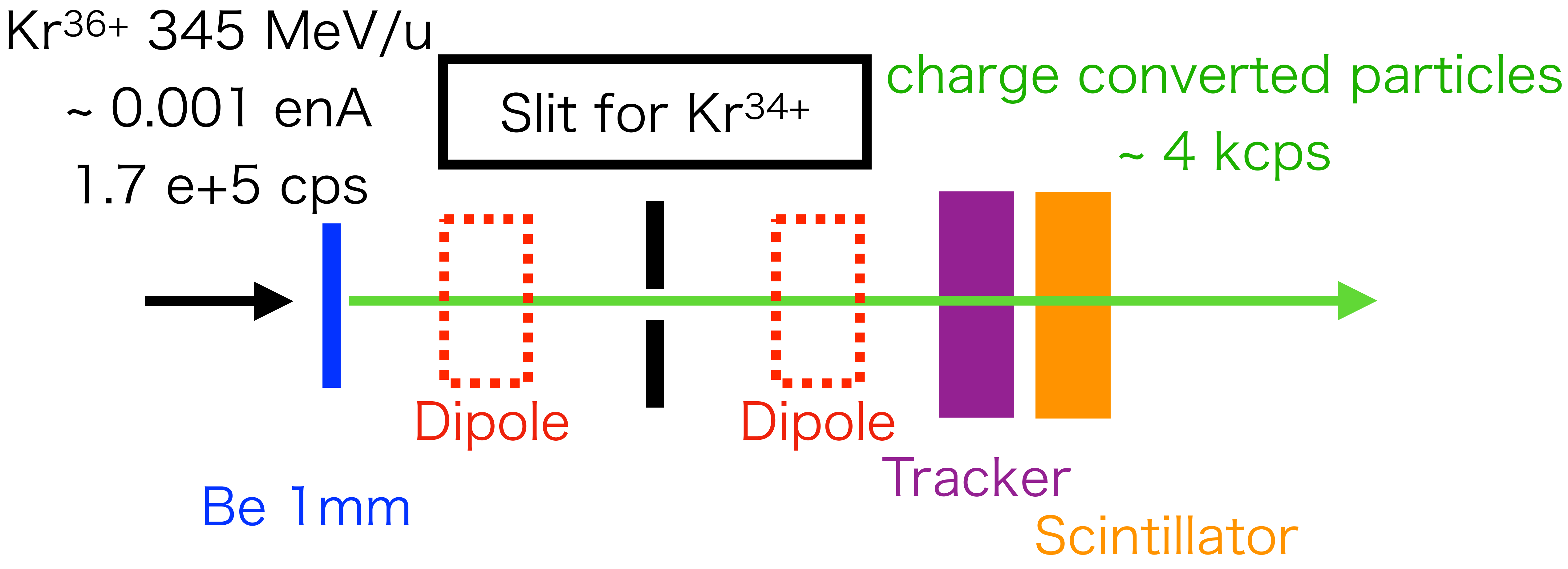
Automatic Optics Optimization ②: Indicator for High Intensity Beam

charged particle dist. at **Slit**
in downstream of 1st target



Intensity : Scintillator in downstream
Spot : Tracker in downstream

Detector in downstream
PPAC (tracker) / Scintillator



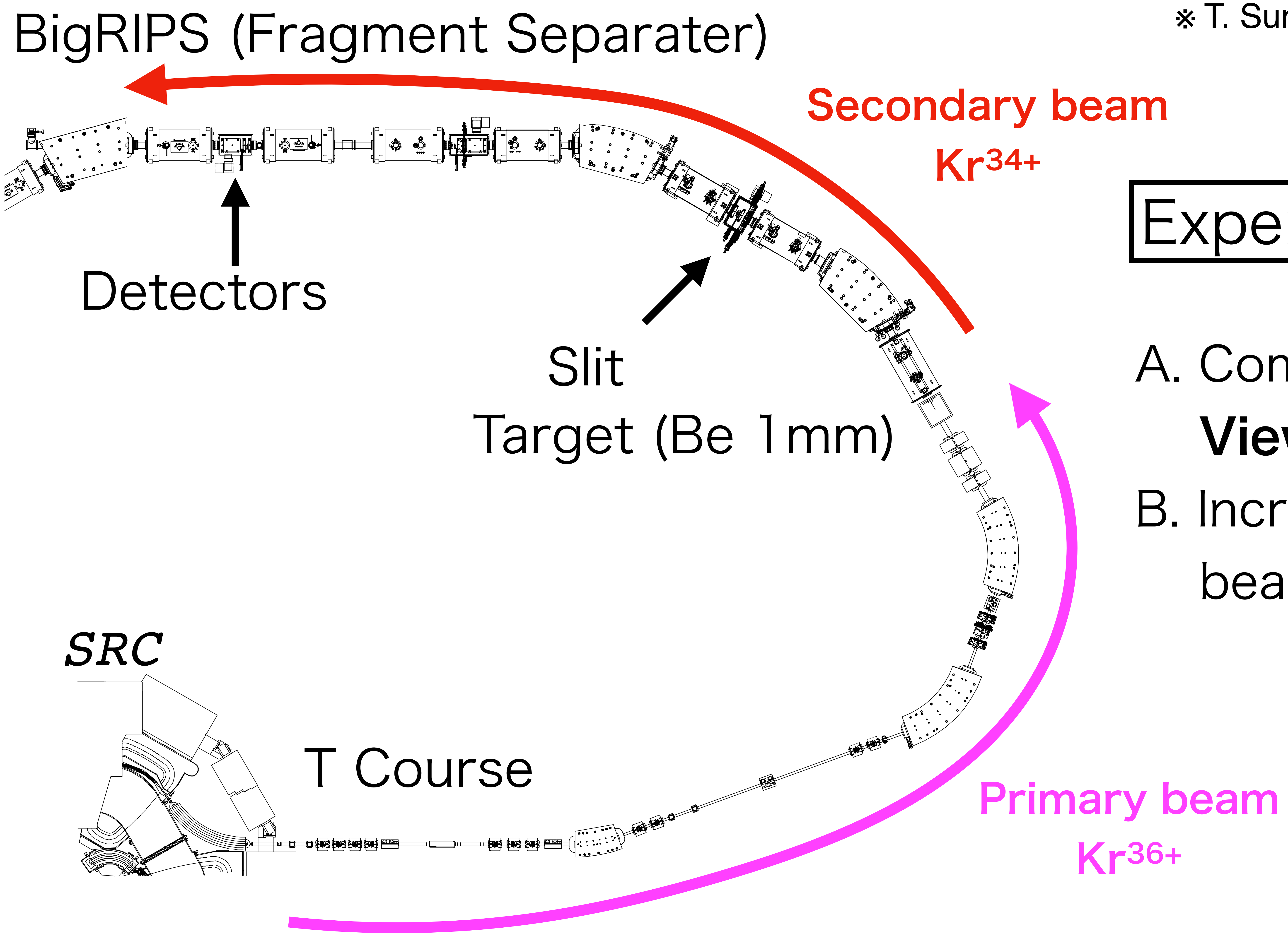
Convert charge state of beam by Be 1mm target
Make low intense beam

	Primary beam intensity	Expected rate
Kr^{35+}	0.001 enA	1 kcps
Kr^{34+}	1 enA	2 kcps
Kr^{33+}	(300 enA)	1 kcps

4th Exp. : Auto Tuning with High Intensity Beam

Real time tracker system is realized and connected to EPICS※
→ test with ideal situation with Be target and tracker / scintillator

※ T. Sumikama *et al*, RIKEN Accel. Prog. Rep 54, 82 (2021)



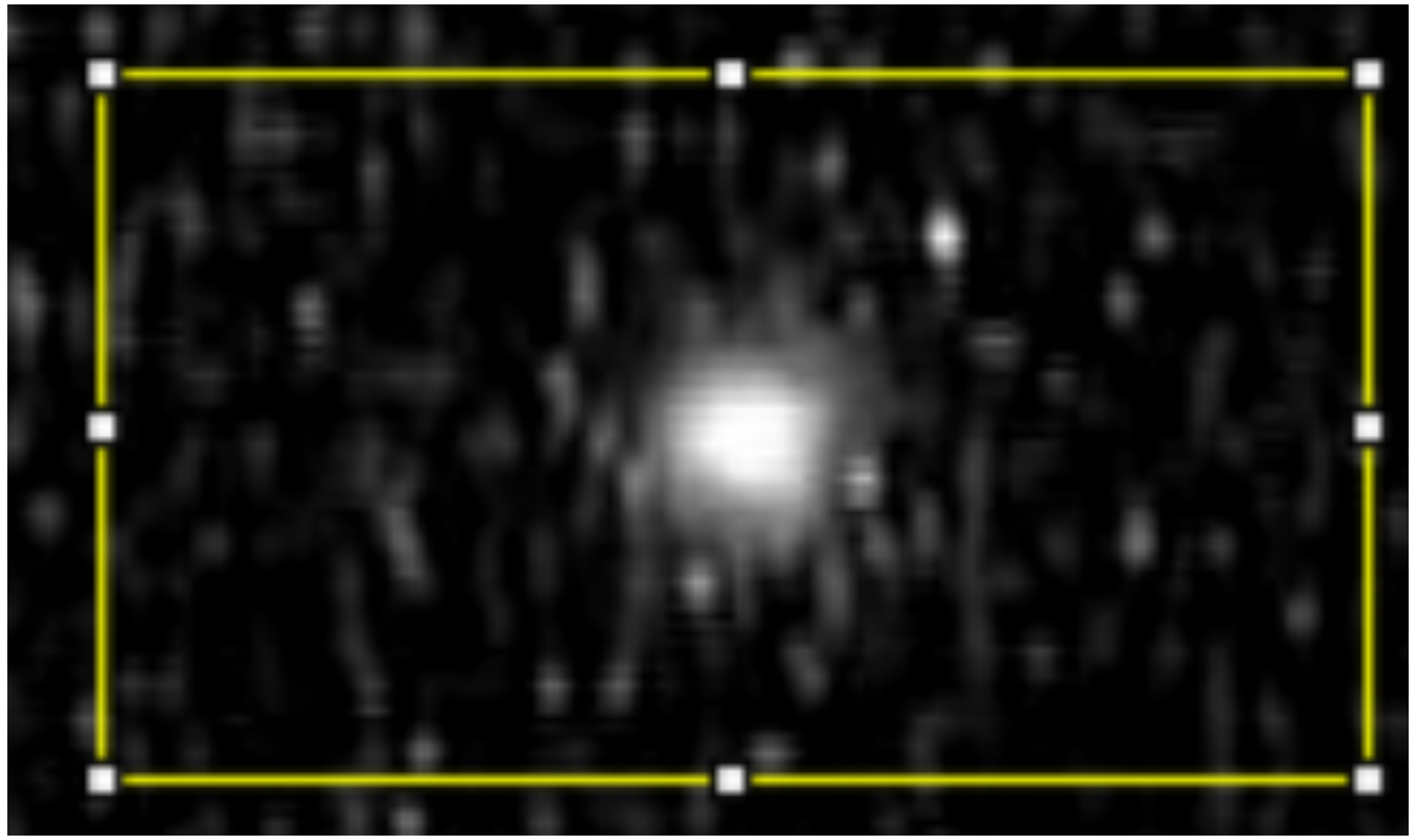
Experiment items

A. Compare beam spot measured by **Viewer** and **Tracker**
B. Increase the beam intensity and optimize beam optics using Tracker / Scintillator

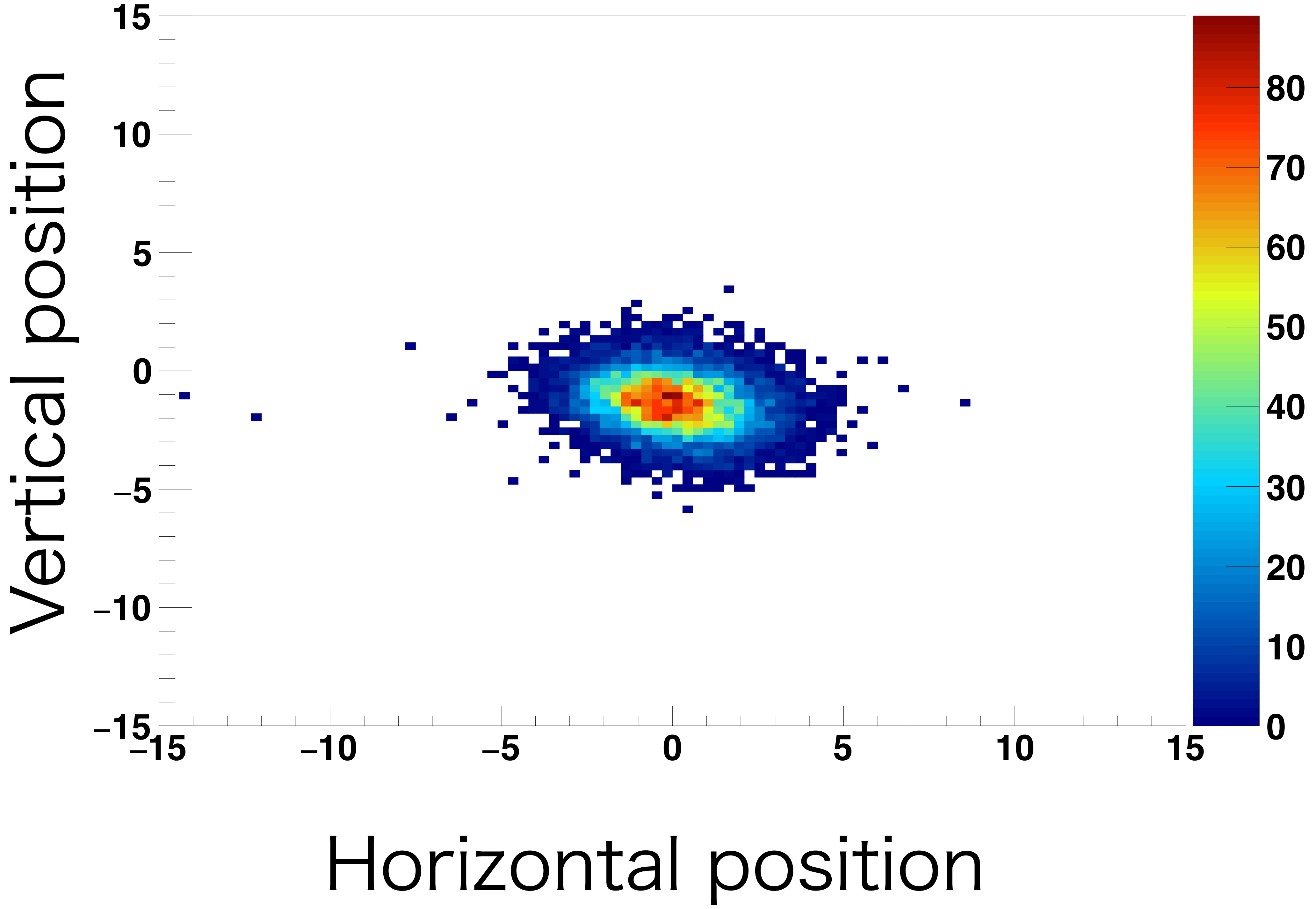
A. Compare Beam Spot measured by Viewer and Tracker

Change optics and compare
- fluorescent viewer image of primary beam (Kr^{36+})
- position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)

Fluorescent viewer (Kr^{36+})



Gas tracker (Kr^{34+})



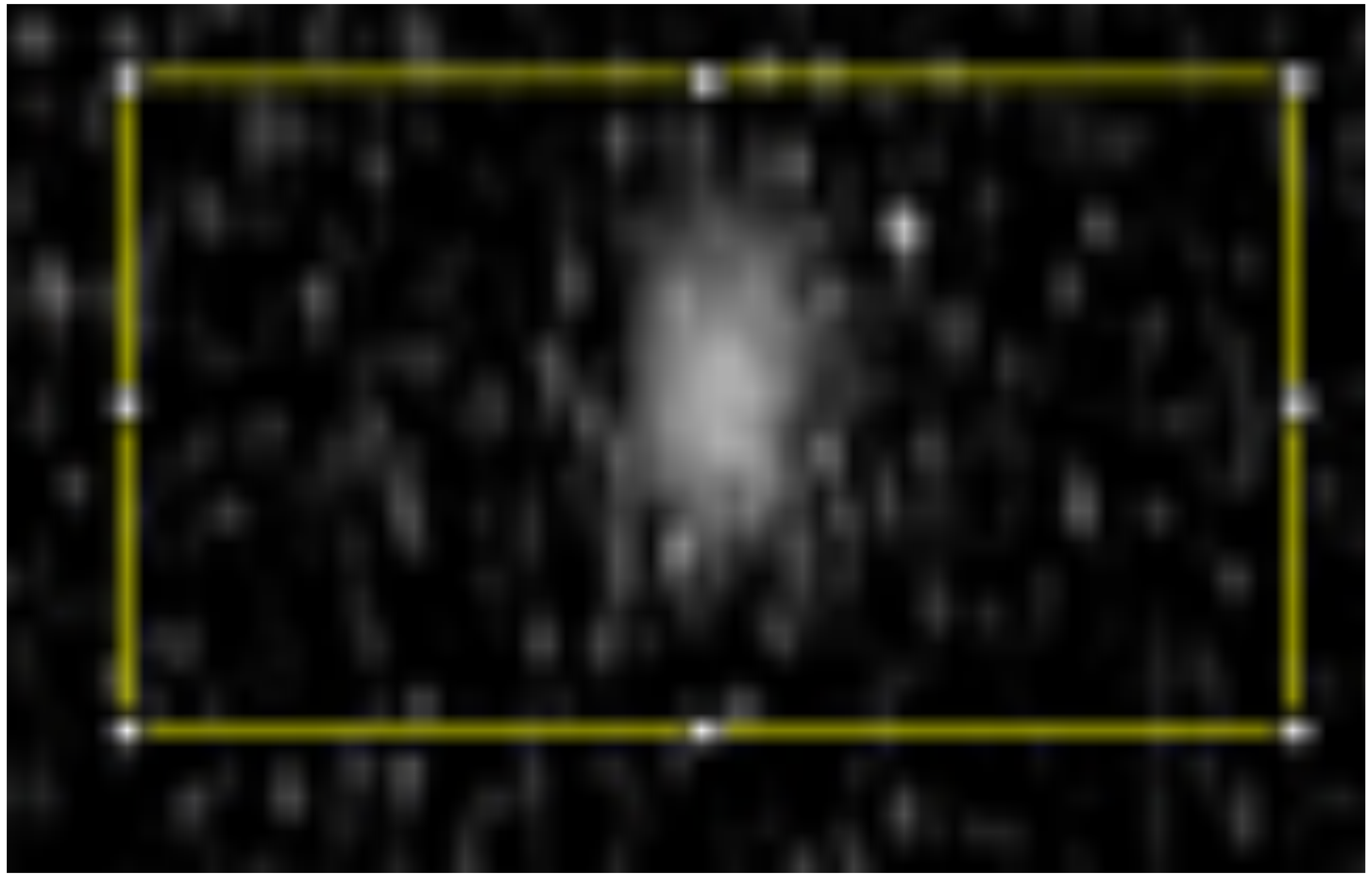
※ In this measurement, Fluorescent viewer was removed and Be 1mm was inserted.

A. Compare Beam Spot measured by Viewer and Tracker

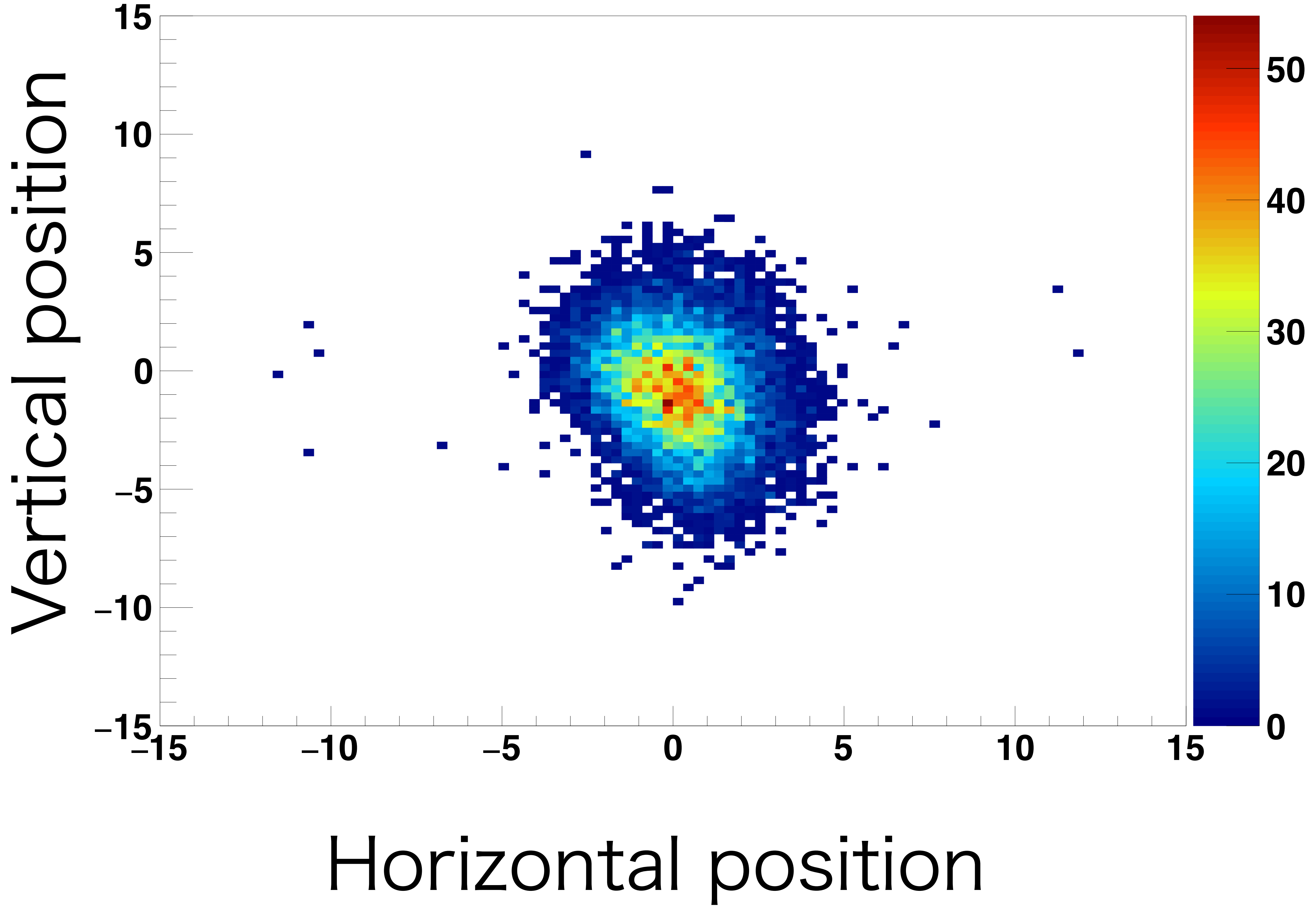
Change optics and compare

- fluorescent viewer image of primary beam (Kr^{36+})
- position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)

Fluorescent viewer (Kr^{36+})



Gas tracker (Kr^{34+})

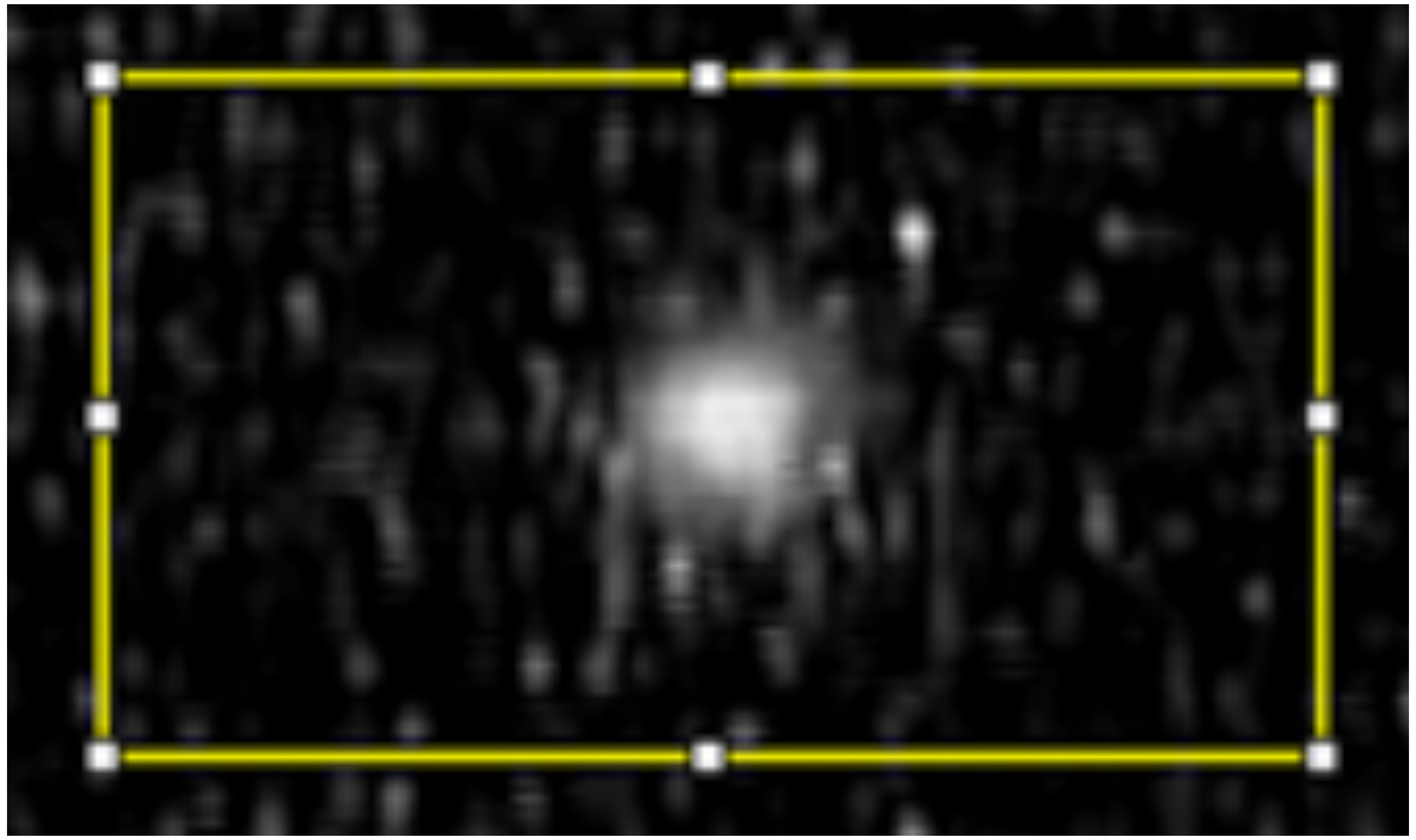


※ In this measurement, Fluorescent viewer was removed and Be 1mm was inserted.

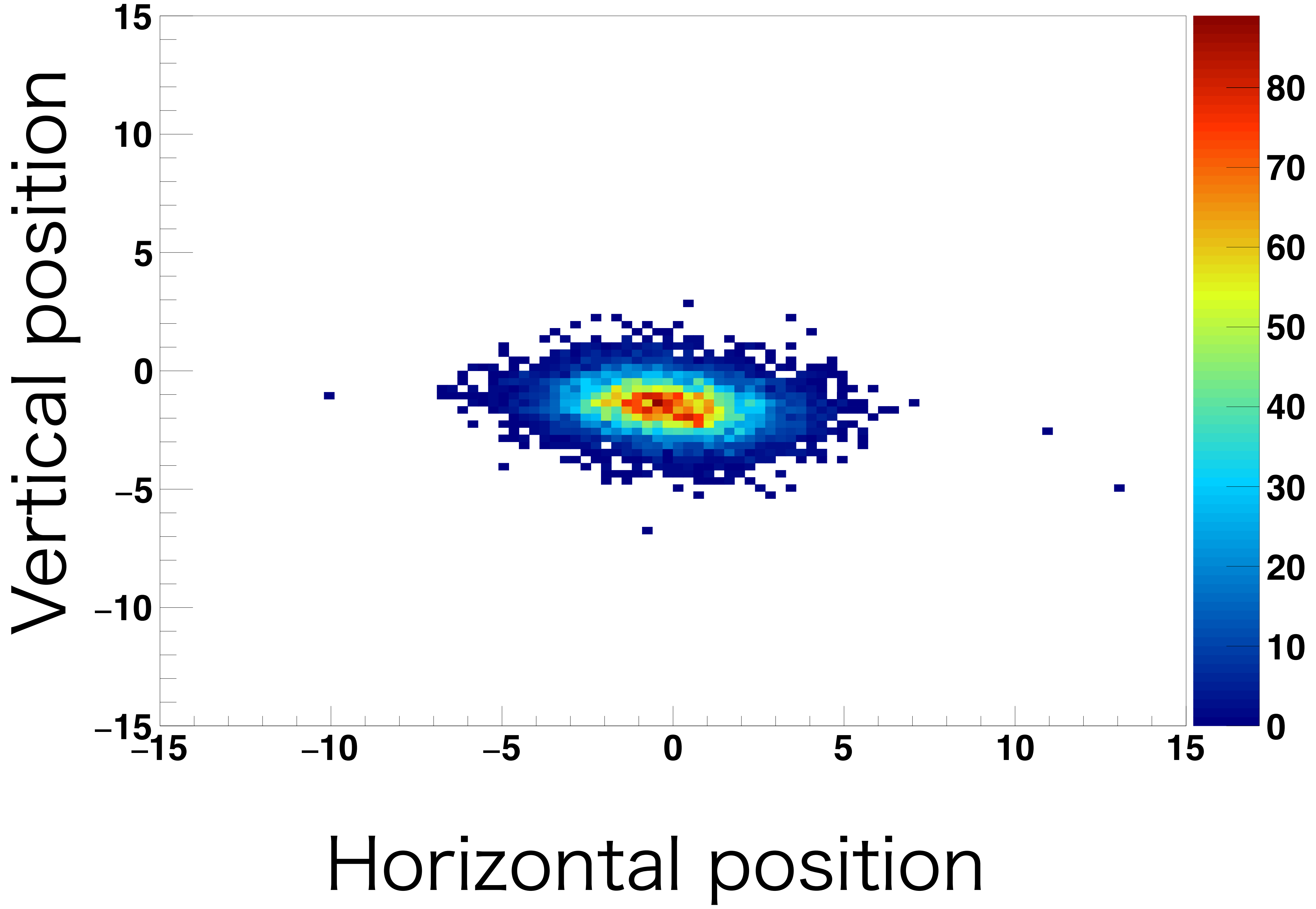
A. Compare Beam Spot measured by Viewer and Tracker

Change optics and compare
- fluorescent viewer image of primary beam (Kr^{36+})
- position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)

Fluorescent viewer (Kr^{36+})



Gas tracker (Kr^{34+})



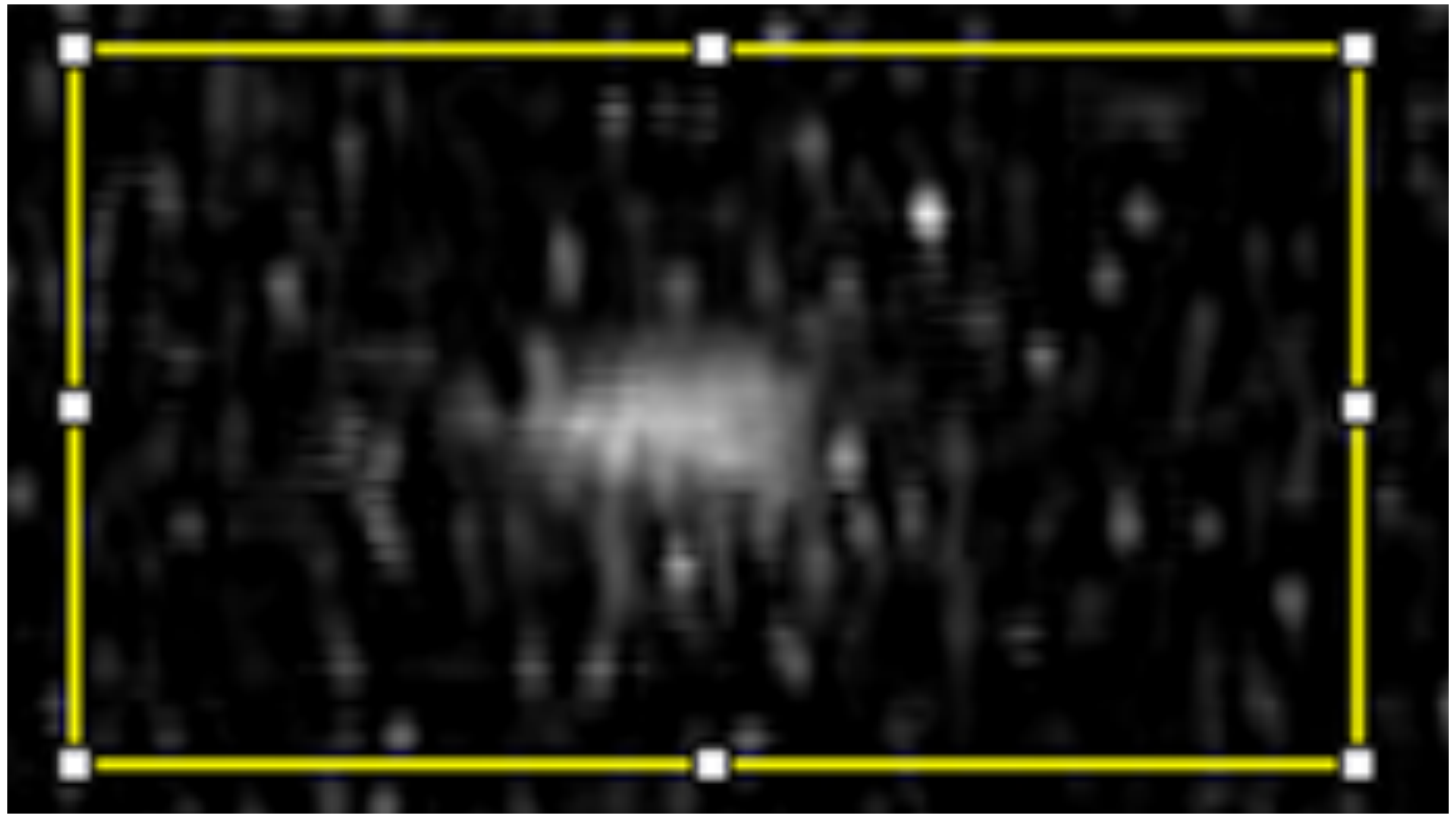
※ In this measurement, Fluorescent viewer was removed and Be 1mm was inserted.

A. Compare Beam Spot measured by Viewer and Tracker

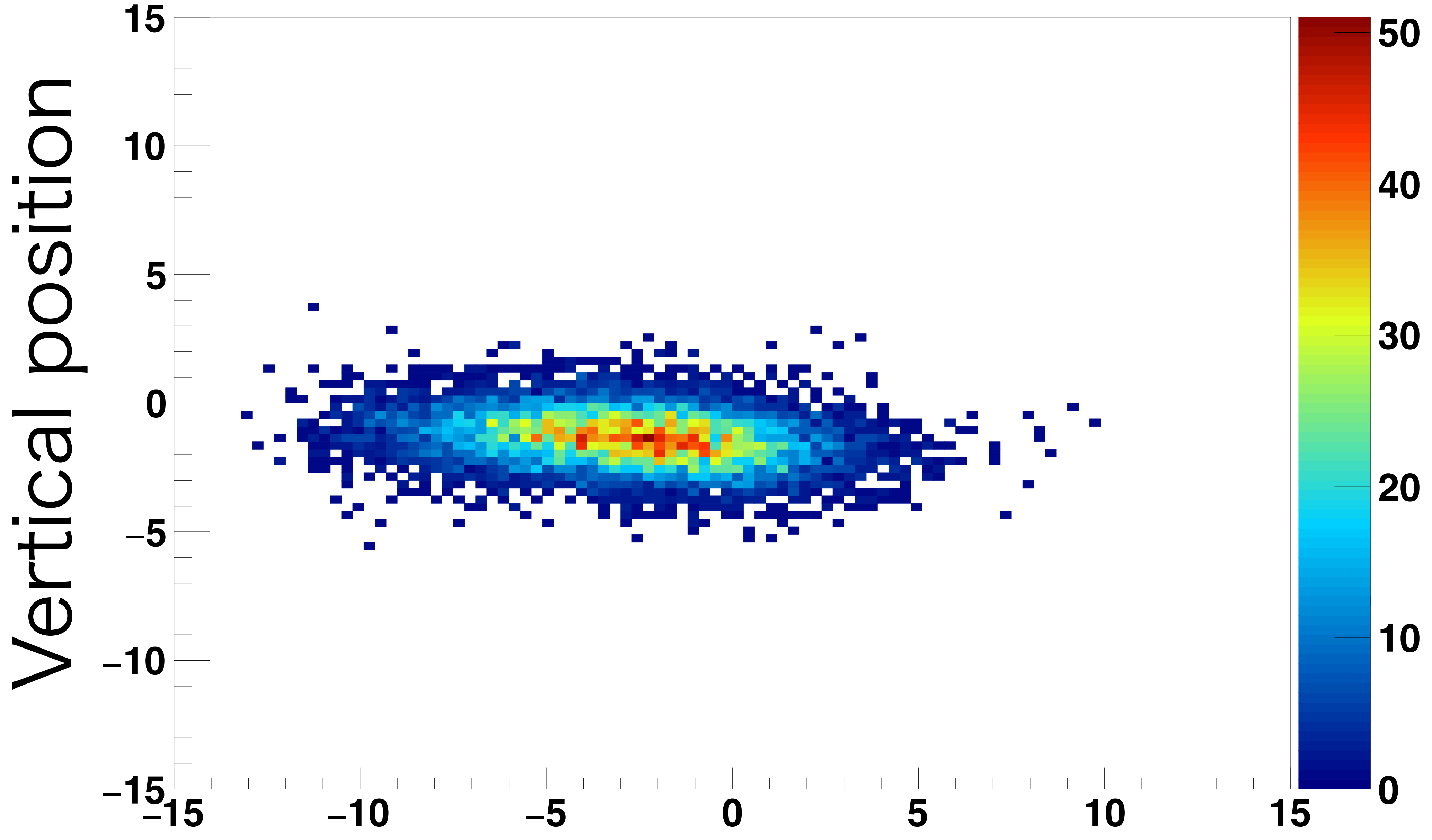
Change optics and compare

- fluorescent viewer image of primary beam (Kr^{36+})
- position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)

Fluorescent viewer (Kr^{36+})



Gas tracker (Kr^{34+})

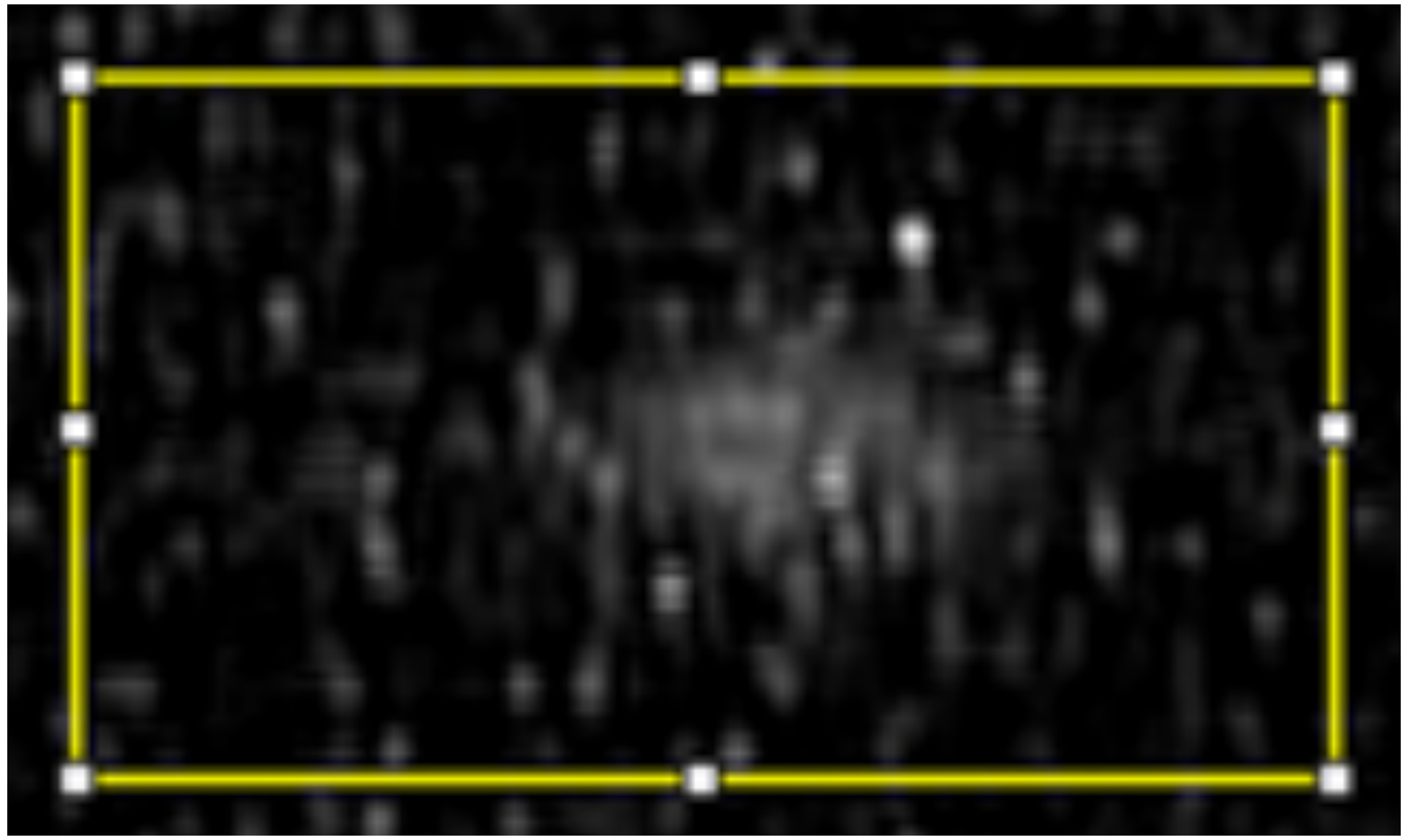


※ In this measurement, Fluorescent viewer was removed and Be 1mm was inserted.

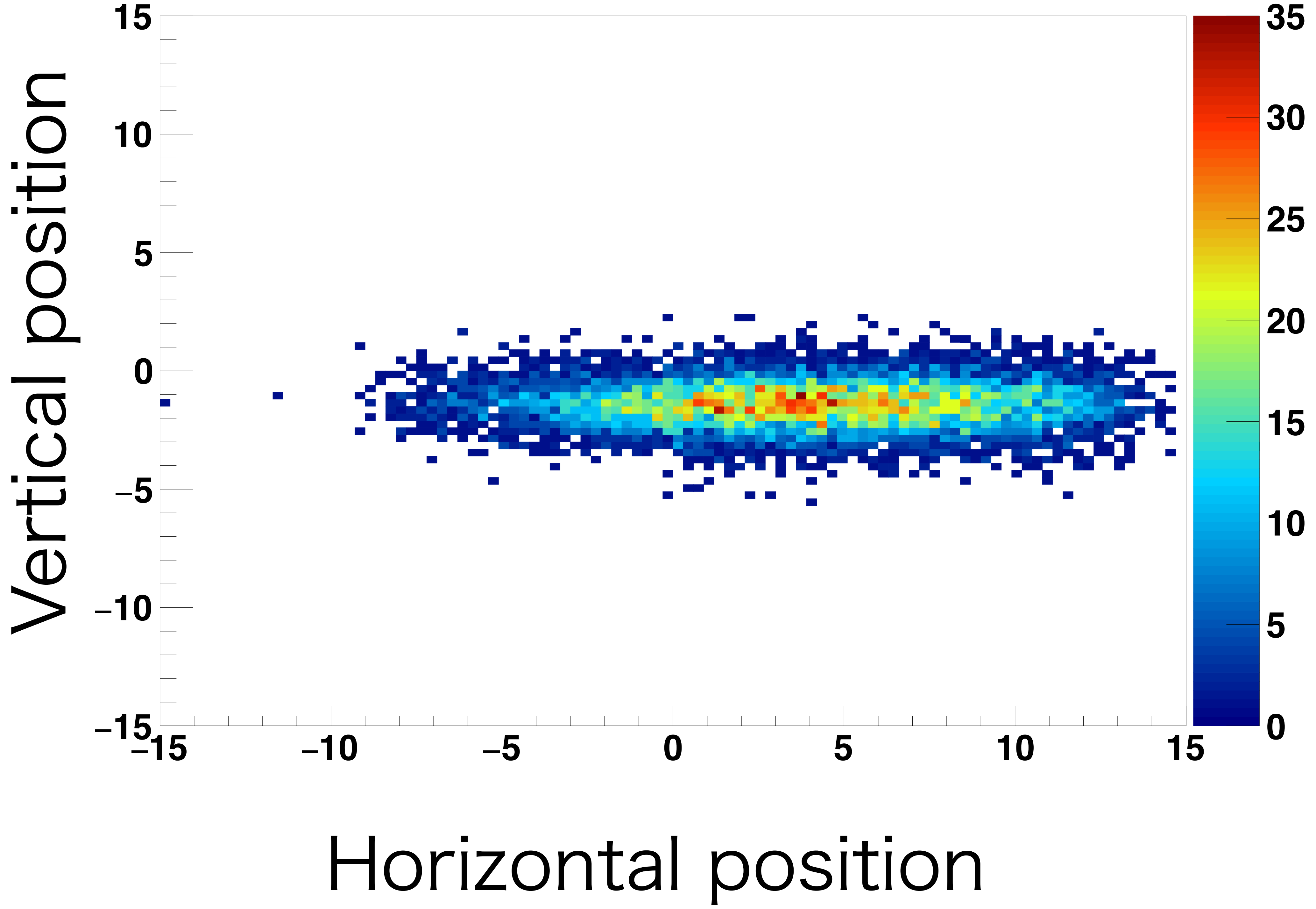
A. Compare Beam Spot measured by Viewer and Tracker

Change optics and compare
・fluorescent viewer image of primary beam (Kr^{36+})
・position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)

Fluorescent viewer (Kr^{36+})



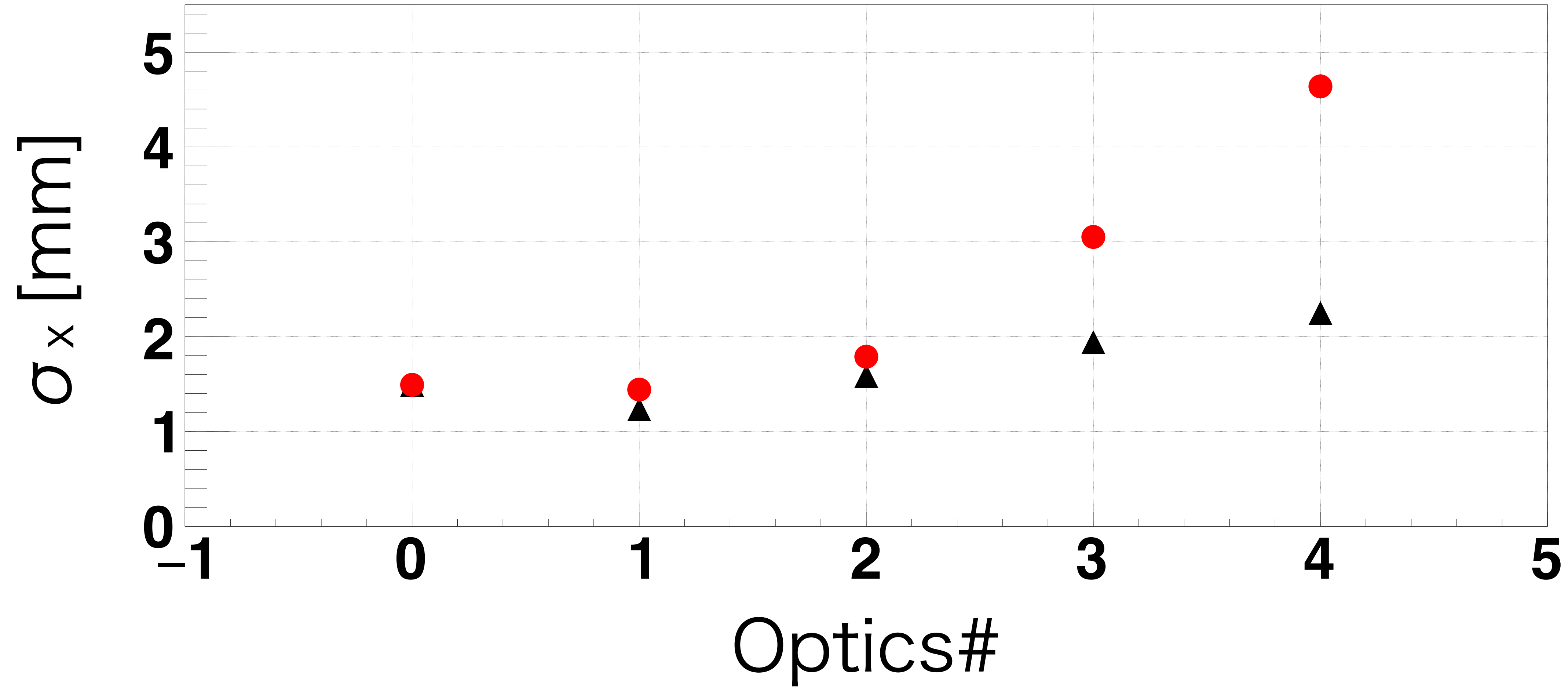
Gas tracker (Kr^{34+})

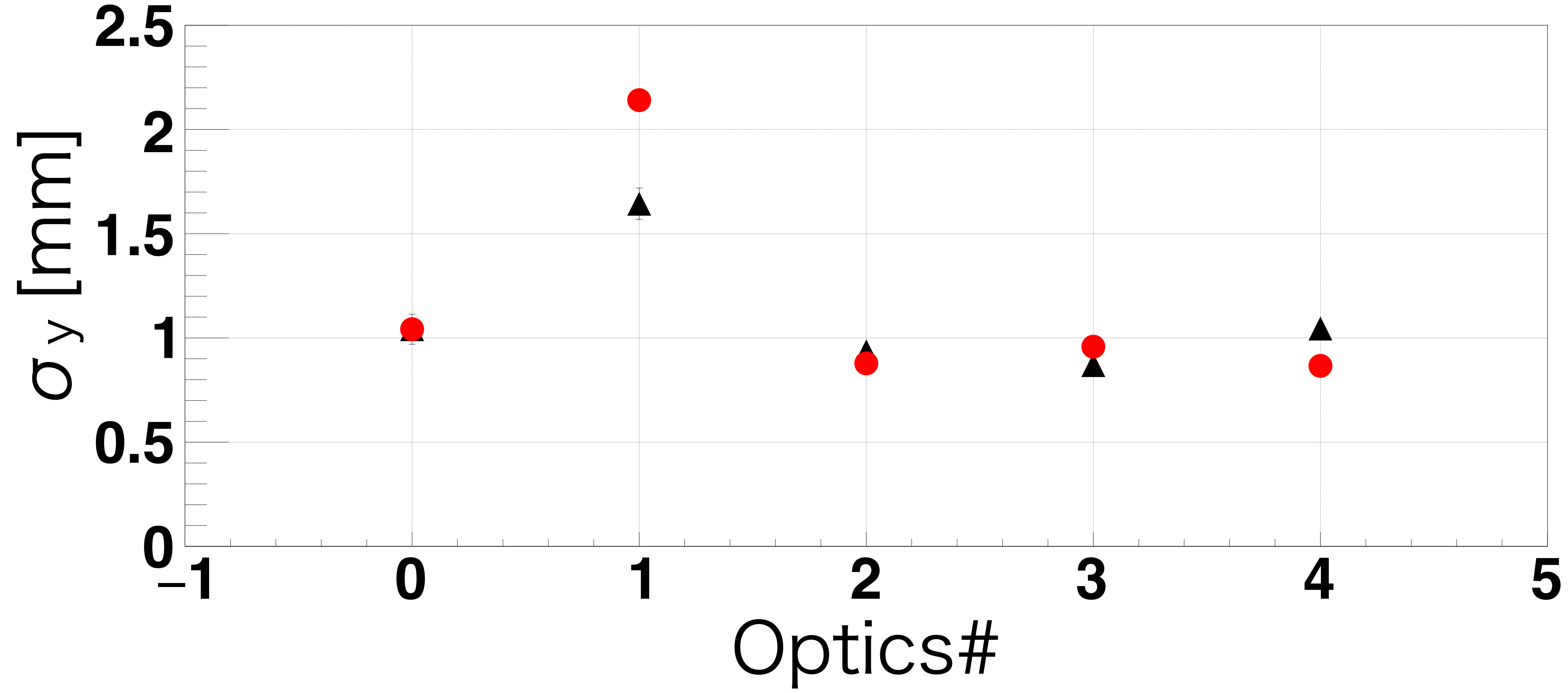


※ In this measurement, Fluorescent viewer was removed and Be 1mm was inserted.

A. Compare Beam Spot measured by Viewer and Tracker

Change optics and compare
- fluorescent viewer image of primary beam (Kr^{36+})
- position distribution of secondary beam (Kr^{34+}) tracked by PPAC (gas detector)



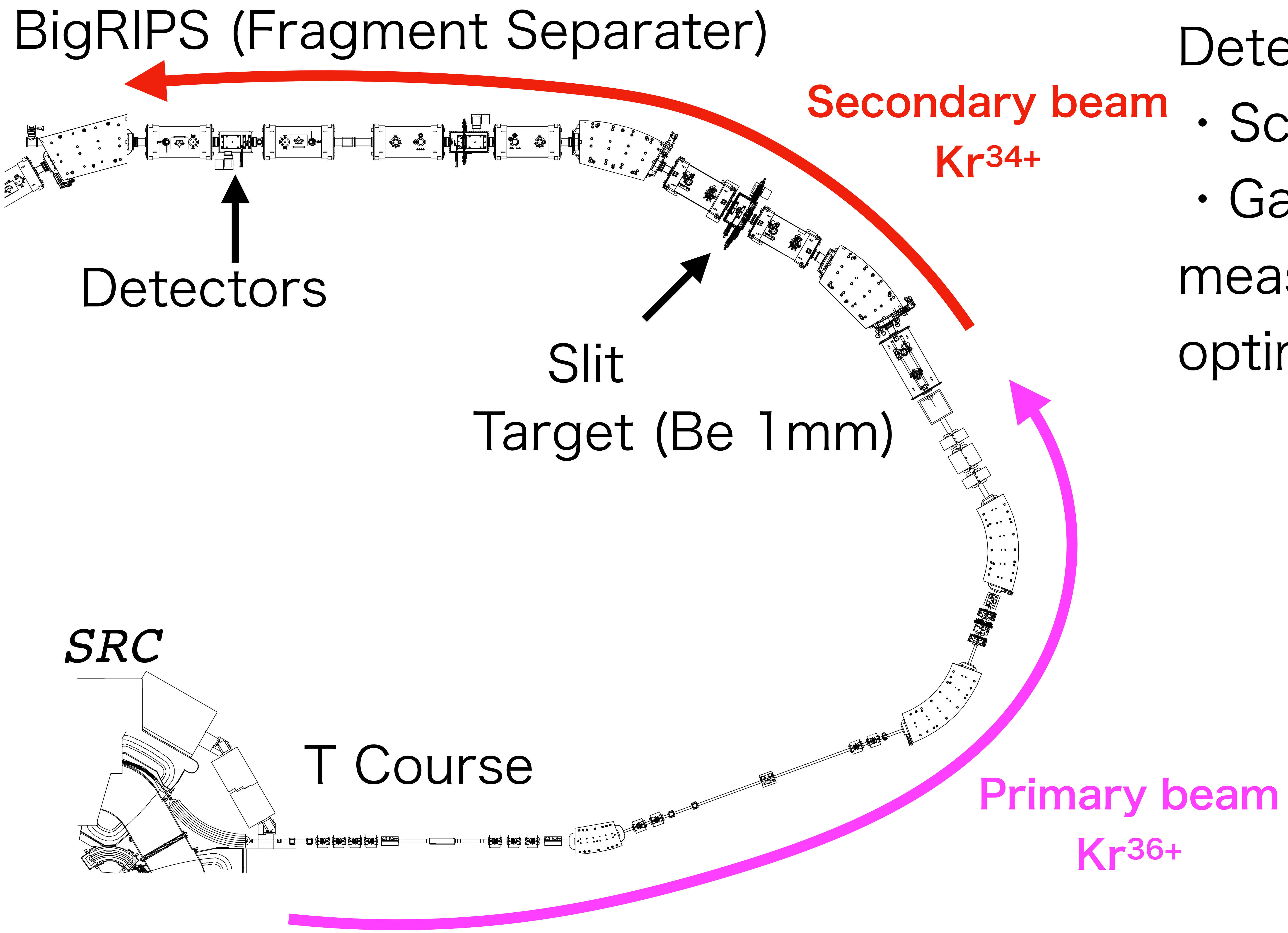


Red : Viewer
Black: Gas tracker
※ Viewer and gas tracker is calibrated by optics 0 data.

Data of viewer and tracker
→ (qualitatively) consistent
※ When the spot become wide,
non linearity may not be negligible

Tracking distribution of secondary beam is good probe for the primary beam spot!!

B. Auto Tuning with “High Intensity“ Primary Beam



Detector in downstream

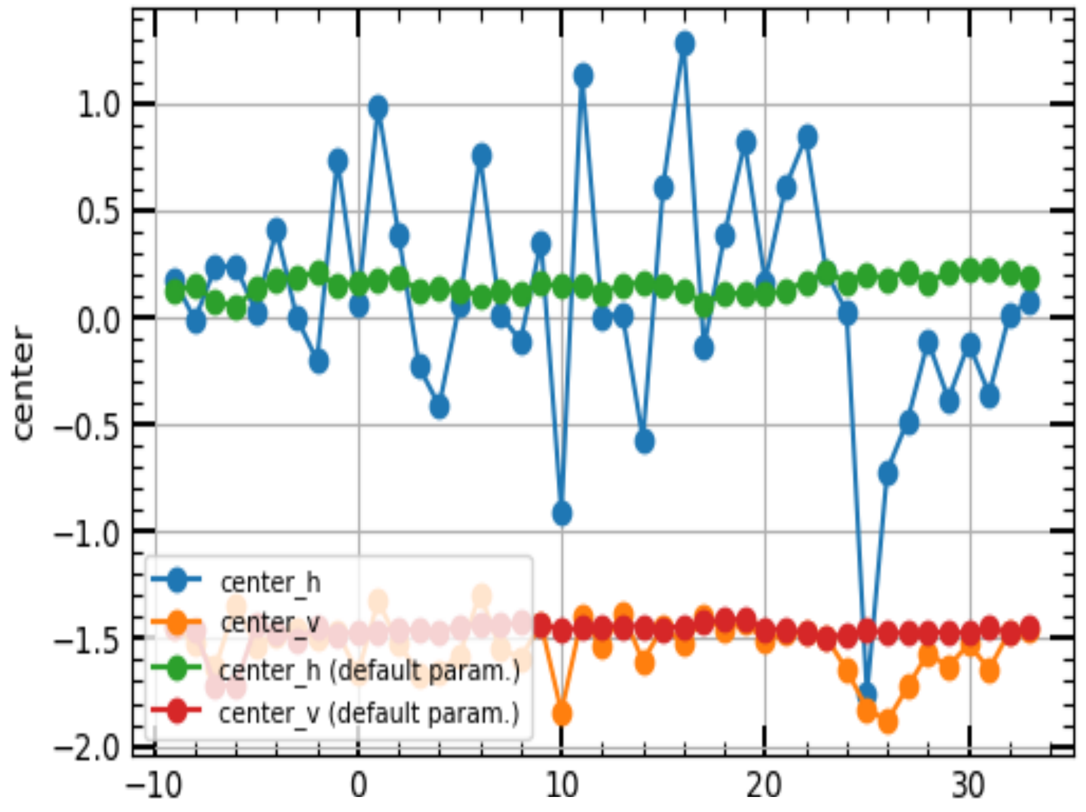
・Scintillator (for Transmission)

・Gas tracker (for spot)

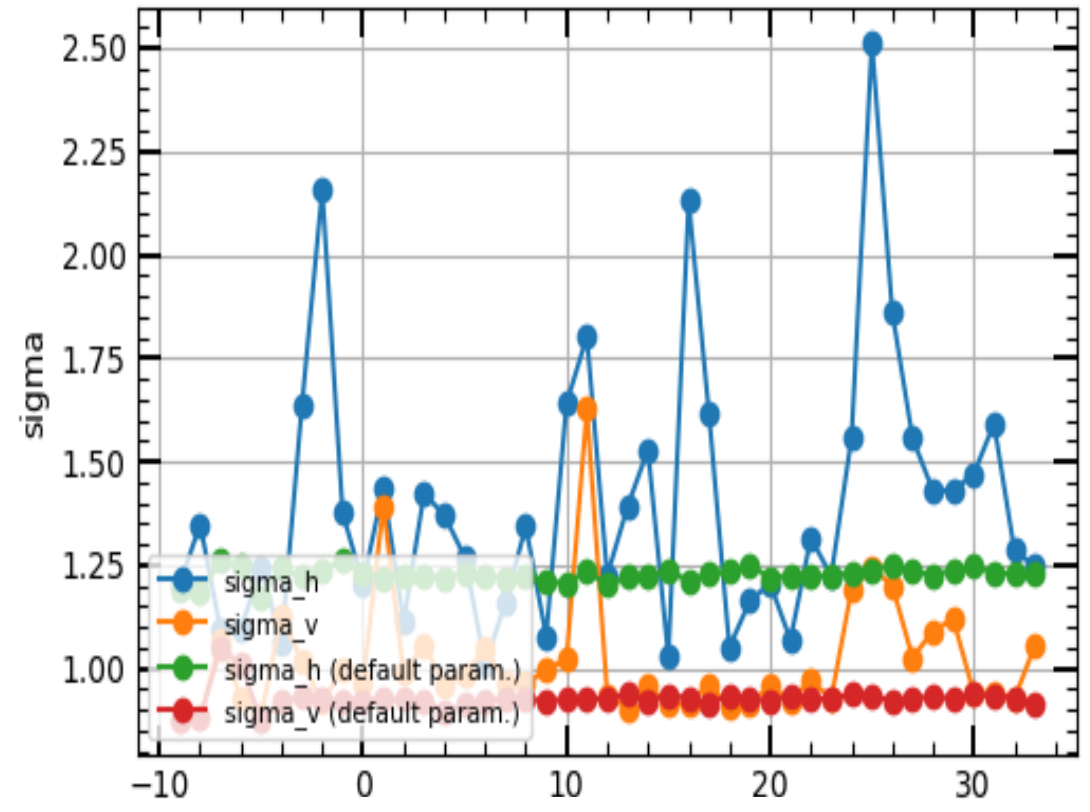
measure Kr^{34+} **10 kcps** / 0.0001 enA

optimize **26 enA** primary beam Kr^{36+}

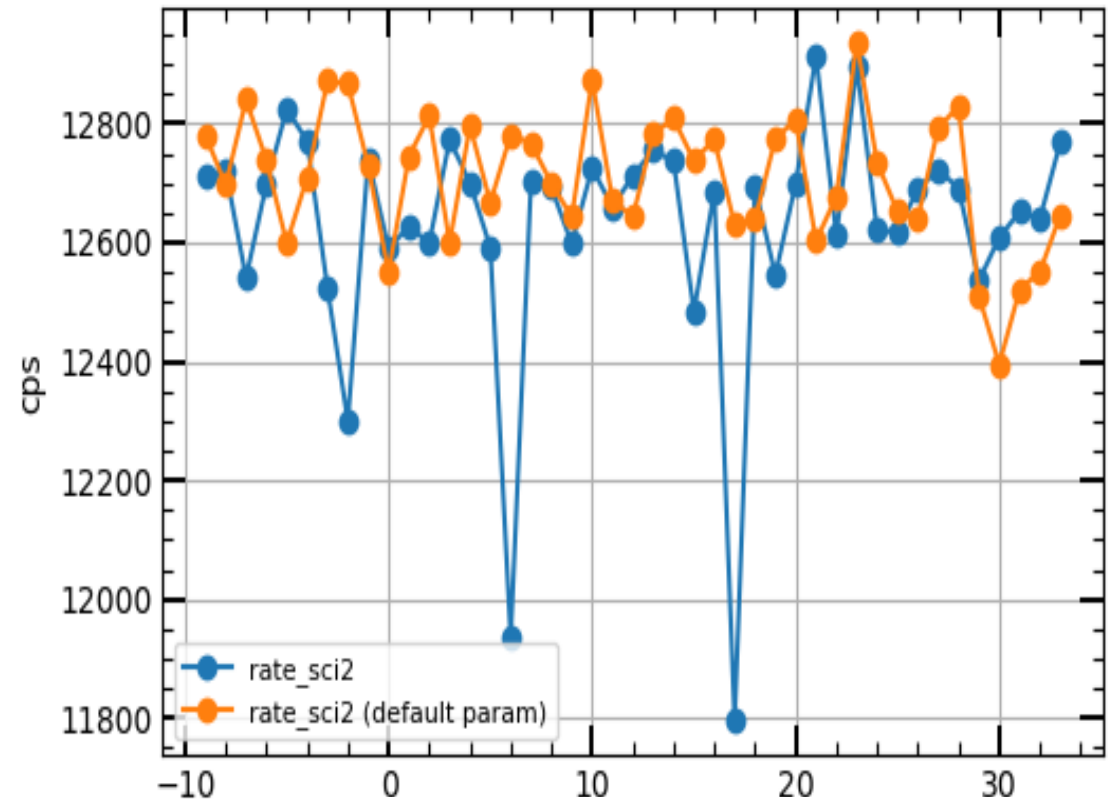
B. Auto Tuning with “High Intensity“ Primary Beam



$Pos_{h,\,v}$ vs Epoch

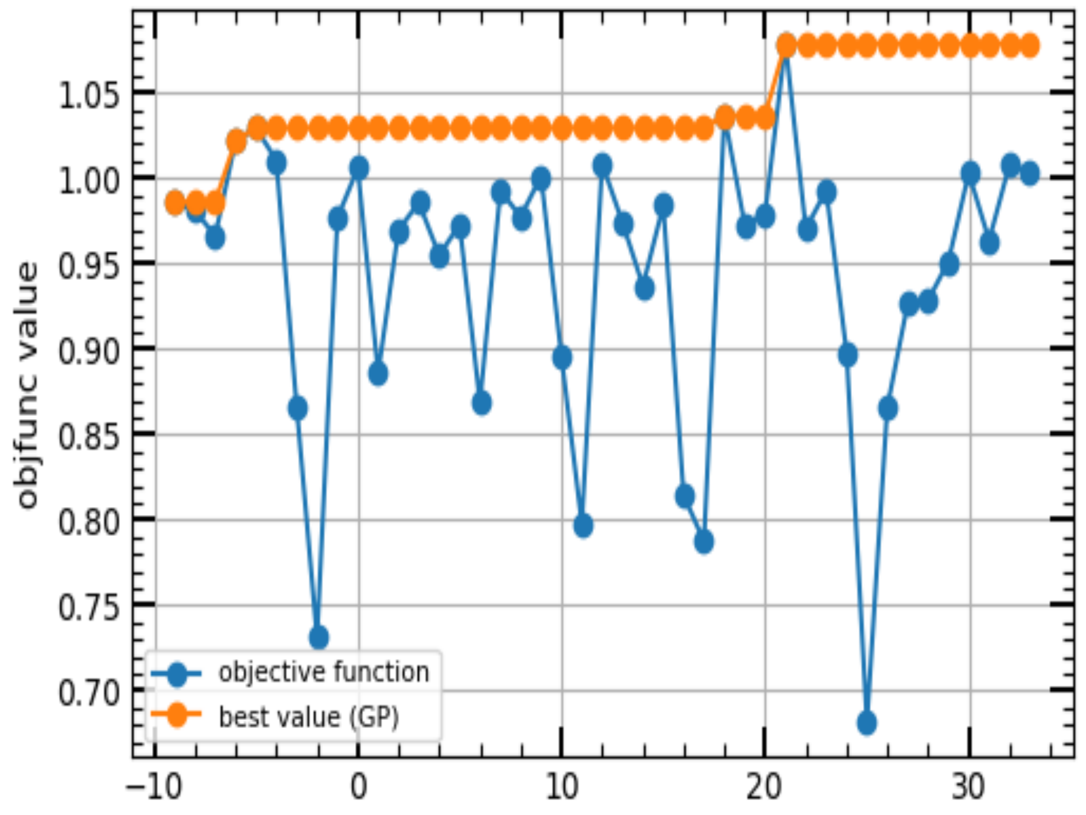


$\sigma_{h,\,v}$ vs Epoch

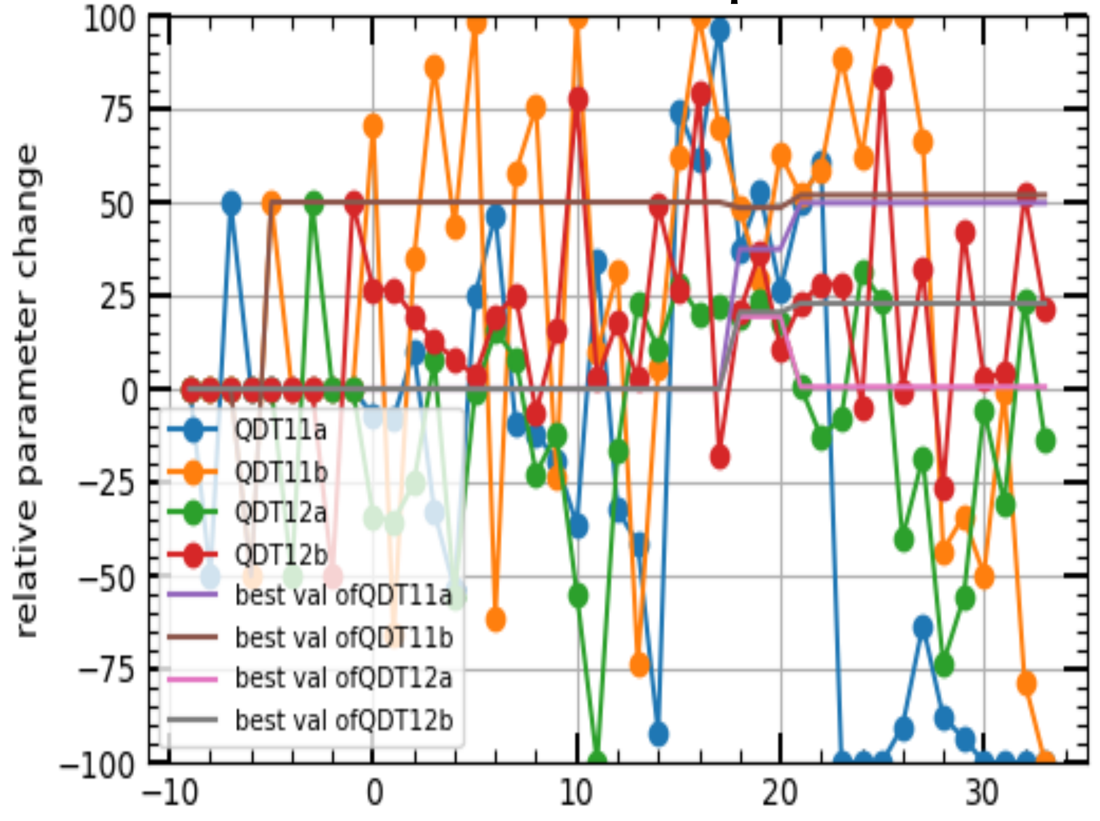


Count rate vs Epoch

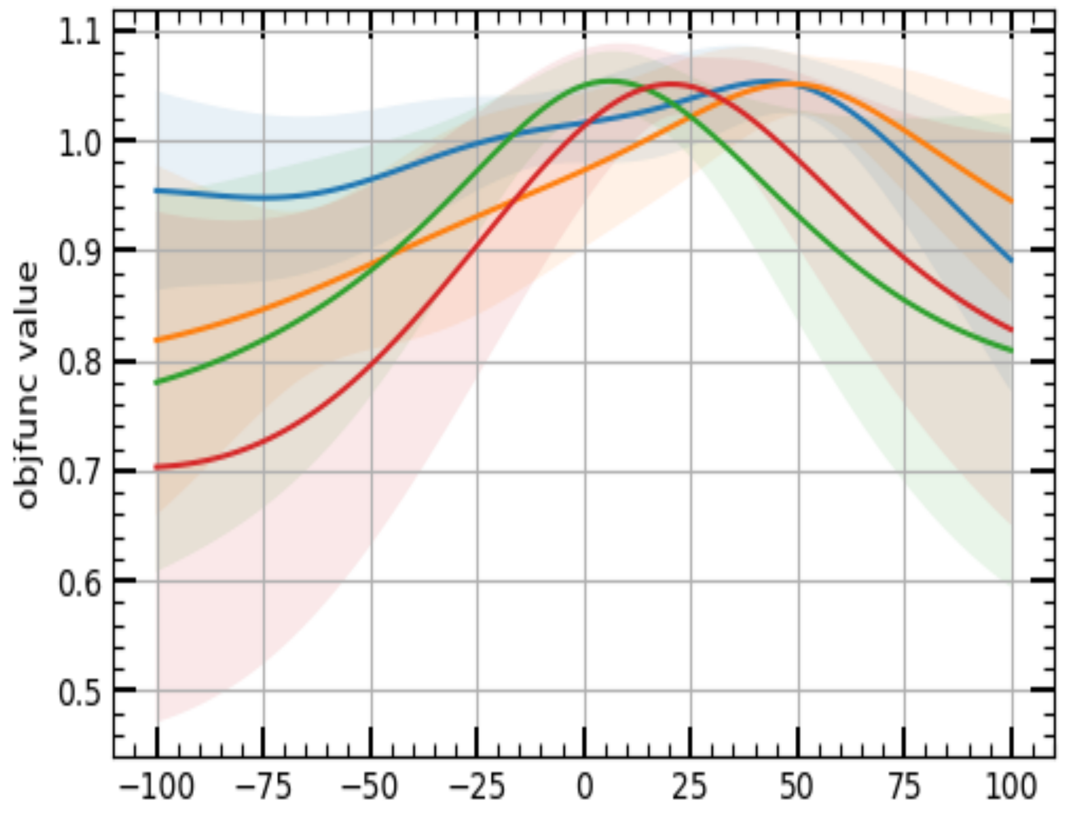
- ~ 4 params. (Quadrupoles)
- 1 epoch ~ 25 s
- 1 try ~ 30 min
- ~ 26 enA



Objective function vs Epoch



Current of Q vs Epoch

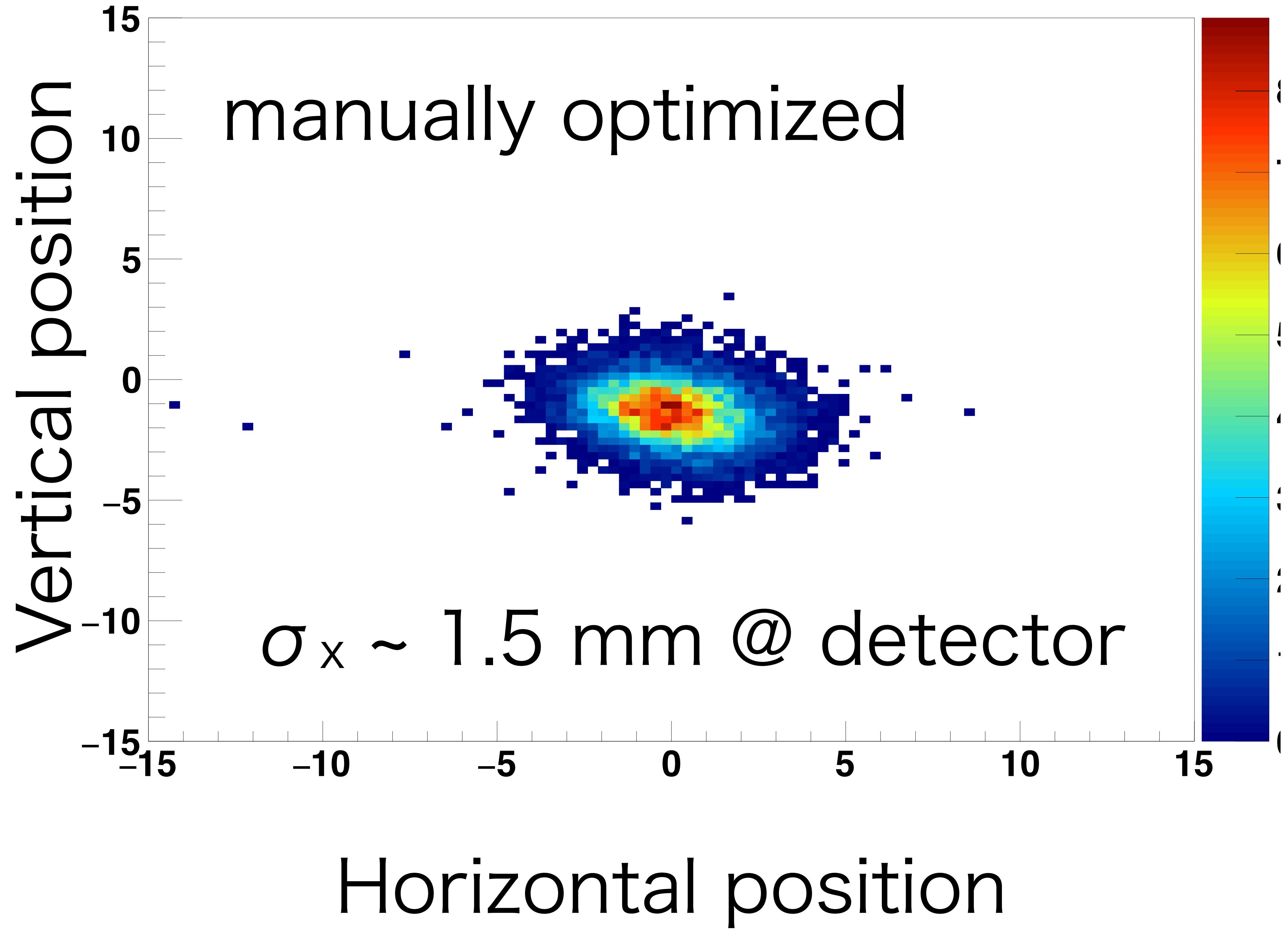


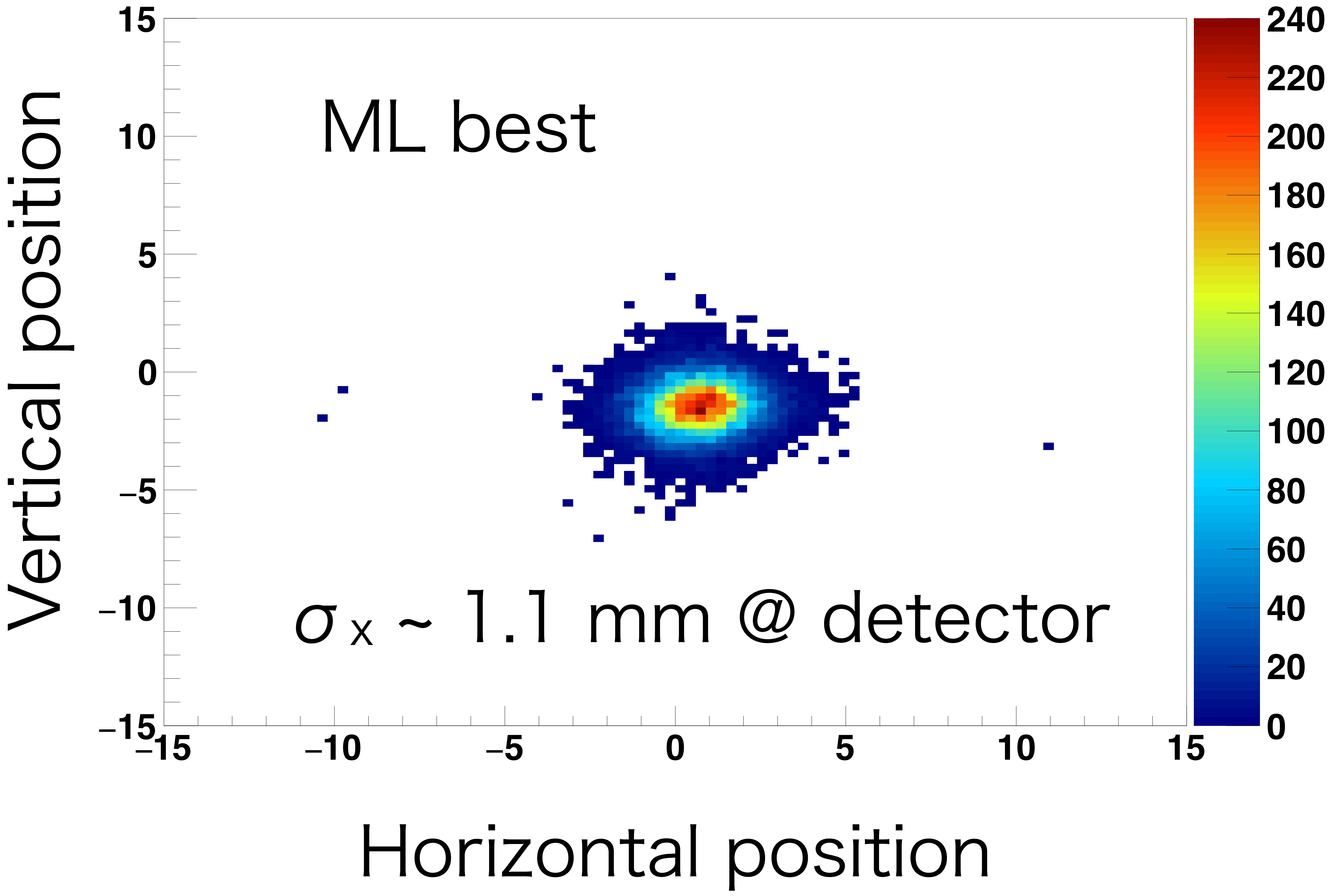
Objectives vs Current of Q

※ time of one epoch is limited by measurements by detectors

4th Exp. : Result of Auto Tuning with High Intensity Beam

Compare with manually optimized vs ML optimized optics
→ **Succeed to reduce spot size by 30%, keeping the transmission**





Achievements so far and Next Challenges

Achievements

- Framework with auto tuning based on ML algorithm
- Auto tuning of primary beam line optics using GPR
- Simultaneous indicator of beam spot and intensity for **high intensity** primary beam

~ 26 enA

Next Challenges

- How to confirm the result is “best solution”, instead of the local minimum？
- More efficient algorithm (c.f. more than 30 min. for 4 param. now)
- New scheme to control parameter range in the safety region

Algorithms to be tested

A. Physics informed Gaussian Process

→ Learning from simulations in advance to perform optimize more efficiently even for multidimensional parameters.

B. Safe optimization using LineBO

→ Prepare distinct GP models for both the objective and the constraints.

C. Adaptive control (Extremum Seeking Control, etc…)

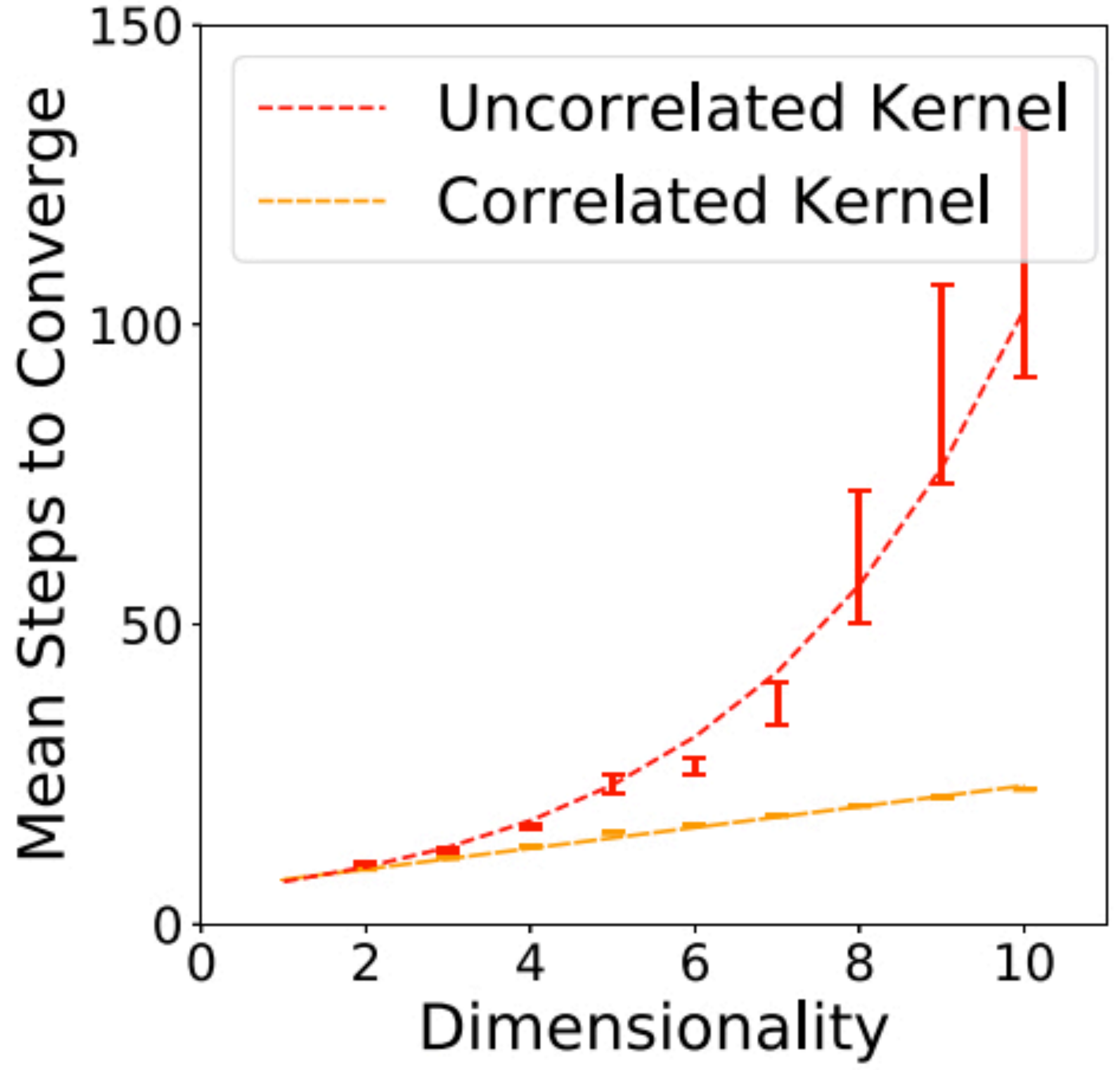
→ Continuously learning and optimizing parameters to keep up with (continuous) changes in the environment.

※ These ideas are based on a proposal by our member, Hiroki Fujii in RIKEN

Algorithms to be tested

A. Physics informed Gaussian Process

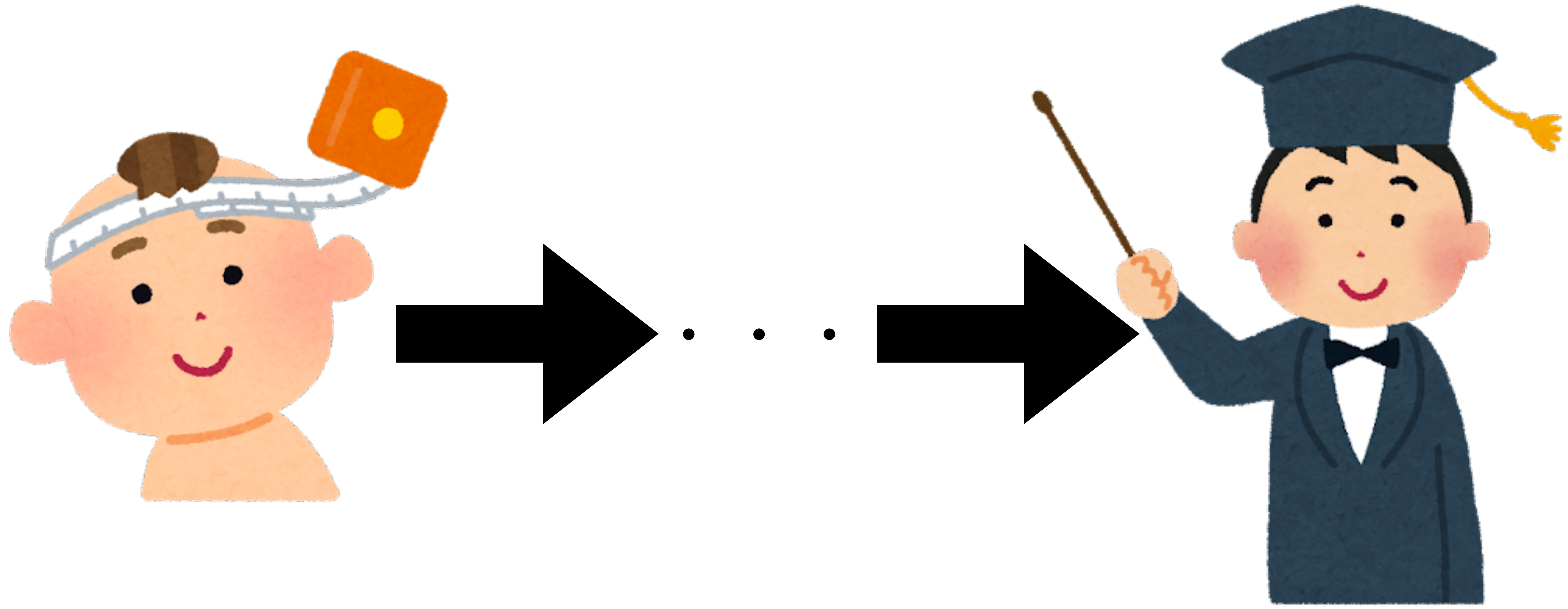
Modify kernel functions by simulations in advance
→ Start with more "physical" knowledge.
Optimize from real data + prior knowledge.



(d) Convergence tests

Convergence in
a relatively small epoch,
even with multiple dimensions.

Normal GPR

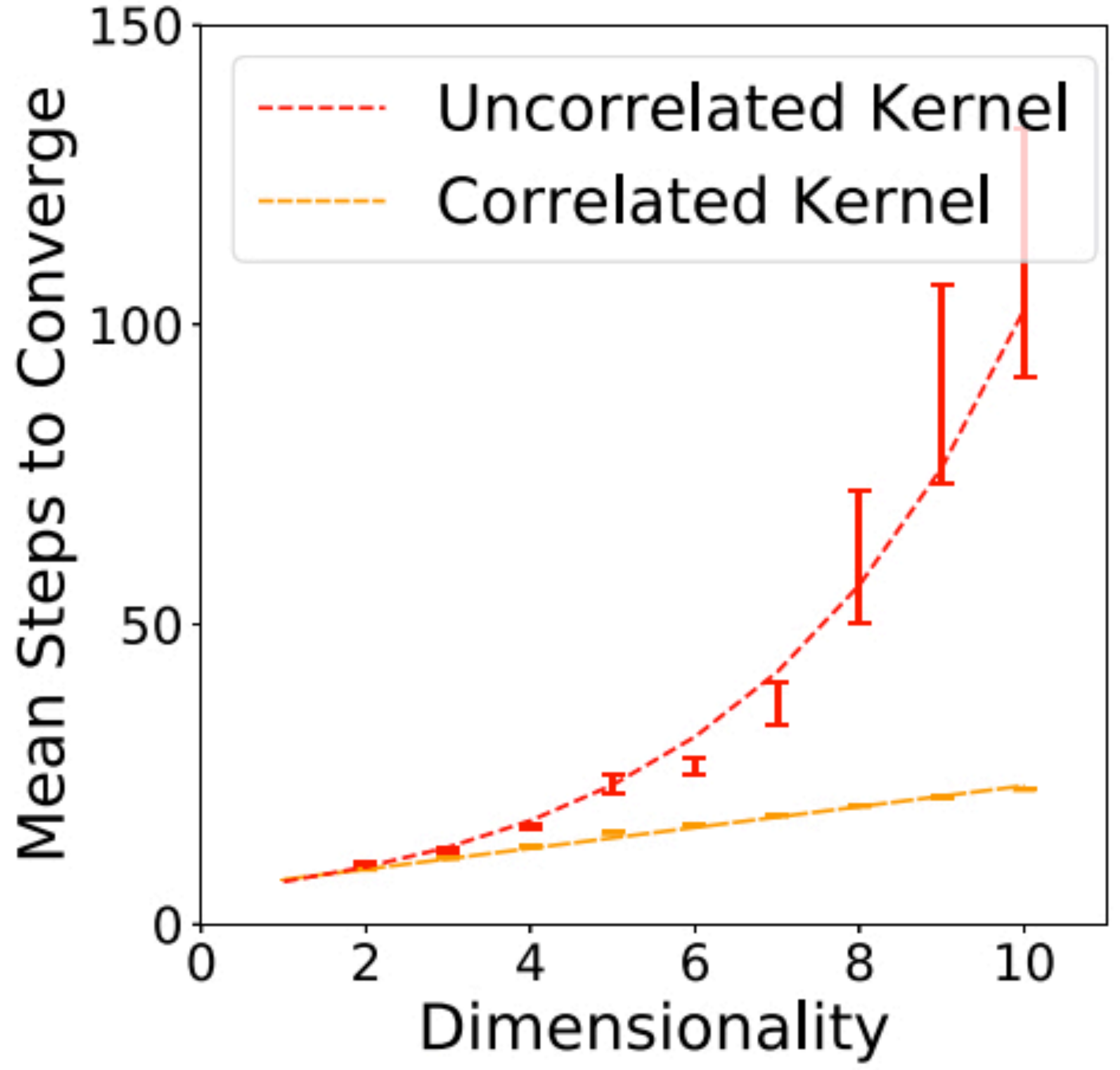


J. Duris, J. et al. Bayesian Optimization
of a Free-Electron Laser. Phys. Rev. Lett. 124, 124801 (2020).

Algorithms to be tested

<u>A. Physics informed Gaussian Process</u>

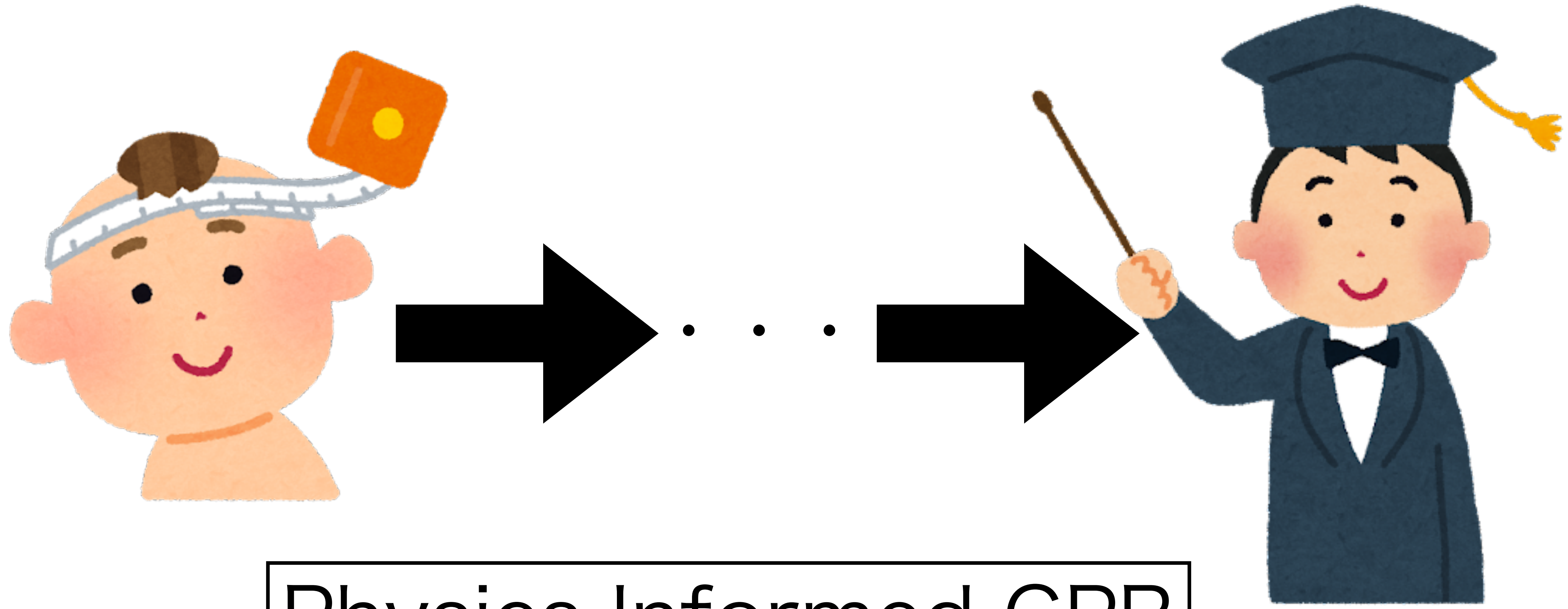
Modify kernel functions by simulations in advance
→ Start with more "physical" knowledge.
Optimize from real data + prior knowledge.



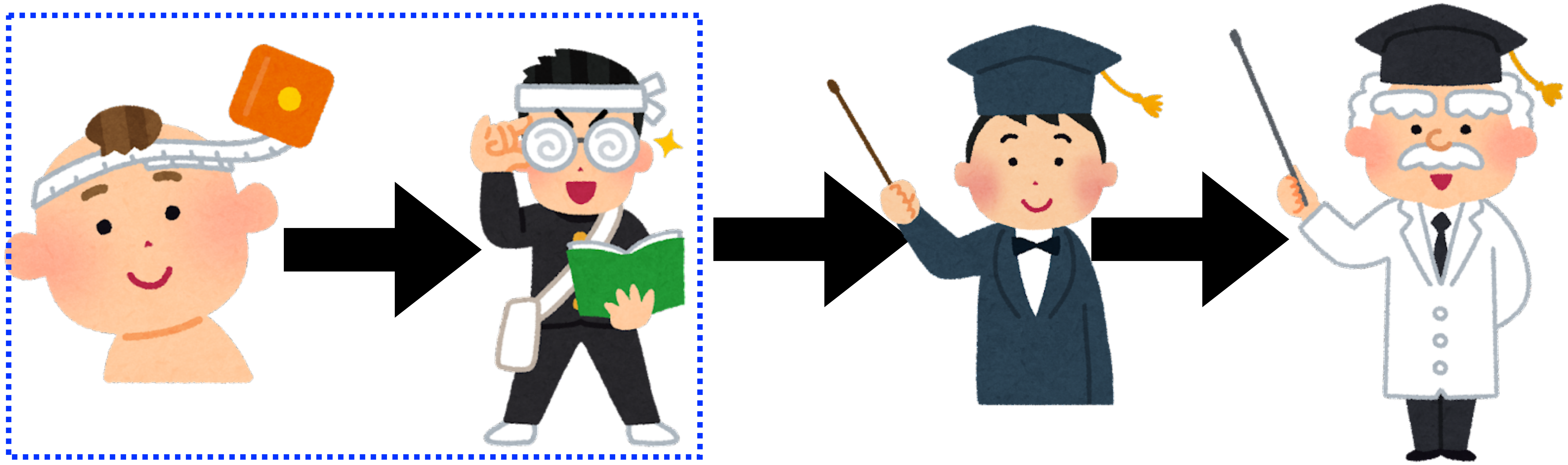
(d) Convergence tests

Convergence in
a relatively small epoch,
even with multiple dimensions.

Normal GPR



Physics Informed GPR



Simulation

J. Duris, J. et al. Bayesian Optimization
of a Free-Electron Laser. Phys. Rev. Lett. 124, 124801 (2020).

Algorithms to be tested

B. Safe optimization using LineBO

Objective

$$\max_{x\in\chi} f(x)\ (\text{or}\ \min f(x)) \quad \text{s.t.} \begin{cases} g_1(x) \le 0 \\ \vdots \\ g_l(x) \le 0 \end{cases}$$

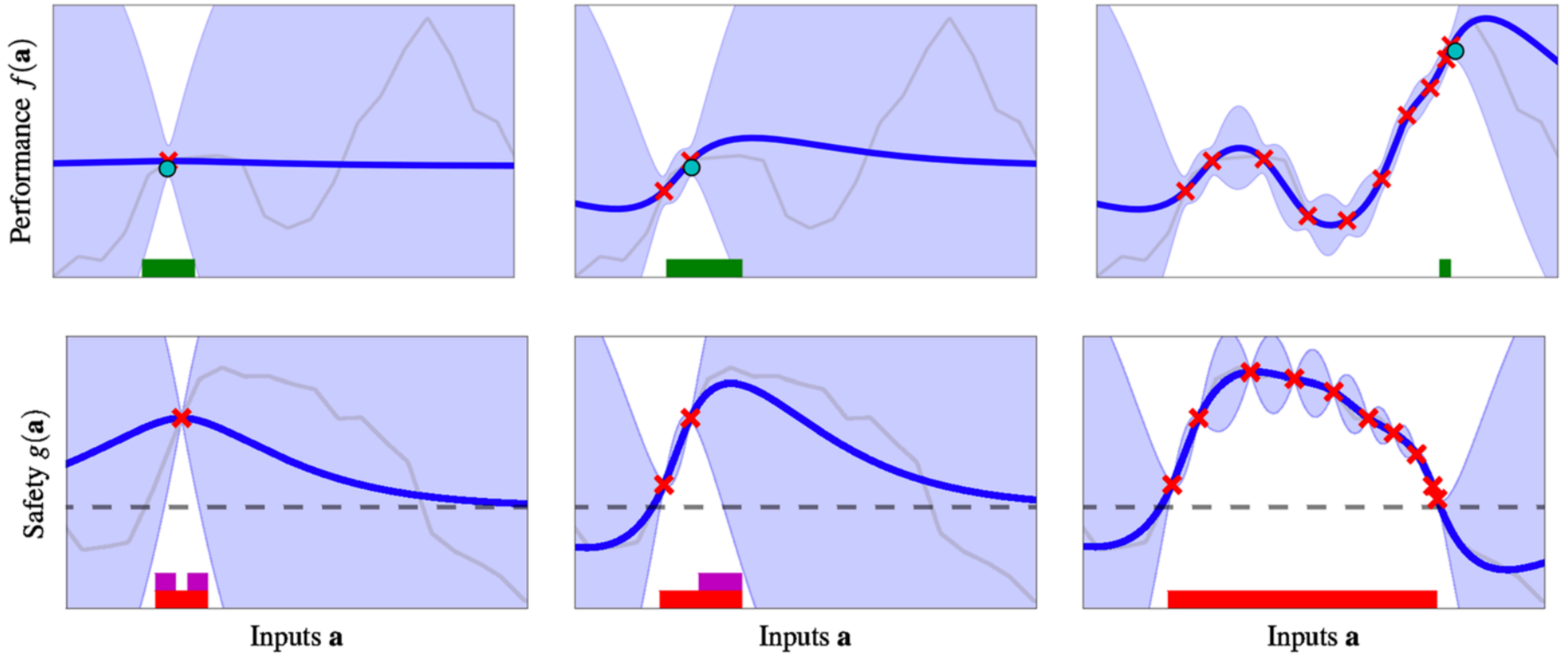
Approaches in the safe optimization community

Prepare distinct GP models for both the objective and the constraints
Evaluate within the safe set only

$$x_t = \arg\max_{x\in S_t^\tau} \mathrm{UCB}_f(\hat{m}_t, x, \delta)$$

$$s.t.\ S_t^\tau = \{x \in \chi : \max_{i\in[l]} \mathrm{UCB}_{g_i}(\hat{m}_t, x, \delta) \le \underline{-\tau}\}$$

Margin



(a) Initial, safe parameters. **(b)** Safe exploration. **(c)** After 10 evaluations: safe maximum found.

Felix Berkenkamp (2021)

Algorithms to be tested

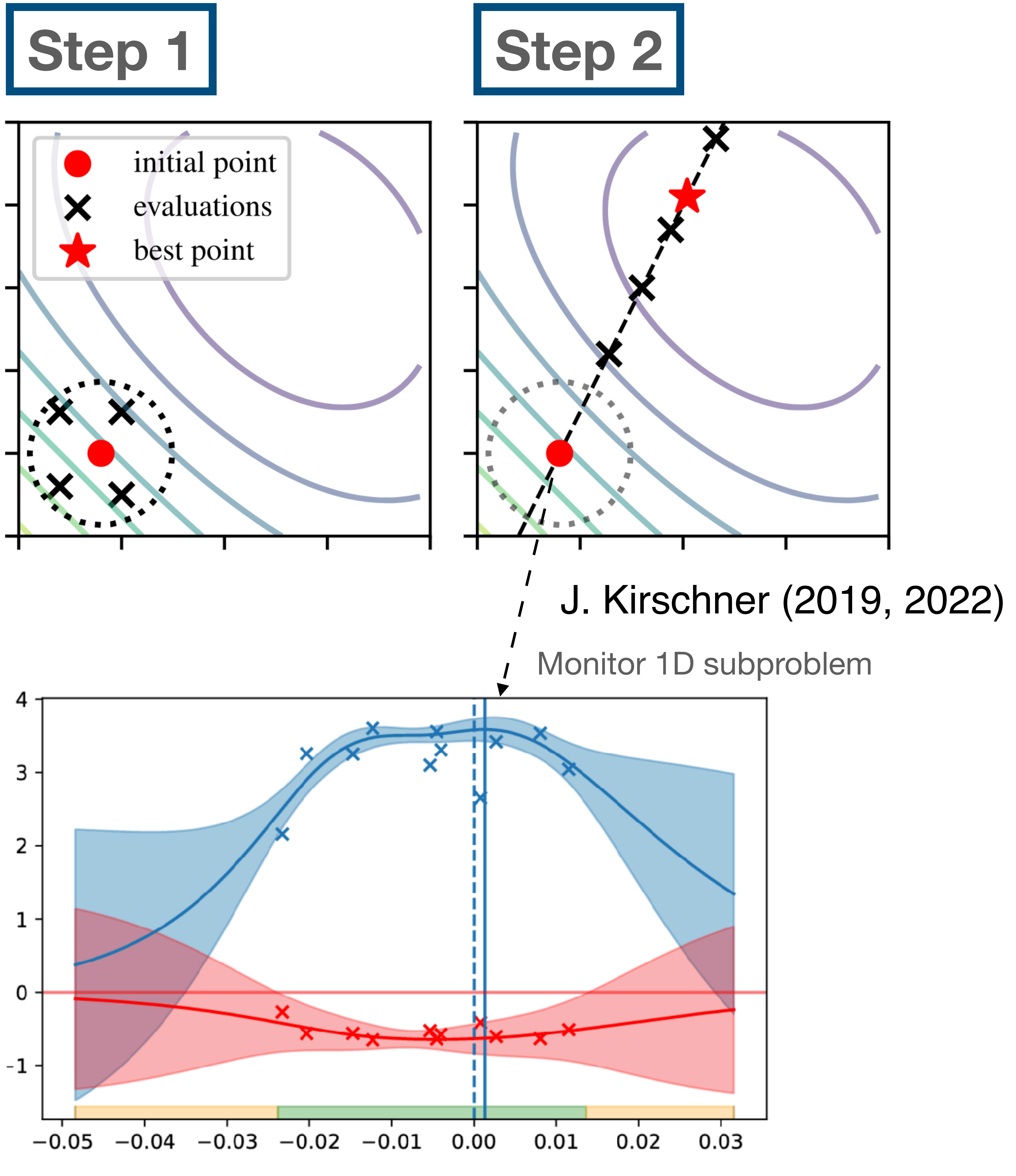
B. Safe optimization using LineBO

Solve a sequence of one-dimensional Bayesian Optimization problems on one dimensional subspaces.

Can be easily applied to higher dimensional systems.

Each subproblem is efficient

Flexible choice of acquisition function (e.g., can use 1-dimensional SafeOPT for high-dimensional systems)

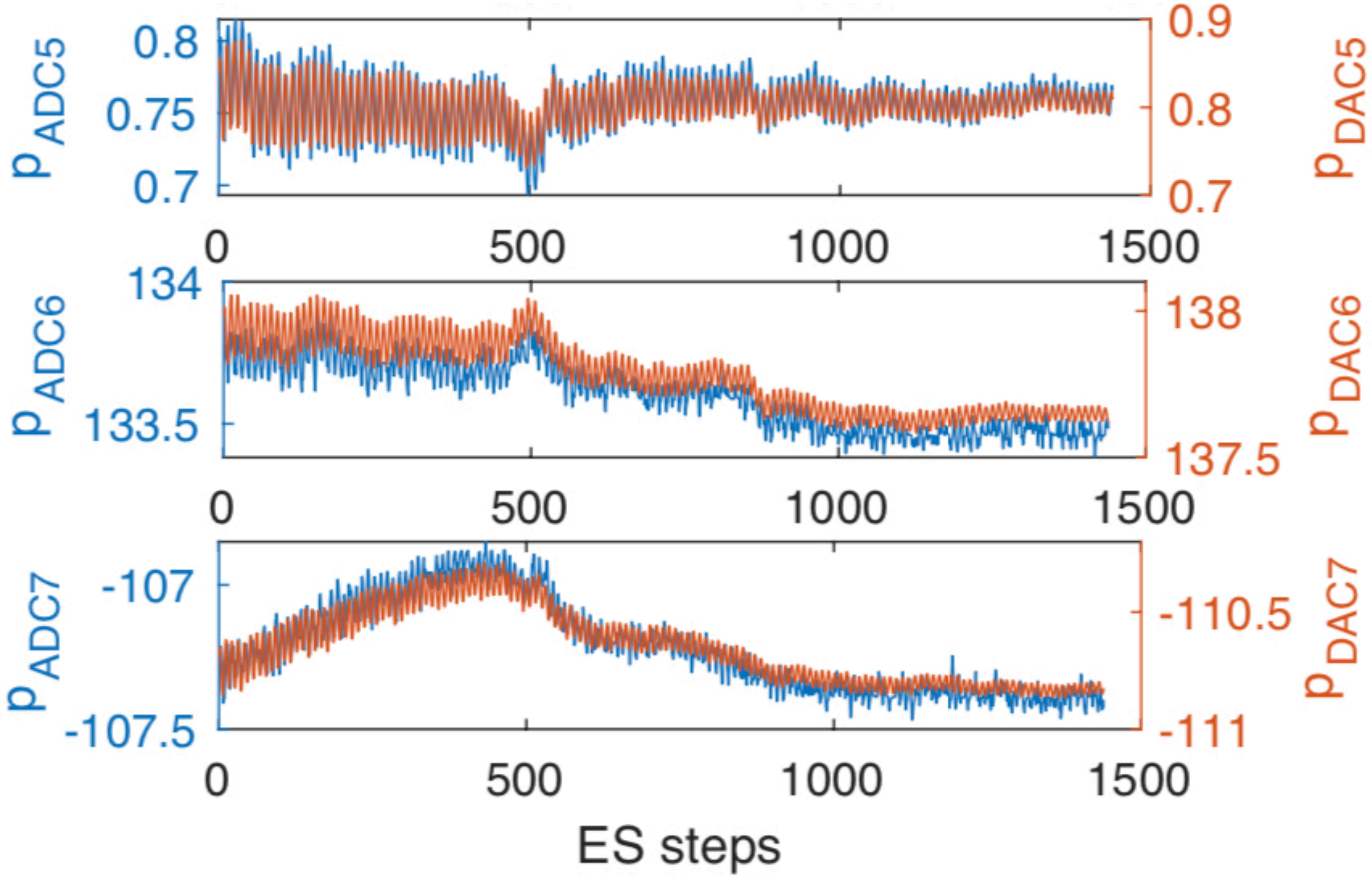


(c) User feedback

Algorithms to be tested

C. Adaptive control by Extremum Seeking Control

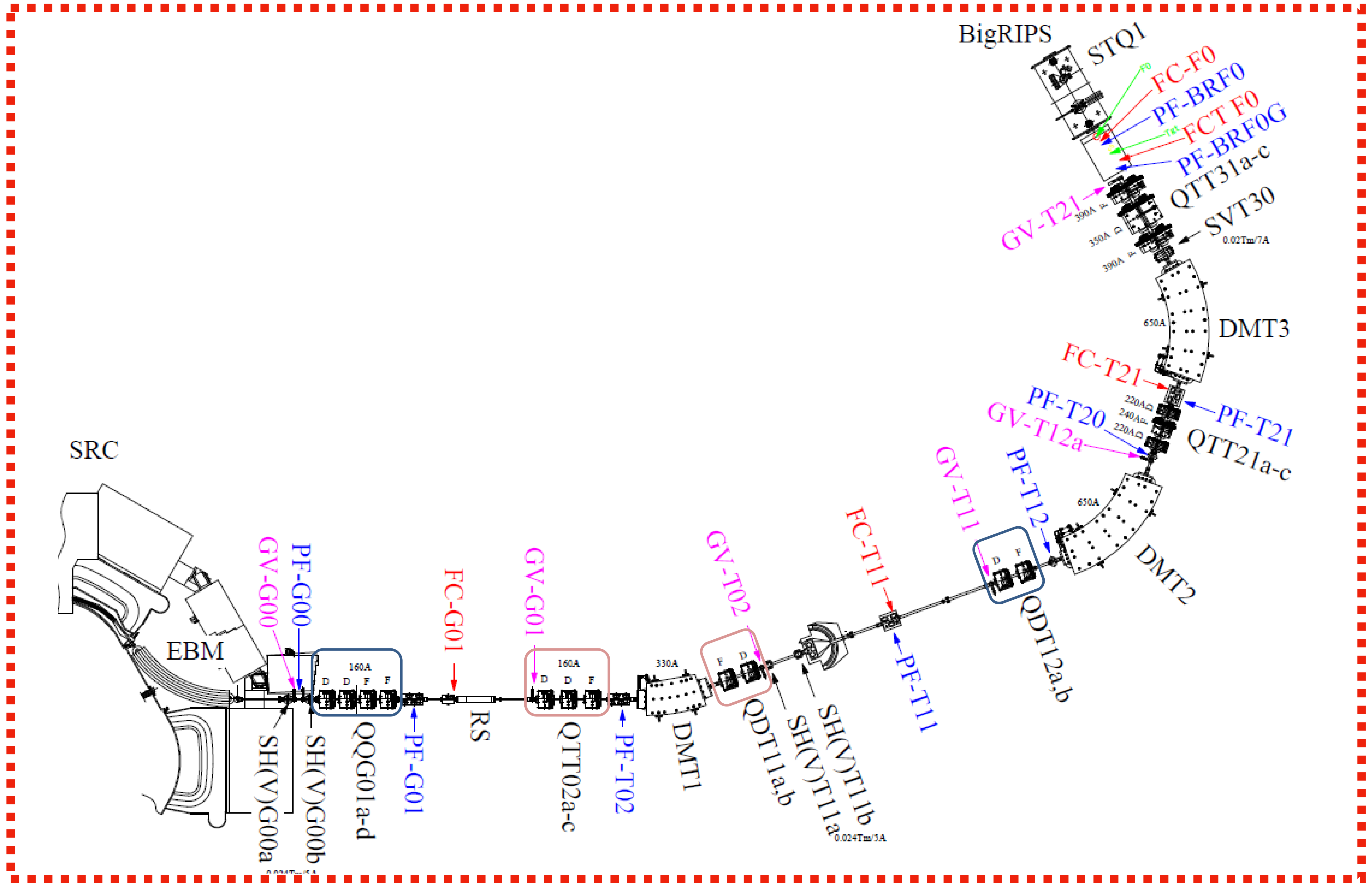
Extremum Seeking Control:
A method of continually optimizing from the response of the target function by constantly varying the parameters in minute increments. The response is separated by Fourier transform response can be separated from the noise, enabling optimization with high accuracy.



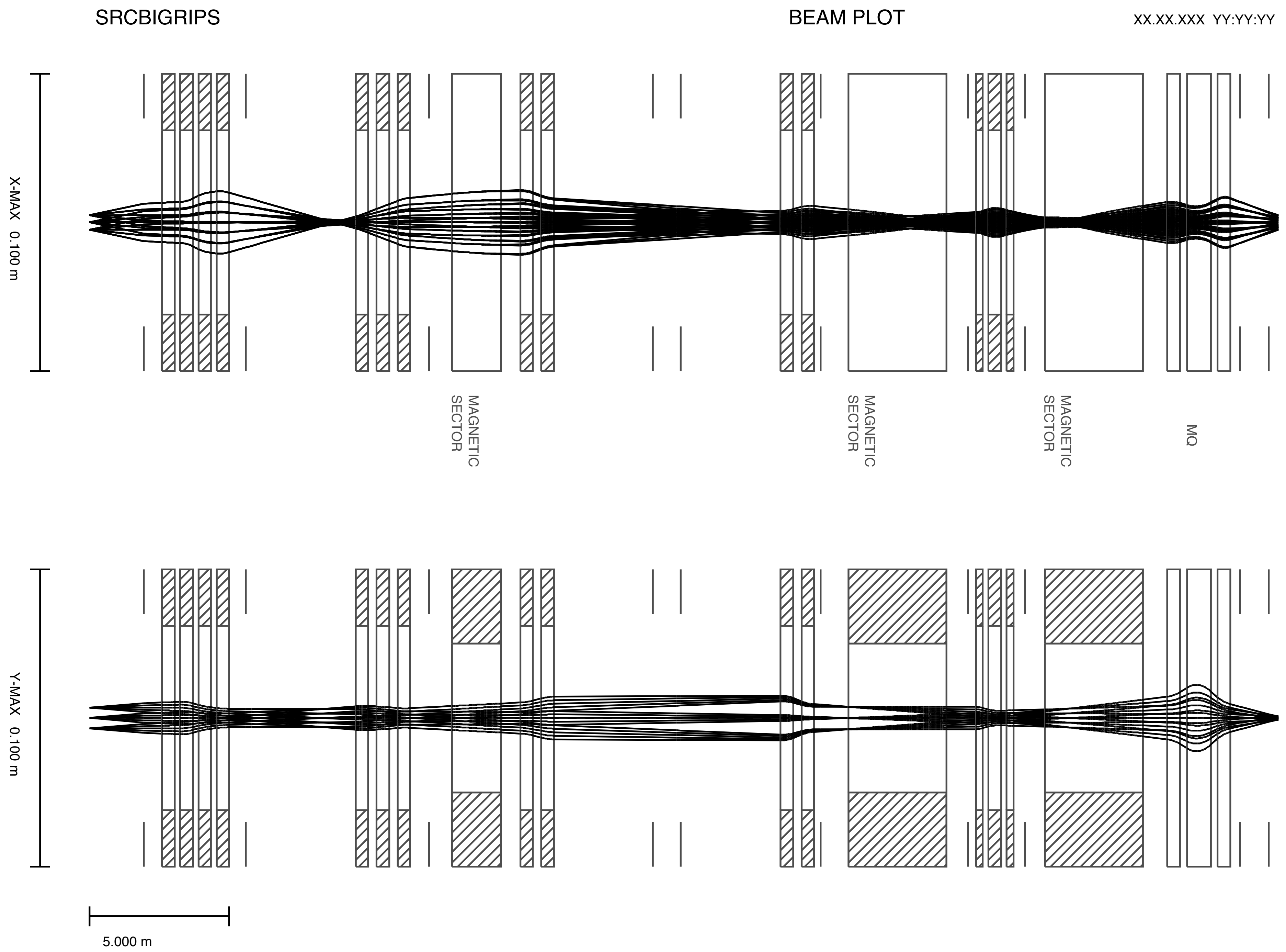
H. Fujii et al, Proc. of PASJ 2021 WEOB02

Data at a beamline in RIBF
Parameters: Quadrupoles / Steerers
Indicator : beam loss at slit / baffle

Recent activities…: Development with Simulations



Simulation by
gicosy (transfer matrix) + MOCADI (Monte Carlo simulation)

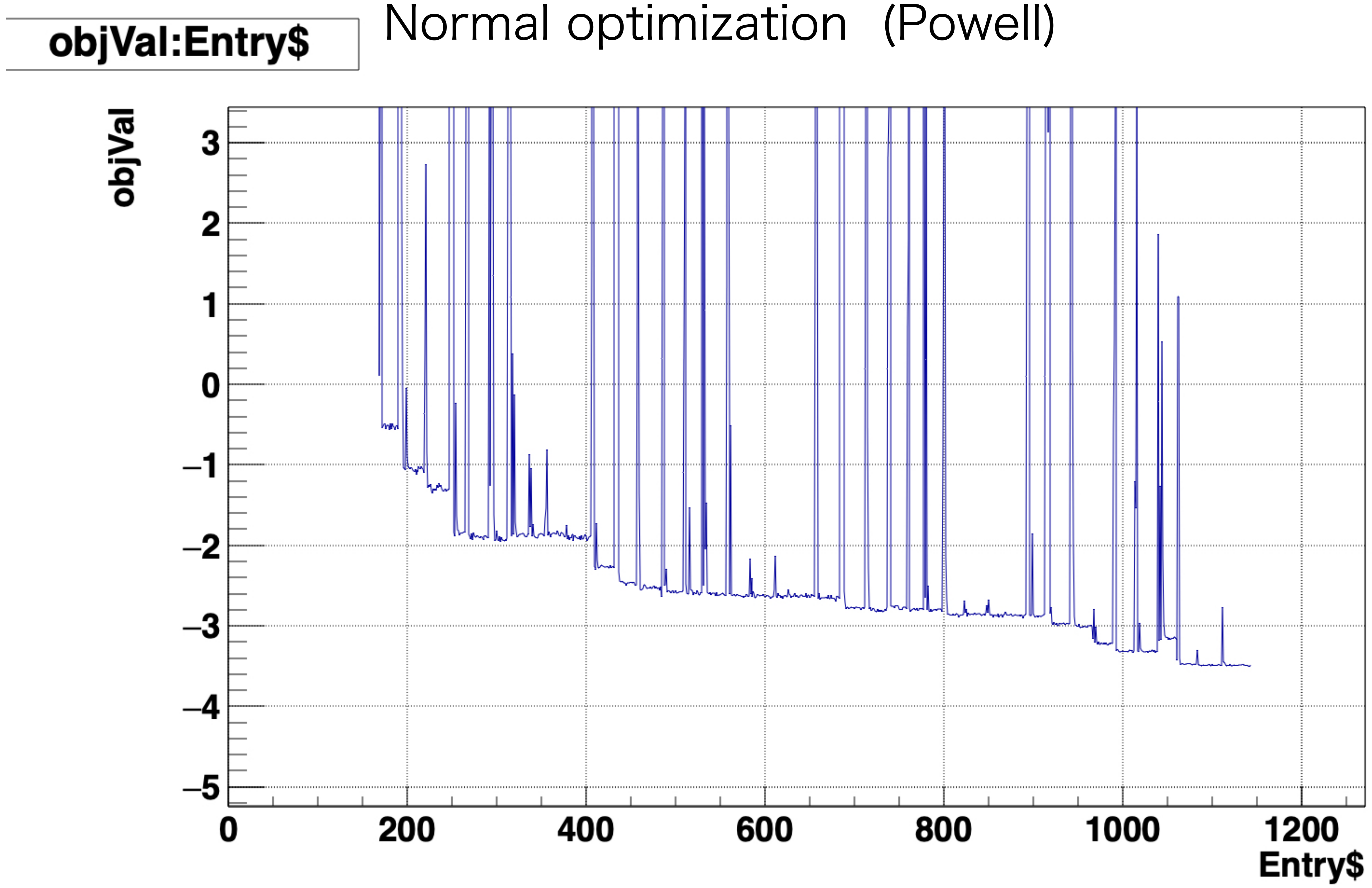


Simulation
includes : higher order matrix effect
not includes: space charge effect / distribution beyond Gaussian

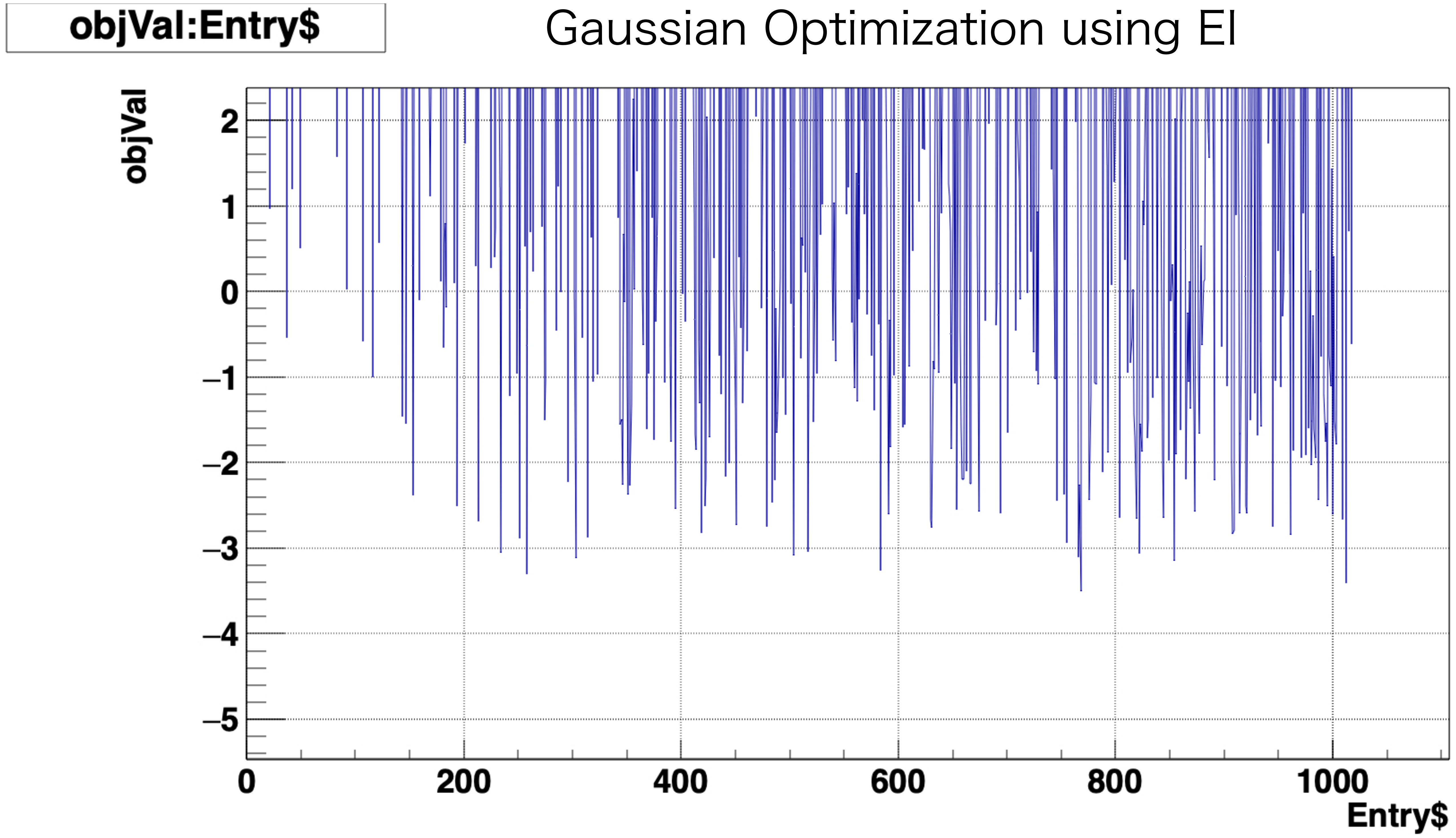
Goal: Development of effective (small epoch) optimization method
check the robustness/safety of the method for beam conditions

Recent activities…: Development with Simulations

Normal optimization (Powell)



Gaussian Optimization using EI



Goal: Development of effective (small epoch) optimization method

check the robustness/safety of the method for beam conditions

Recent activities…: non-destructive monitor

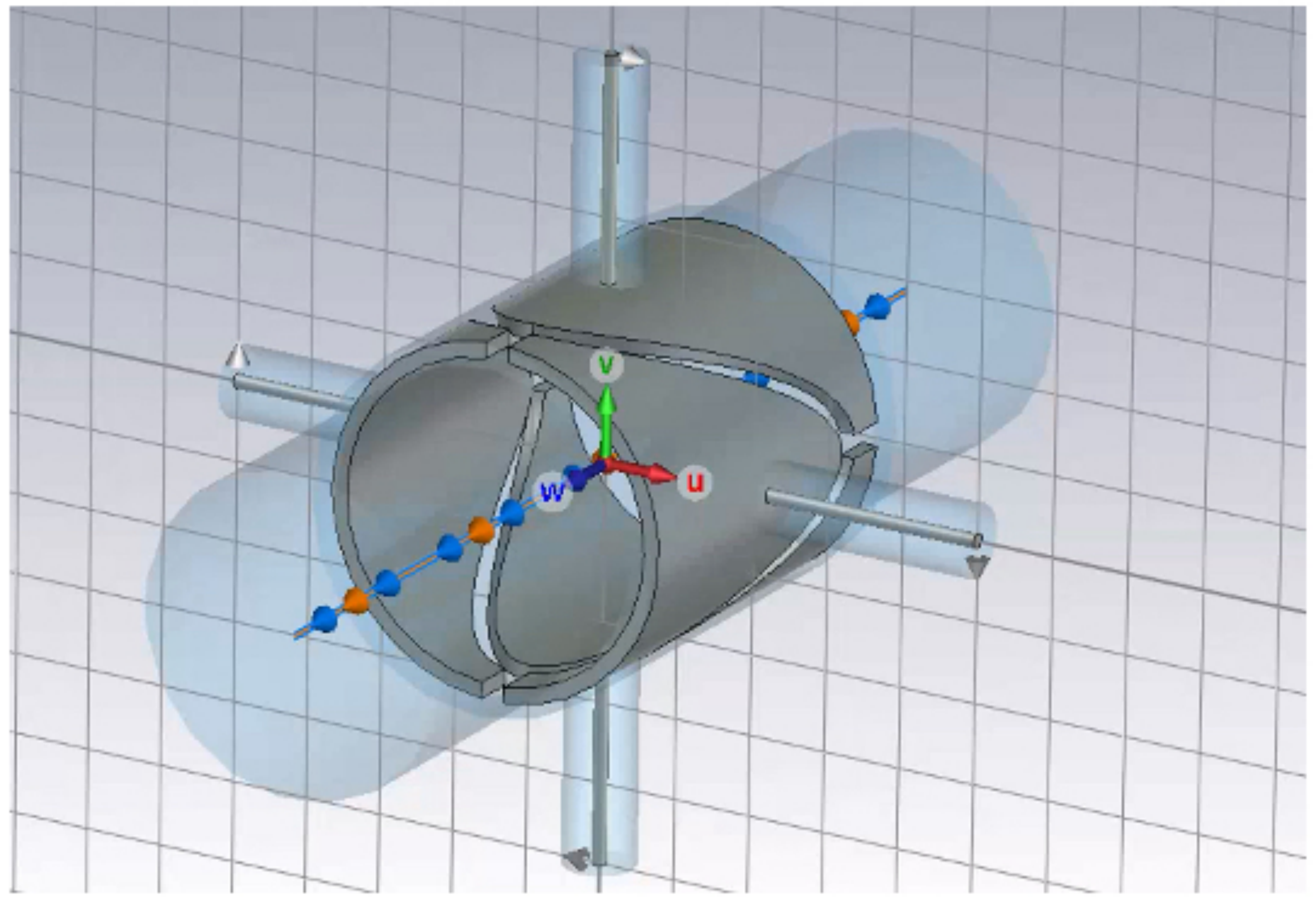


Figure 2: CAD model of type-A BEPM. The beam comes from the upper right corner towards the foreground. Up and down (right and left) electrodes are represented as upstream (downstream).

$$Q \equiv \sigma_x^2 - \sigma_y^2 = k_q \frac{V_L + V_R - V_U - V_D}{V_L + V_R + V_U + V_D} - <x>^2 + <y>^2, \quad (4)$$

$$\begin{pmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_8 \end{pmatrix} = (\mathbf{H}, \mathbf{V}) \begin{pmatrix} \sigma_{xx}(0) \\ \sigma_{xx'}(0) \\ \sigma_{x'x'}(0) \\ \sigma_{yy}(0) \\ \sigma_{yy'}(0) \\ \sigma_{y'y'}(0) \end{pmatrix} \quad (1)$$

where

$$\mathbf{H} \equiv \begin{pmatrix} (M_{11}^{01})^2, 2M_{11}^{01}M_{12}^{01}, (M_{12}^{01})^2 \\ \vdots \\ (M_{11}^{08})^2, 2M_{11}^{08}M_{12}^{08}, (M_{12}^{08})^2 \end{pmatrix}, \quad (2)$$

$$\mathbf{V} \equiv \begin{pmatrix} -(M_{33}^{01})^2, -2M_{33}^{01}M_{34}^{01}, -(M_{34}^{01})^2 \\ \vdots \\ -(M_{33}^{08})^2, -2M_{33}^{08}M_{34}^{08}, -(M_{34}^{08})^2 \end{pmatrix}. \quad (3)$$

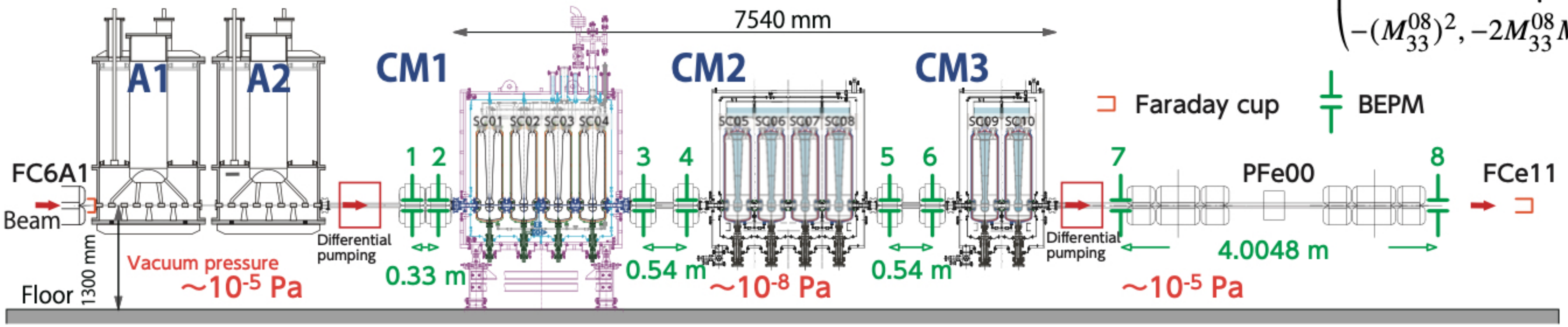


Figure 1: Schematic of beamline including SRILAC. Green numbers denote Beam Energy Position monitors, and PFe00 denotes a wire scanner.

T. Nishi Proc. of SRF2023 /
to be presented in HB2023

Recent activities…: non-destructive monitor

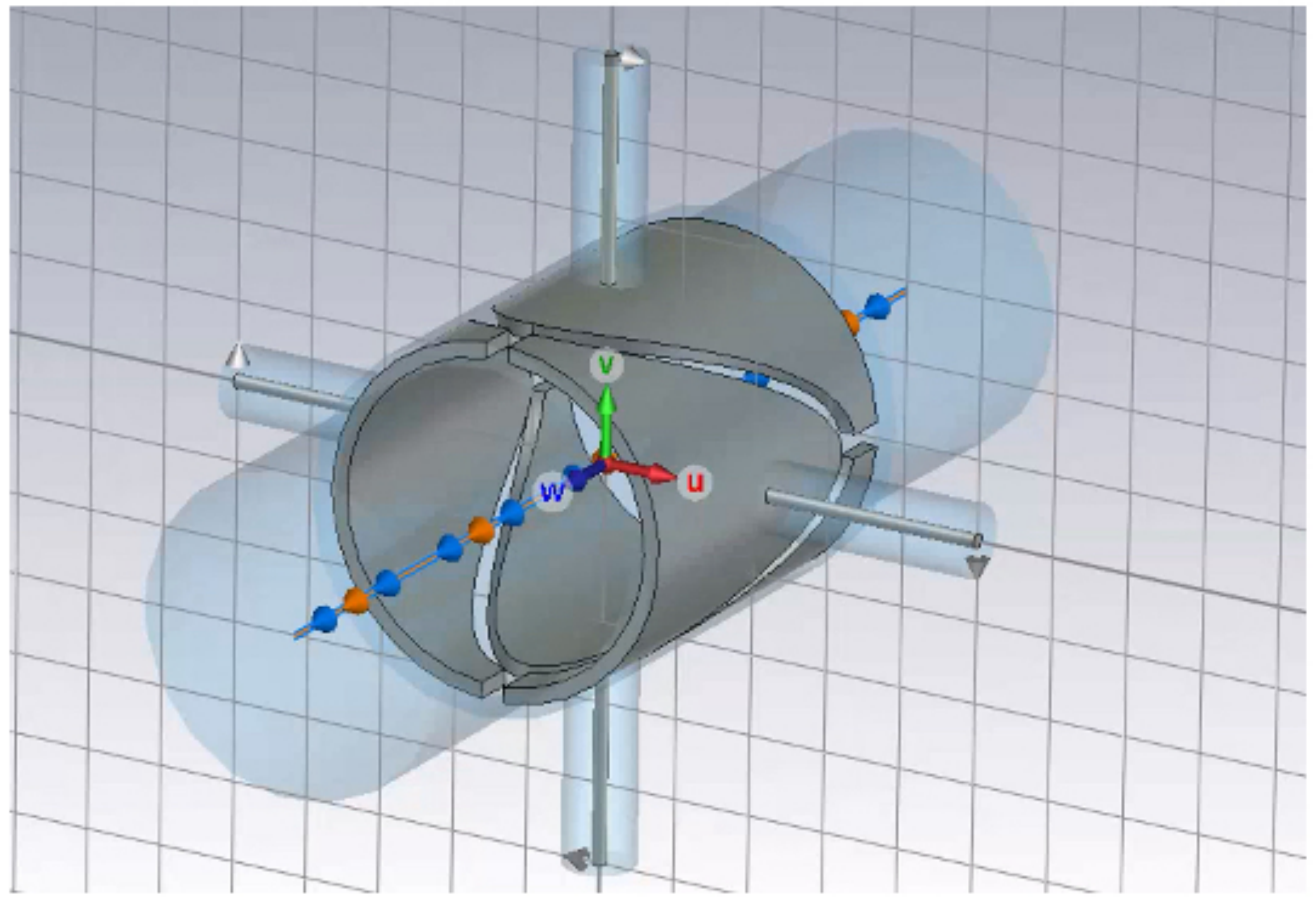
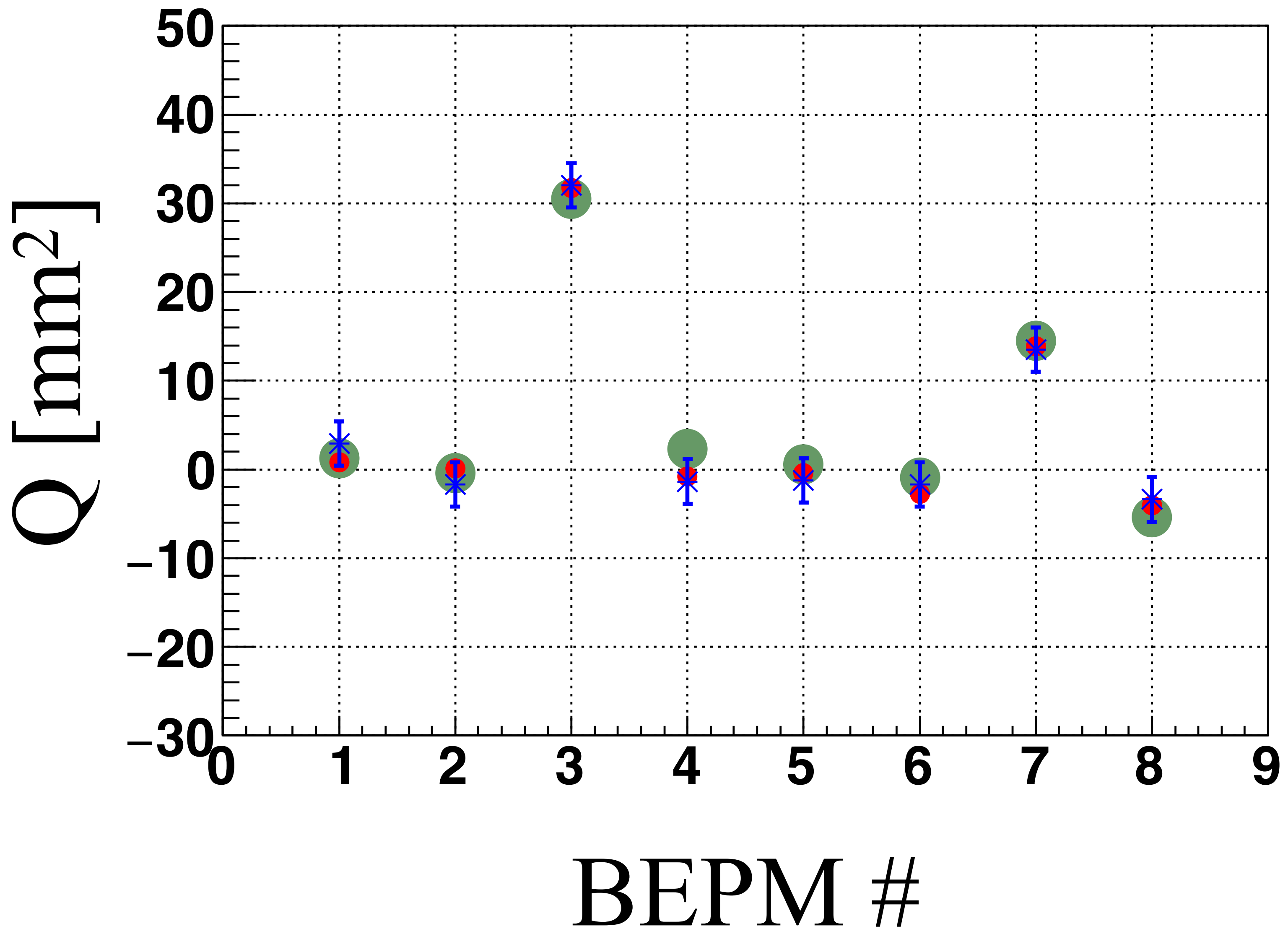
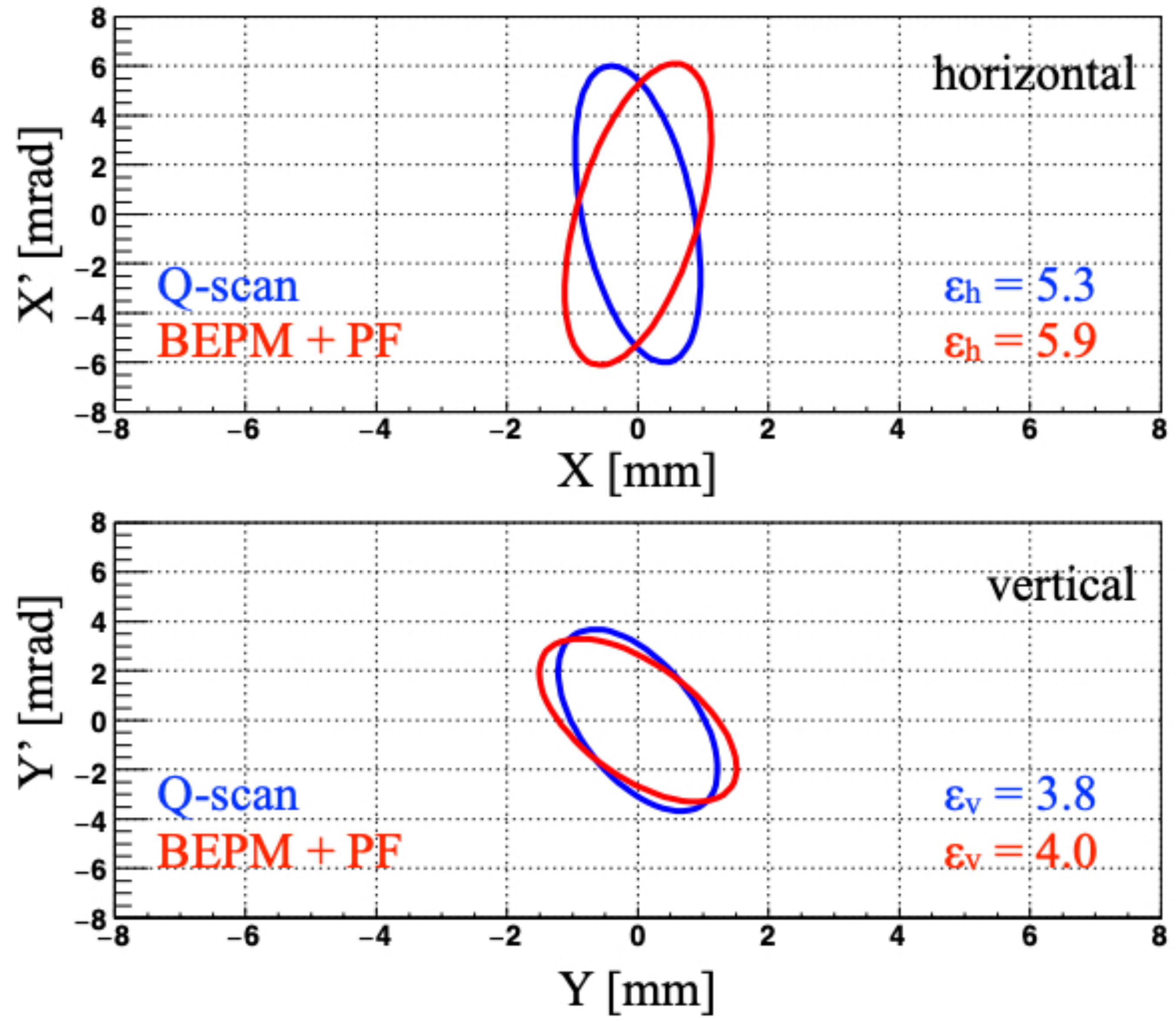


Figure 2: CAD model of type-A BEPM. The beam comes from the upper right corner towards the foreground. Up and down (right and left) electrodes are represented as upstream (downstream).





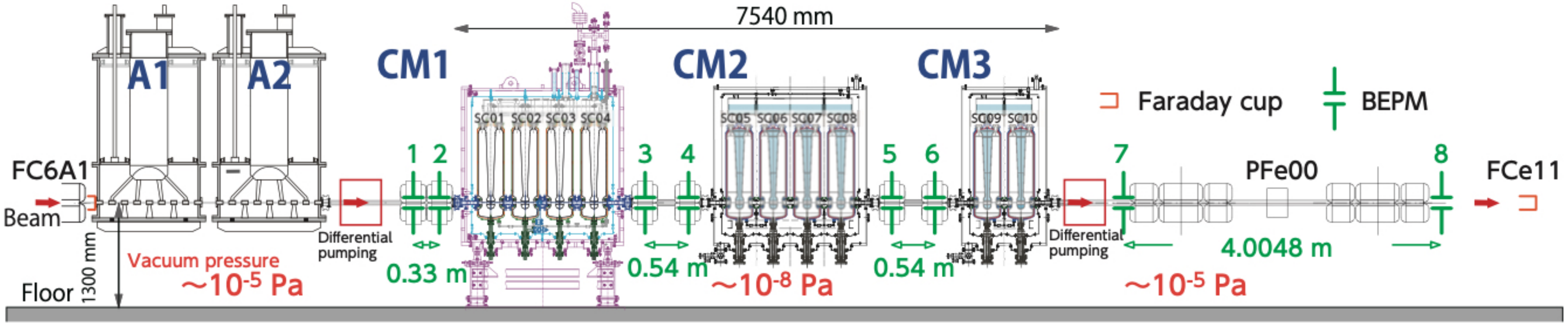
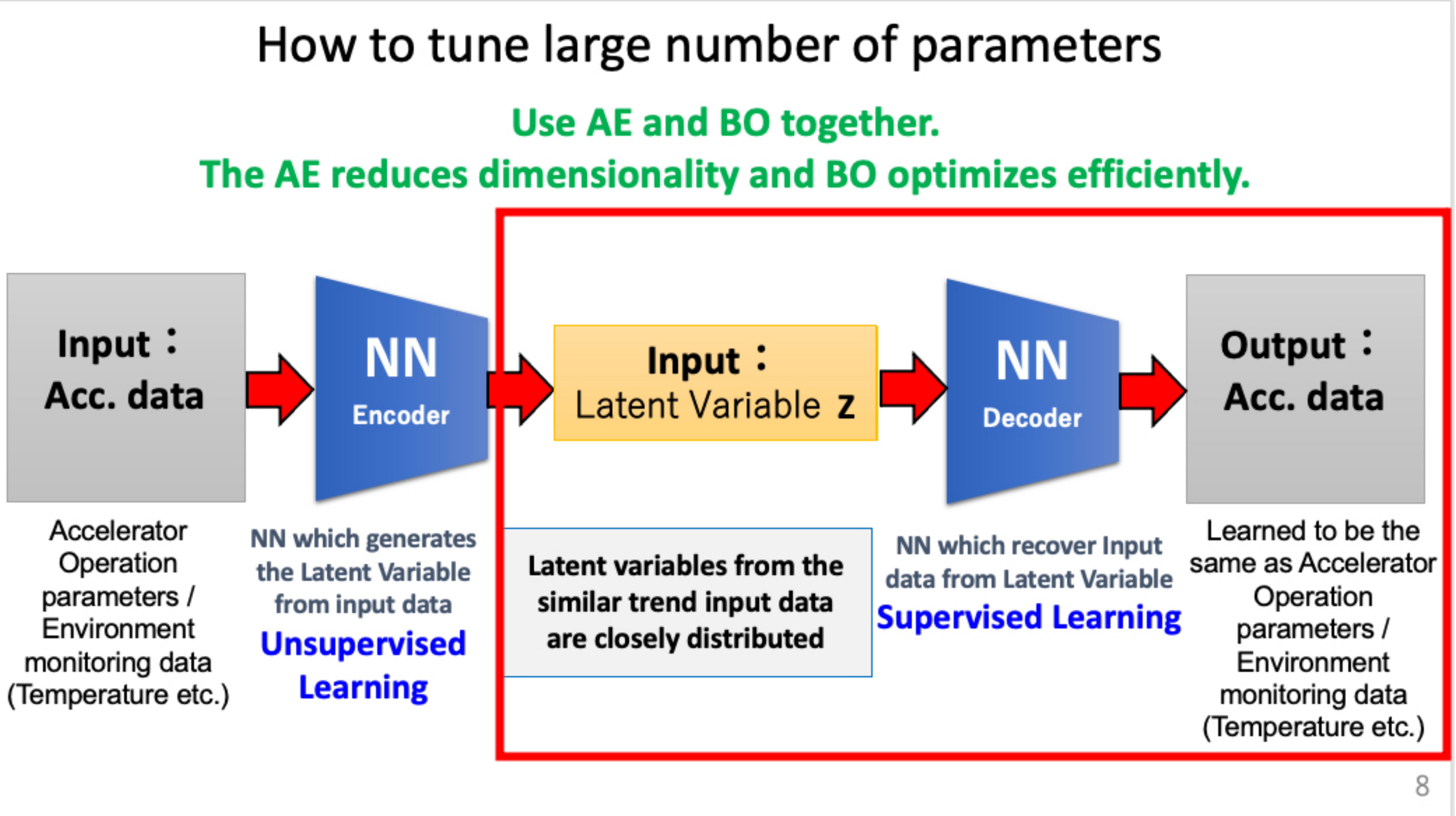


Figure 1: Schematic of beamline including SRILAC. Green numbers denote Beam Energy Position monitors, and PFe00 denotes a wire scanner.

T. Nishi Proc. of SRF2023 / to be presented in HB2023

Recent activities…: new method to reduce parameters



Slide of Dr. Morita