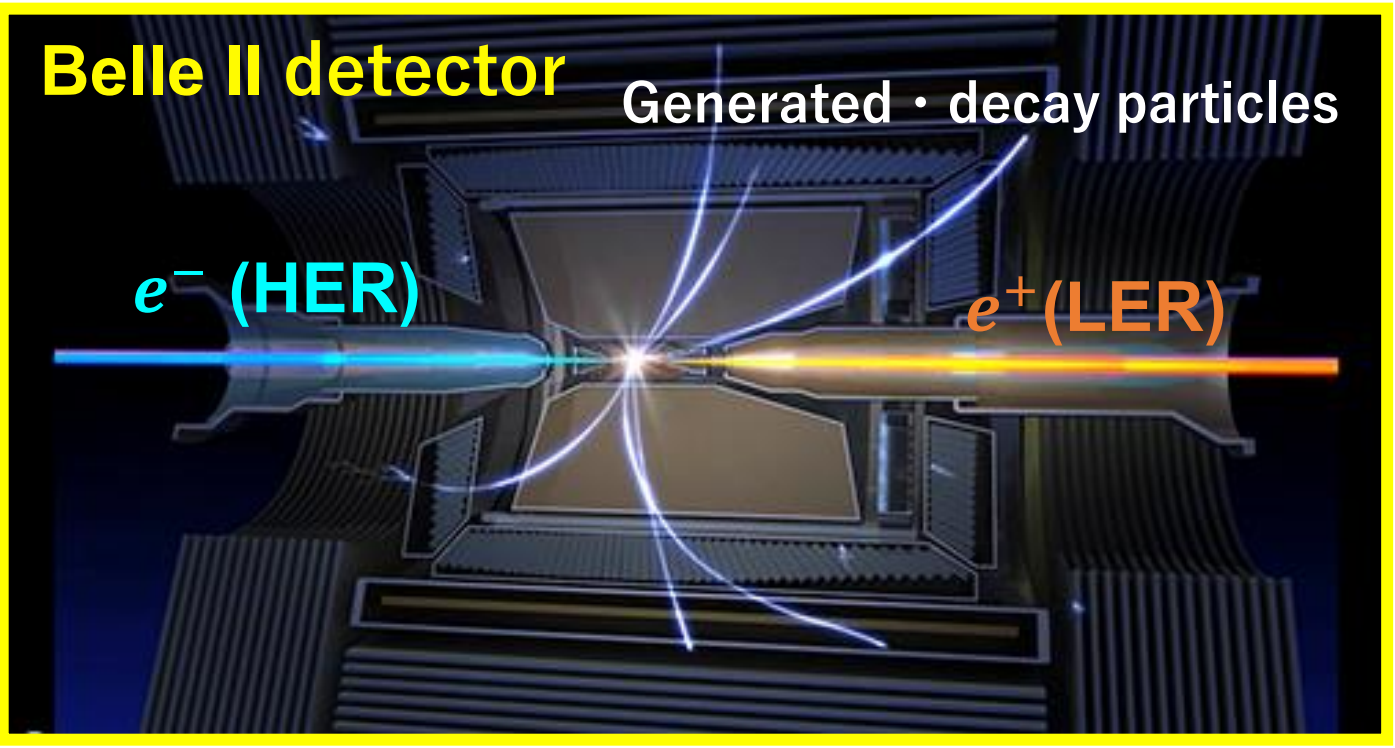
Machine Learning-Assisted Beam Operation at SuperKEKB and Linac at KEK

Shinnosuke Kato (The University of Tokyo, KEK)

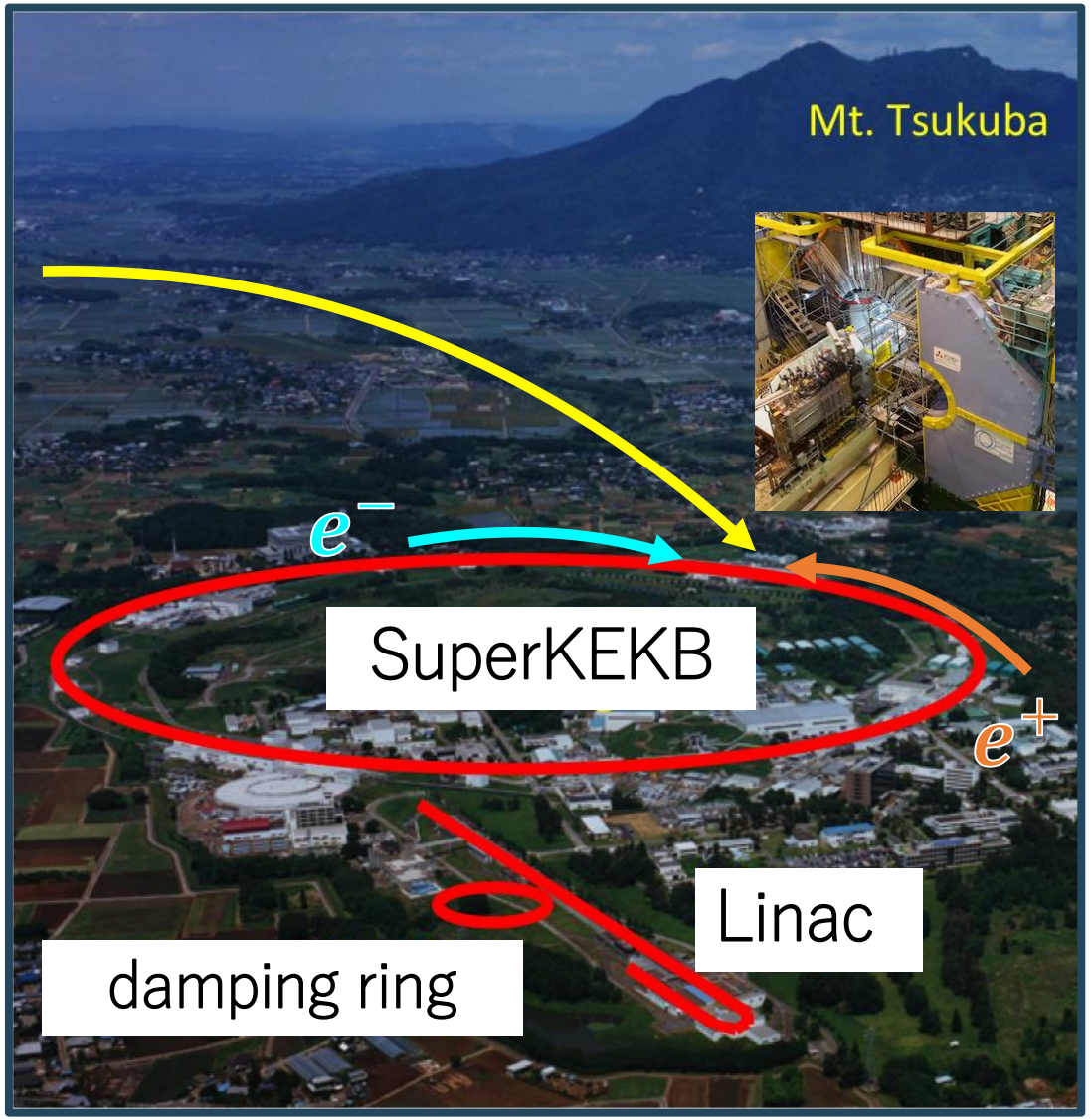
Gaku Mitsuka , Naoko Iida,
Takuya Natsui , Masanori Satoh (KEK Acc.)

Luminosity is important for a new physics



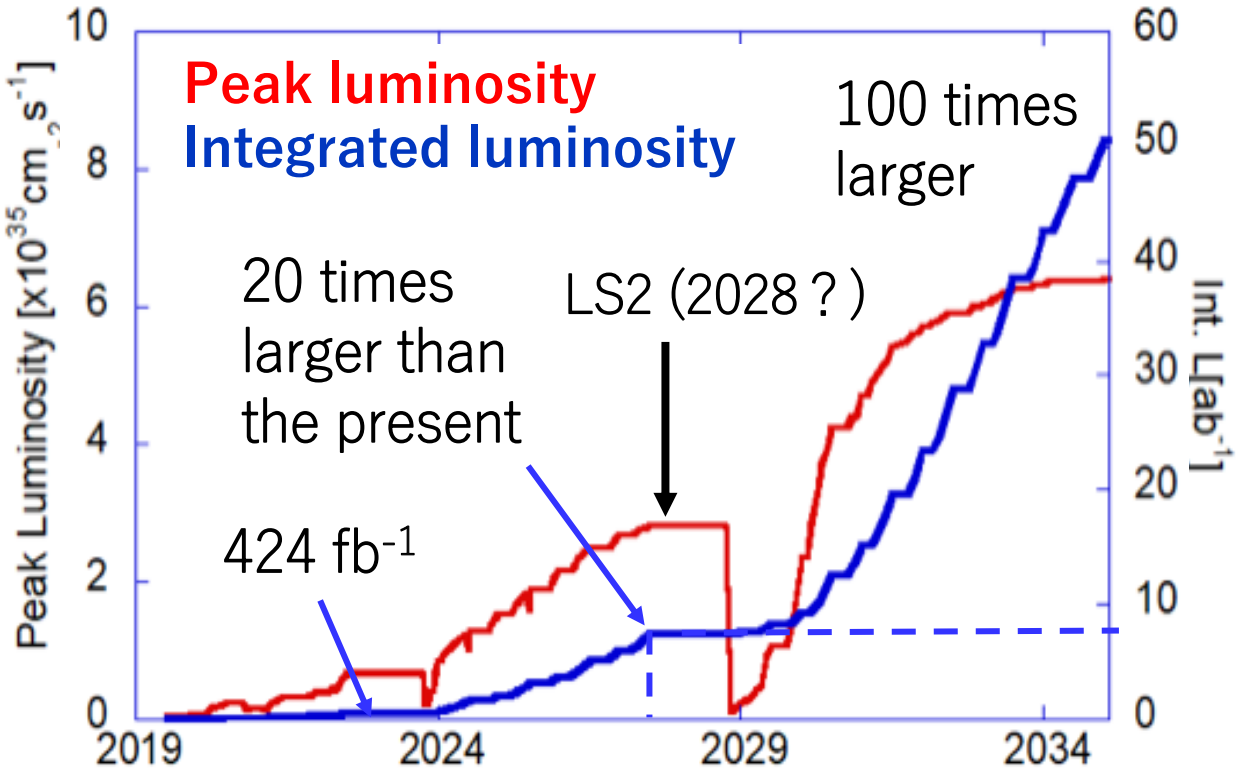
$$N = \sigma\,[\mathrm{cm}^2] \int L\,[\mathrm{cm}^{-2}\mathrm{s}^{-1}]dt\,[s]$$

Accumulating statistics N to search for a new physics
→Requires high peak luminosity and long-term stability

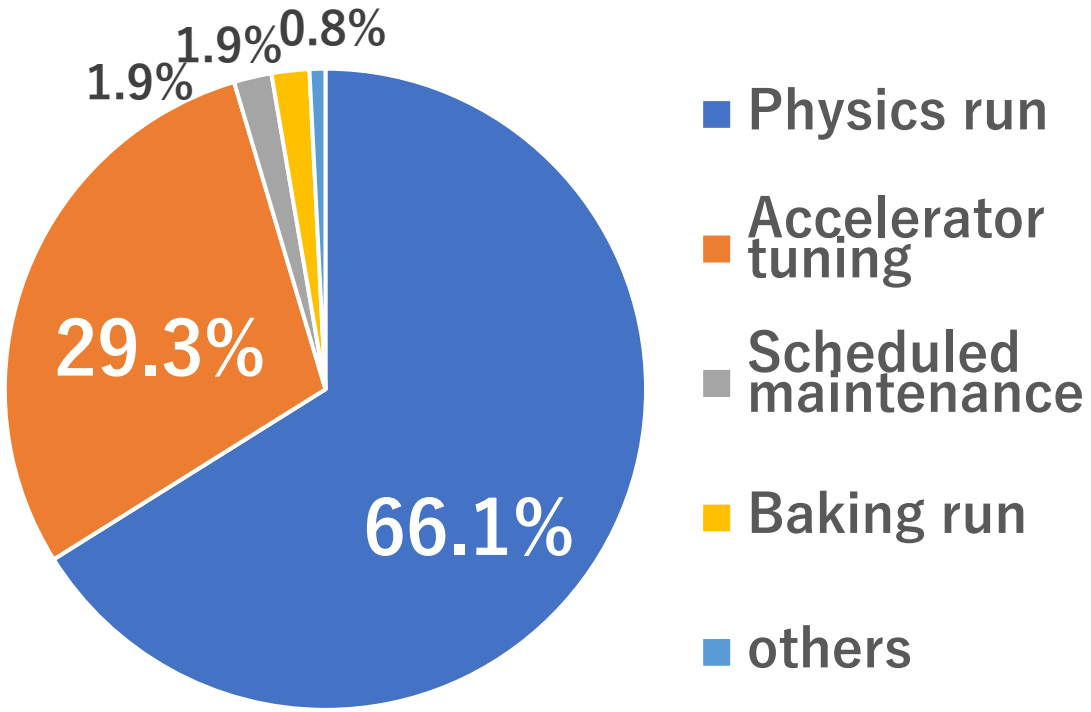


Motivation to introduce machine learning

Target profile of integrated luminosity



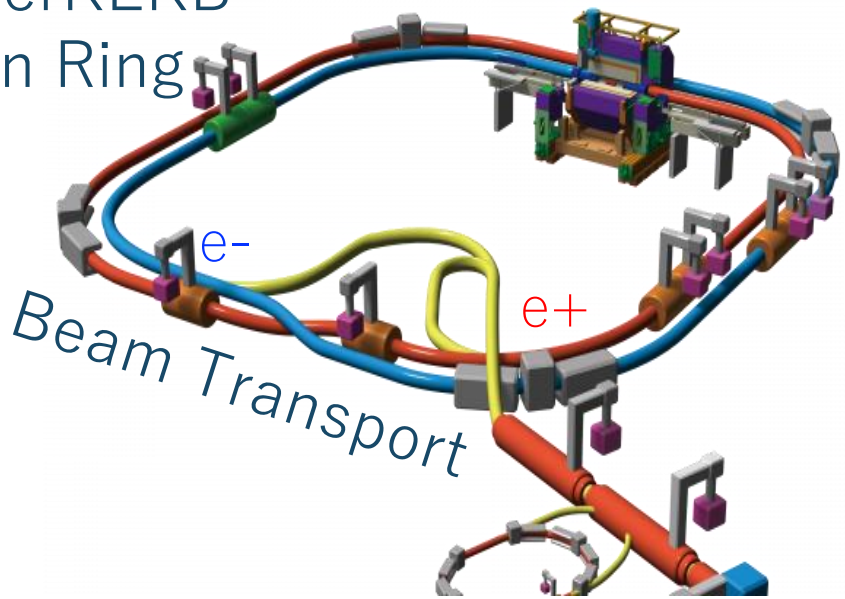
Belle II /SuperKEKB Run time breakdown (Feb. 21 - Jun. 22, 2022)



We want to make accelerator tuning efficient using machine learning

Linac study for SuperKEKB injection tuning

SuperKEKB Main Ring



2023								2024
4	5	6	7	8	9	10	11	12	1
Long shutdown (We don't test ML-assisted injection tuning)								SuperKEKB LER operation restart	SuperKEKB HER operation restart

Linac study → **Linac study** → **SuperKEKB Injection tuning**

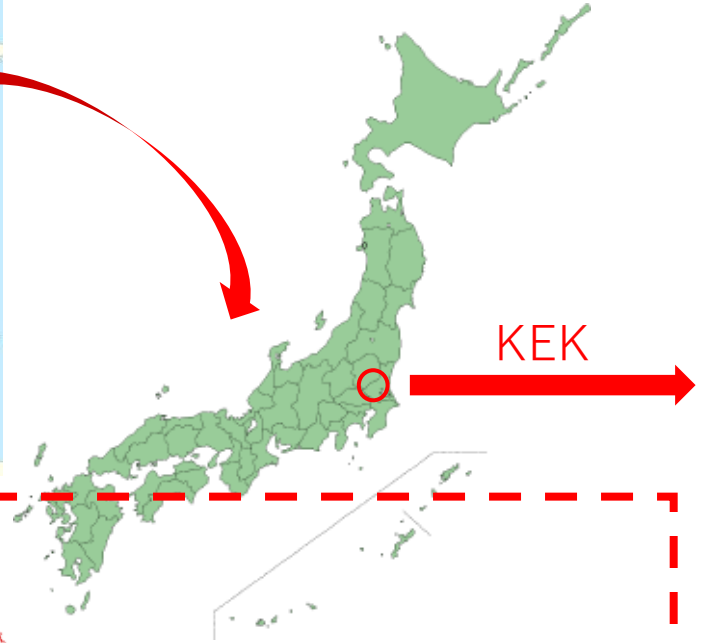
For machine learning……

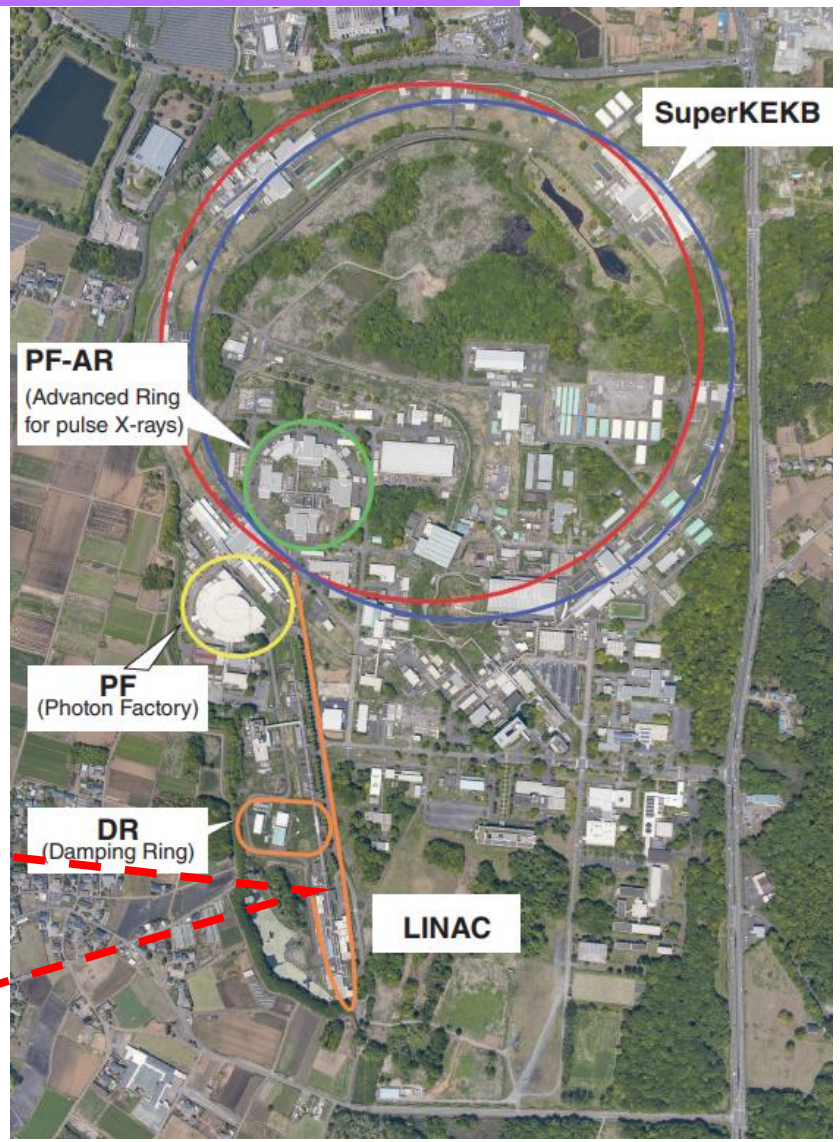
- **How efficiently and quickly can we optimize the beam?**
- **What characteristics does the optimization have?**
- **Which parameters are important to optimize ?**

Outline

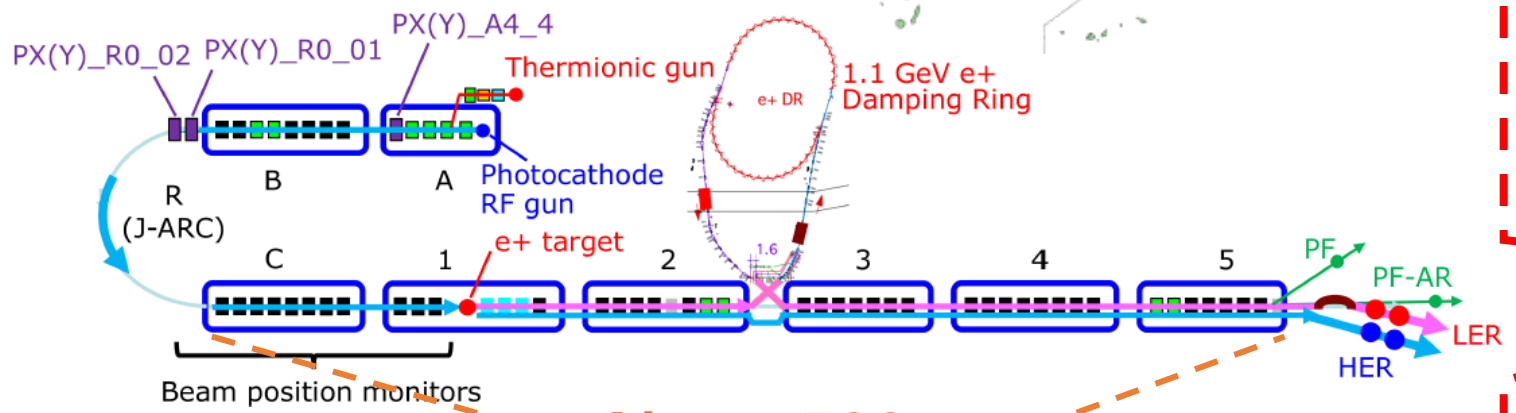
- Introduction
  - Experiment setup of this study
  - How to optimize the machine parameters at the KEK LINAC
  - Two types of optimization algorithms
- Results of optimization (beam charge maximizing)
  - e^- beam
  - e^+ beam
- Summary and prospects

Introduction to the KEK LINAC





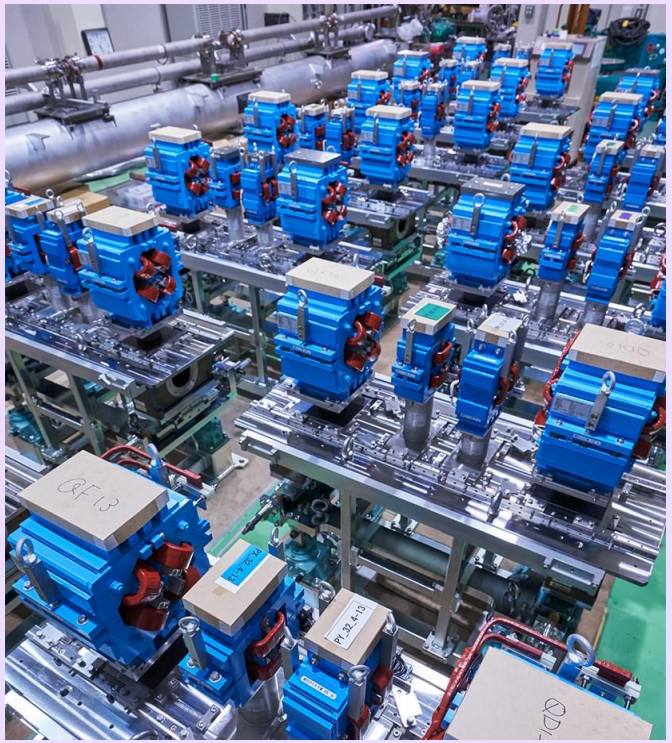
Outline of the KEK LINAC

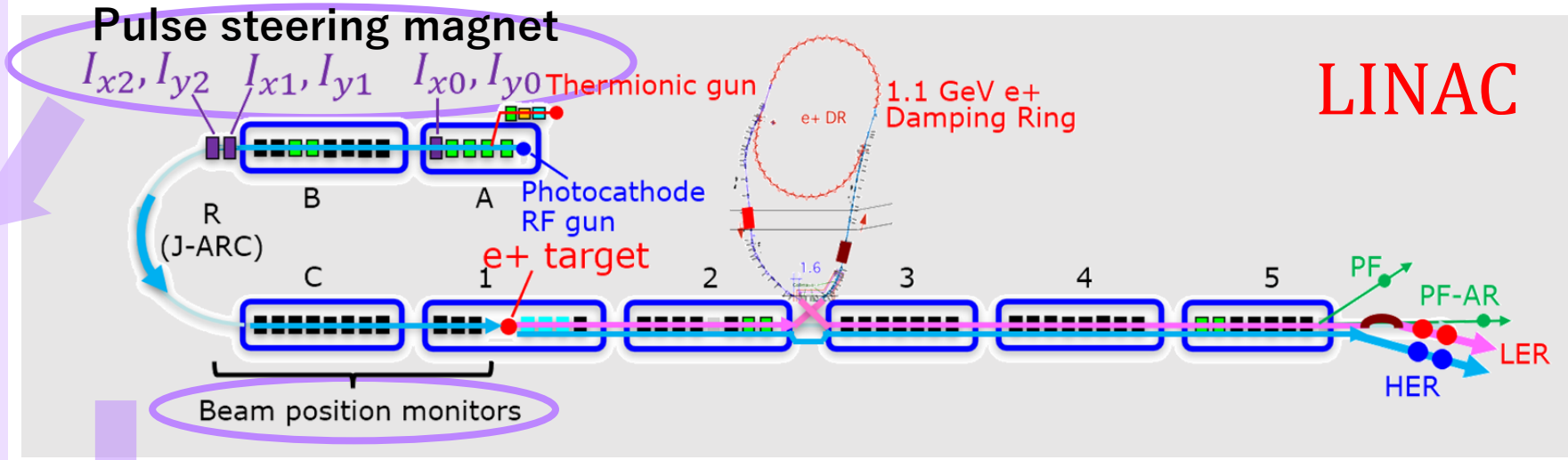


About 500m

Components used in this study

Pulse magnets ready for installation (taken in 2017)





Ordinary BPM used in the KEK LINAC

BPM signal processing system

1% uncertainty

Experiment setup of this study

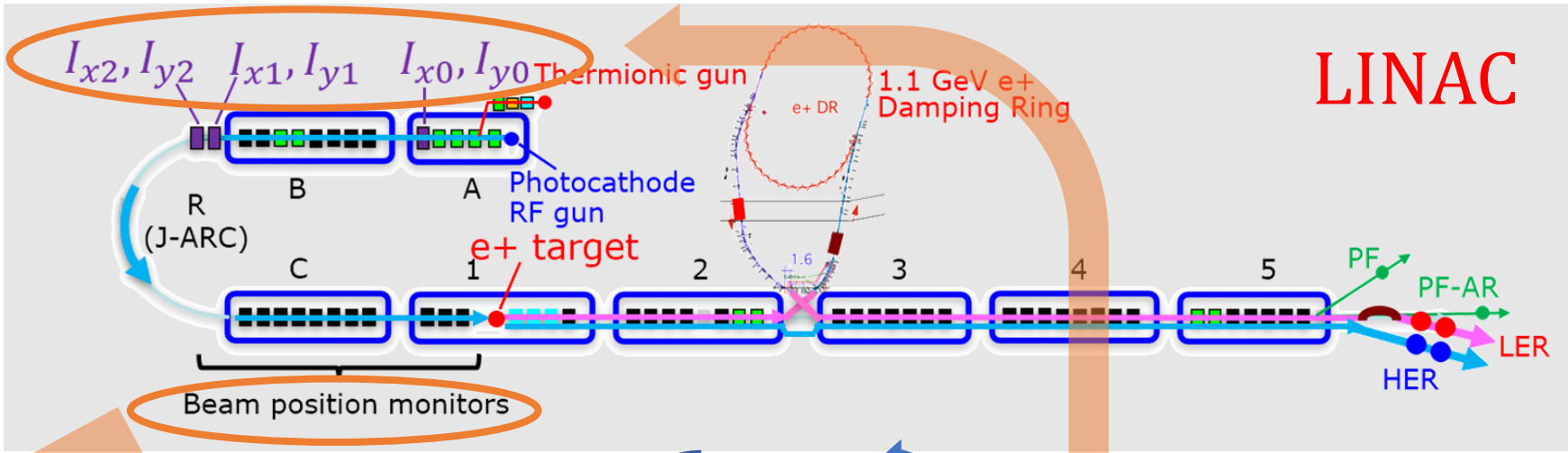
Tuning parameters

Applied currents to 6 steering magnets:

$I_{x0}, I_{y0}, I_{x1}, I_{y1}, I_{x2}, I_{y2}$ (A)

Evaluation parameter

Beam charge of 14 BPMs' average: Q (nC)



Get beam charge

Loop again !

Put current of magnets

Machine learning

Accumulate data

Trial1 $I_{x0} \sim I_{y2} \leftrightarrow Q$

Trial2 $I'_{x0} \sim I'_{y2} \leftrightarrow Q'$

⋮ ⋮ ⋮

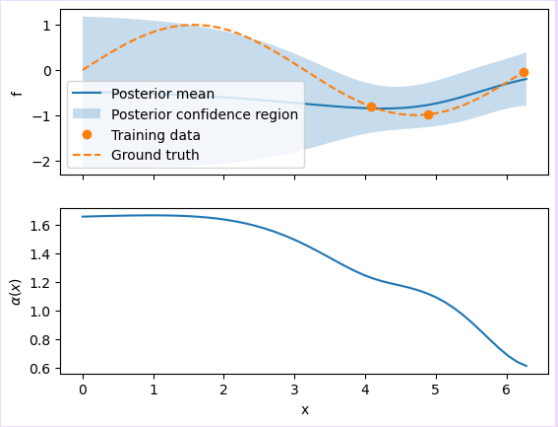
Optimization algorism (next page in detail)

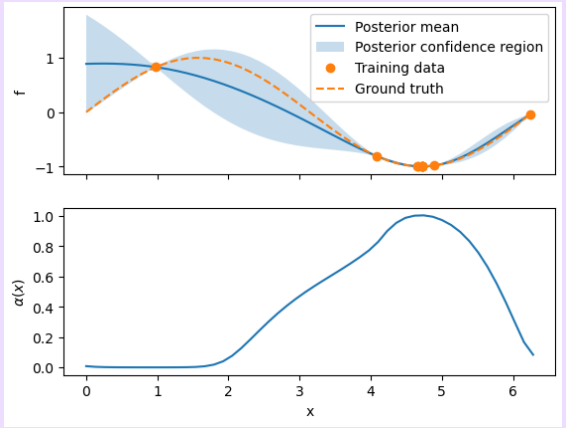
Estimate best current at that time to maximize charge

Two types of algorithm to optimize

(a)Bayesian optimization (BoTorch)

Optimize black-box functions using Gaussian process. Below figure is looking for the minimum.



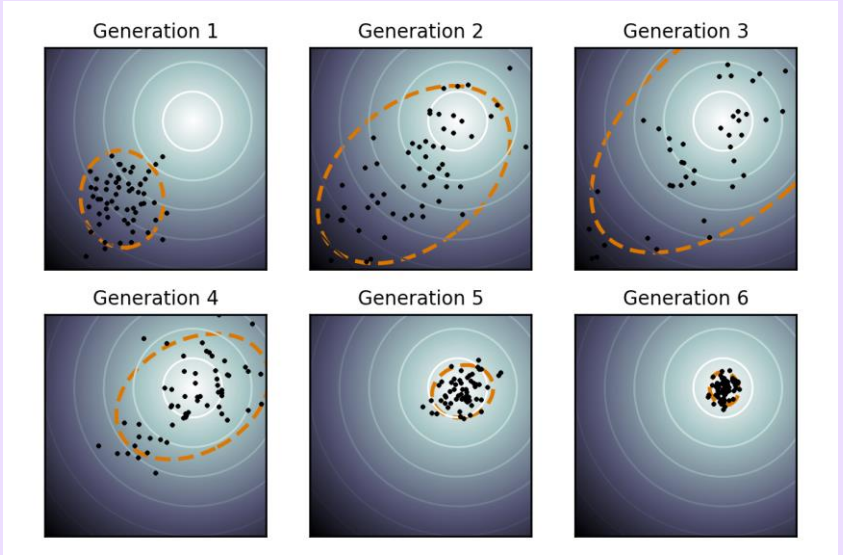


1st trial **10th trial**

Conventional

(b)CMA-ES

One of evolutionary computation algorithms.



https://en.wikipedia.org/wiki/CMA-ES#/media/File:Concept_of_directional_optimization_in_CMA-ES_algorithm.png

Recently proposed

We tested the both algorithms in this study.

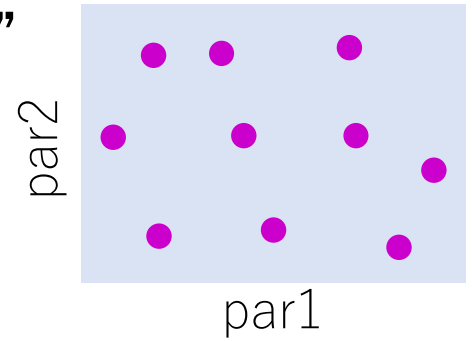
Detail of optimization for e^- beam

	Date	Expert tuned value
(a)Bayesian	June 2nd 11 am to 3 pm	9.3 nC
(b)CMA-ES	June 12th 11 am to 2 pm	8.9 nC

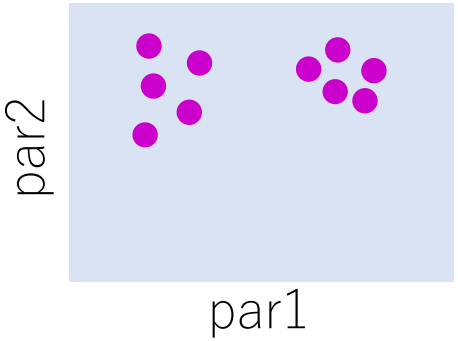
“Random” = First 10 trials are randomly initialized.

“Previous best value”= First 10 trials initialized using the top 10 results giving the best beam charge taken from the last run

“Random”



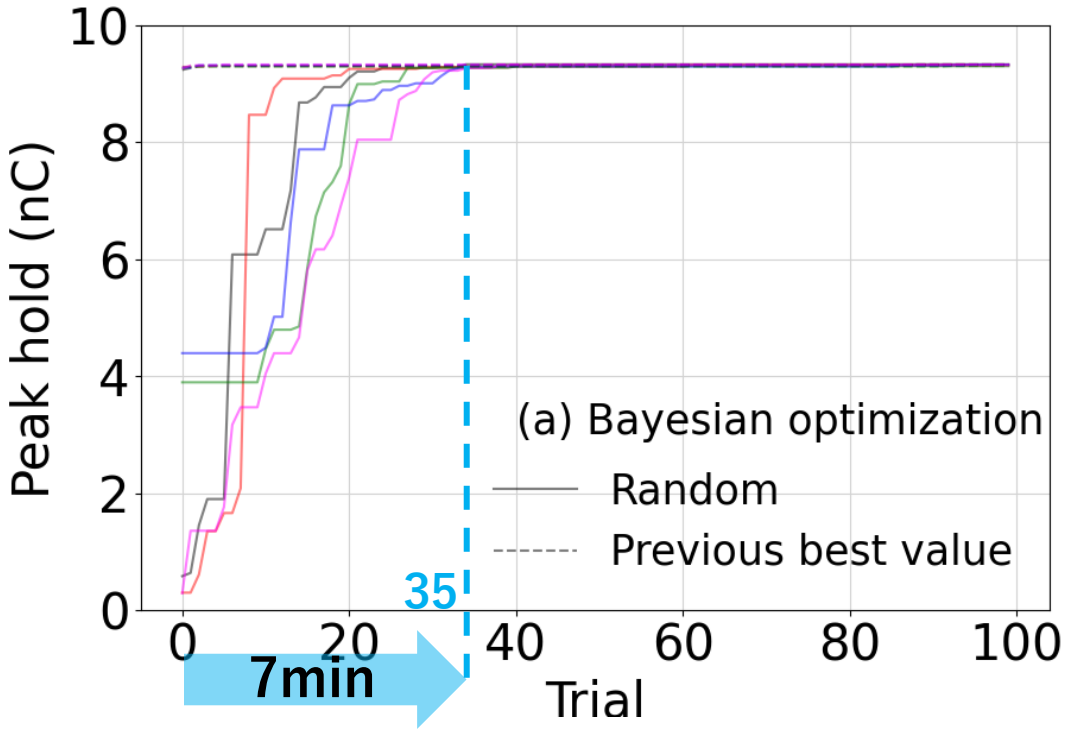
“Previous best value”

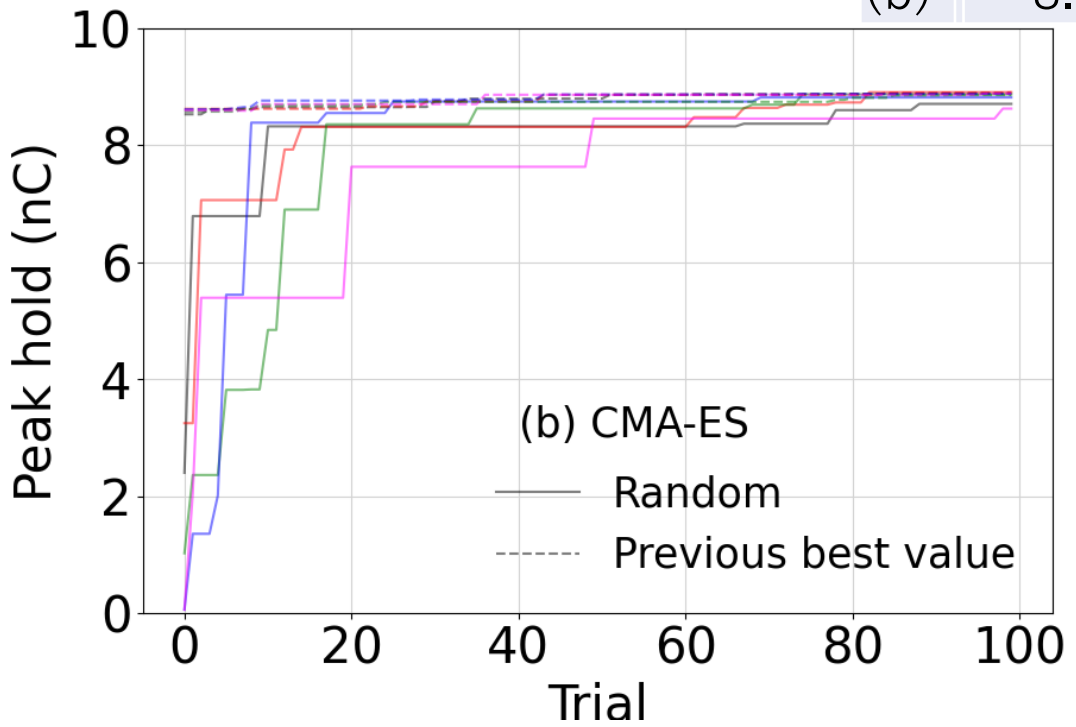


To what level and how quickly can algorithms optimize ?

- Peak hold charge is the maximum value of the beam charge.
- Iterating 100 trials and repeated the runs 5 times (identified by color)

	Expert tuned value
(a)	9.3 nC
(b)	8.9 nC





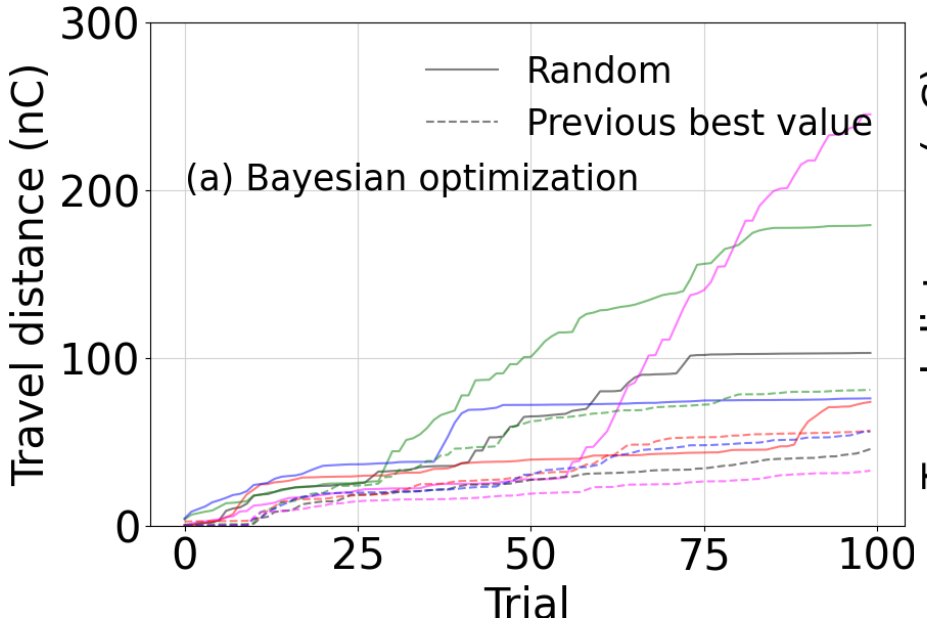
Both algorithms were able to optimize even “Random” !
In Bayes opt., all runs are maximized in about 35 trials (7 minutes) !

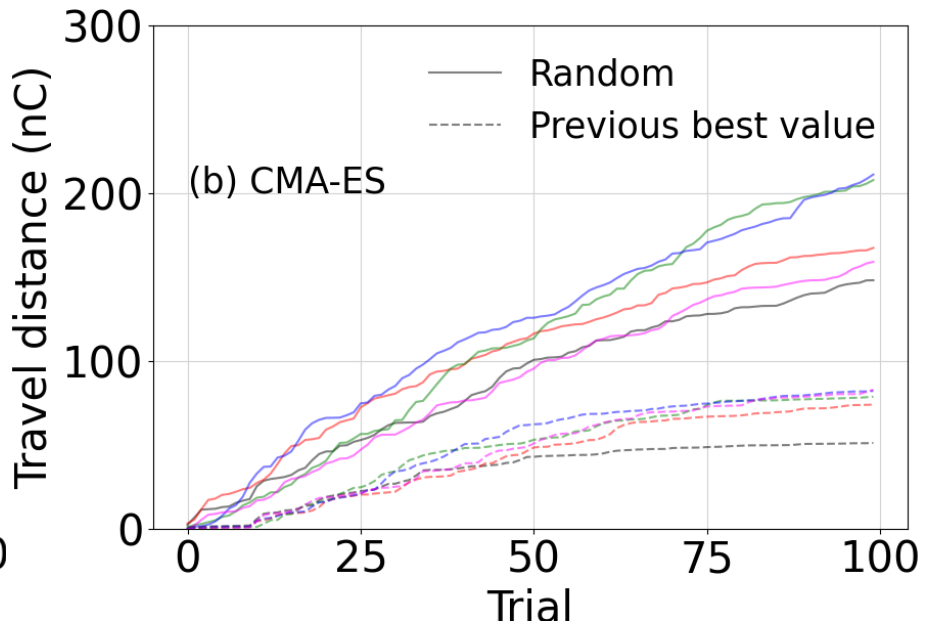
What characteristics do they have?

Travel distance is formulated as $Q_{td} = \sum_{1}^{t} |q(t) - q(t-1)|$

Q_{td} represents how exploitative or exploratory per trials.

steep slope → exploratory shallow slope → exploitative





Too exploratory

↳ Search for unnecessary regions

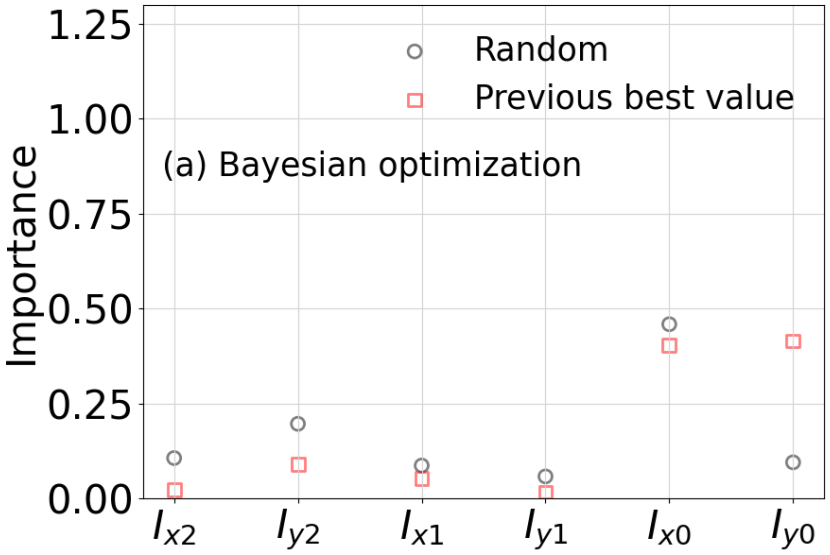
Too exploitative

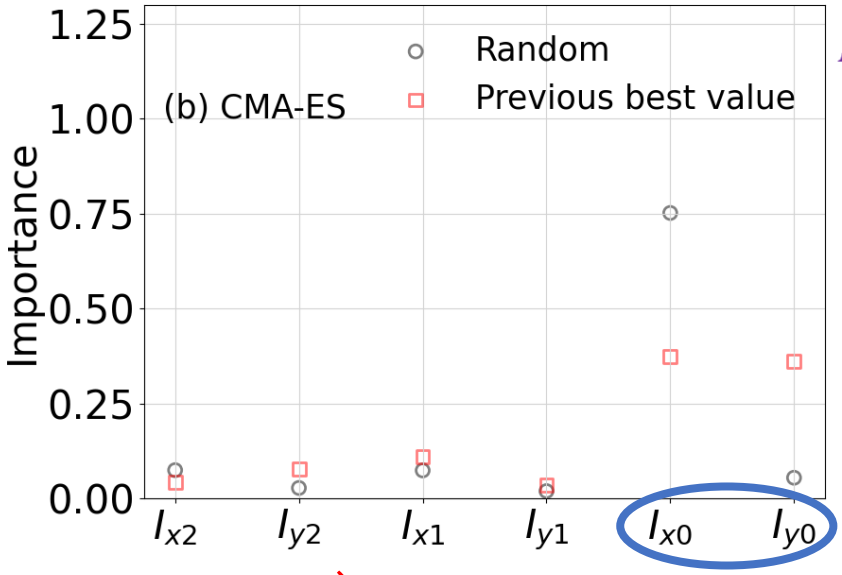
↳ Stuck in local maximum

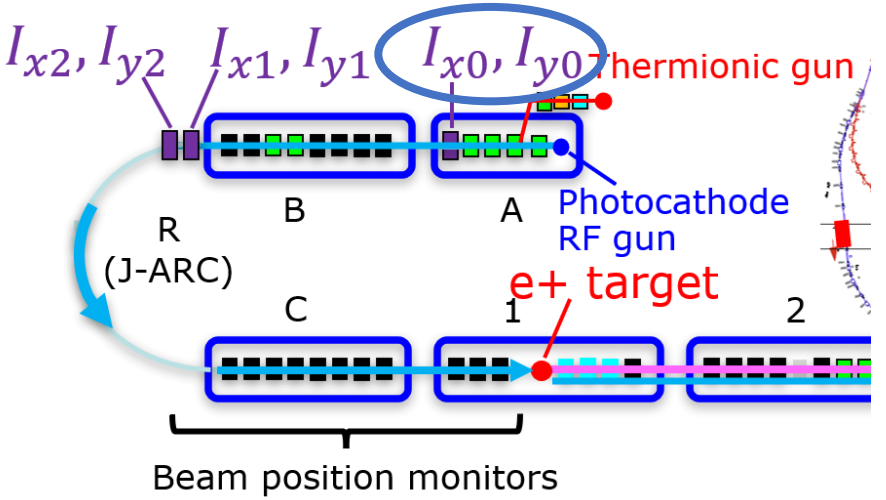
In “Random”, characteristics of Bayes opt. are influenced by the initial values.
In “Previous best value”, no big difference between Bayes opt. and CMA-ES.

Which parameters are important to optimize ?

- "Importance" describes which parameters give the significant change in charge.
- It can be quantified using the fANOVA method. [Hutter, ICML 2014]
- Importance sum is normalized to 1. (Σ (Importance) = 1)







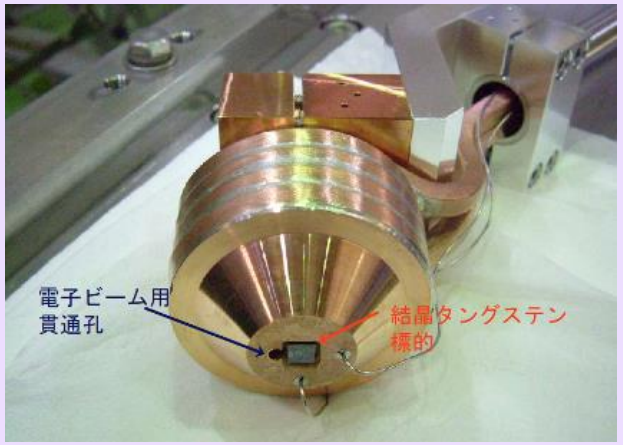
I_{x0}, I_{y0} (most upstream magnets) have higher importance in the both.
Agree with experts' rule of thumb *"Sequentially tuned from upstream"*

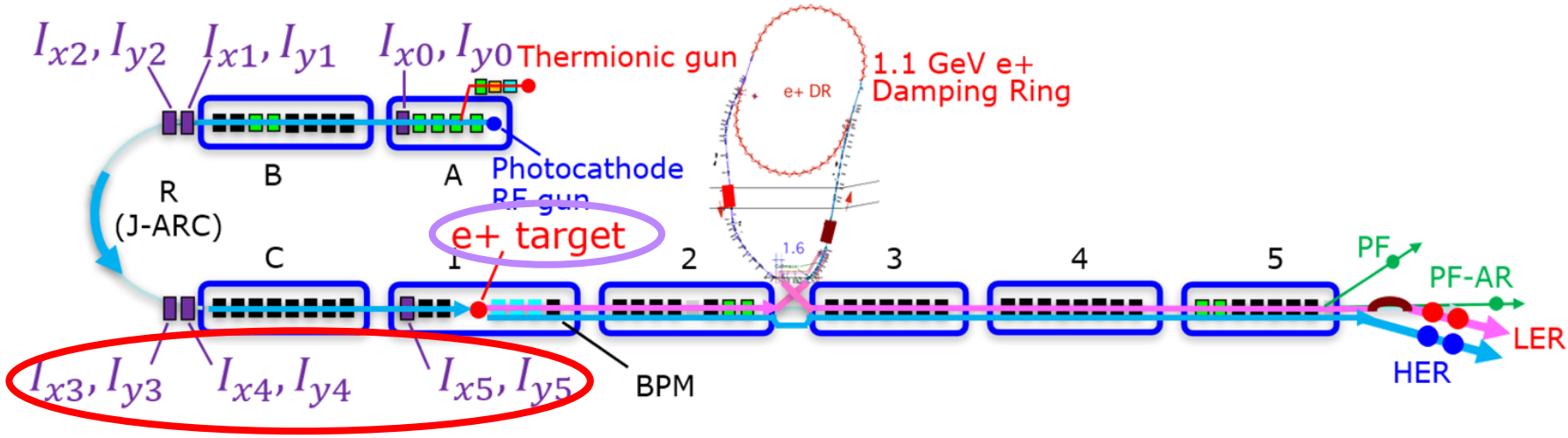
Horizontal collimator in the R-sector → High importance of I_{x0}

Detail of optimization for e^+ beam

- Maximize the **positron beam** charge by adjusting the steering magnets close to the tungsten target (only **"Random"**)
- Optimize total 6 parameters using the Bayesian optimization
- Iterate 100 trials and repeat the runs 5 times

Tungsten target

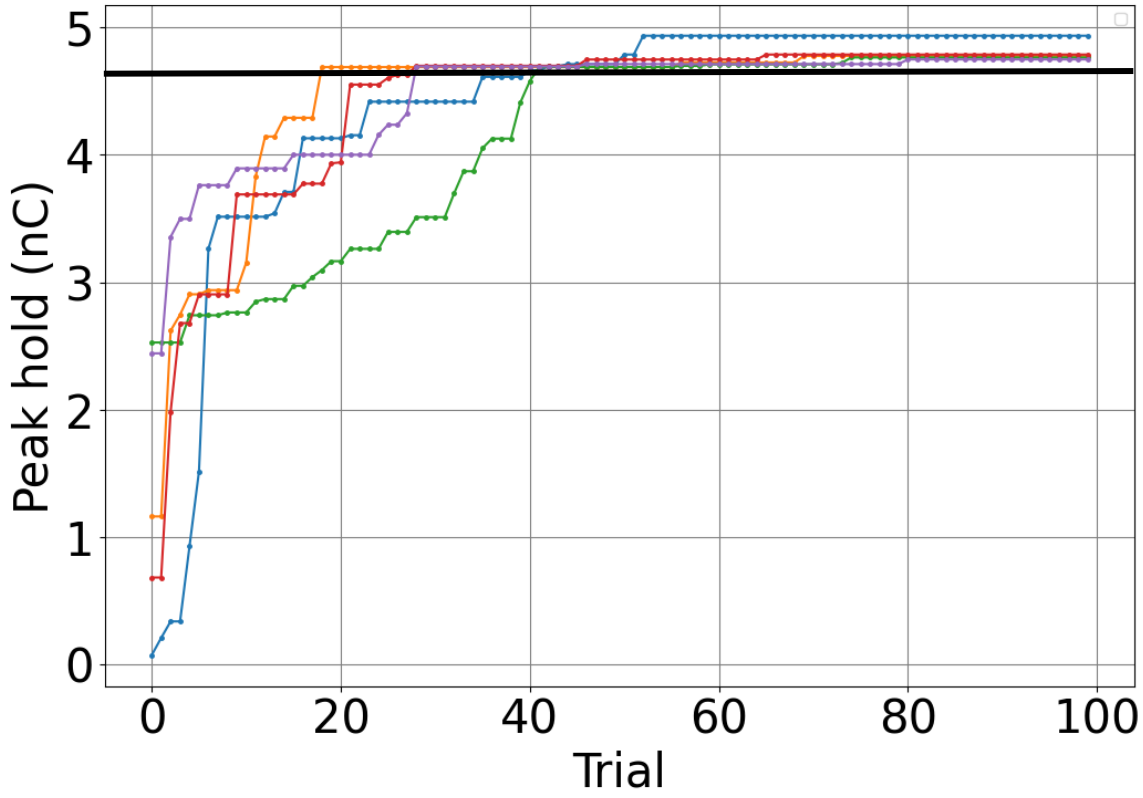




6 magnets (3 X + 3 Y) before e^+ target and 1 BPM after e^+ target

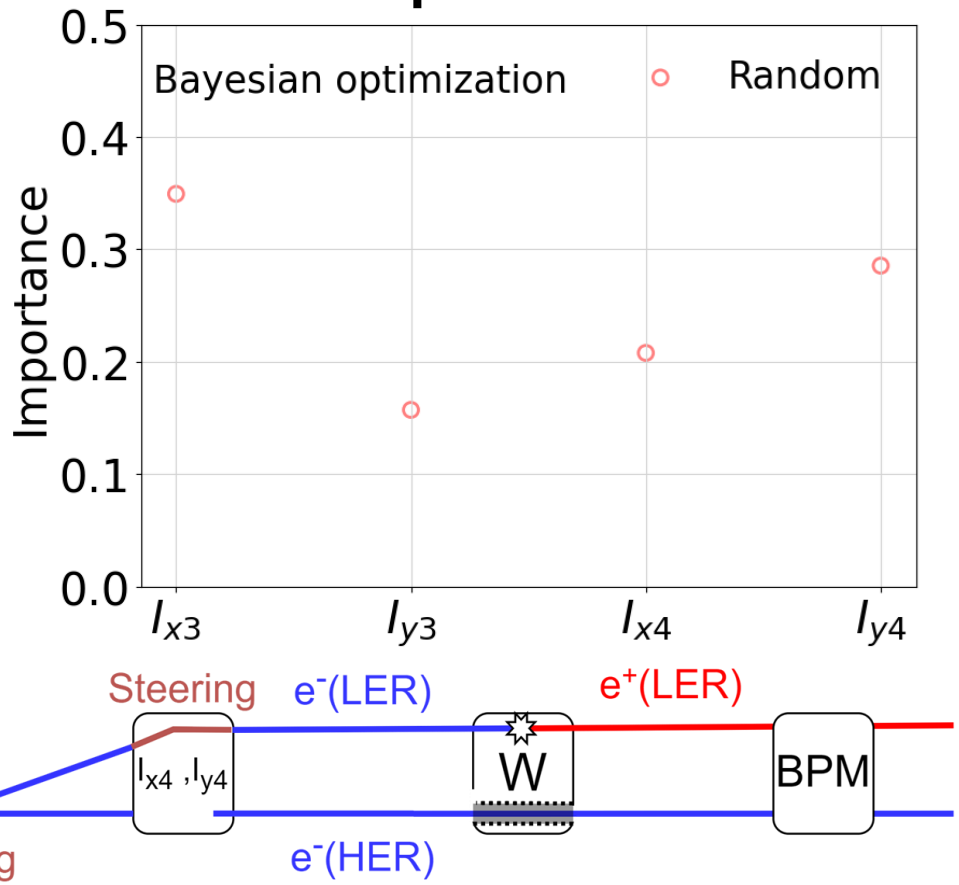
Result of optimization for e^+ beam

Peak hold



Expert tuned
Value (4.6 nC)

Importance



The e^- beam has been optimized also for e^+ production!

Agree with experts' rule of thumb *"Both I_3 and I_4 are equally important"*

Summary and prospects

- Using the Linac electron beam, we studied the possibility of using machine learning to adjust the SuperKEKB beam.
- Bayesian optimization achieved the maximum charge in 7 minutes, while experts took 30 minutes.
  We also clarified the characteristics and important parameters.
- Optimization worked in the two different tests, so it would work for SuperKEKB beam-injection tuning.
- In this December after restarting operation, we will apply ML-assisted methods to SuperKEKB.