

PAUL SCHERRER INSTITUT



Nicole Hiller :: Operation :: Paul Scherrer Institut

Machine Learning at PSI

10th Workshop on Longitudinal Electron Bunch Diagnostics

wikipedia.org: **Machine learning (ML)** is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence.



Samuel, Arthur L. (1959).

"Some Studies in Machine Learning Using the Game of Checkers". *IBM Journal of Research and Development*. **44**: 206–226.

[CiteSeerX 10.1.1.368.2254](#). [doi:10.1147/rd.441.0206](#)

- **Learning:** (un-)supervised, reinforcement, self-learning, feature learning, anomaly detection, clustering...
- **Models:** neural networks, decision trees, support vector machines, ...
- **Strong ties to:** Optimisation, data mining, statistics -> data science

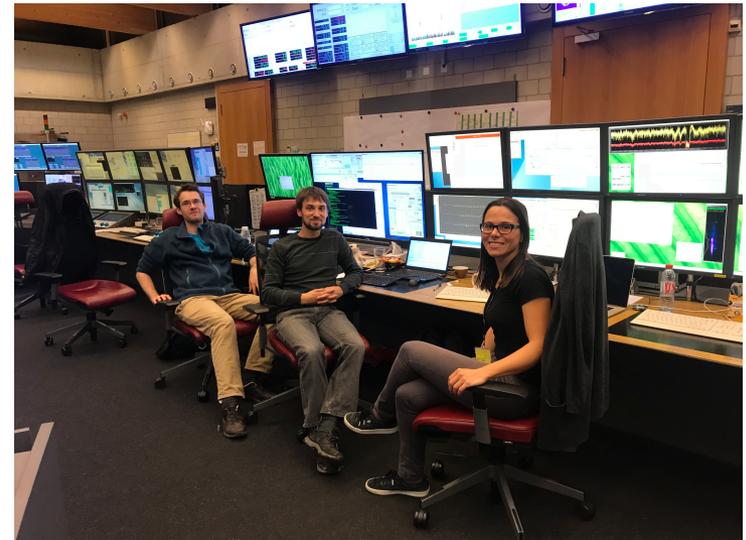
ML at PSI - What for?

- **Tuning / Optimisation / Control**
(e.g. Bayesian optimisation at SwissFEL & HIPA)
- **Prognostics / alarm handling / anomaly-breakout detection**
(e.g. interlock prediction at HIPA, virtual diagnostics -> Photon diagnostics looking into this)
- **Data analysis**
(e.g. clustering, pattern recognition -> beamline people are looking into this)
- **Simulations / Modeling**
(e.g. surrogate models when simulations are very time-consuming or there is no clear model -> predicting detector calibration factors; training a fast “online” model of an accelerator)
- **Other projects in the planning**
(e.g. efficient patient scheduling for proton therapy, ...)

Goal: Optimise the FEL pulse intensity
-> Many Photons = Happy Users

Challenges:

- Exponential dependence of FEL output on many, coupled machine parameters
- Manual tuning is time consuming & inefficient
- Use many knobs (~40)
- Don't drive the machine into the wall...

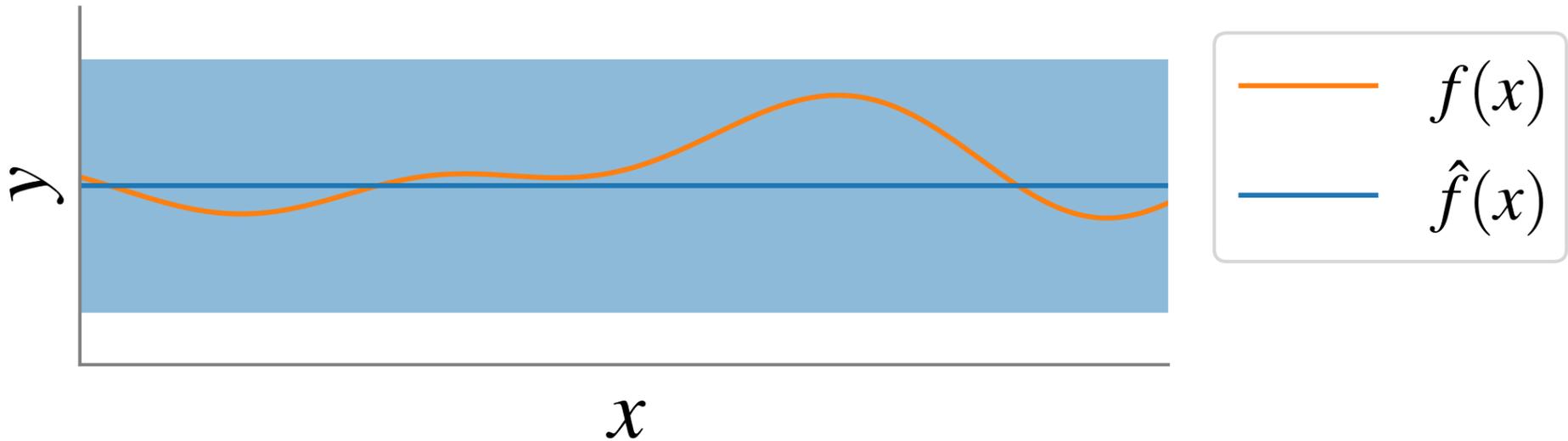


J. Kirschner¹, M. Nonnenmacher¹, M. Mutny¹, A. Krause¹, N. Hiller², R. Ischebeck², A. Adelman²

¹ETH Zürich, Department of Computer Science, Zürich, Switzerland

²Paul Scherrer Institut (PSI), CH-5232 Villigen PSI, Switzerland

- The ML Experts from ETHZ suggested Bayesian Optimisation!
-> How does this work?

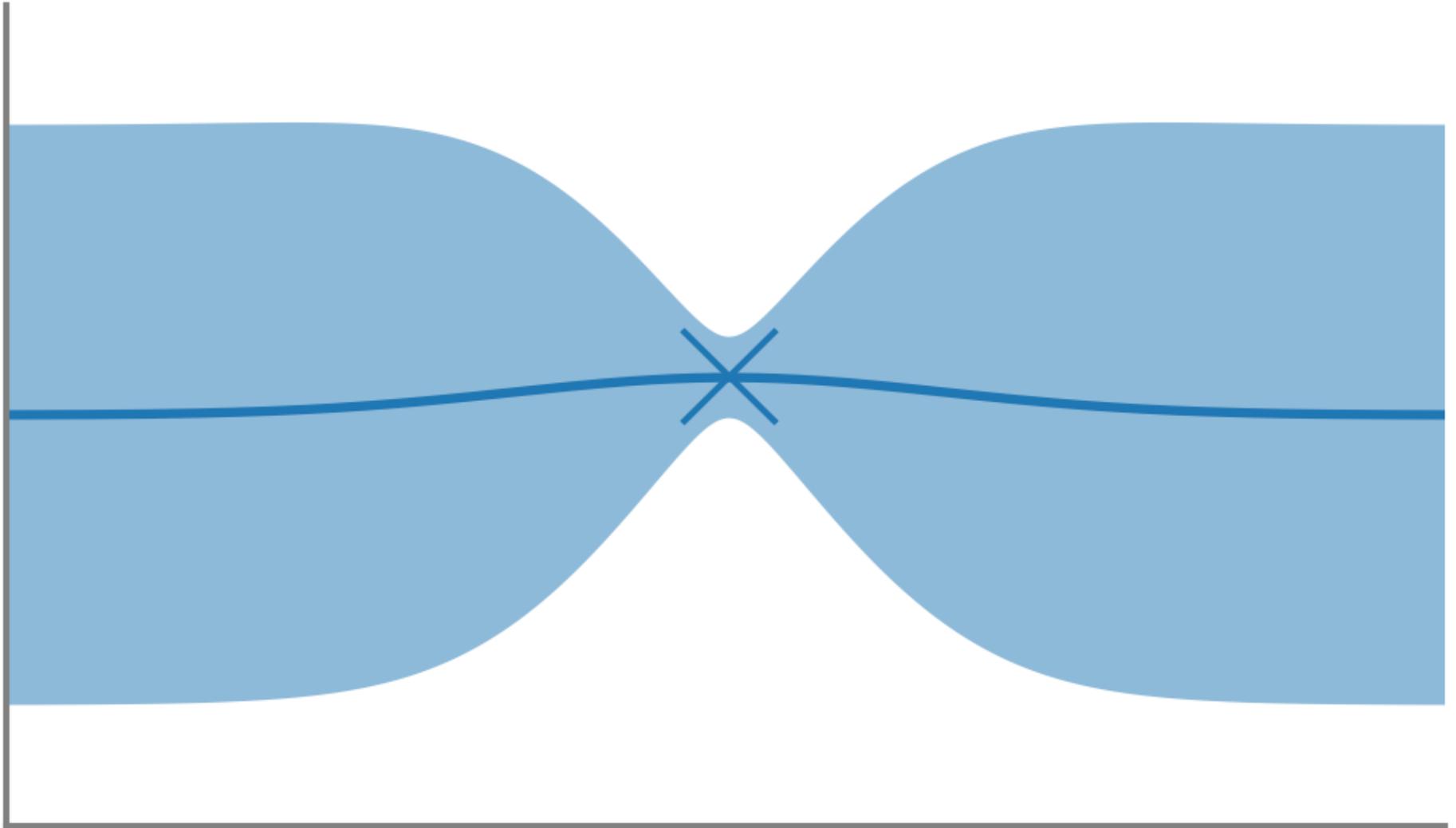




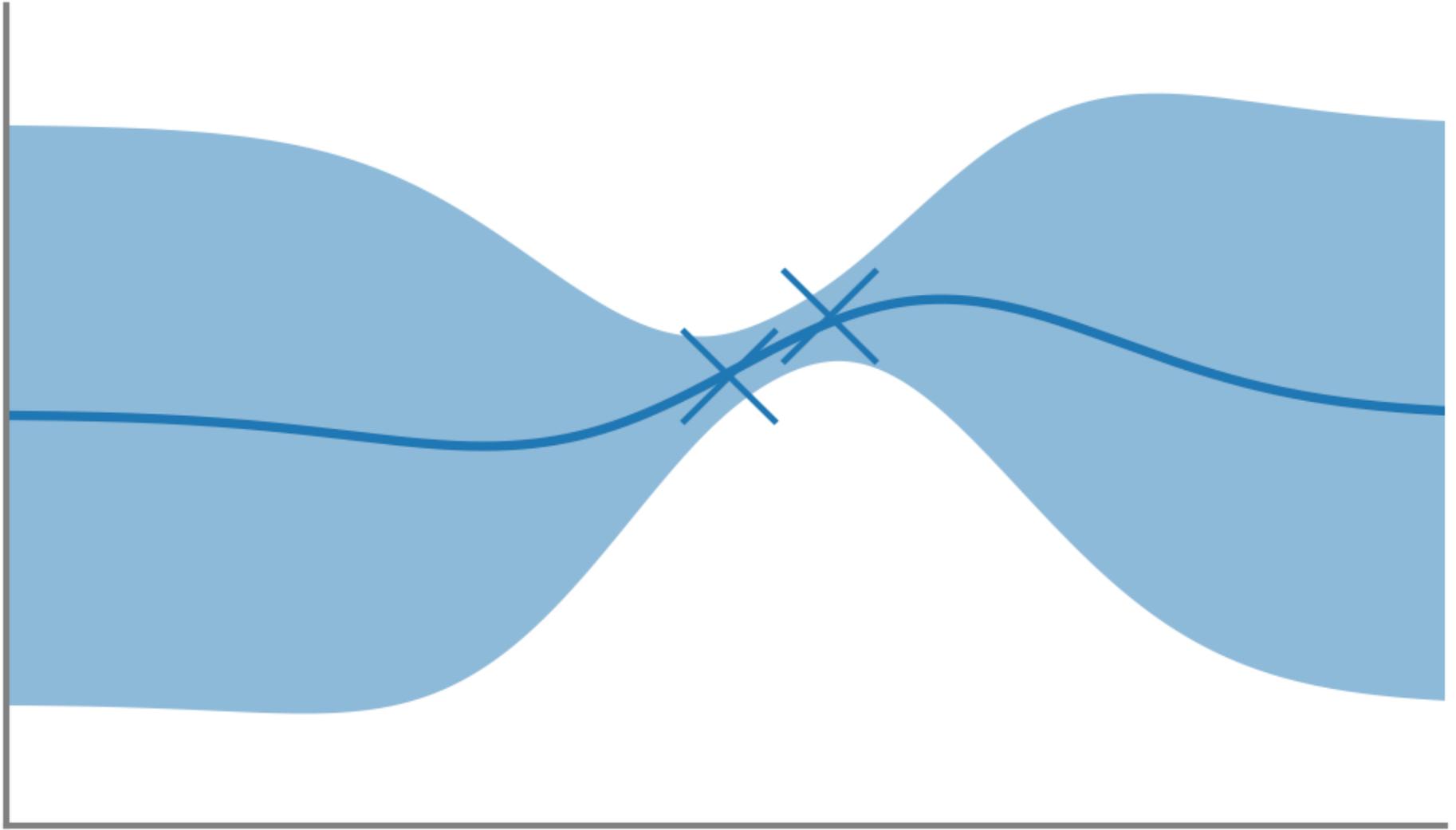
Bayesian Optimisation at SwissFEL

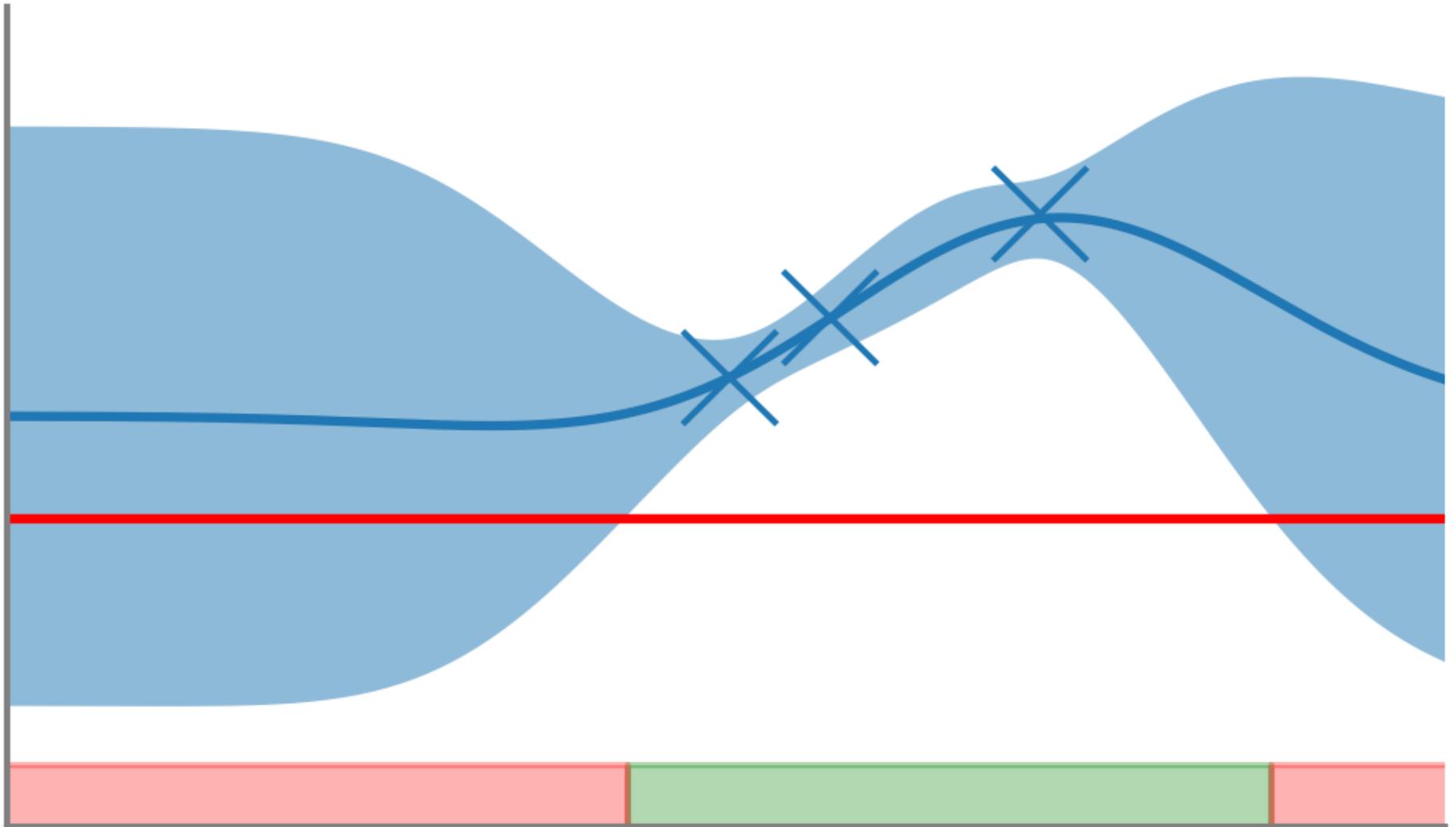


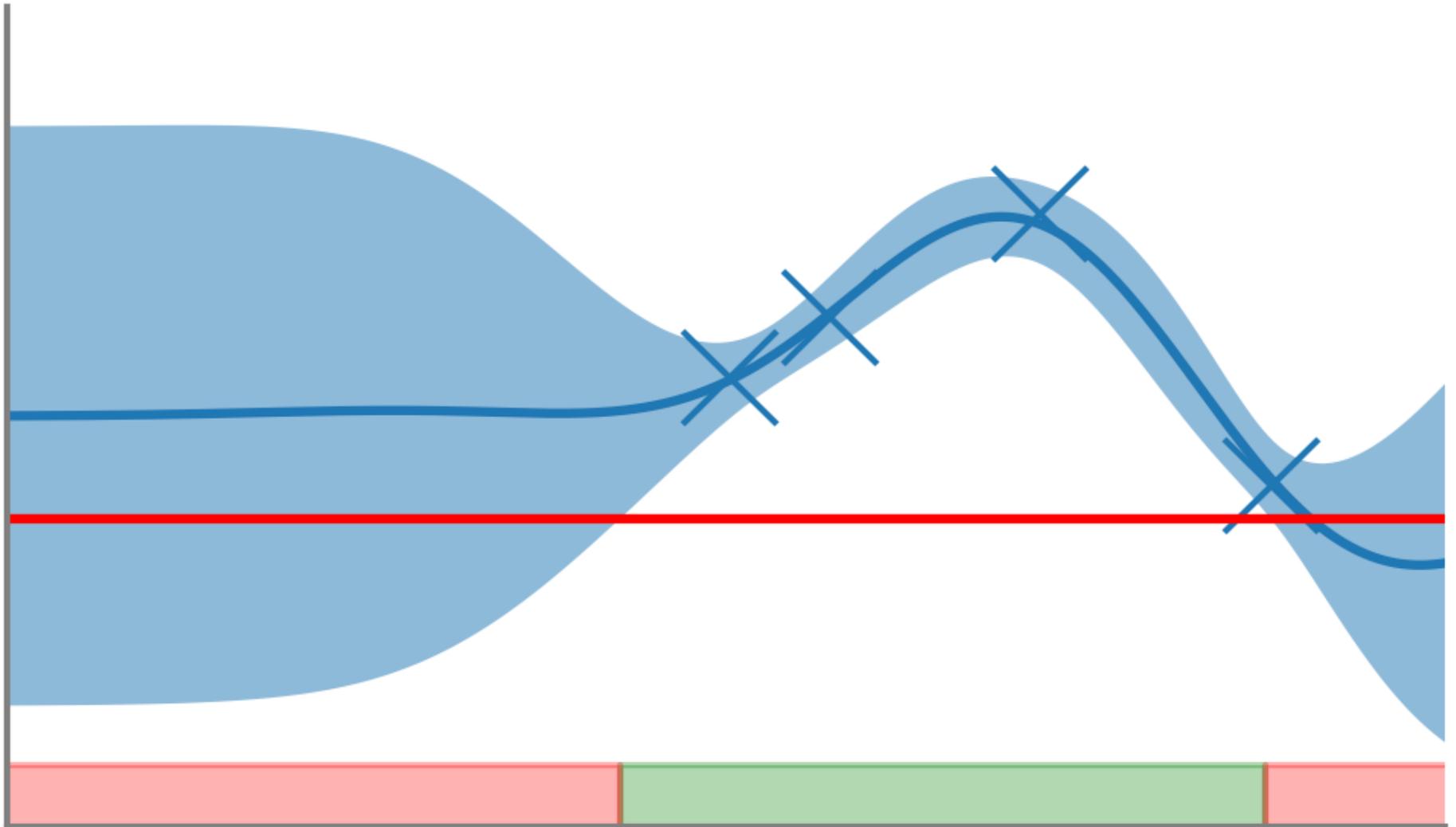
Bayesian Optimisation at SwissFEL

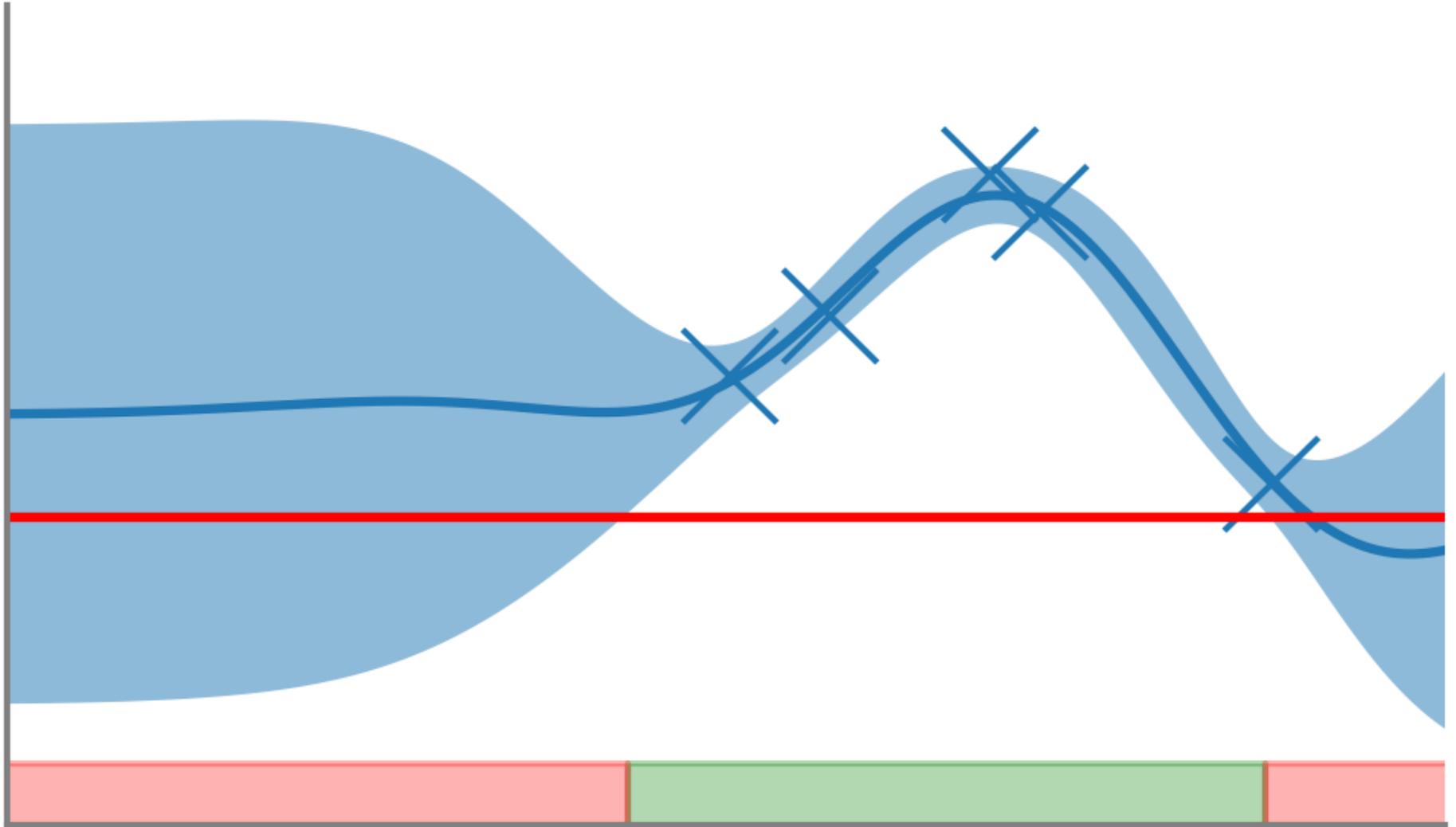


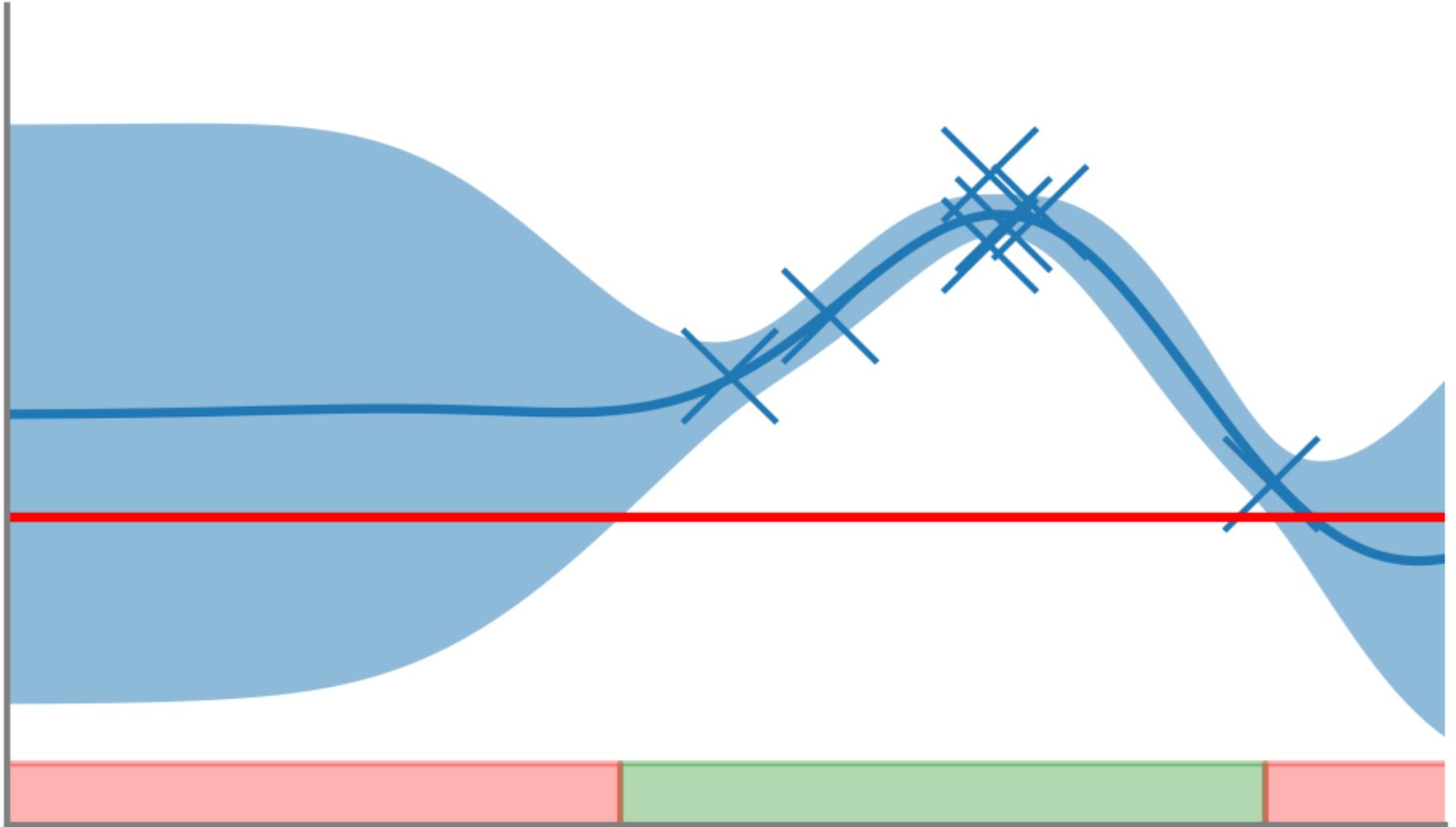
Bayesian Optimisation at SwissFEL

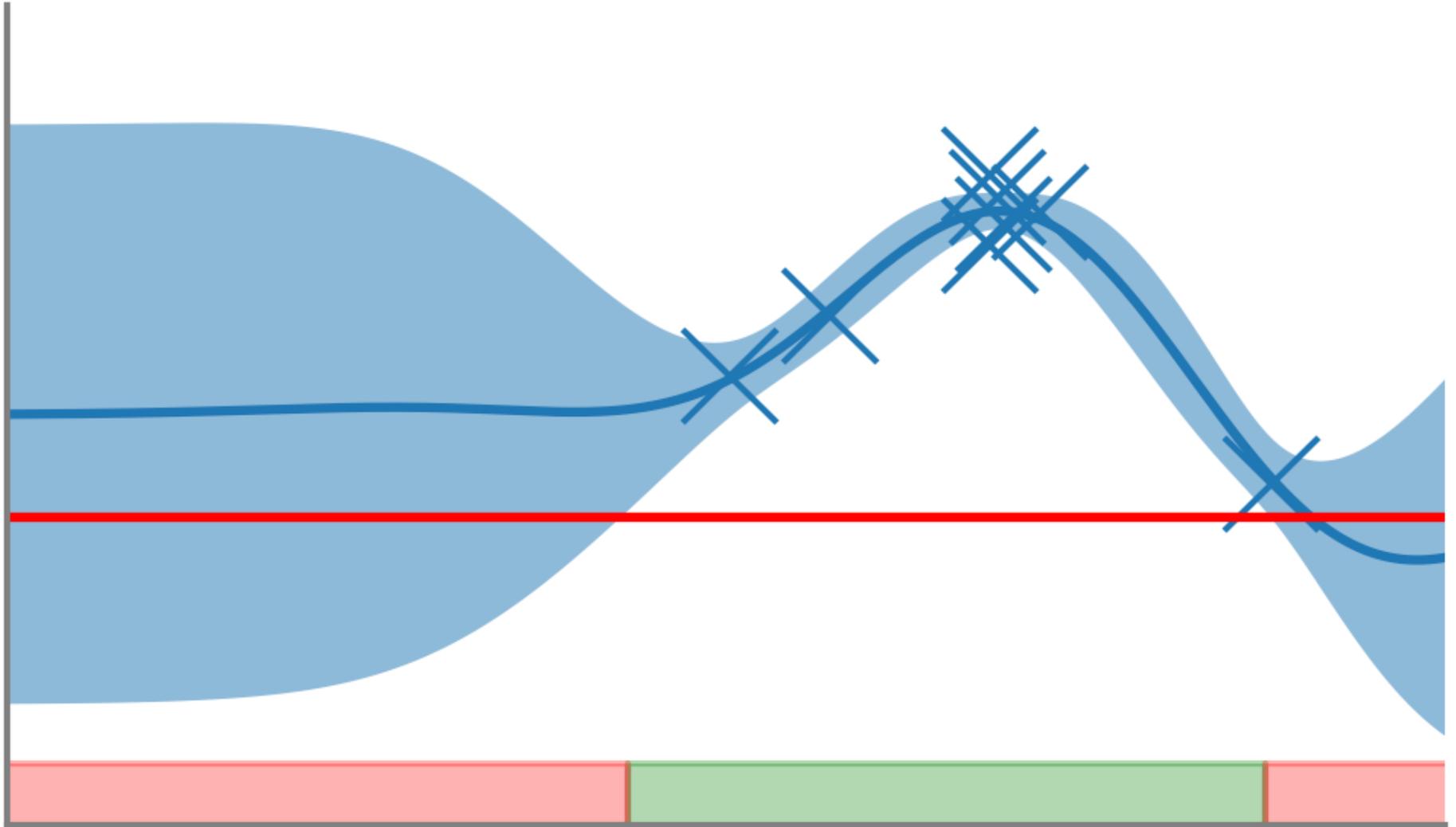


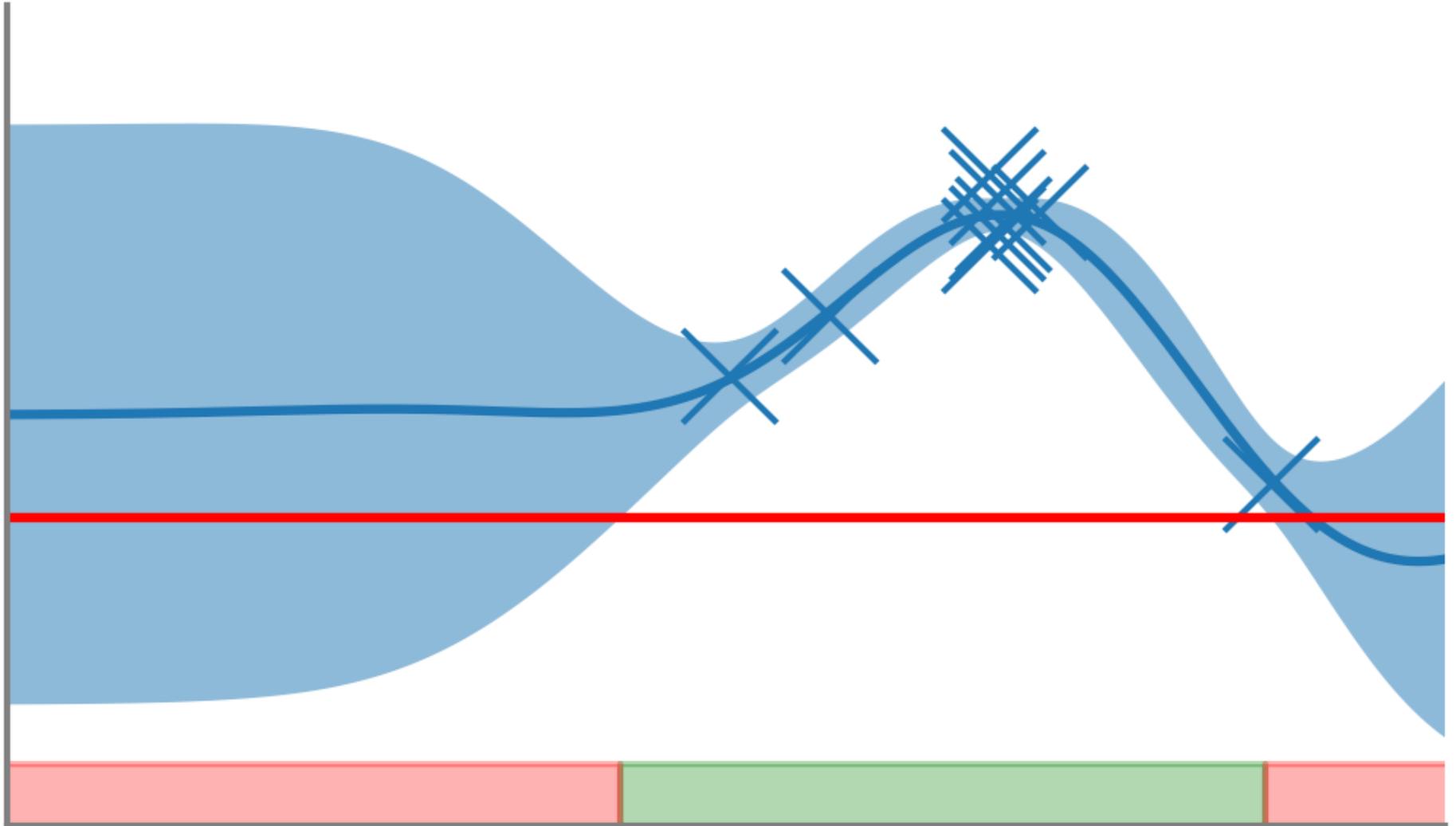












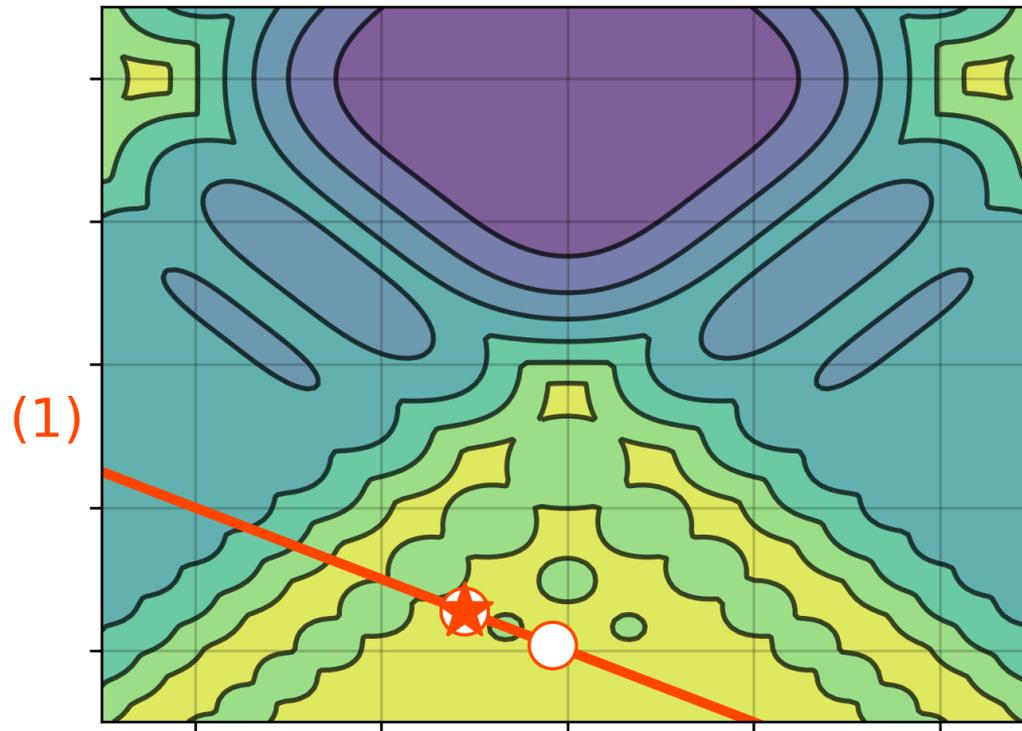


Bayesian Optimisation at SwissFEL

- How does it scale for many parameters? -> Poorly...
- What do we do? -> Combination of gradient method (local sampling) and 1D line approach over several lines

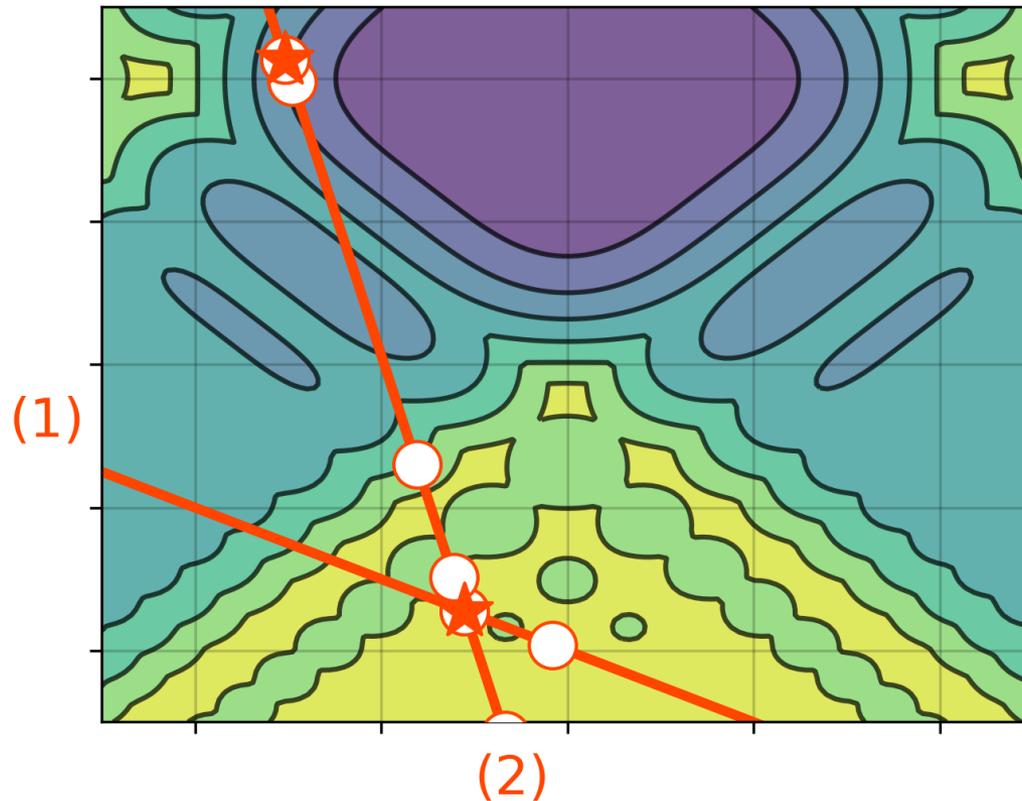
Bayesian Optimisation at SwissFEL

- How does it scale for many parameters? -> Poorly...
- What do we do? -> Combination of gradient method (local sampling) and 1D line approach over several lines



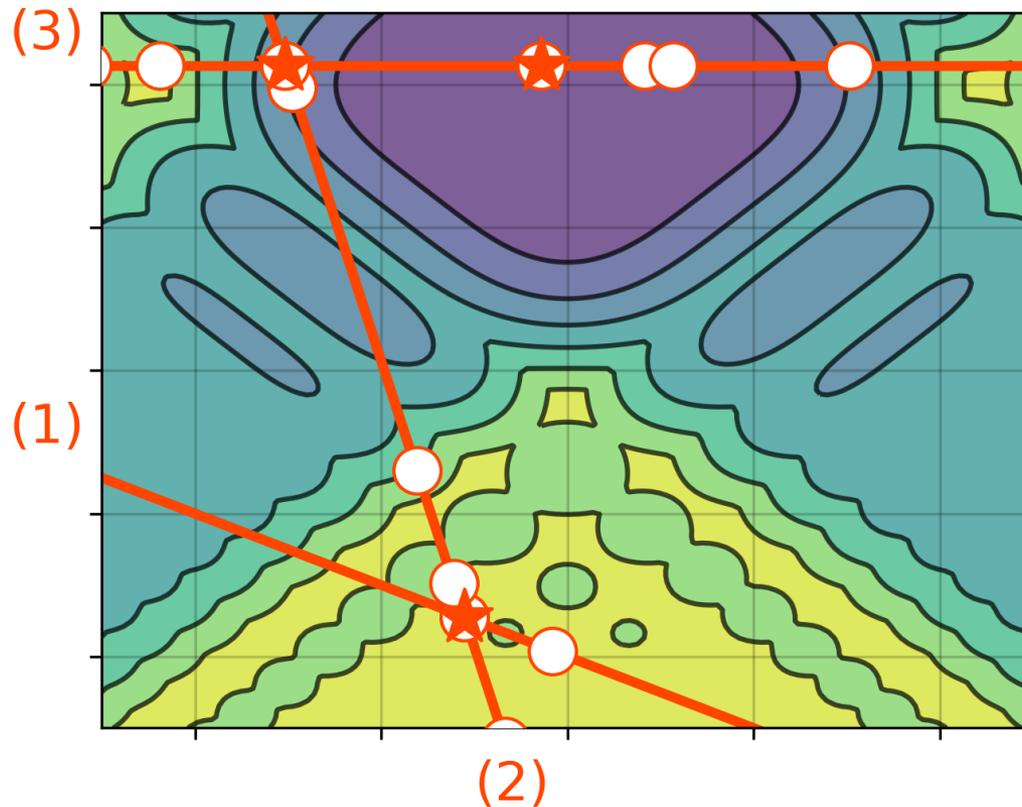
Bayesian Optimisation at SwissFEL

- How does it scale for many parameters? -> Poorly...
- What do we do? -> Combination of gradient method (local sampling) and 1D line approach over several lines



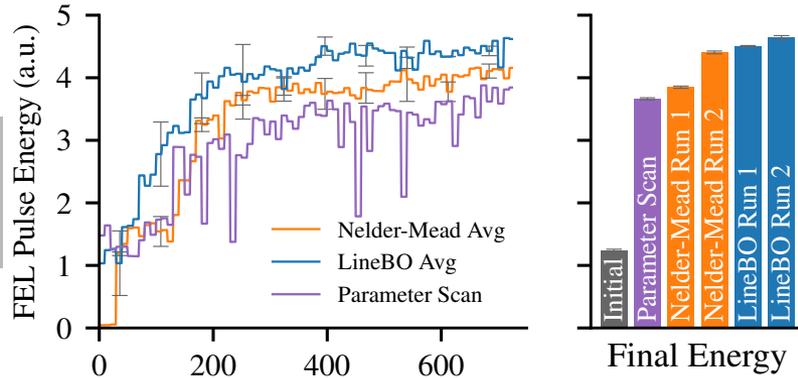
Bayesian Optimisation at SwissFEL

- How does it scale for many parameters? -> Poorly...
- What do we do? -> Combination of gradient method (local sampling) and 1D line approach over several lines



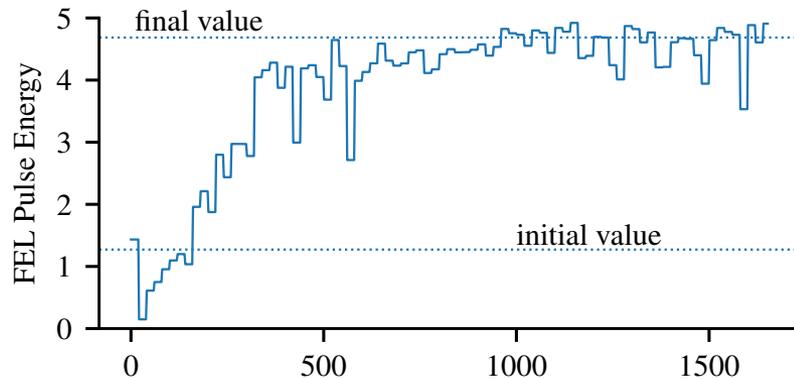
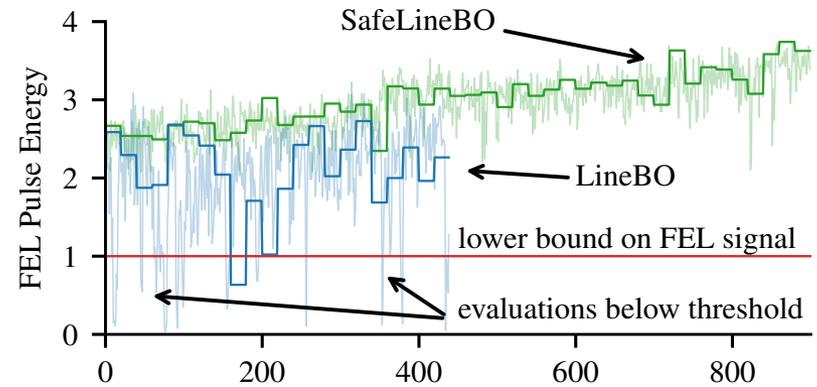


Bayesian Optimisation at SwissFEL



Benchmarking against other algorithms!
(24 Parameters)

Adding Safety Constraints
(24 Parameters)

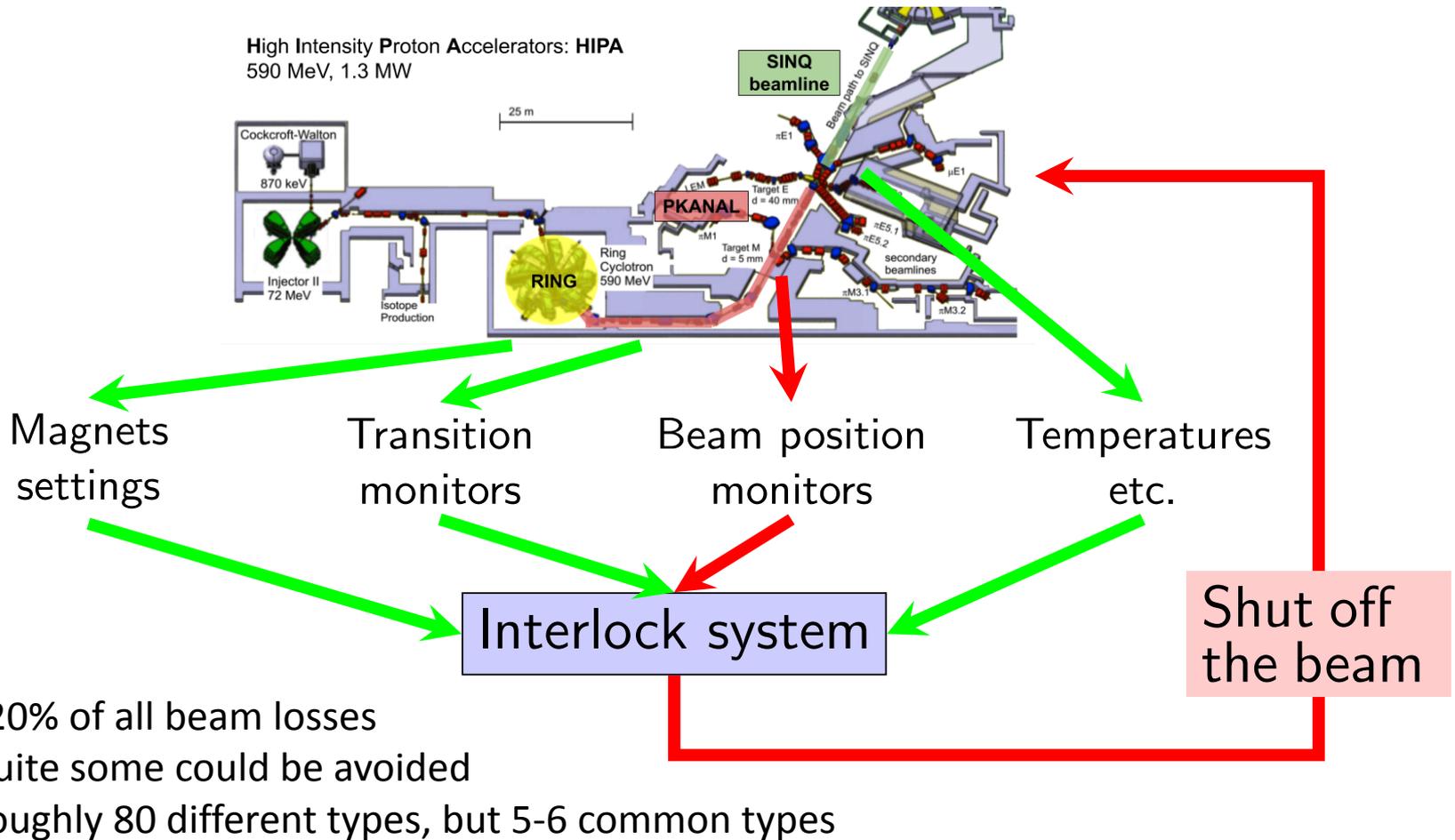


Tuning at about 1 Hz: 500 steps = 8 min

Still works fast with 40 Parameters!

- Where does the journey go to?
 - Versatile Framework with a GUI that
 - can be used by operators
 - can be used by other accelerators -> we are also tuning HIPA with it now!

Interlock Prediction at HIPA



Andreas Adelman, Jaime Coello, Davide Reggiani, Jochem Snuverink, Melissa Zacharias (PSI); Anastasia Pentina (Swiss Data Science Center)

- What would be your approach?

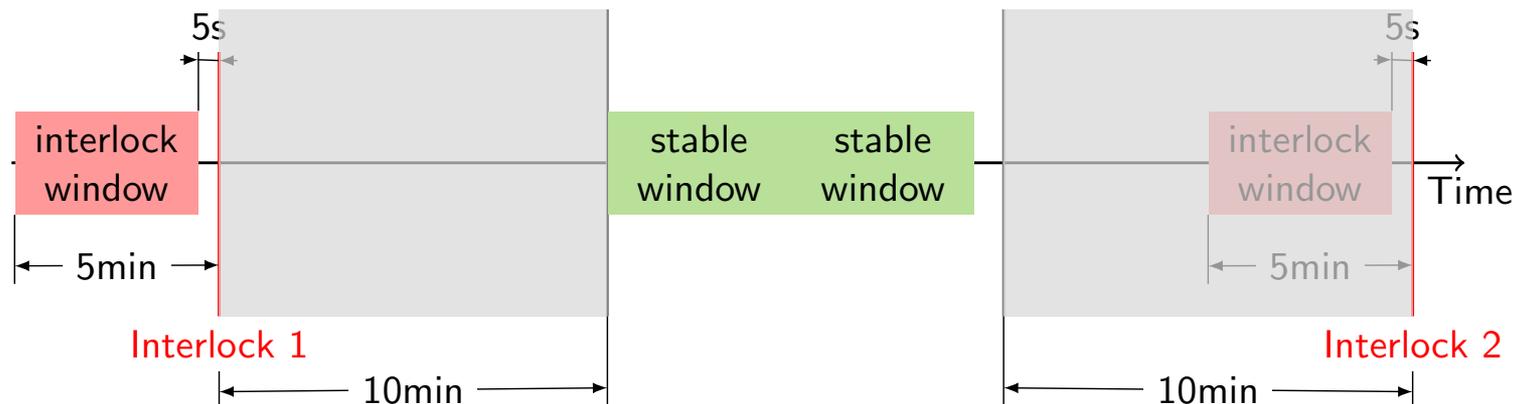
Interlock Prediction at HIPA

Goal: Train a model (supervised learning) that can then predict interlocks and prevent them by lowering the beam current temporarily

Approach:

- Training on archived data (450 channels) -> lots of data preparation needed
- Finding the best model

- **Interlock-window:** interlock happens in 5 seconds
- **Stable-window:** interlocks are far from happening



Interlock Prediction at HIPA

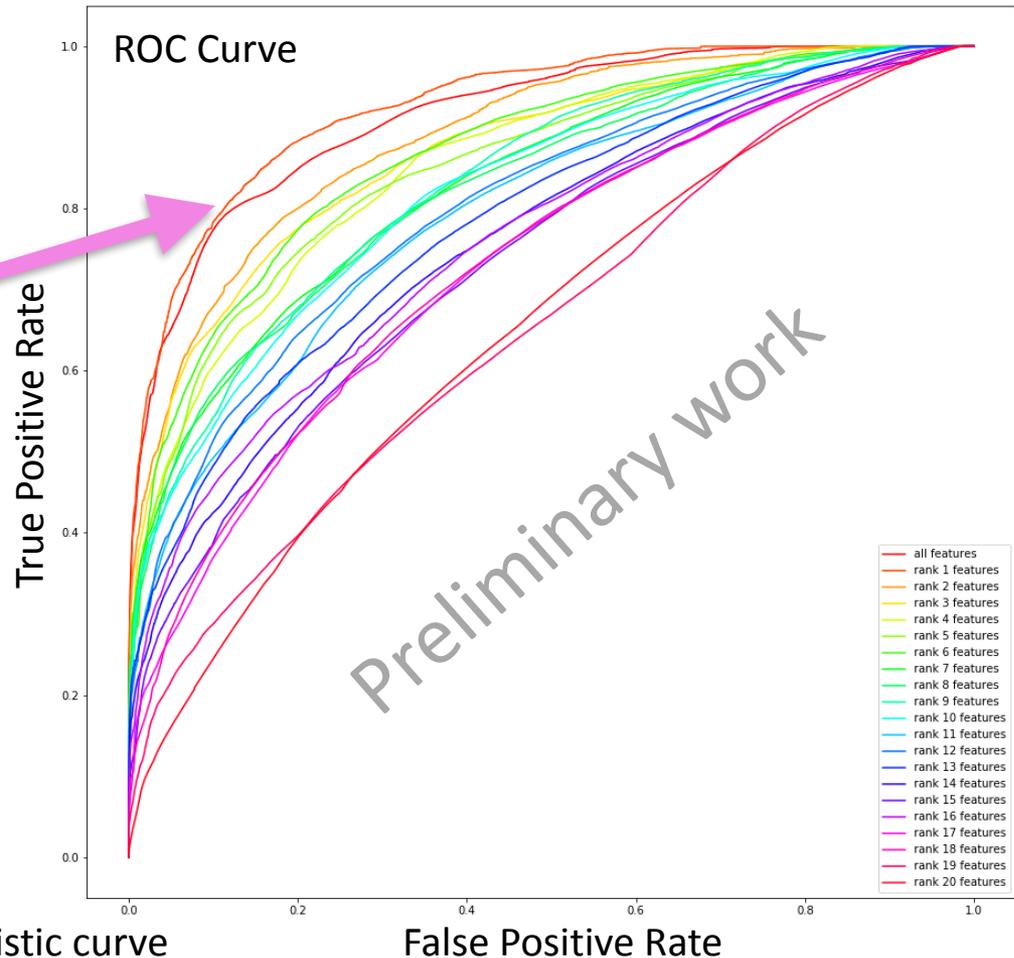
In an interlock happens: 25 s no beam

If the charge is temporarily lowered: corresponds to 6 s no beam

-> **We want a high positive true prediction, but a very low positive false one!**

**So far best results with:
Random Forest**

Best results with
50/450 features!
-> Maybe we can
also learn some
physics from that

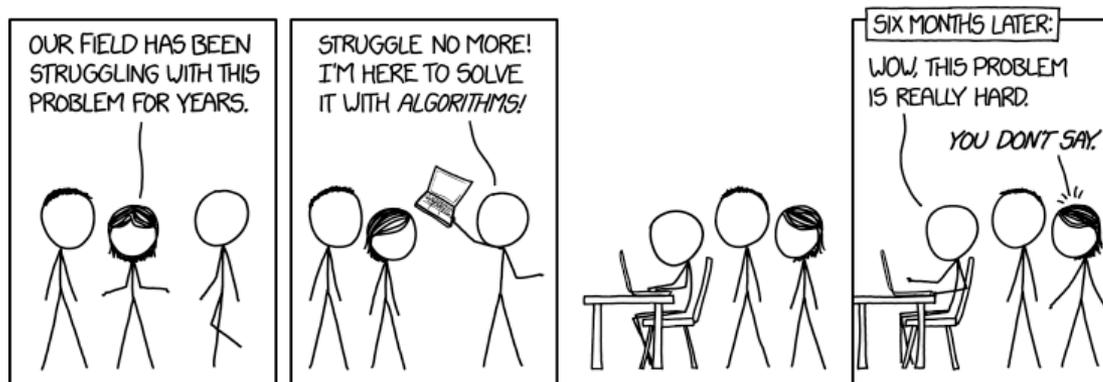


ROC curve = receiver operating characteristic curve

False Positive Rate

Lessons learned so far

- **Machine learning techniques can be great tools especially when**
 - Your problem is hard to formulate analytically (pattern recognition,...)
 - You have a lot of data
 - You would have to frequently use very time-consuming simulations
 - Your “standard tools” do not work very well
 - You need a high degree of automation
- **but**
 - You still need to understand your problem! -> it does not do that for you
 - You still need to prepare your data (data science skills are crucial) -> do not underestimate this part
 - Quite often a lot of “hyper-parameters” are involved -> you will need to learn the “tricks of the trade”



- Regular (every 6 weeks) discussion rounds: ML Luncheons (11:30-13:00 with free sandwiches) -> open discussions, informal presentations
- ML Awareness event -> Half day event with invited speakers
- ML Mailing List & SLACK channel for announcements & discussions
- Collaboration with the Adaptive Systems Group at ETHZ (Prof. Andreas Krause):
 - PhD & Master Students
- Project grants for the Swiss Data Science Center (SDSC)
 - www.datascience.ch -> professional support for projects (e.g. data scientist is at PSI once per week)
 - 3rd ICFA Mini-Workshop on Machine Learning Applications for Charged Particle Accelerators 18-21 Feb 2020 at PAL, Korea <https://www.indico.kr/event/5/>

What are your
opinions/
experiences?



“The machine learning algorithm
wants to know if we’d like a
dozen wireless mice to feed the
Python book we just bought.”

THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG
PILE OF LINEAR ALGEBRA, THEN COLLECT
THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL
THEY START LOOKING RIGHT.



How are
things at your
institute?

www.xkcd.com

OH, HEY, YOU ORGANIZED
OUR PHOTO ARCHIVE!

YEAH, I TRAINED A NEURAL
NET TO SORT THE UNLABELED
PHOTOS INTO CATEGORIES.

WHOA! NICE WORK!



ENGINEERING TIP:
WHEN YOU DO A TASK BY HAND,
YOU CAN TECHNICALLY SAY YOU
TRAINED A NEURAL NET TO DO IT.