Jochem Snuverink :: Paul Scherrer Institut

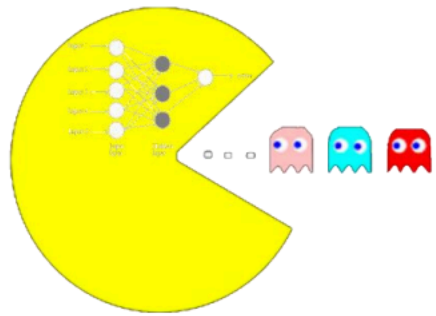
Automation of accelerator tuning with safety constraints using Bayesian Optimisation

LMU seminar – 28.01.2022

Outline

- PACMAN
- (beamline) optimisation
- optimisation GUI
- Examples PSI
  - SwissFEL
  - HIPA
  - PiE1 - MIXE
- Practicalities & Next steps
- Conclusions

PACMAN



- PACMAN (Particle Accelerators & Machine Learning) project
- Swiss Data Science Center (SDSC), the École Polytechnique Fédérale de Lausanne (EPFL), CERN, ETH-Z and PSI
- Particle accelerators are complex facilities that produce and handle large amounts of data
- The project target is to explore the possibilities to use machine learning to improve the performance of particle accelerators
- March 2019 - June 2021

Why automate beamline tuning?

- Time-consuming repetitive task for operators
- (potentially) faster
  - Quick iterations, no dead time
  - Can change many parameters at once
  - Evaluate noise
- (potentially) safer
  - No “human” mistakes
  - Can supervise all safety constraints simultaneously

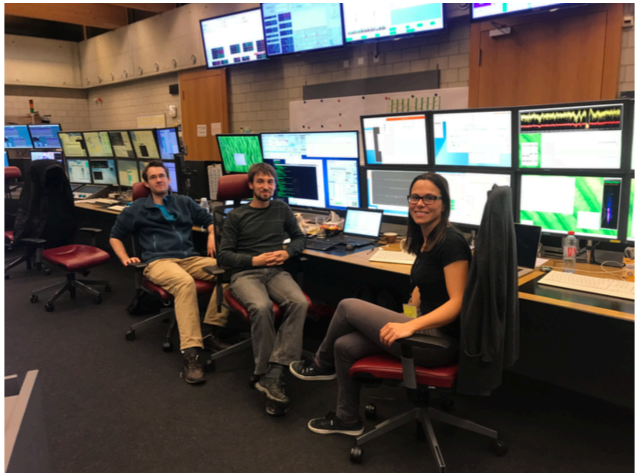
Optimisation at HIPA and SwissFEL

HIPA

- HIPA beam power (1.3 MW) limited to the beam losses
  - Reduce damage and activation
- Optimisation is now mostly done empirically
- Large potential for automated optimisation and surrogate model construction
  - No accurate and fast physics model available
  - Needs to be *safe*

SwissFEL

- Goal: Optimise the FEL pulse intensity
  -> Many Photons = Happy Users
- Challenges:
  - Manual tuning is time consuming & inefficient
  - Use many knobs (~40)
  - Don't drive the machine into the wall...
  - Exponential dependence of FEL output on many, coupled machine parameters



ML experts from ETHZ suggested Bayesian Optimisation

- J. Kirschner[1], M. Nonnenmacher[1], M. Mutny[1], A. Krause[1], N. Hiller[2], R. Ischebeck[2], A. Adelmann[2]
- [1] ETH Zürich, Department of Computer Science, Zürich, Switzerland
- [2] Paul Scherrer Institut (PSI), CH-5232 Villigen PSI, Switzerland

- The work is supported by the Swiss National Science Foundation under grant 200020_159557

Why Bayesian Optimisation?

Human Optimisation ≠ Numerical Optimisation

Human Optimisation

- Life-long learning
- Experience
- Limited working memory
- (relatively) Slow decisions

Numerical Optimisation

- Bulk learning
- Cannot estimate uncertainty
- Juggle many things at once
- Fast decisions

PAUL SCHERRER INSTITUT PSI

ETH zürich

Learning & Adaptive Systems

Why Bayesian Optimisation?

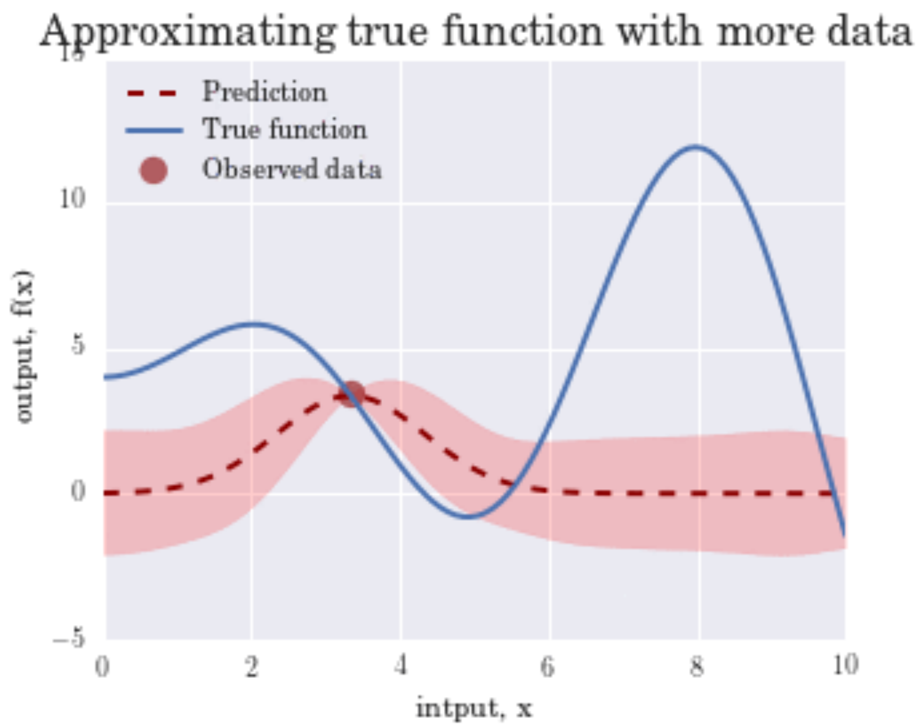
Bayesian Optimisation

- ~~Life-long~~ learning
- Experience
- ~~Limited working memory~~
- ~~(relatively) Slow decisions~~

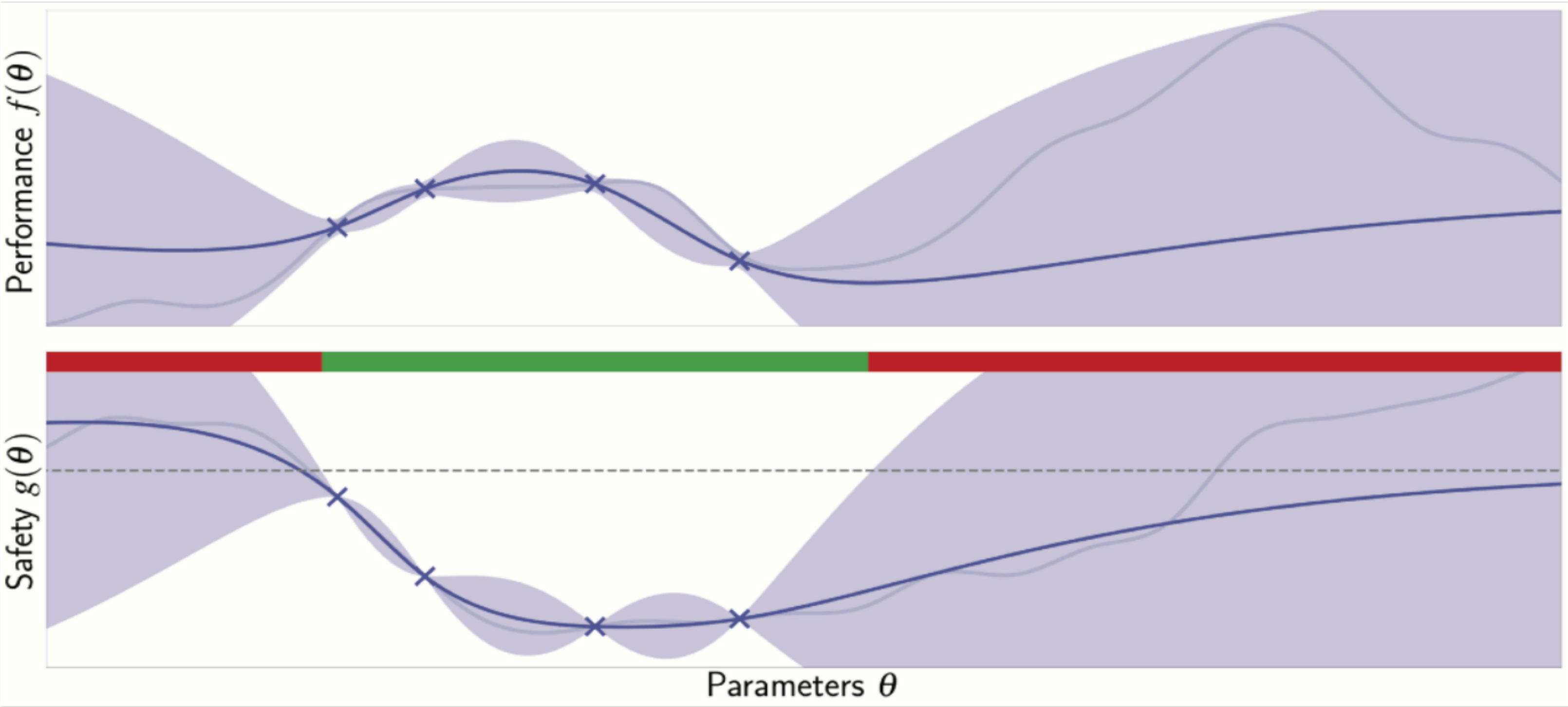
- ~~Bulk learning~~
- ~~Cannot estimate uncertainty~~
- Juggle many things at once
- Fast decisions

Bayesian Optimisation

- Bayesian Optimisation: flexible, data-driven approach for global optimisation with noisy feedback.
- Fit a statistical regression model of the data, typically a gaussian process.
- On each iteration the model looks for evaluation points that reduce the uncertainty of the unknown target function.

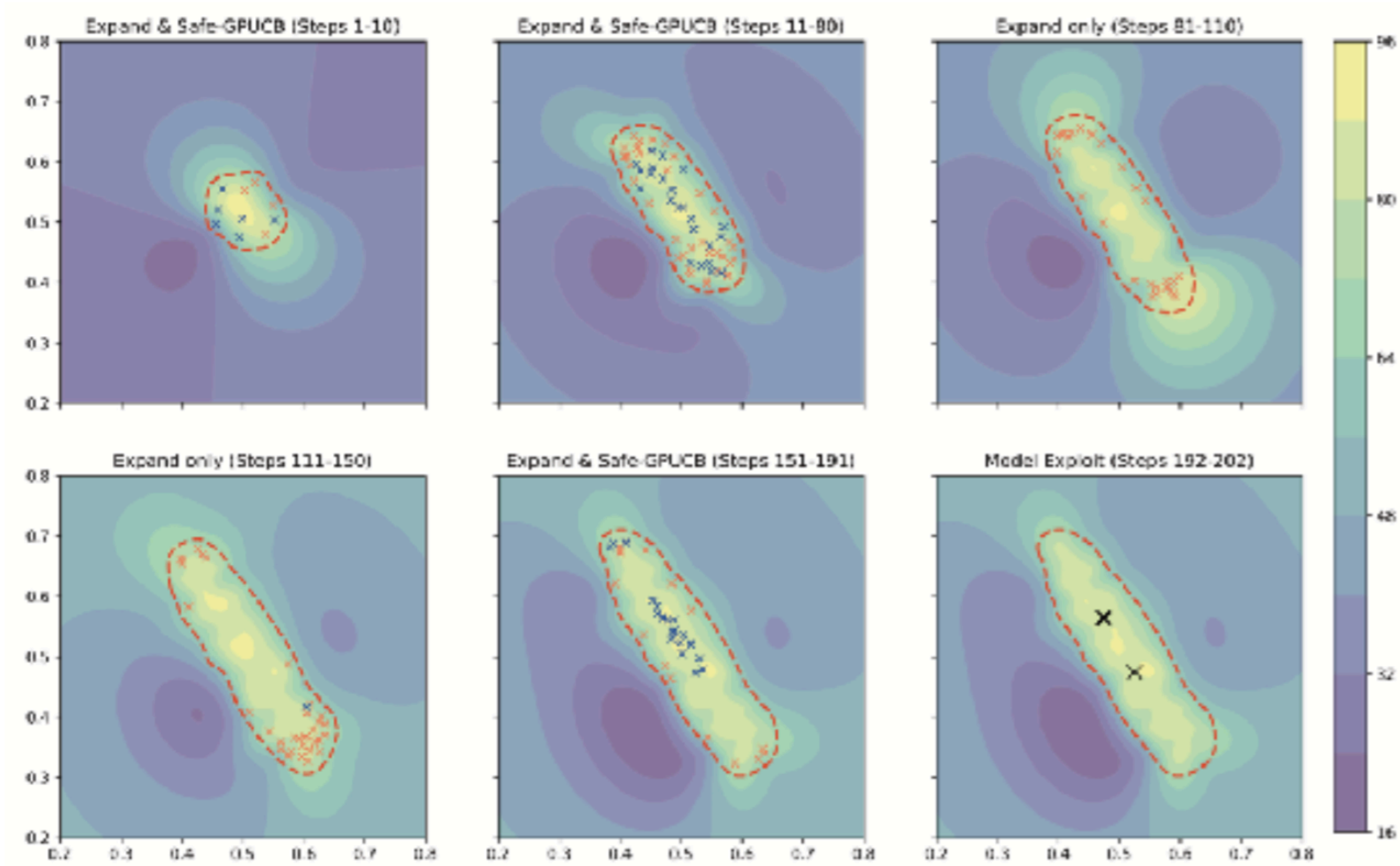


Safe Bayesian Optimisation



- X are evaluation points, for each evaluation point the performance (e.g. FEL pulse energy) is measured and the model is updated
- Additionally, the safety function (e.g. losses recorded by various detectors) is evaluated and its model is also updated ➜ taking into account the uncertainties of the safety function this defines the safe-set marked in green
- Combines both exploration and exploitation

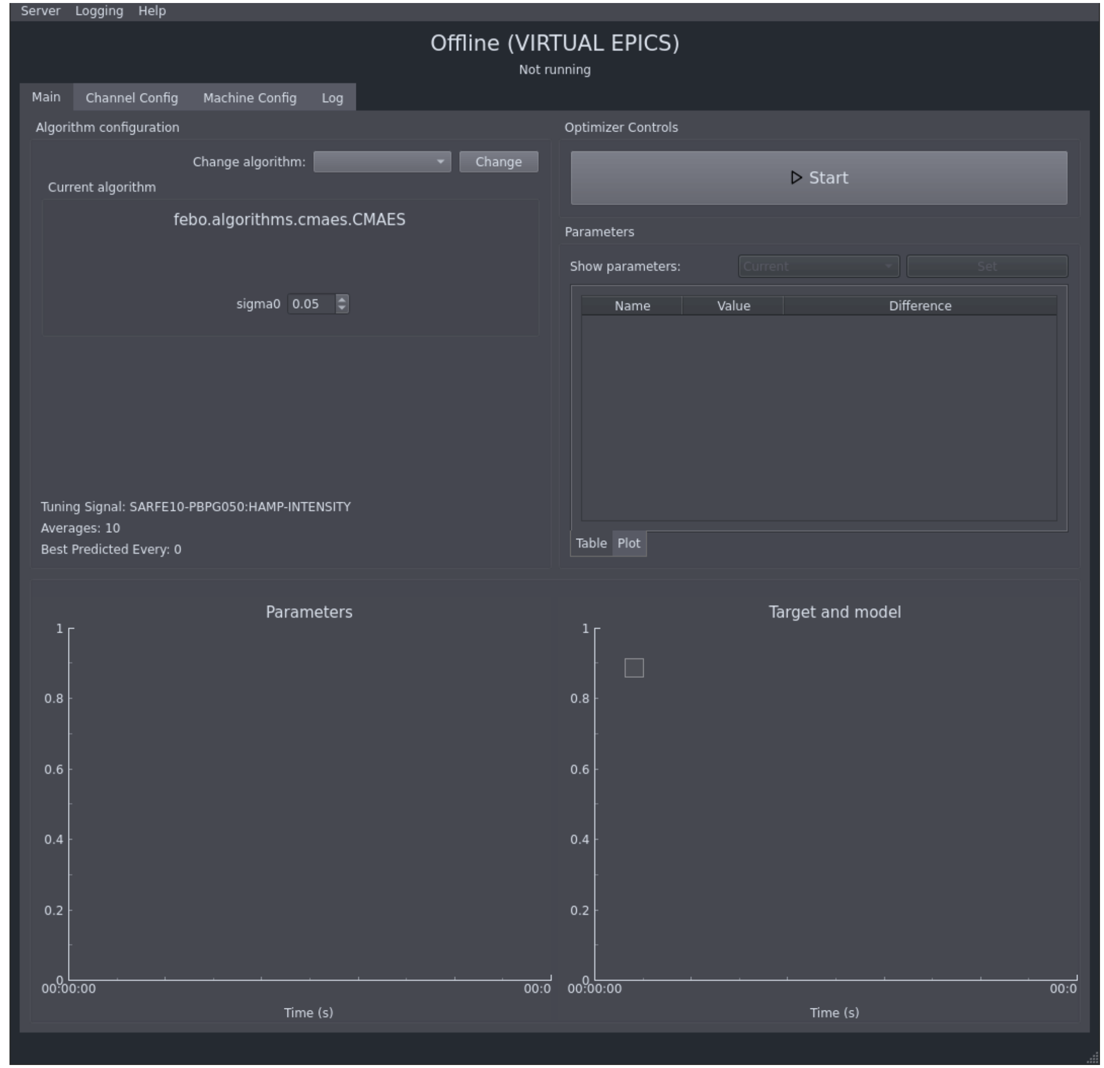
SwissFEL example



- Orange crosses are expanding steps, blue crosses are exploitation steps, red dashed line marks safe-set boundaries, contour plot shows model of the performance function (modeled with Gaussian processes)

Optimisation GUI

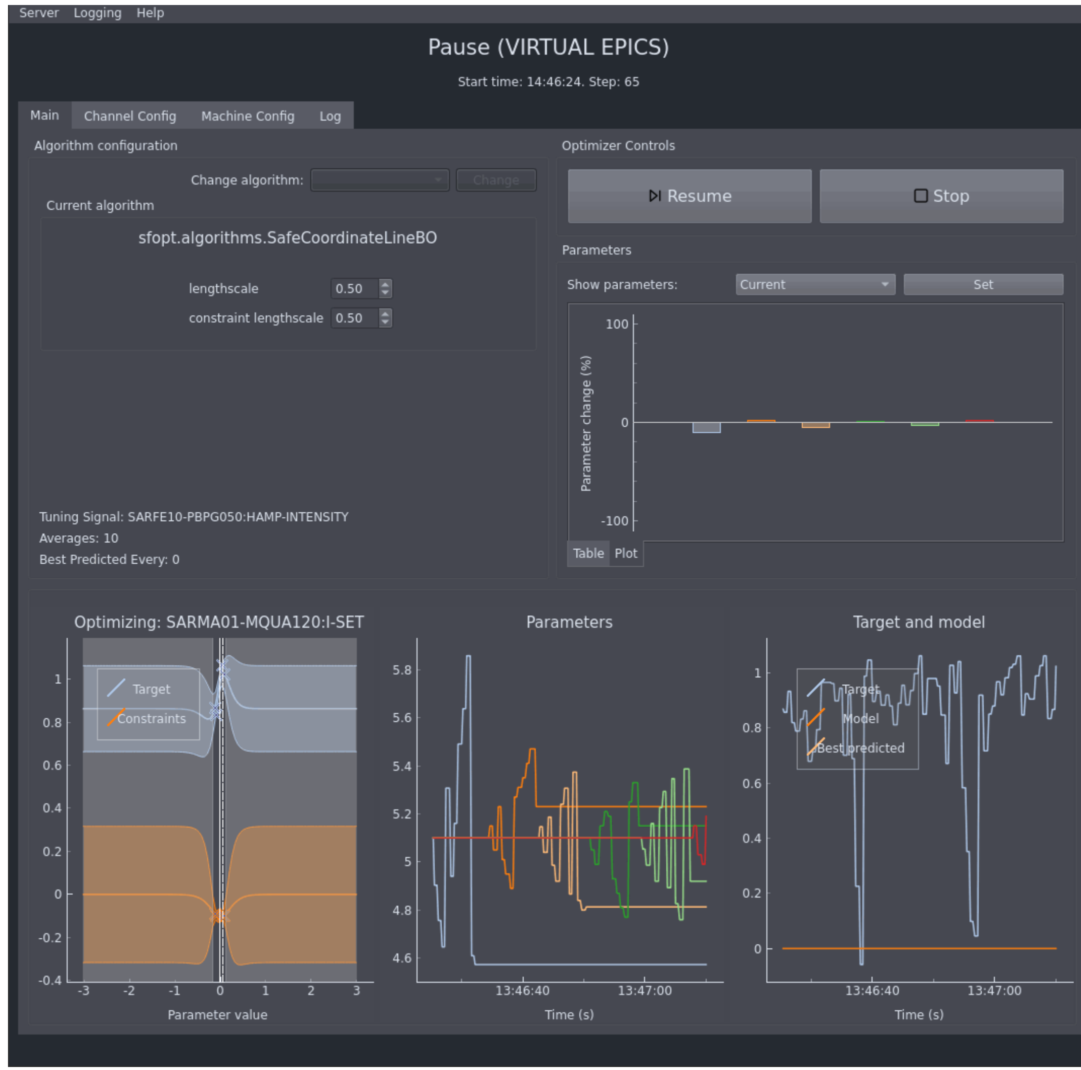
- Server application
  - Python
  - Optimiser runs on a server
  - client GUI through REST API
- Status data in EPICS
- Additional algorithms
  - Extremum seeking (CMA-ES)
  - Nelder Mead (Simplex)
  - Line Scan
  - Easy to implement more
- Few settings
- Machine related beam checks (e.g. good-flags from the feedbacks, beam detection)
- Live plotting & analysis
- Accelerator independent
  - used at SwissFEL, HIPA, PiE1



PAUL SCHERRER INSTITUT

GUI

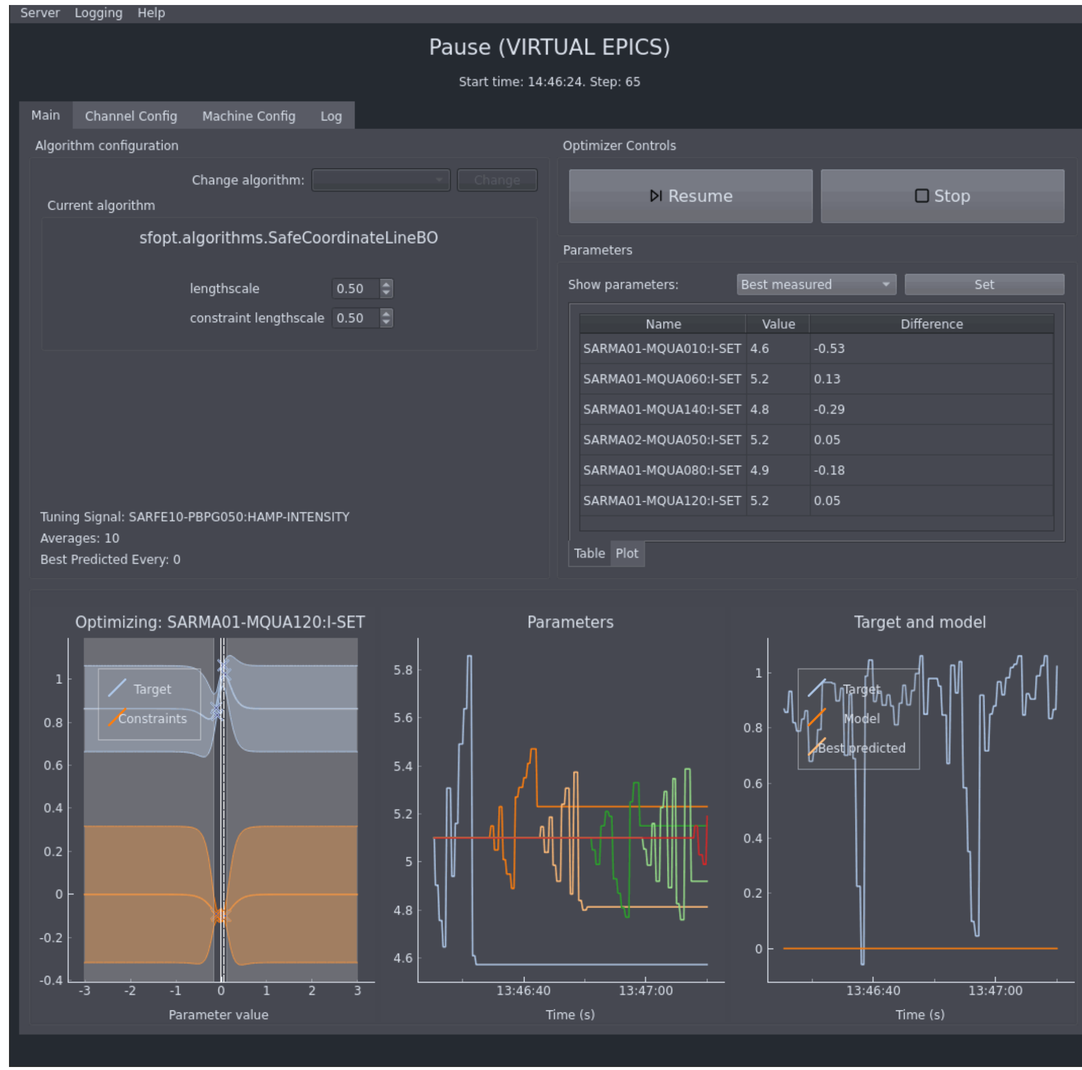
- Possible to pause
  - Load different results
  - Set machine to these states
- Resume



PAUL SCHERRER INSTITUT

GUI

- Another example: CMA-ES
- virtual mode ("virtual EPICS") available for testing

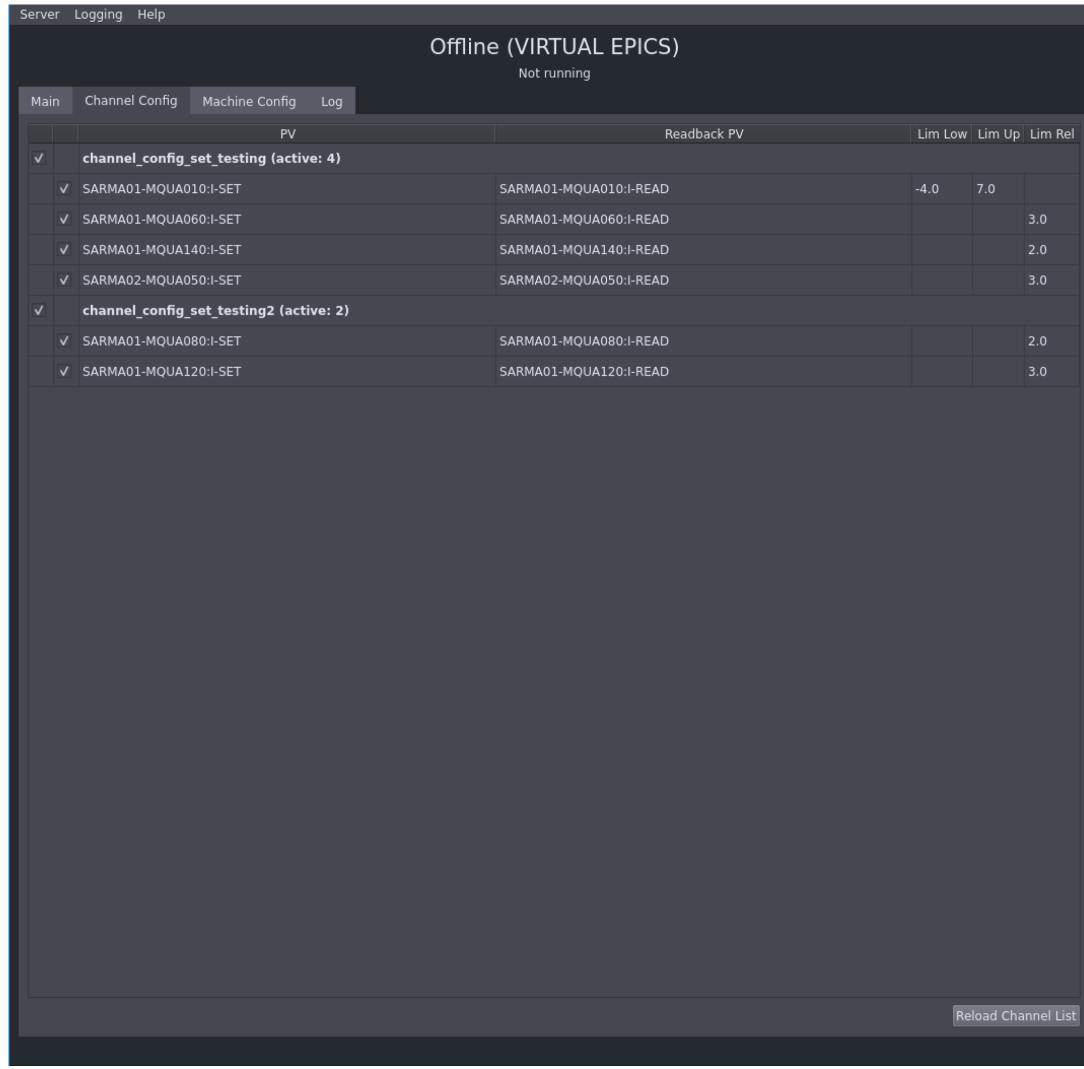


PAUL SCHERRER INSTITUT PSI

GUI

- Easy selection of the parameters (channels) to tune during the optimisation.

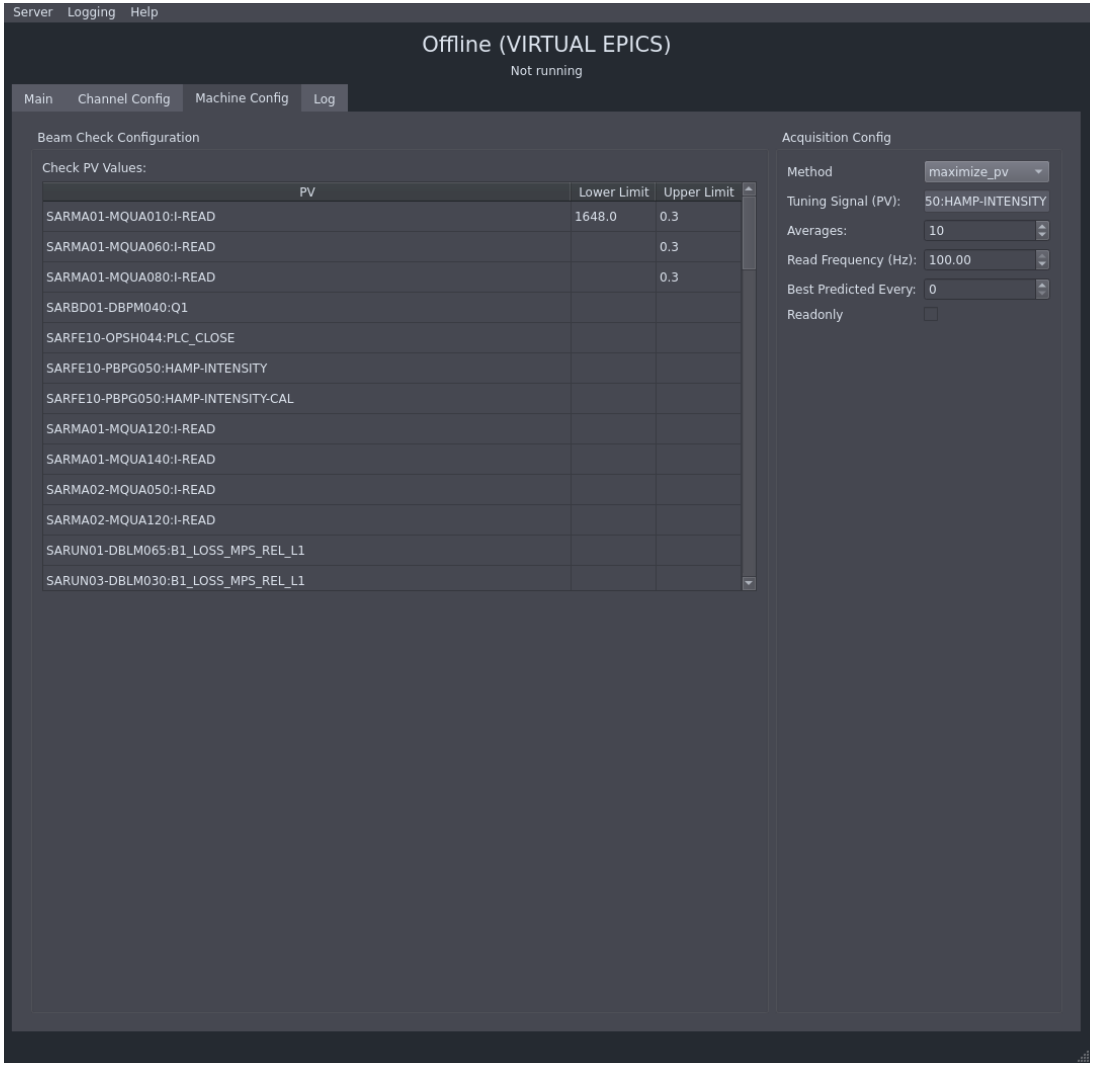
- Individual limits can be set for each channel
  - machine check
  - optimisation constraint



 14

GUI

- Beam limits are checked before moving to the next step
  - Stability conditions
- Possible to investigate and go back to each previous step



Results

Variants

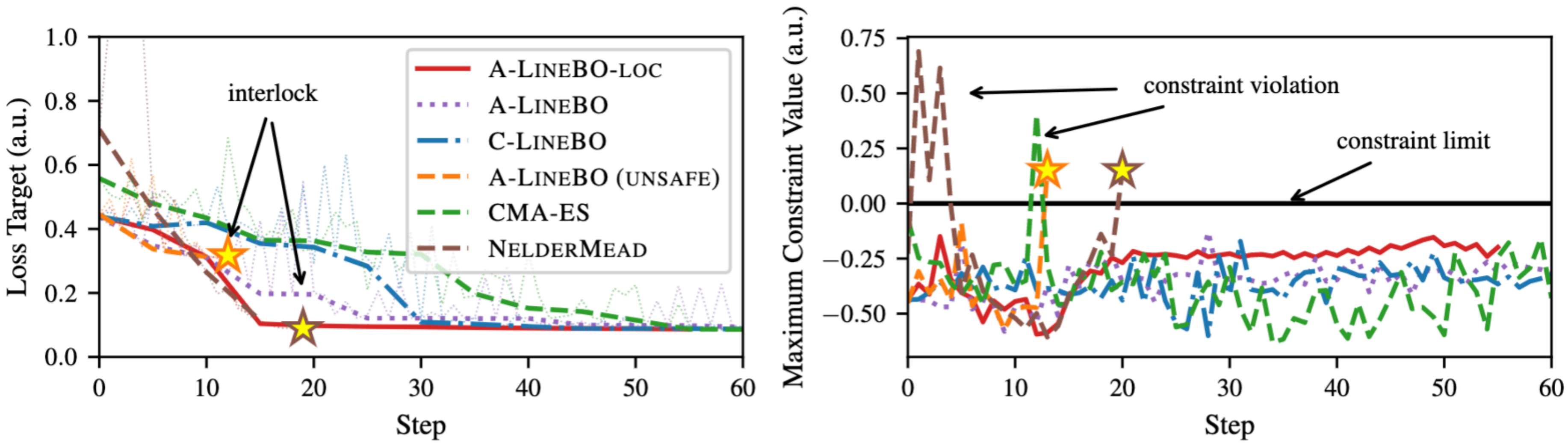
- Variants:
  - C(oord)-LineBO: Find each parameter target optimum individually.
  - A(scent)-LineBO: Move each parameter to find the “best-direction” and then search for the best point in that direction.
  - *-Loc: “Localised” Maximum step size is 10% of the available range.
  - CMA-ES: Similar to swarm optimisation.
    - Does not have safety constraints.
    - Standard well developed global optimisation evolutionary algorithm.
    - Used here as baseline to compare the rest of the method

HIPA

- Objective: Minimize combined losses (M4HIPA:VERL:2)
  - weighted average of 60 loss monitors
- Tuning Parameters: 5-16 Quadrupole Magnets
- Constraints: About 200 loss monitors with individual warning levels
- Beam checks:
  - Beam on
  - Orbit feedbacks stable
- Effective control rate: ~ 5 seconds / step

HIPA

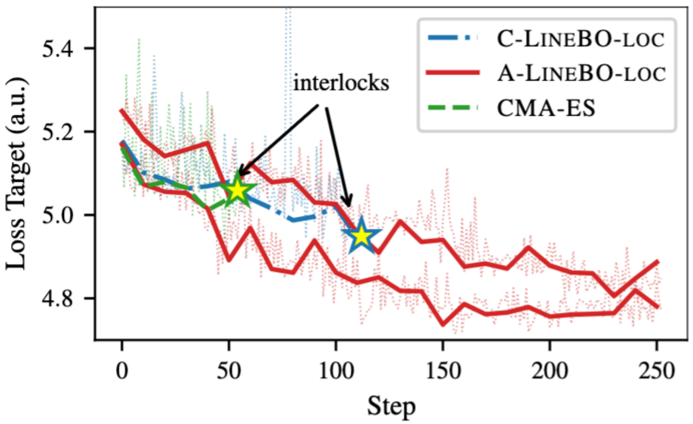
- Low Intensity with manually detuned machine
- 5 parameters

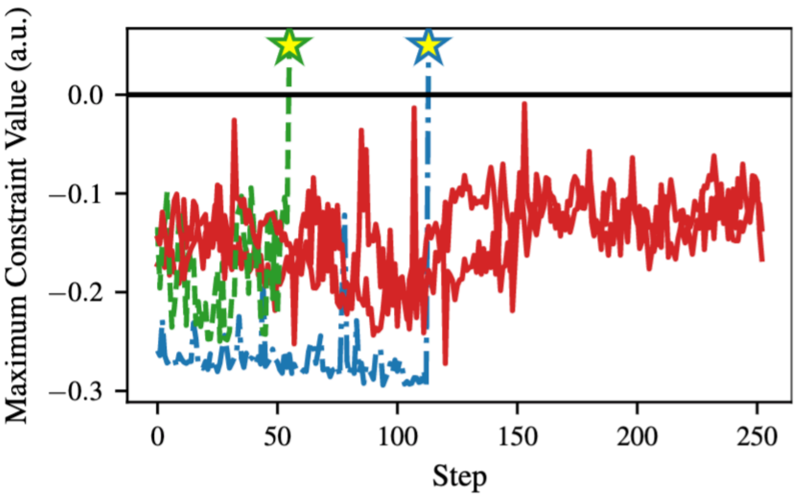


- safe variants competitive
- non-safe methods create interlocks (violate constraints)
  - proves constraints are working

HIPA

- High intensity
- 16 parameters

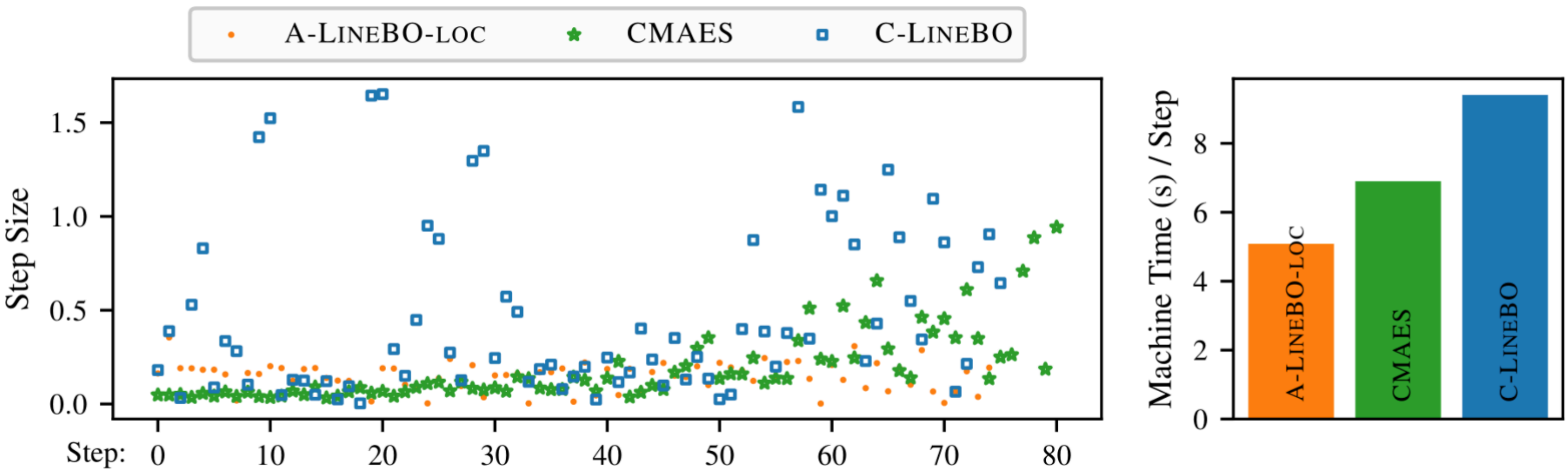




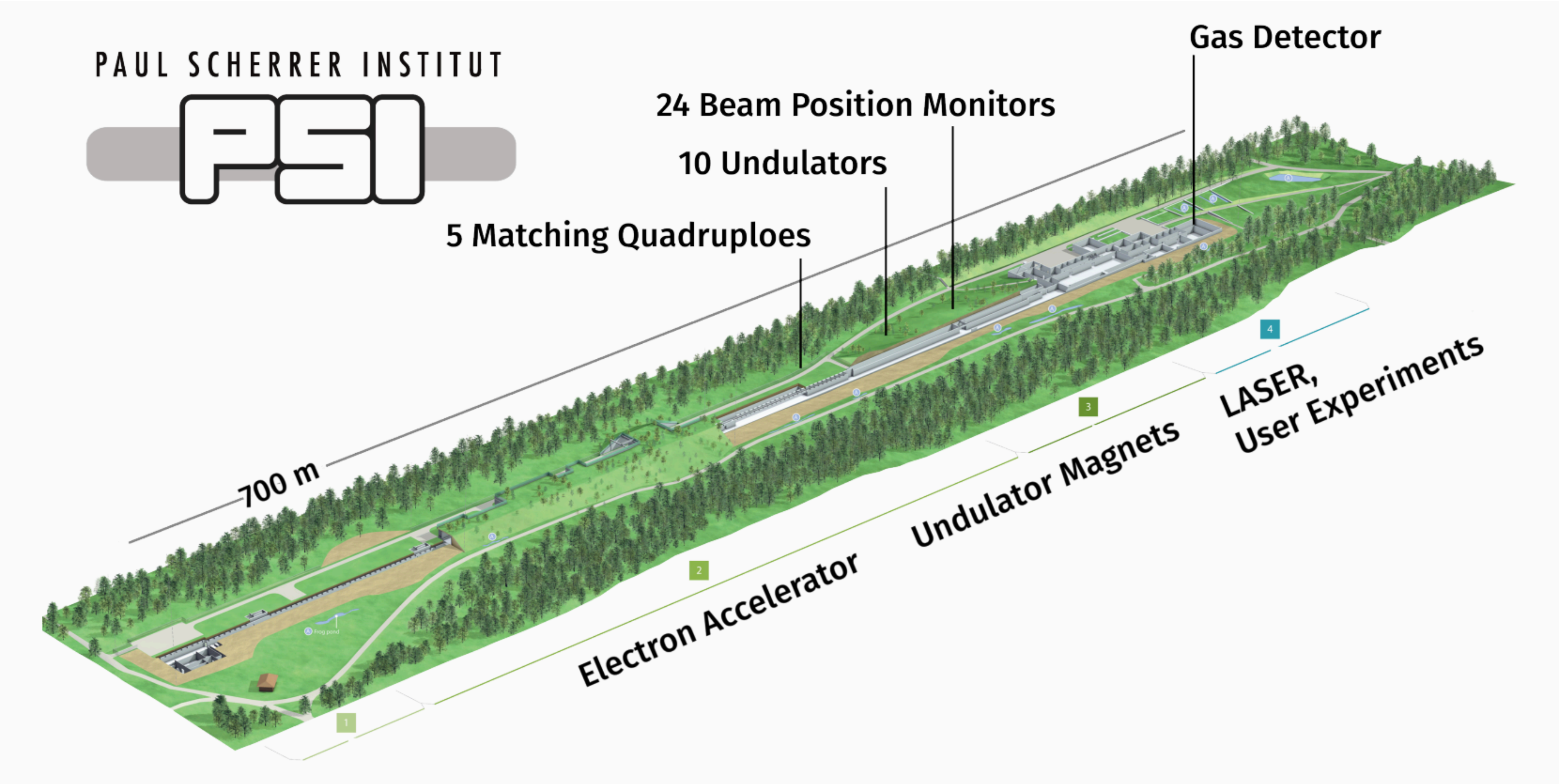
- Also competitive and safe at high current

HIPA

- Localised option (Maximum step size is 10% of the available range)
  - faster to set
    - at HIPA we don't set parameters instantly
  - quicker for feedbacks to stabilise
  - similar solution

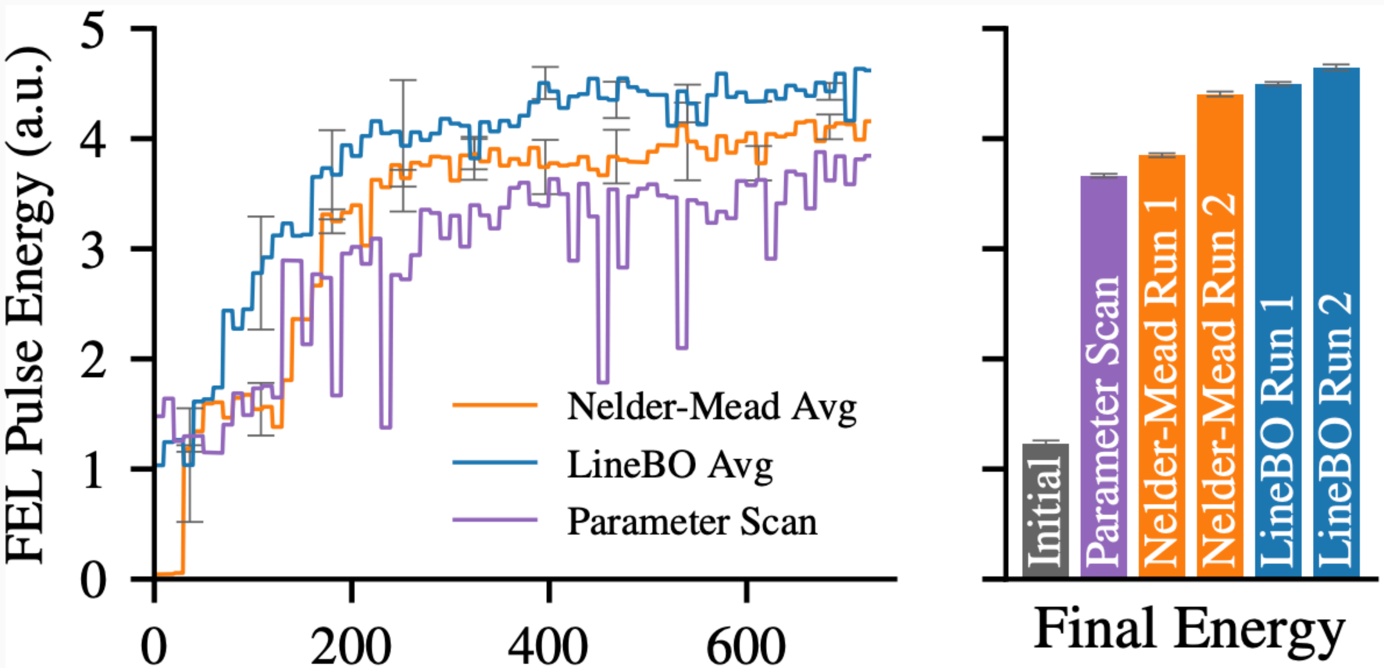


SwissFEL

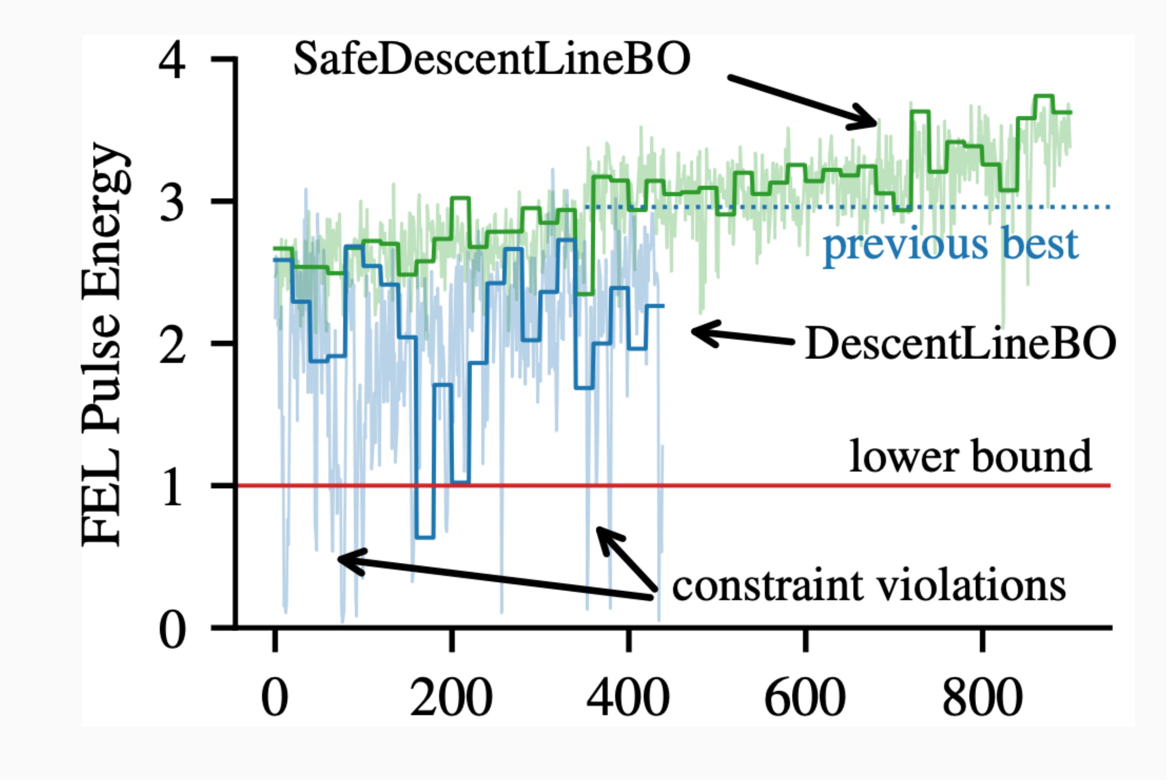


PAUL SCHERRER INSTITUT
PSI
Gas Detector
24 Beam Position Monitors
10 Undulators
5 Matching Quadruploes
700 m
1
2
3
4
Electron Accelerator
Undulator Magnets
LASER,
User Experiments

- Objective: Shot by shot FEL intensity
- Tuning Parameters (39): Quadrupole Magnets, Beam position, Undulator settings
- Constraints:
  - Lower bound on intensity
  - Loss monitors
- Effective control rate: ~ 0.5 seconds / step



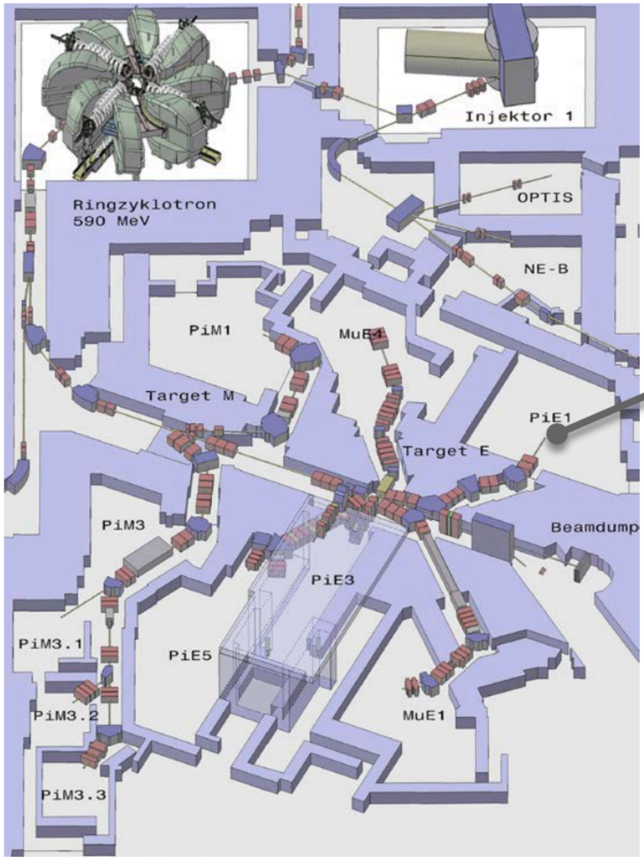
- Bayesian Optimisation shows best result

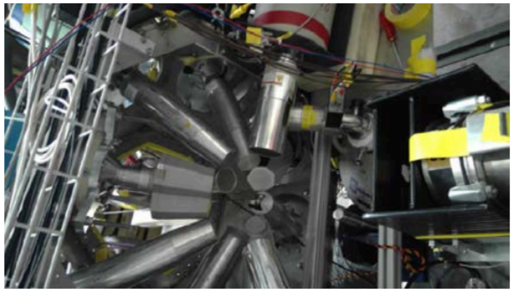


- Constraint on minimum energy improves average and even optimum

MIXE - PiE1

- Elemental composition analysis with negative muons
  - non-destructive and depth sensitive
  - Implant muons at different depths by selecting muon energy





Courtesy Alex Amato

PAUL SCHERRER INSTITUT

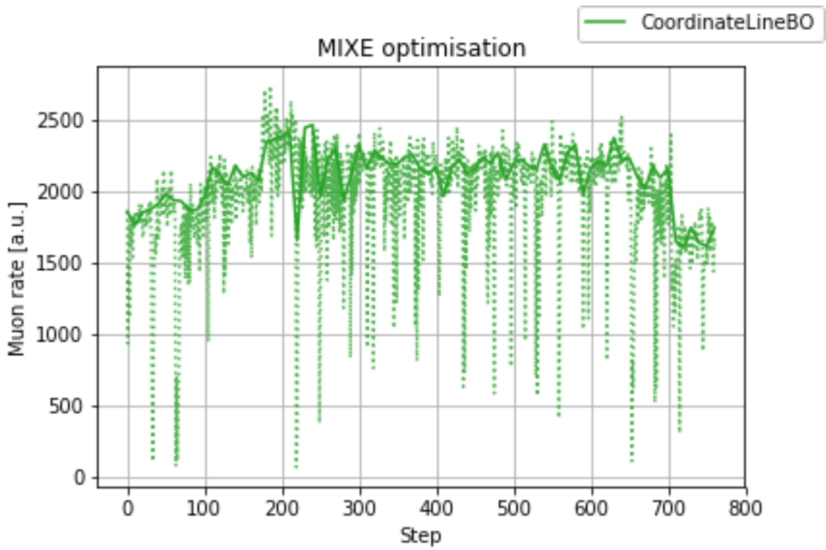
MIXE - PiE1

- probe scanned at several layers
  - different muon momenta about every hour
  - —> tuning manually every hour(!)

- Objective: muon rate (ZPIE1CNT:INP3)
  - average rate —> reduce exploration
- Parameters: 20 quads, magnets and spin rotator (same as in setpoint)
  - 'ASY51:SOL:2', 'ASL51:SOL:2', 'ASK51:SOL:2', 'SPIN2:SOL:2', 'QSN54:SOL:2', 'QSN55:SOL:2', 'QSN56:SOL:2', 'QTH51:SOL:2', 'QTH52:SOL:2', 'QTB51:SOL:2', 'QTB52:SOL:2', 'QSL51:SOL:2', 'QSL52:SOL:2', 'QSL53:SOL:2', 'QSL54:SOL:2', 'QSE51:SOL:2', 'QSE52:SOL:2', 'QSN51:SOL:2', 'QSN52:SOL:2', 'QSN53:SOL:2'
- Constraints: none
  - Could set a minimum muon rate as constraint
- Machine checks
  - Beam on (MHC4 current monitor > 1800)

MIXE - PiE1

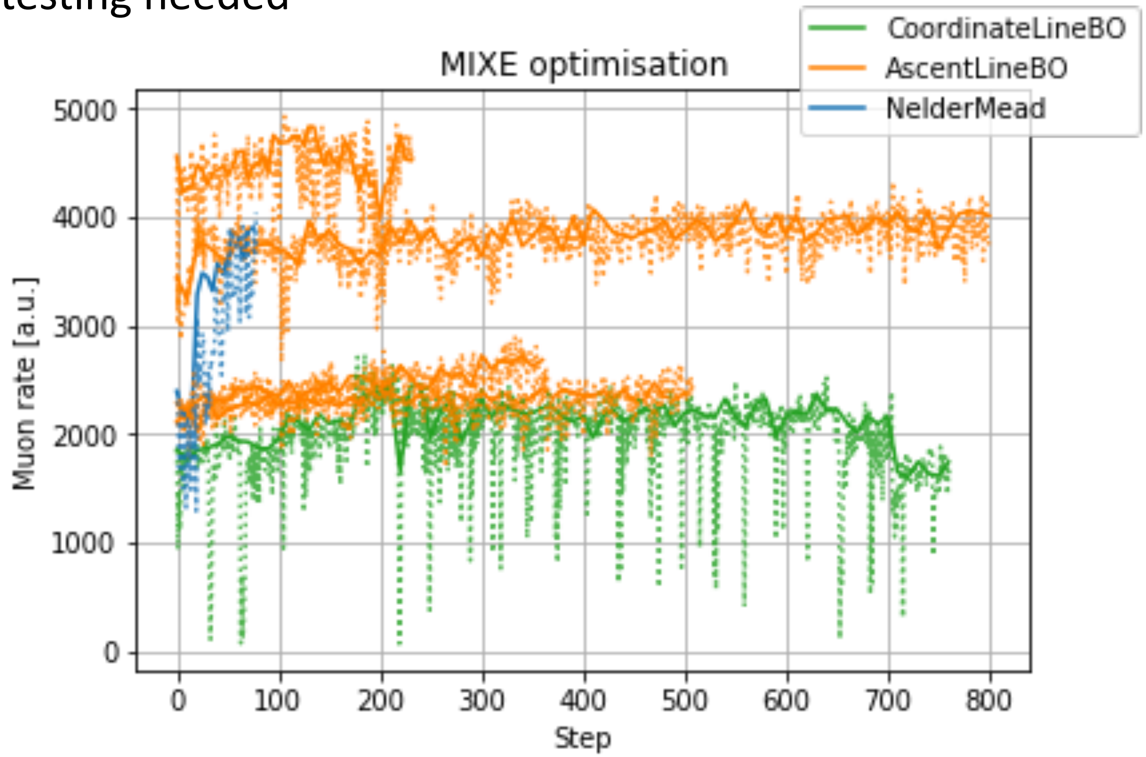
- Tested 13 December 2021 w Lars Gerchow
- First try: CoordinateLine (one parameter line scan)



- Slow noisy progress
  - Too many evaluations per line scan
- Sometimes signal went down
  - Noise parameters might have set wrong
  - Too much exploration?
- CoordinateLine slow with many parameters that are not orthogonal

MIXE - PiE1 - AscentLine

- AscentLine (multiple parameters at once)
  - Steady constant increase, even small increases after one hour
- Note that expected muon rate dependent on muon energy
- NelderMead also promising result
  - More testing needed



PAUL SCHERRER INSTITUT

Practicalities

- Every problem unique
  - Algorithms and parameters should be benchmarked for optimal results
- Results stored
  - Analysis framework available (Jupyter notebook)
- Similar tools exist (OCELOT - DESY, SLAC)
- Possible improvements
  - Installation tricky
  - More automatisation
    - Operator is still needed
    - More robust
  - CPU usage and memory usage large
    - especially for long optimisation runs

Example configuration - PiE1

- api
  - pie1.yaml

```
sfopt.environment:
  tuning_signal_pv: ZPIE1CNT:INP3
  outlier_level: 3.0
  outlier_detection: true
  num_repetitions: 10
sfopt.interface:
  set_max_trials: 600
  set_tolerance: 0.05 # sets tolerance of difference of SOL to IST value
  frequency: 10.0 # machine read/set interval time
sfopt.checks:
  min_value_pvs:
    MHC4:IST:2: 1800.
```

  - algorithm_ascent-linebo.yaml

```
algorithm.gp:
  beta: 1.
  noise_std: 40.0
  kernel: rbf
  scale: 4000.0
  bias: 40.0
  lengthscale: 0.3
algorithm.constraint_gp:
  beta: 1.
  noise_std: 0.1
  kernel: matern52
```

  - algorithm_neldermead.yaml

```
algorithm.nelder_mead:
  contraction_factor: 0.8
  initial_stepsize: 0.1
  restart_threshold: 0.0
  adaptive: True
```

- config
  - channel_config_set.txt

```
pv,lim_low,lim_high,lim_rel,wait_const,wait_linear,wait_min,wait_max,active,step_size
ASY51:SOL:2,,,0.5,,,,,1,
ASL51:SOL:2,,,0.5,,,,,1,
ASK51:SOL:2,,,0.5,,,,,1,
SPIN2:SOL:2,,,10.0,,,,,1,
QSN54:SOL:2,,,2.0,,,,,1,
```

Next steps

- Experiments planned end April (HIPA startup commissioning phase)
  - Systematic algorithm comparison
- MIXE First two weeks of May

PAUL SCHERRER INSTITUT

Conclusions & References

- Ready to use powerful automated tool
  - Very flexible
  - https://gitlab.psi.ch/ext-kirschner_j/swissfel-opt
- Results shown on several beamlines
- More tests in planning at PiE1 - MIXE
- Please contact me if interested

- Papers:
  - Tuning Particle Accelerators with Safety Constraints using Bayesian Optimization - Phys. Rev. Accelerator and Beams (2022)
  - Bayesian optimisation for fast and safe parameter tuning of SwissFEL - FEL 2019
  - Adaptive and Safe Bayesian Optimization in High Dimensions via One-Dimensional Subspaces - ICML 2019
- Presentation:
  - Bayesian Optimization for Safe & Efficient FEL Tuning at SwissFEL & Loss Optimisation at HIPA , N. Hiller, OWLE seminar (2020)

Many thanks to

- Jaime Coello de Portugal (ex-GFA)
- Johannes Kirschner (ETHZ)
- Lars Gerchow (NUM)
- Nicole Hiller (GFA-operation)

CMA-ES Algorithm

- Covariance Matrix Adaptation – Evolution Strategy.
- Similar to swarm optimization.
- Does not have safety constraints, we used a very small step size.