Machine Learning in accelerators operation and design

Elena Fol

CERN, Beams Department
(BE-ABP-LAF)

I.FAST 2nd Annual Meeting, April 20th, Trieste, Italy

Why applying ML to accelerators?

Accelerators

- Operation
- Diagnostics
- Beam Dynamics Modeling

Which limitations can be solved by ML with reasonable effort?

➢ large amount of optimization targets
➢ computationally expensive simulations
➢ direct measurements are not possible
➢ previously unobserved behaviour
➢ non-linear interacting sub-systems, rapidly changing environment.

Why applying ML to accelerators?

Accelerators

- Operation
- Diagnostics
- Beam Dynamics Modeling

Which limitations can be solved by ML with reasonable effort?

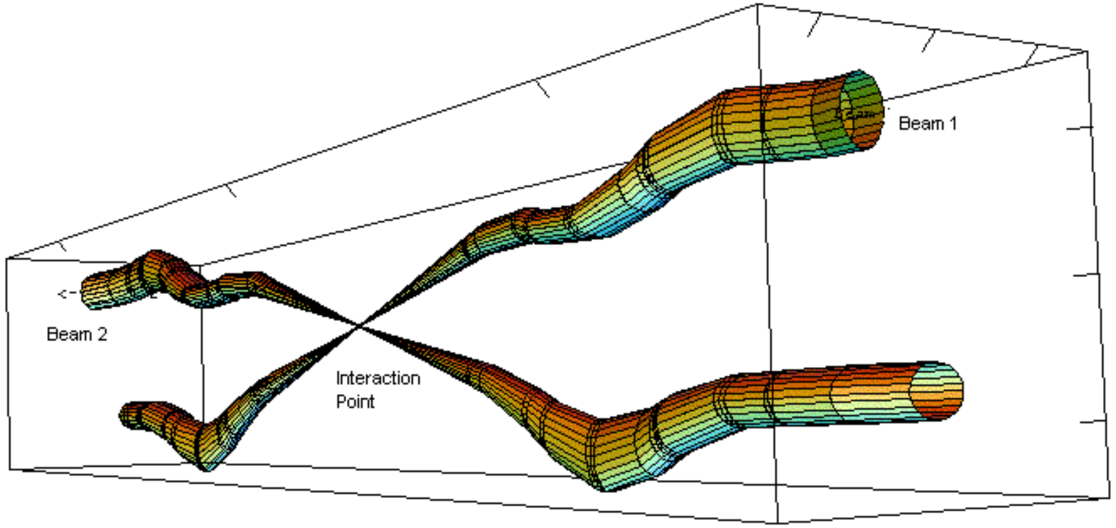
➢ large amount of optimization targets
➢ computationally expensive simulations
➢ direct measurements are not possible
➢ previously unobserved behaviour
➢ non-linear interacting sub-systems, rapidly changing environment.

Machine Learning:

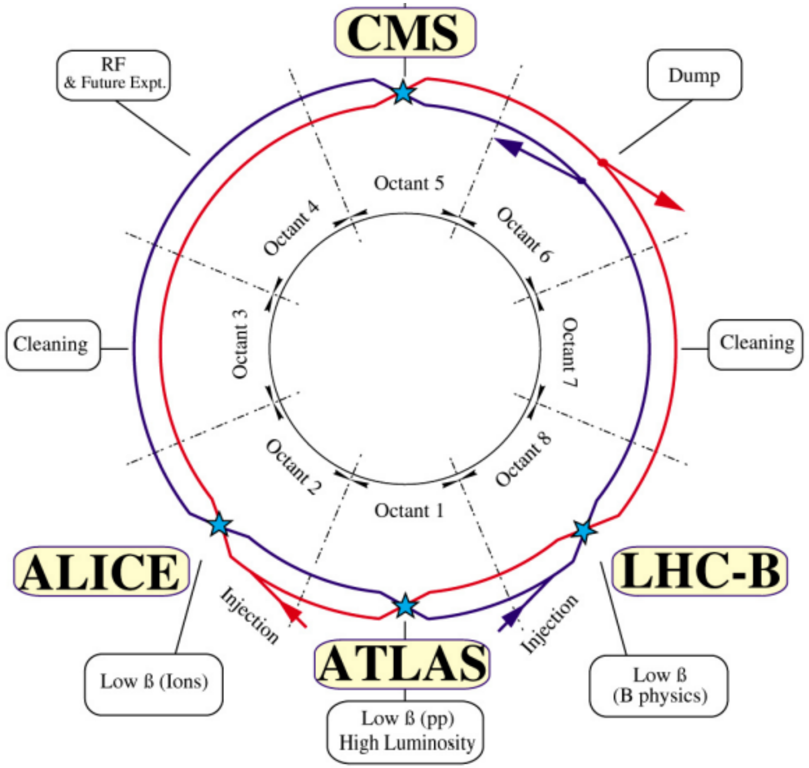
✓ ***Learn arbitrary models***
✓ ***Directly from provided data***

Machine Learning for beam optics control

Beam optics control at the LHC



Relative beam sizes around IP1 (Atlas) in collision



Large Hadron Collider:

- 9300 magnets for bending and focusing the beam.
- Main experiments: ALICE, ATLAS, CMS, LHCb
- Collision rate: sufficient and balanced between experiments —> **Luminosity**

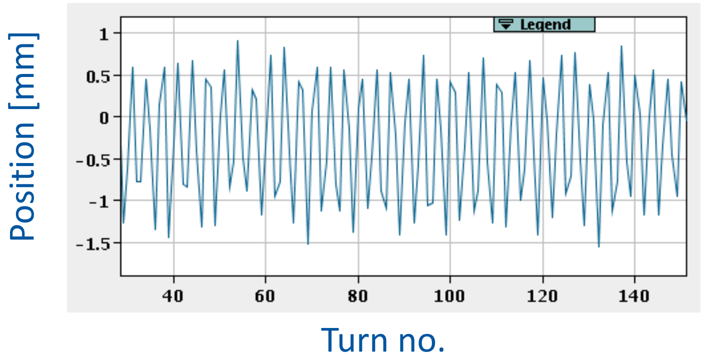
➢ How to increase chances of collisions?

➢ How to ensure machine protection?

→ **Beam Optics control**

Measuring the optics: instrumentation faults detection

Turn-by-turn beam position

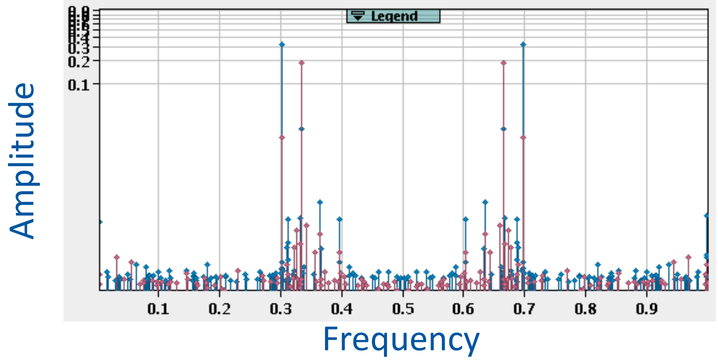


- Excite the beam to perform transverse oscillations.

→ **Beam Position Monitors (BPMs) to measure the beam centroid turn-by-turn**

Denoising (SVD)
Signal cuts

Spectrum

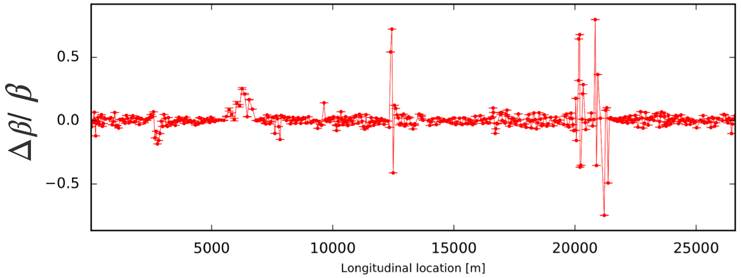


- Harmonic analysis using Fast Fourier Transform (FFT)

Semi-automatic and manual cleaning of outliers

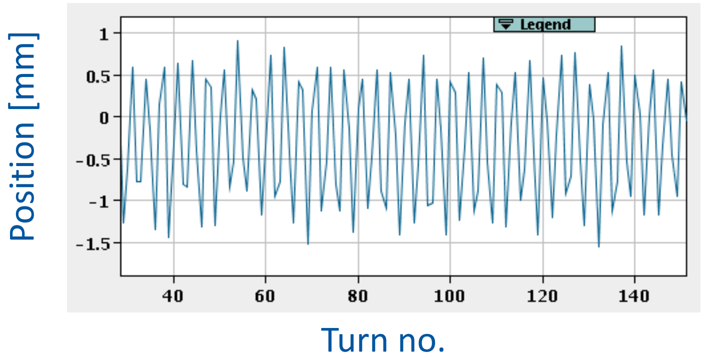
Optics

(beta-beating and other optics functions)



Measuring the optics: instrumentation faults detection

Turn-by-turn beam position

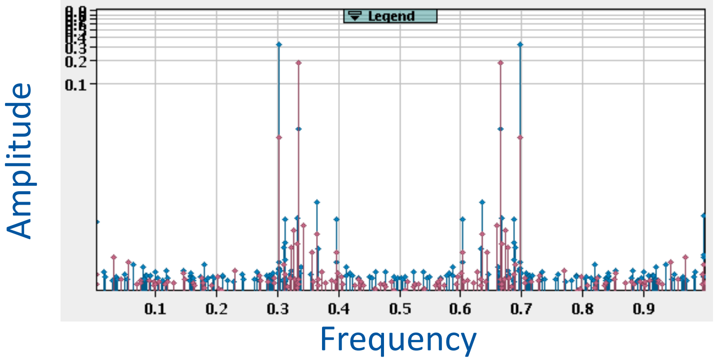


- Excite the beam to perform transverse oscillations.

→ **Beam Position Monitors (BPMs) to measure the beam centroid turn-by-turn**

Denoising (SVD)
Signal cuts

Spectrum

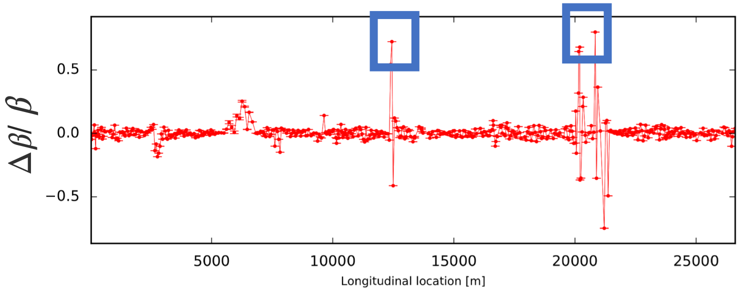


- Harmonic analysis using Fast Fourier Transform (FFT)

Semi-automatic and manual cleaning of outliers

Optics

(beta-beating and other optics functions)



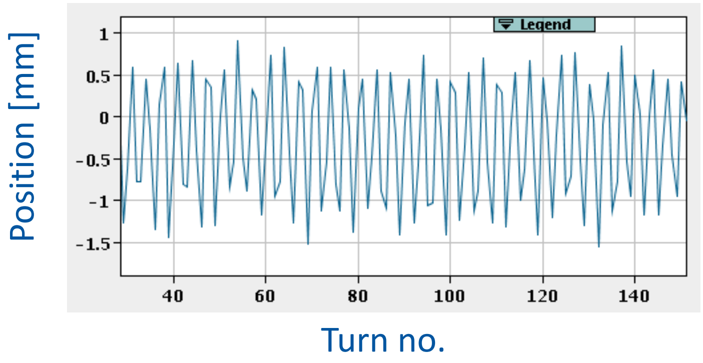
Unphysical values still can be observed

Faulty BPMs **are a-priori unknown**:

- cause erroneous computation of optics functions
- manual cleaning is required
- repeating optics analysis after manual cleaning

Measuring the optics: instrumentation faults detection

Turn-by-turn beam position

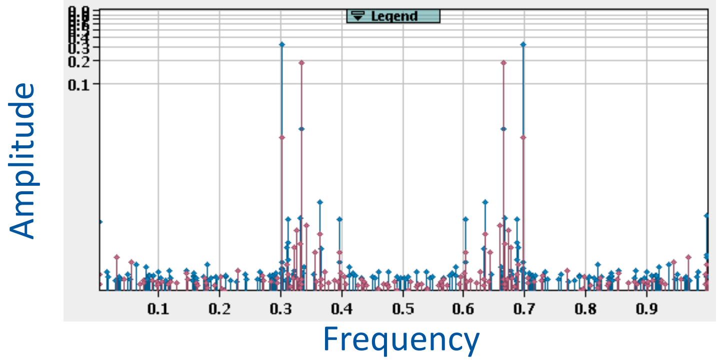


- Excite the beam to perform transverse oscillations.

→ **Beam Position Monitors (BPMs) to measure the beam centroid turn-by-turn**

Denoising (SVD)
Signal cuts

Spectrum

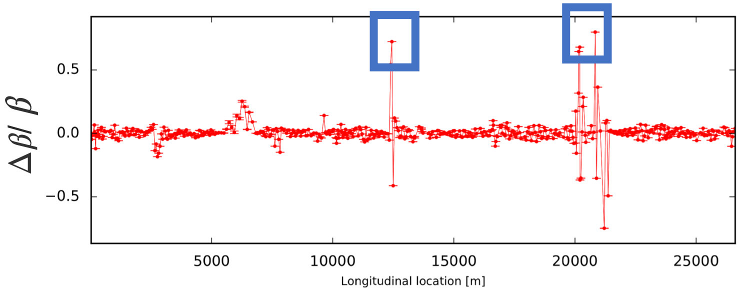


- Harmonic analysis using Fast Fourier Transform (FFT)

Semi-automatic and manual cleaning of outliers

Optics

(beta-beating and other optics functions)



Unphysical values still can be observed

Faulty BPMs **are a-priori unknown**:

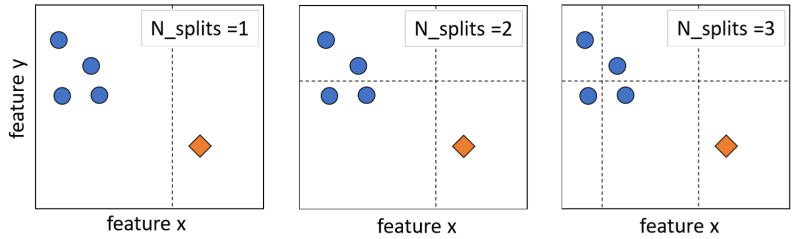
- cause erroneous computation of optics functions
- manual cleaning is required
- repeating optics analysis after manual cleaning

Anomaly detection using Unsupervised Learning

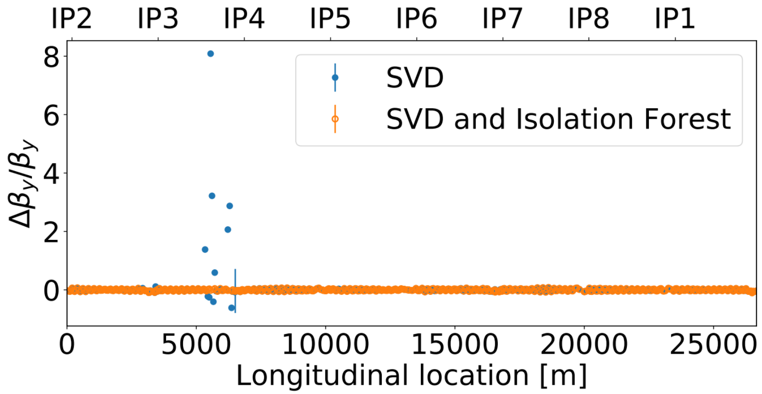
CERN

Instrumentation faults detection: operational results

➡ **Detect anomalies with Isolation Forest algorithm:**



- **Unsupervised Learning**
- No training, **immediate results** applying Isolation Forest algorithms on FFT properties of BPM signal



✓ Successfully used in **LHC beam measurements since 2018**

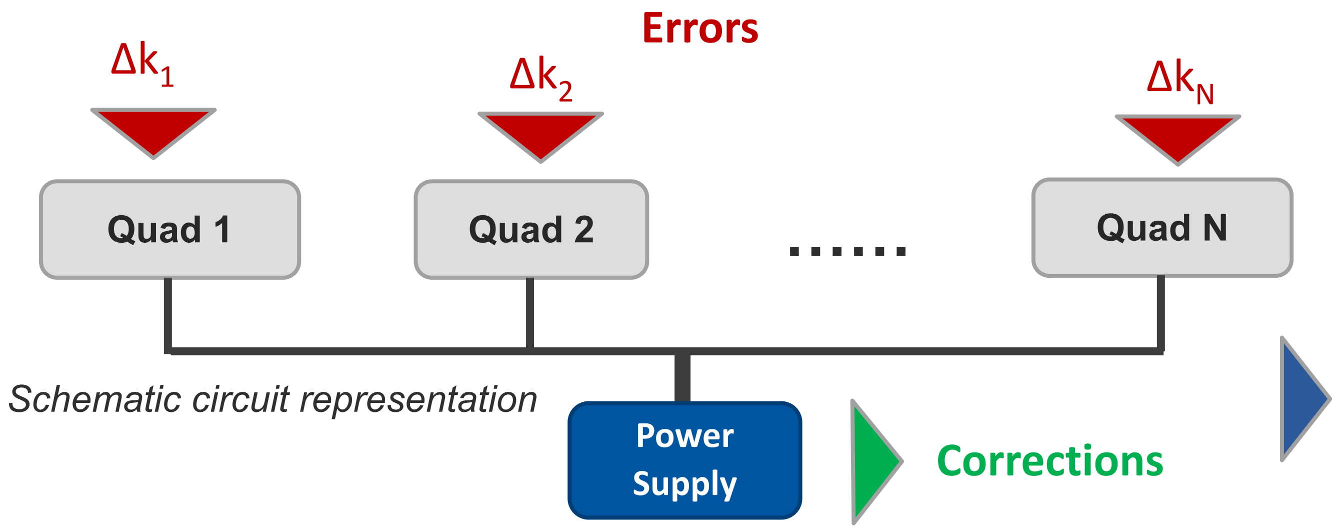
➢ **Providing information to BI experts:**

✓ IF- algorithm: Identify **dominant signal properties for faults classification**

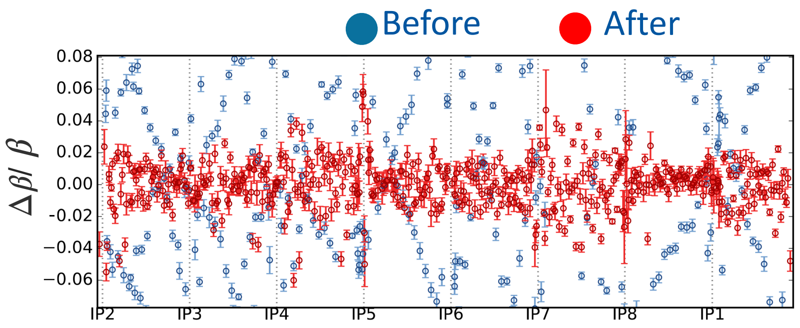
✓ Identified **116 critical faulty BPMs out of more than a thousand BPMs** in the LHC.

Thanks to ML: ➡ **Detection of otherwise unexplored hardware and electronics problems in beam instrumentation**

Optics corrections in the LHC using Supervised Learning

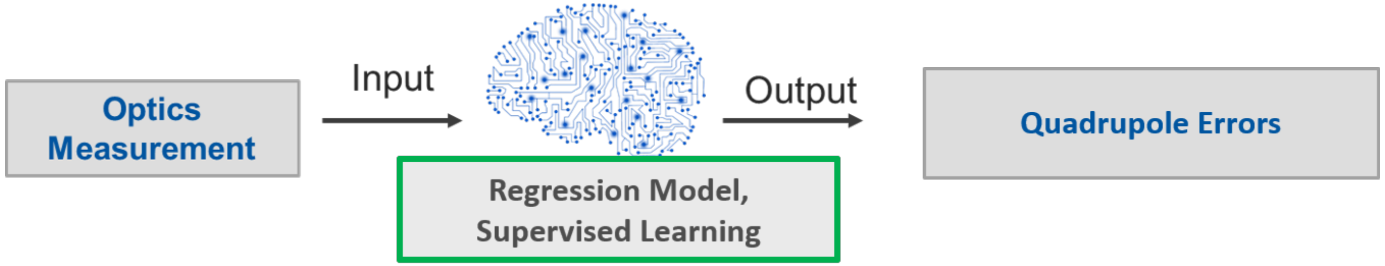


β → Determined by **quadrupole arrangement and powering:** $\dfrac{\Delta\beta}{\beta} = \dfrac{\beta_{meas} - \beta_{model}}{\beta_{model}}$



- Access to the magnets for direct measurements is not possible during operation.
  - ➡ Beam-based measurements and corrections of lattice imperfections.
- Computed corrections provide **circuit settings to compensate measured beta-beating**
  - ➡ What are the **actual individual magnet errors?**

Optics corrections: prediction of magnets errors

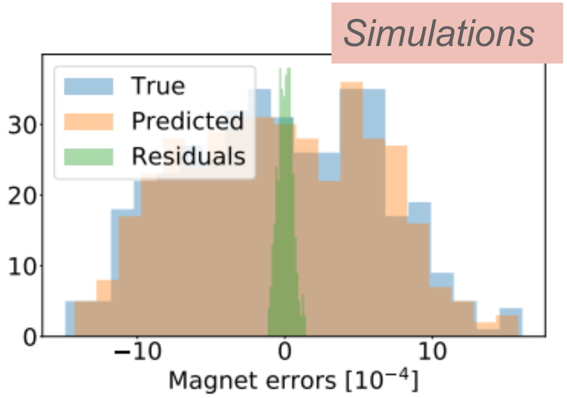


<u>Random Forest Regressor:</u>

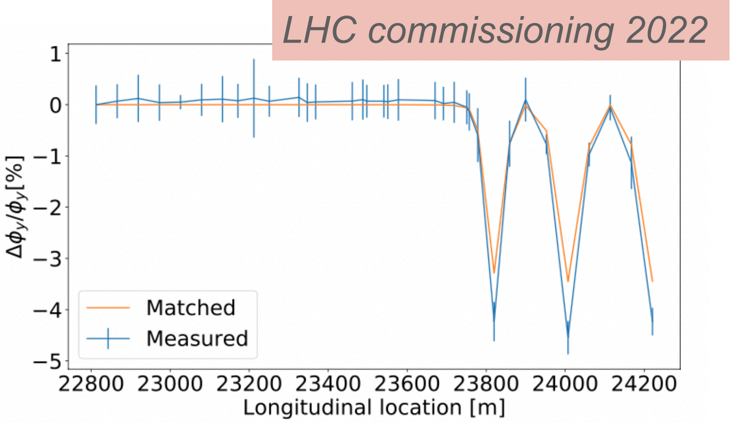
- Ensemble of decision trees: lower complexity vs. NN
- **1256 target** variables, **2048 input** variables
- Tested on simulations, historical data and **LHC commissioning**

Optics corrections: prediction of magnets errors

Optics Measurement → Input → **Regression Model, Supervised Learning** → Output → **Quadrupole Errors**



True magnet errors are known in the simulations:
=> direct comparison true vs. Predicted



Measurements: beam-based verification:
Measured phase errors vs. Matching with predicted magnet errors

Random Forest Regressor:

- Ensemble of decision trees: lower complexity vs. NN
- **1256 target** variables, **2048 input** variables
- Tested on simulations, historical data and **LHC commissioning**

✓ Corrections by **applying the predicted magnet errors** with opposite sign as correction settings

✓ Simultaneous local correction in all Interaction Regions **within seconds.**

→ Potential to **save operational time**

E. Fol et al. The European Physical Journal Plus volume 136, Article number: 365 (2021) , "Supervised learning-based reconstruction of magnet errors in circular accelerators".

E.Fol et al.,"Experimental Demonstration of Machine Learning Application in LHC optics commissioning",IPAC'22 MOPOPT047

Beam optics corrections: Reinforcement Learning approach

The **High Luminosity Large Hadron Collider (HL-LHC)**: upgrade of the LHC

- aims to achieve luminosities a factor of 5 to 7.5 larger than the LHC
- enabling the experiments to enlarge the data volume by one order of magnitude

➡ **Precise optics control required:** traditional techniques show limitations in challenging HL-LHC optics control (simulations)

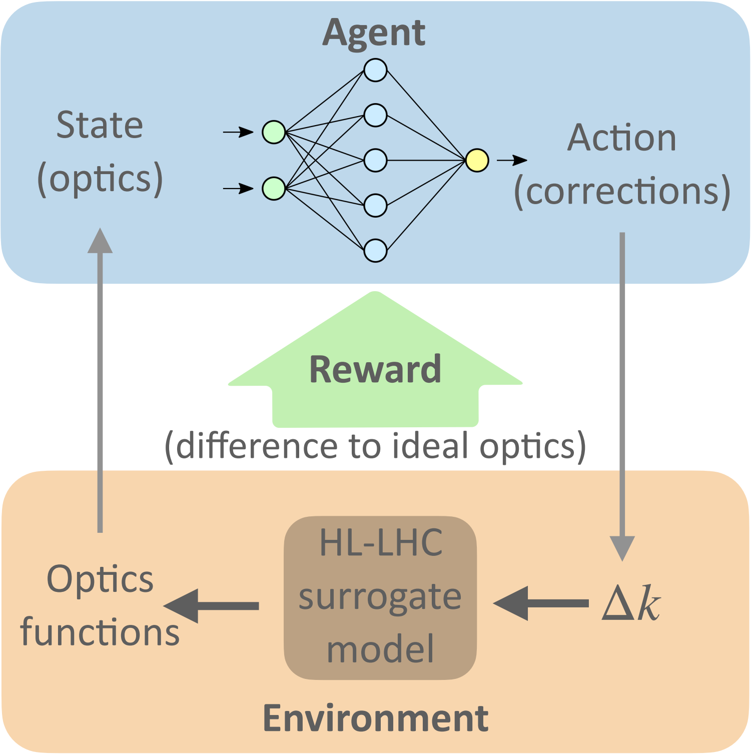
➡ **Reinforcement learning based optics control:**

- ✓ Concept and prototype of **RL formalism for optics control**
- ✓ Implementation of **training data generation**
- ✓ Built **surrogate model** to be used as **environment**

➡ **Preliminary results:**

Demonstration of order of magnitude improved optics corrections in HL-LHC simulations.

optimise NN to minimise optics errors



Random Forest regression trained on simulations

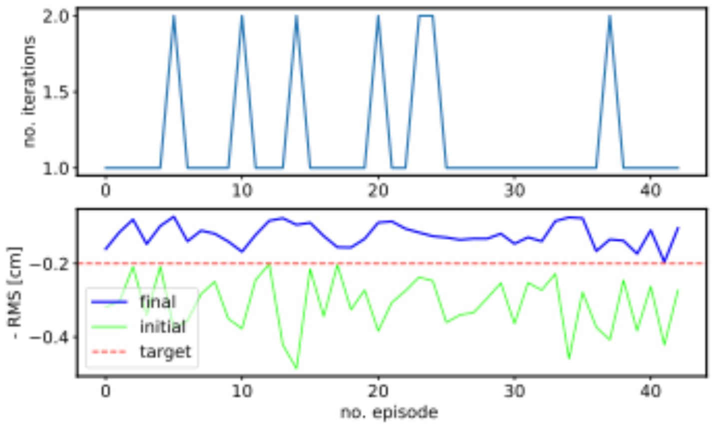
(E.Fol, H. Garcia-Morales, BE-ABP)

Further examples from CERN

Reinforcement Learning in Controls:

- RL agents were trained for **trajectory correction on the AWAKE electron line**
- Offline training on simulations, **short online re-training** of the agent in operation

➡ Trajectory correction results better than the set target



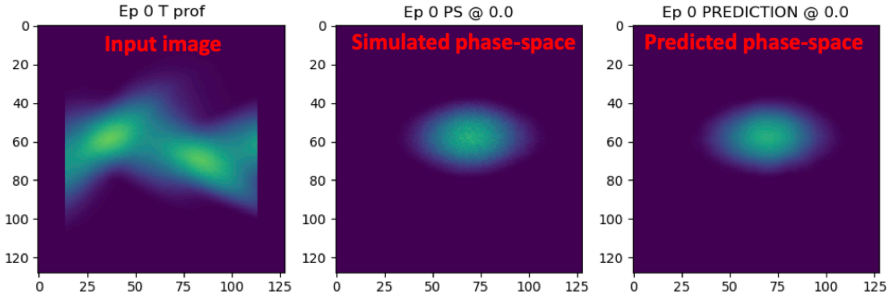
"Sample-efficient reinforcement learning for CERN accelerator control"
V. Kain et al., Phys. Rev. Accel. Beams **23**, 124801

Image-based beam diagnostics:

- Reproduce the **longitudinal beam parameters** at the LHC injection, **given as an input the longitudinal beam profiles**

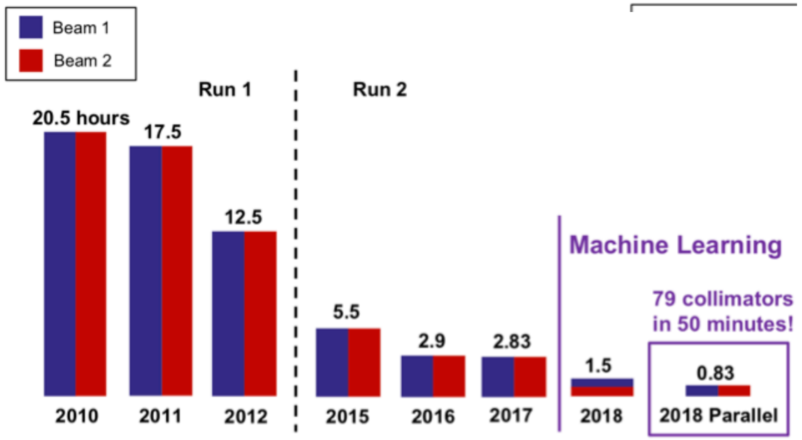
✓ Can be used as **online monitoring tool during operation**

✓ Potential to **reduce time** for the data analysis needed **to extract the same information.**



"Artificial Intelligence-Assisted Beam Distribution Imaging Using a Single Multimode Fiber at CERN", Trad, Georges , Burger, Stephane
JACoW IPAC 2022 (2022) 339-342

Supervised Learning: Collimators alignment



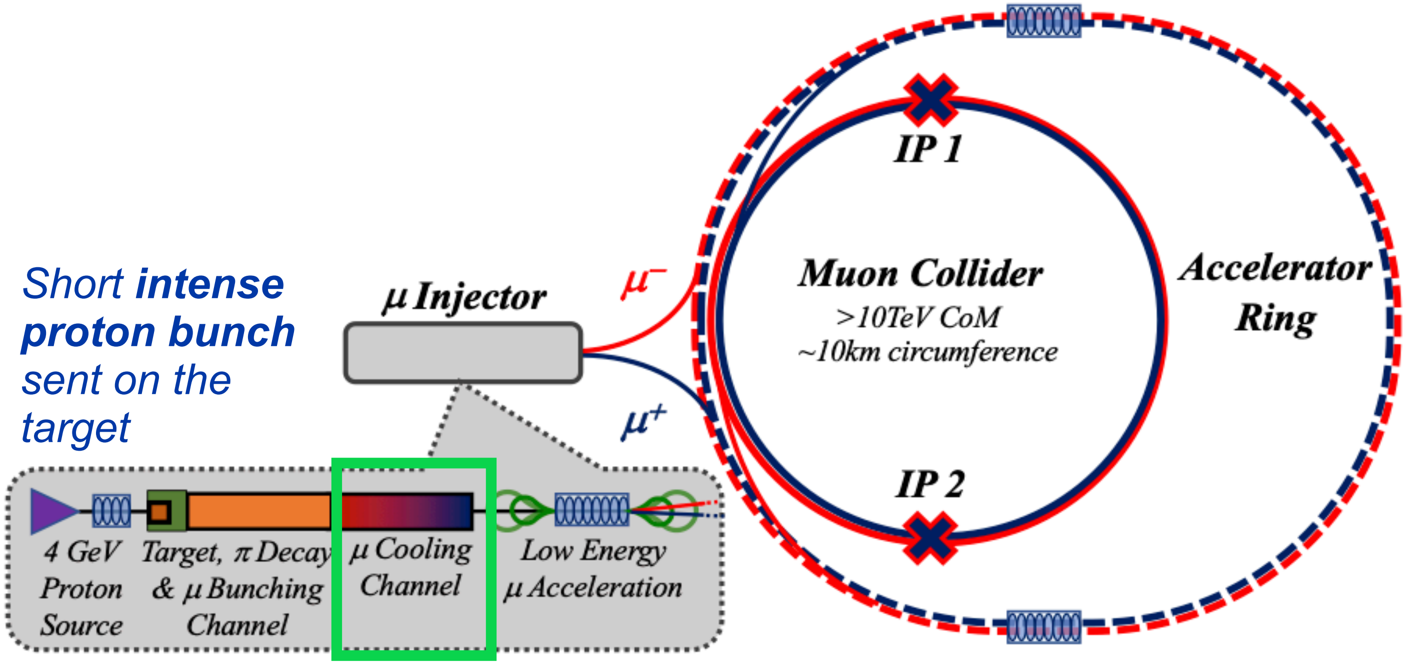
- **Collimators have to be realigned** during operation due to orbit shifts and beam parameter changes
- **Order of magnitude speed up** of collimators alignment, reduction of manual effort

"Operational results on the fully automatic LHC collimator alignment"
G. Azzopardi, Phys. Rev. Accel. Beams **22**, 093001

Machine Learning in design of new facilities:
Muon Collider design study

Muon Collider: overview

Short **intense proton bunch** sent on the target



Interaction with the target produces pions
➡ **decay into muons**

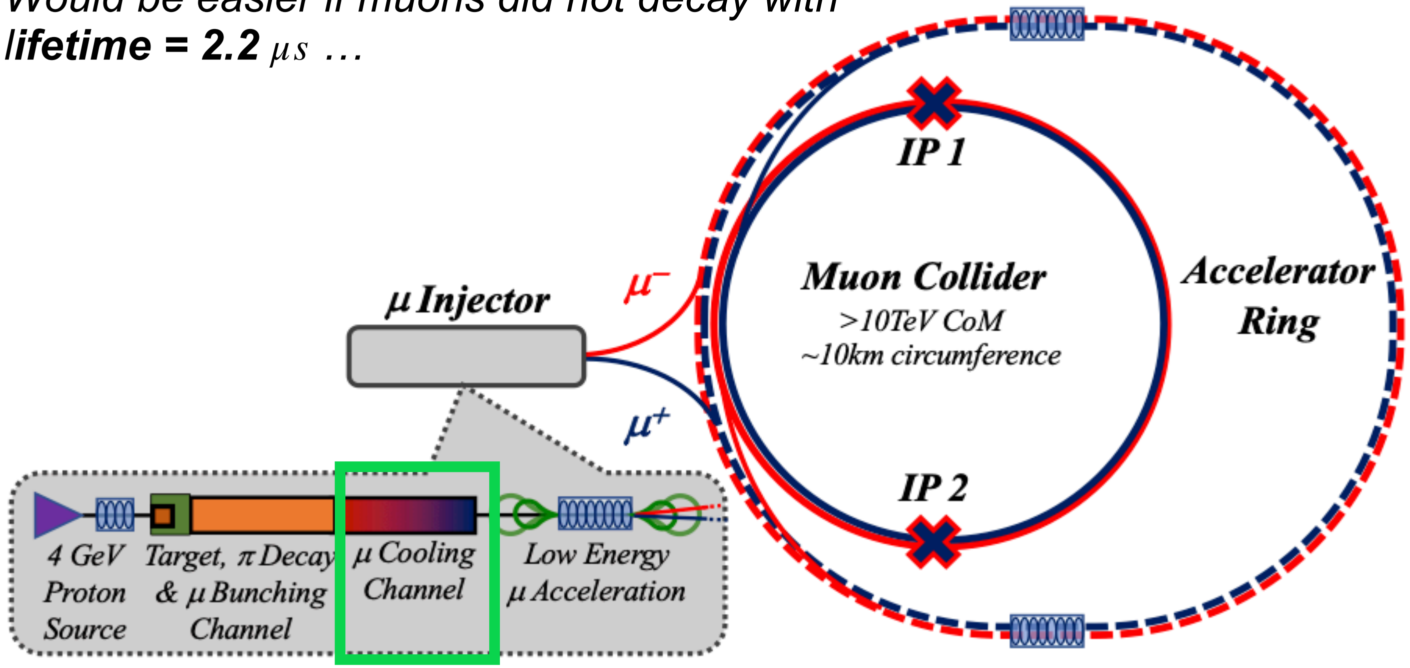
Muons are captured and **cooled to lower emittance**

Acceleration to **high collision energy**

[1]: https://muoncollider.web.cern.ch

Muon Collider: overview

Would be easier if muons did not decay with ***lifetime = 2.2*** μs …

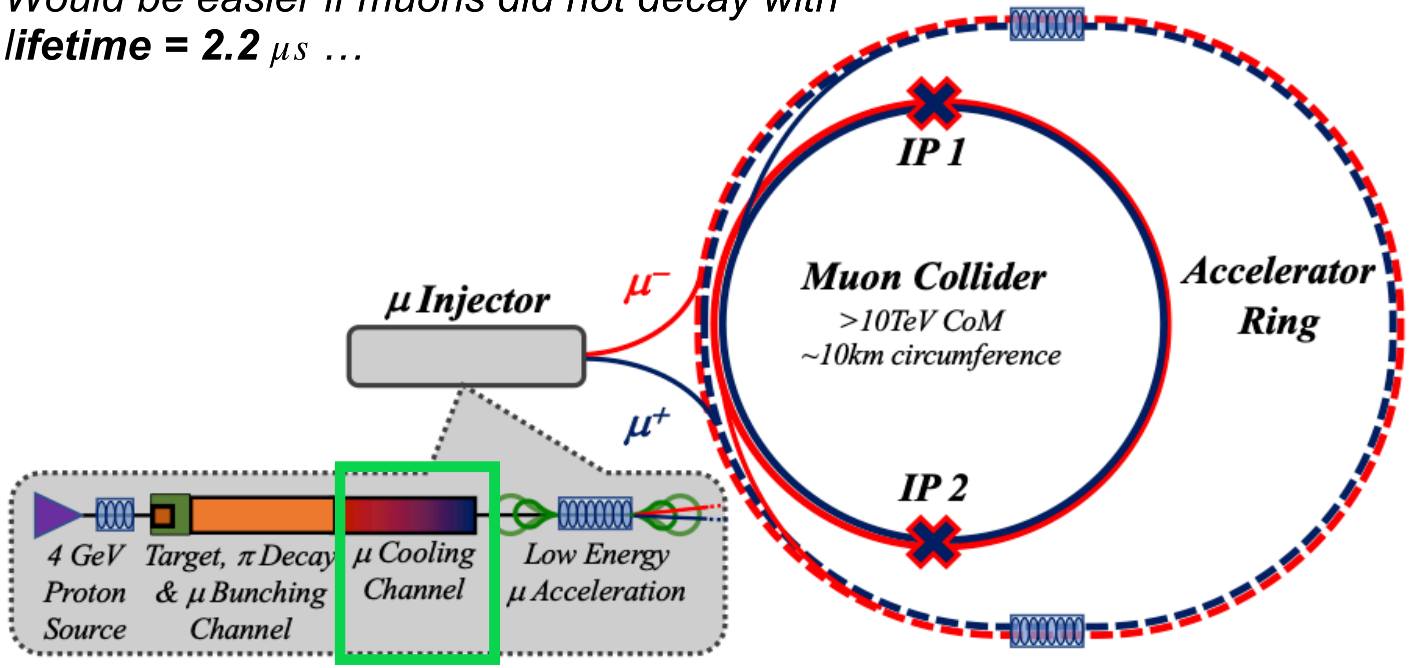


Muons are created as pions decay products and form a beam with a **huge emittance (~beam size)**

- Ionisation cooling (the reduction of occupied phase-space by muons) is required**:** novel technique, demonstrated by MICE collaboration
- Design of cooling channel: numerical optimization, particles tracking simulations

Muon Collider: overview

Would be easier if muons did not decay with ***lifetime = 2.2*** μs …



Muons are created as pions decay products and form a beam with a **huge emittance (~beam size)**

- **Ionisation cooling** (the reduction of occupied phase-space by muons) is required**:** novel technique, demonstrated by MICE collaboration
- Design of cooling channel: numerical optimization, particles tracking simulations

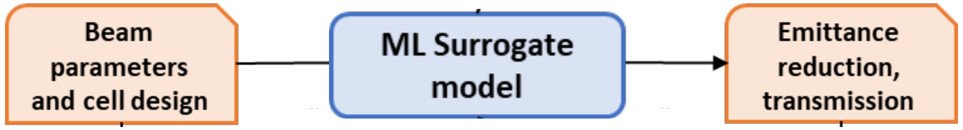
- How to **speed up** tracking simulations?
- How to **estimate initial optimization parameters**?
- Analytical **models** combined with **data-driven** approach

➡ Surrogate models
➡ Bayesian Optimization
➡ Feature Importance Analysis with Decision Trees

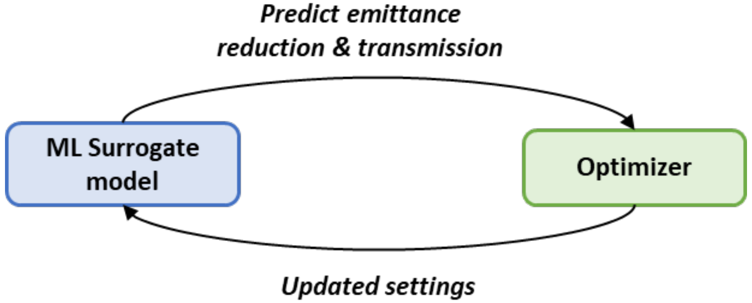
Optimization speed-up: supervised learning

1. Run traditional numerical optimisers, systematically saving the data

2. Train a surrogate model (Random Forest Regressor): predict parameters of interest



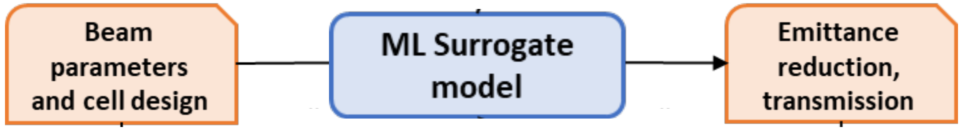
3. Replace time-costly simulations with ML model



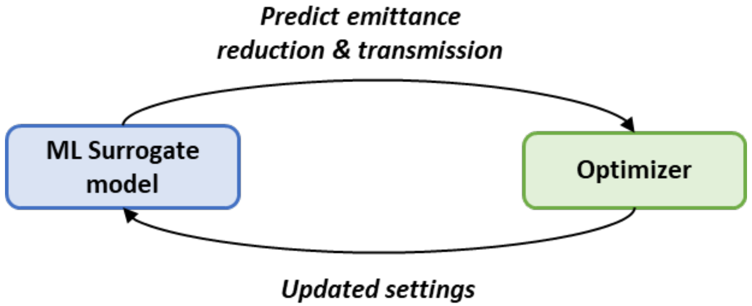
Optimization speed-up: supervised learning

1. Run traditional numerical optimisers, systematically saving the data

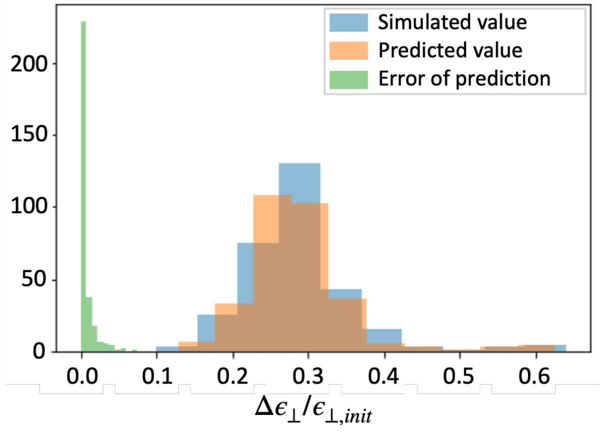
2. Train a surrogate model (Random Forest Regressor): predict parameters of interest

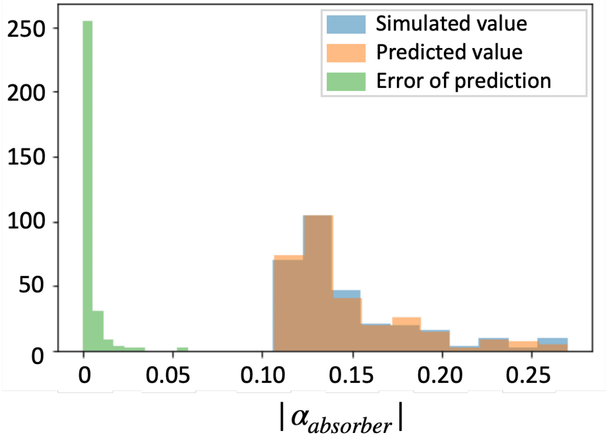


3. Replace time-costly simulations with ML model



Predicting beam properties included in optics optimisation:





- ✓ Compute optimization function from ML-model prediction
- ✓ Optimization in a **few minutes instead of ~1.5 hours** for 200 steps using tracking simulations

* work with D. Schulte, C. Rogers, B. Stechauner

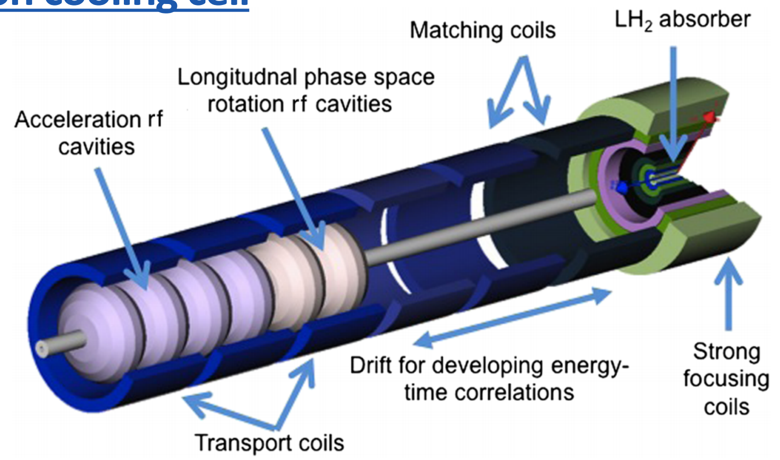
Inverse models for fast design parameters estimate

Desired cooling performance → **ML Surrogate model** → **Required beam parameters and cell design**

Result is as required?

➡ Fast design estimate

Design of a muon cooling cell



- 25 parameters to optimize in each cell
- Expected to need ~15 cells

Inverse models for fast design parameters estimate

Desired cooling performance → ML Surrogate model → Required beam parameters and cell design

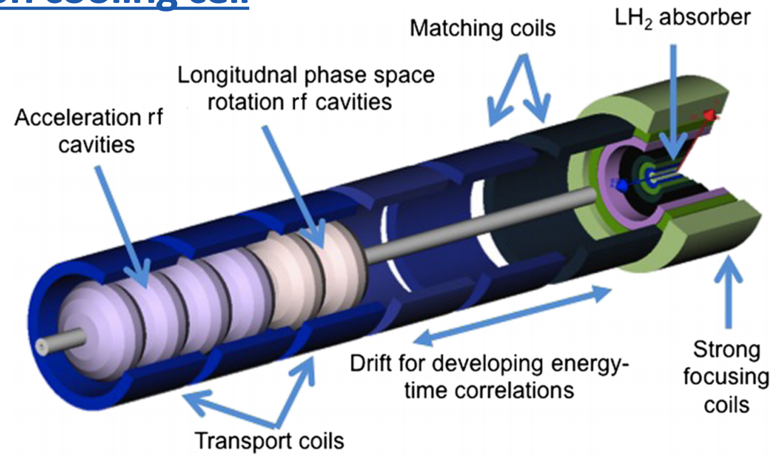
Result is as required? ➡ Fast design estimate

1. A few optimization steps using e.g. genetic algorithms
2. Supervised Learning - based Random Forest model to predict cell design based from required cooling performance

Cell	Emittance Tr. [mm	Emittance Long. [mm]	Bunch length	Pz [MeV/c]	Pz spread
	Beam parameters (end of the cell)				
	300.0	1.5	50.0	135	3.5
1	295	1.7	79	125	3.6
2	283	2.2	61	118	4.6
3	270	2.3	128	105	2.4
4	**255**	**4.8**	**210**	**95**	**4.1**
	260	**16.5**	**715**	**93**	**7.1**

Vs. Results obtained by traditional optimization approach

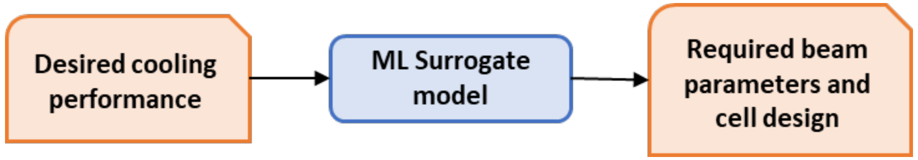
Design of a muon cooling cell



- 25 parameters to optimize in each cell
- Expected to need ~15 cells

✓ Better trade-off between longitudinal and transverse emittance
✓ Flexible automatic optimization framework

Combining optimisers and surrogate models



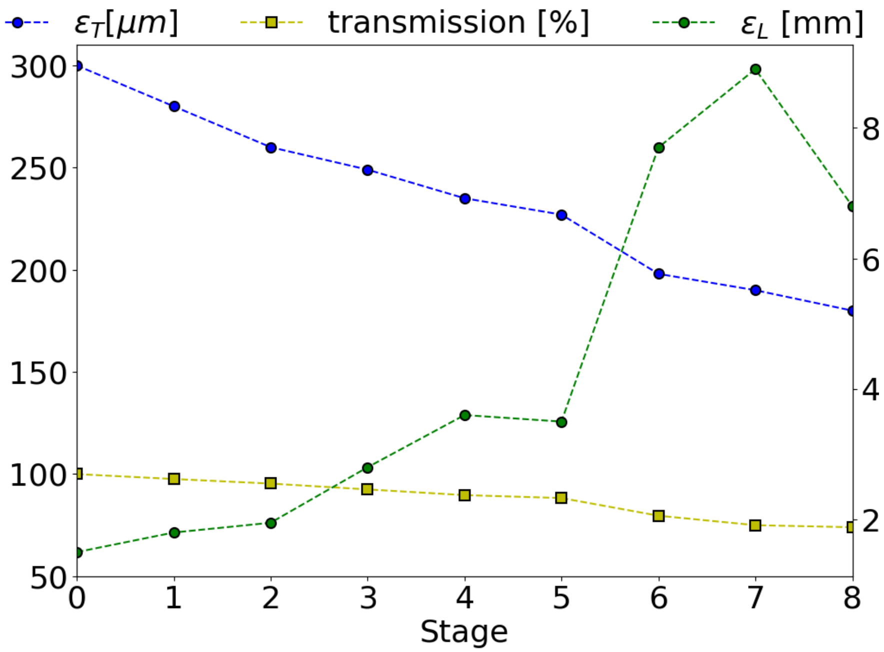
Result is as required? ➡ Fast design estimate

Further optimization needed? ➡ **Use as initial guess** for optimisation algorithms (optimal solution is found within fewer steps)

Applied optimisation techniques:

- **Bayesian optimization** using Gaussian Process*
- Inverse **surrogate models for initial parameters estimation.**

- * **Update probabilistic model** based on function evaluation
- Optimise an acquisition function (e.g. probability of improvement) for sampling the new optimisation step
- Balance exploration and exploitation by controlling parameters of acquisition function



***Note**: previous optimisation achieved **255** mm mrad after **4th stage**, here: **230** mm mrad*

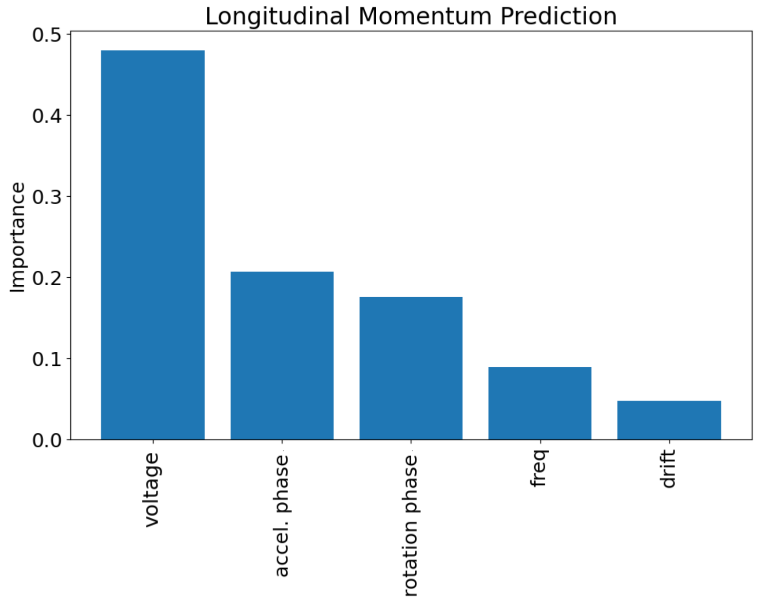
Model interpretability: permutation features importance

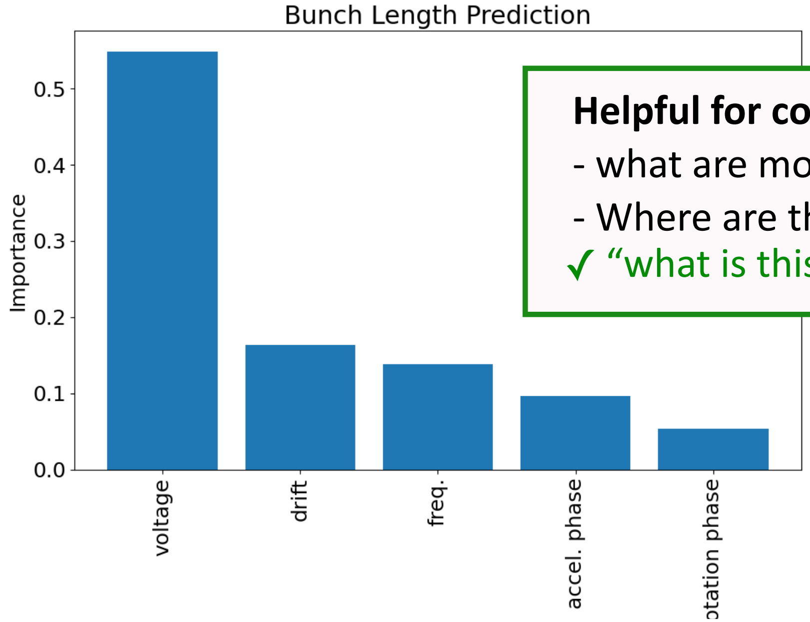
Feature permutation

- Measuring how much **model's performance decreases** when each **feature is randomly shuffled**
- Identify **which features have greatest impact** on model's output
- Applying **Random Forest** algorithm: **automatically computed while training** each tree on a subset of features and minimising the loss function

Example: longitudinal beam parameters:

- Collect data during optimisation,ML-model: Input: **cell set up**, output: **beam parameters** at the end of a cooling cell





Helpful for complex models:

- what are most critical parameters to be optimised?
- Where are the bottle necks?

✓ "what is this model actually learning?"

Further potential ML applications in Muon Collider Design

Sample-efficient optimization:

- Classify a small number of simulation setups based on tracking results
- Find a boundary for desired cooling performance
- Run optimization exploring parameter space within this boundary
- Demonstrated e.g. Dynamic Aperture optimization for HL-LHC using Support Vector Machine Classifier
  (*F.F. Van der Veken, et al., "Determination of the Phase-Space stability border with ML", IPAC'22*)

Emittance computation for non-gaussian beams:

- Tackles limitations of traditional n-"sigma" threshold cuts
- Detecting "lost" particles based on the whole 6D phase space
- e.g. density-used clustering methods: unsupervised learning, fast executable

Integrated model of muon collider complex:

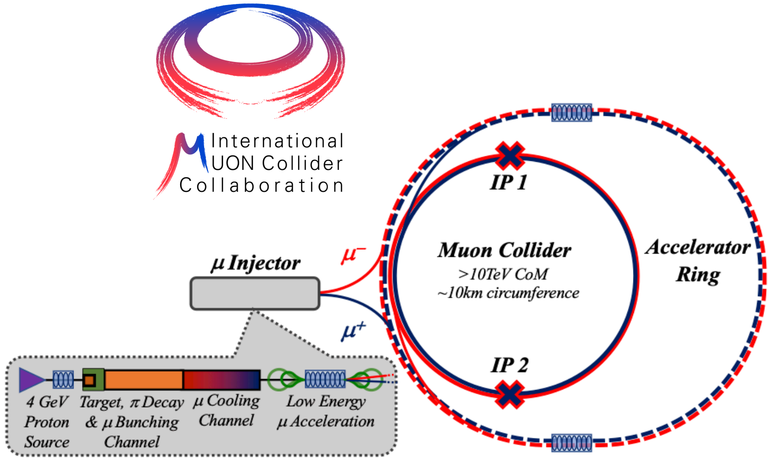
- Optimization routines is a typical instrument across different collider sub-systems
- Systematically saving the data

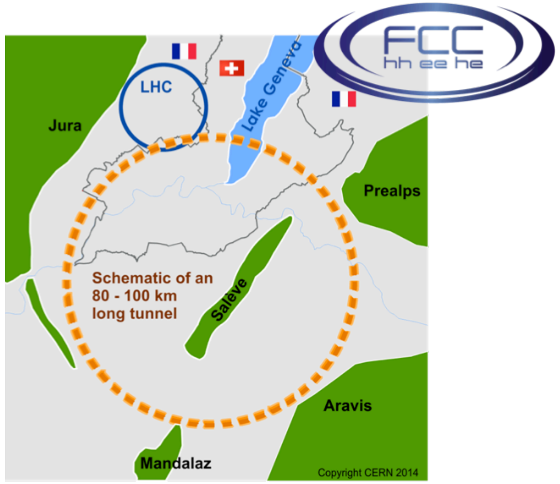
  ➡ *Collecting data from otherwise non-compatible simulations tools*

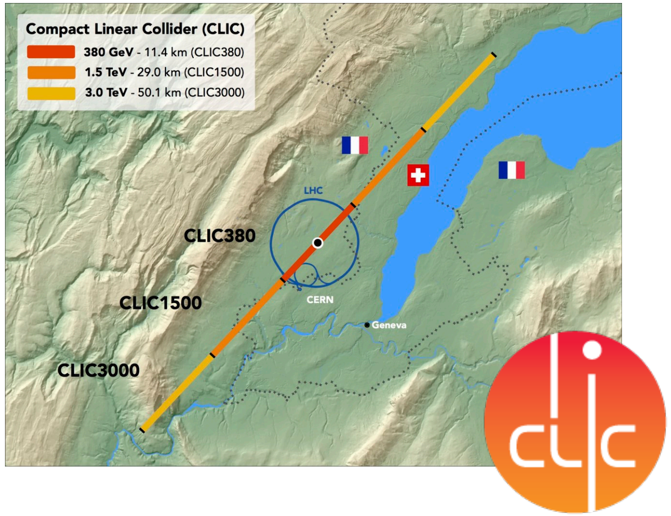
✓ **Opens several opportunities:** identification of most critical parameters for collider performance (e.g. feature importance analysis, but also dimensionality reduction techniques)

✓ **Fast-executable model** for changing requirements as design evolves

Potential ML applications in Collider Design







Several collider projects are considered for the future:

- Large scale facilities: **thousands of parameters to be optimised**, interacting sub-systems
- New simulation tools required to **model complex physical processes**
- Tight tolerances for **beam control** (beam focusing, losses, lifetime)
- **Cost**-effectiveness
- **Energy** efficiency

→ **AI can be a crucial component of design studies to push towards optimal solutions**

Thanks a lot for your attention!

ML in accelerators: summary

Accelerator Problem	ML methods	Benefits	To be considered
• Automation of particular components	Supervised techniques for classification: Decision Trees, SVR, Logistic Regression, NN	Saving operation time, reducing human intervention, preventing subjective decisions	Dedicated machine time usually required to collect training data and to fine tune developed methods.
• Online optimization of several targets which are coupled • Unexpected drifts, continuous settings readjustment needed to maintain beam quality	Reinforcement Learning, Bayesian optimization, Gaussian Process, Adaptive Feedback	Simultaneous optimization targeting several beam properties, automatically finding trade-off between optimization targets, allows faster tuning offering more user time.	Ensuring that all important properties are included as optimization targets.
• Detection of anomalies	Unsupervised methods: clustering, ensembles of decision trees (e.g. Isolation Forest), supervised classification, Recurrent NN for time-series data.	Preventing faults before they appear, no need to define rules/ thresholds, no training is needed and can be directly applied on received data	In unsupervised methods, usually no “ground truth” is available → methods can be verified on simulations.

ML in accelerators: summary

Accelerator Problem	ML methods	Benefits	To be considered
• Computationally heavy, slow simulations • Reconstruct unknown properties from measurements	Supervised Regression models, NN for non-linear problems	Learning underlying physics directly from the data, faster execution	100% realistic simulations are not possible → the model performance will be as good as your data is.
• Reduction of parameter space e.g. for optimization	Clustering, Feature Importance Analysis using Decision trees	Speed up of available methods, simpler defined problems, easier to interpret	Parameter selection and combination (feature engineering) can have significant impact on ML methods performance
• Missing or too noisy data	Autoencoder NN	Robust models, data quality	Significant information should not be removed from the signal.