Machine Learning Application in the Large Hadron Collider at CERN

Elena Fol

CERN

Accelerators and Beam Physics Group (BE-ABP-LAF)

2nd Workshop on Artificial Intelligence for the Electron Ion Collider

10-14 Oct 2022

ML and accelerators: motivation

Accelerators

- Operation
- Diagnostics
- Beam Dynamics Modeling

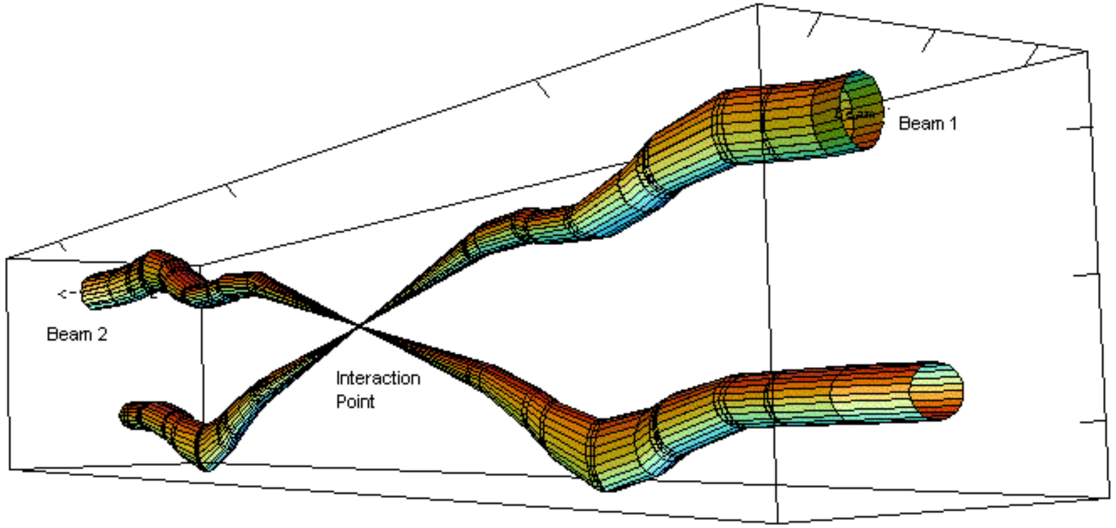
ML is a powerful tool for prediction and data analysis

Which limitations can be solved by ML with reasonable effort?

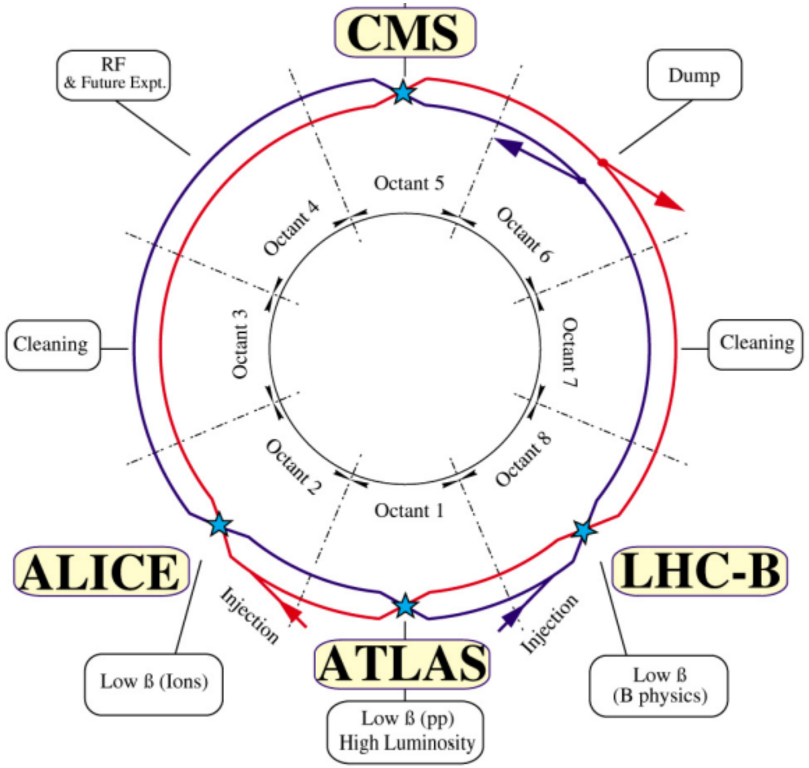
➢ large amount of optimization targets
➢ computationally expensive simulations
➢ direct measurements are not possible
➢ previously unobserved behaviour
➢ non-linear interacting sub-systems, rapidly changing environment.

Machine Learning methods can learn an arbitrary model from given examples without requiring explicit rules.

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**

Why and how is the beam optics controlled in the LHC?

Beam optics control at the LHC

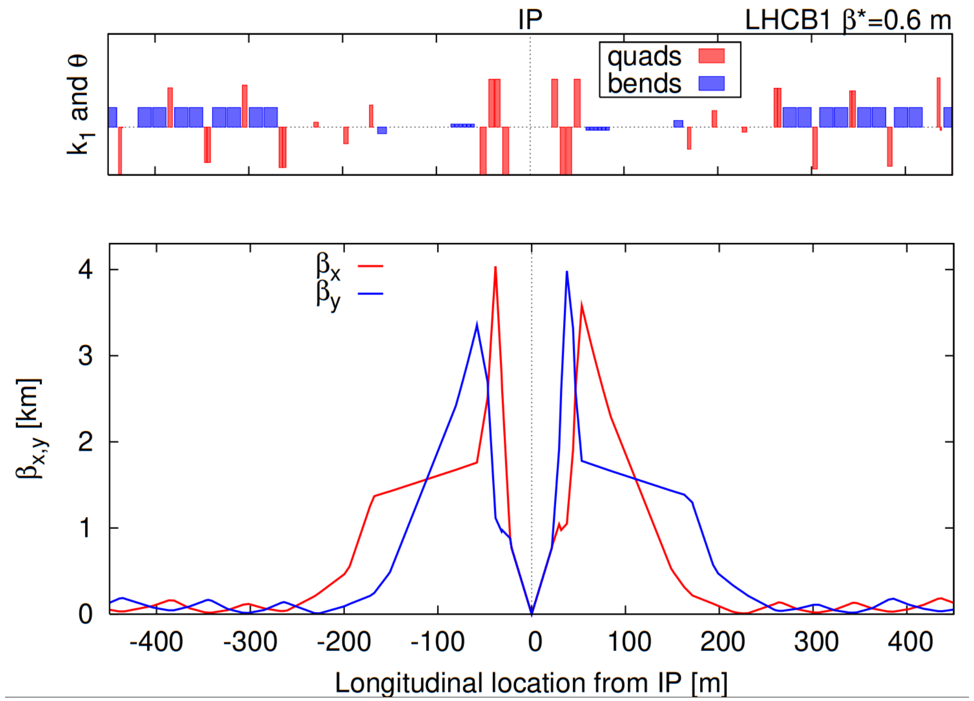
- **Luminosity:** maximize the number of collision events.

$$\boxed{\mathscr{L} \sim \frac{f \cdot N^2}{4\sigma^2}} \qquad \boxed{\sigma = \sqrt{\varepsilon\beta}}$$

ε → Const

β → Determined by **quadrupole arrangement and powering**

Optics



Why and how is the beam optics controlled in the LHC?

Beam optics control at the LHC

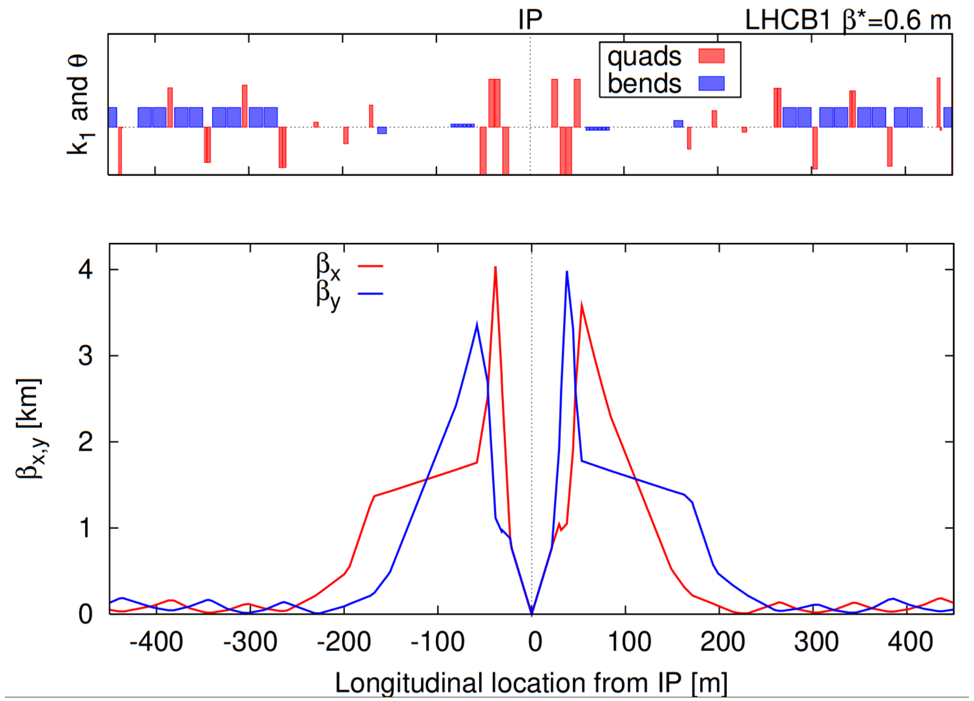
- **Luminosity:** maximize the number of collision events.

$$\mathscr{L} \sim \frac{f \cdot N^2}{4\sigma^2} \qquad \sigma = \sqrt{\varepsilon\beta}$$

ε → Const

β → Determined by **quadrupole arrangement and powering**

Optics ←



$$\frac{\Delta\beta}{\beta} = \frac{\beta_{meas} - \beta_{model}}{\beta_{model}}$$

- Access to the magnets for direct measurements is not possible during operation.

→ **Beam-based measurements** and **corrections of lattice imperfections.**

Overview

Assisting Beam Optics Control with ML

Unreliable measurements of beam properties

Detection of instrumentation faults: Beam Position Monitors

What are the actual magnet errors

Predict optimal settings: local magnet errors corrections

Uncertainties in the measured optics functions

De-noising of beam measurements

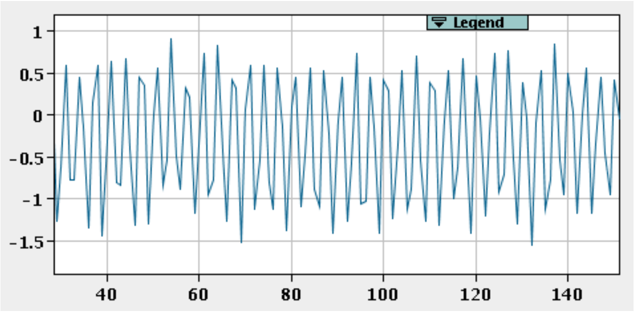
How to reconstruct the missing data?

Virtual diagnostics: obtaining beam properties without direct measurement

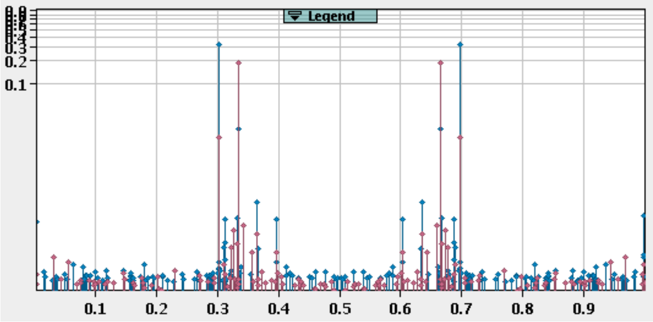
Detection of instrumentation faults

Measuring the optics

Turn-by-turn beam position



Spectrum



- Excite the beam to perform transverse oscillations.

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

- Harmonic analysis using Fast Fourier Transform (FFT)

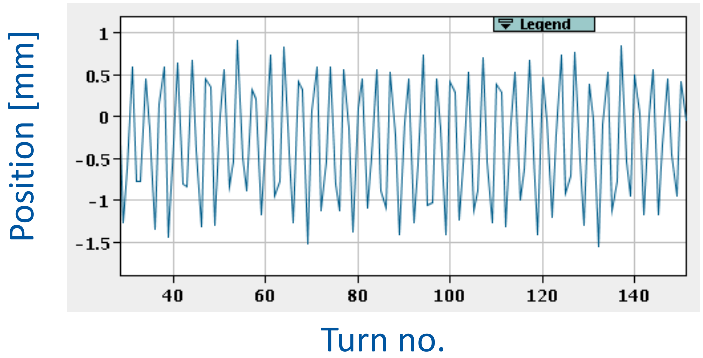
Denoising (SVD)
Signal cuts

Semi-automatic and manual cleaning of outliers

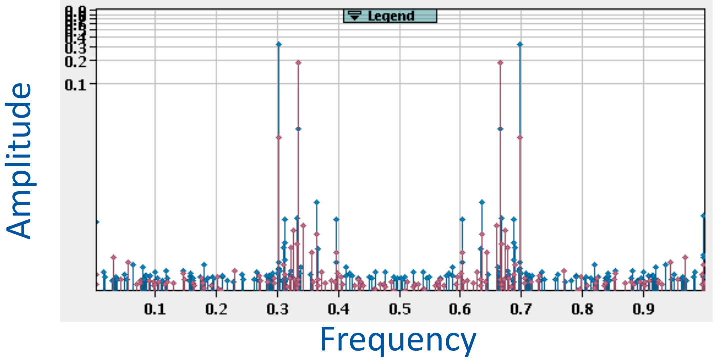
What are the limitations of traditional techniques?

Measuring the optics

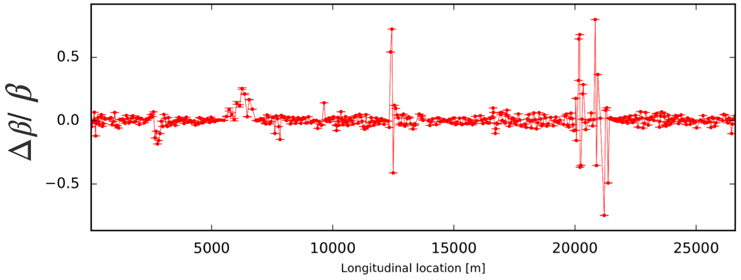
Turn-by-turn beam position



Spectrum



Optics



- Excite the beam to perform transverse oscillations.

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

Denoising (SVD)
Signal cuts

- Harmonic analysis using Fast Fourier Transform (FFT)

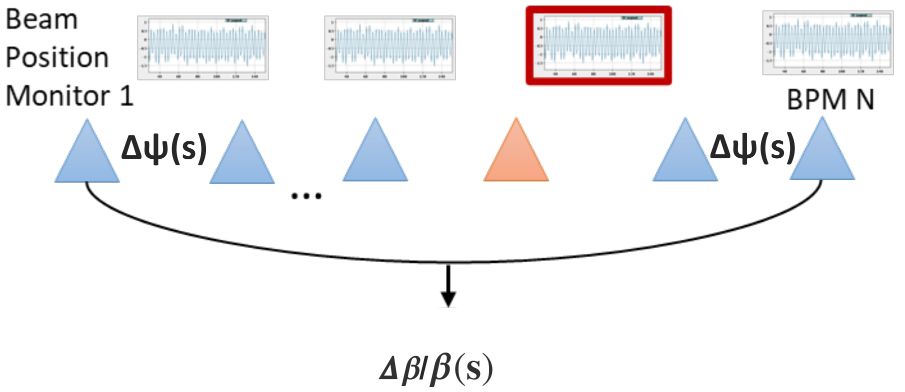
Semi-automatic and manual cleaning of outliers

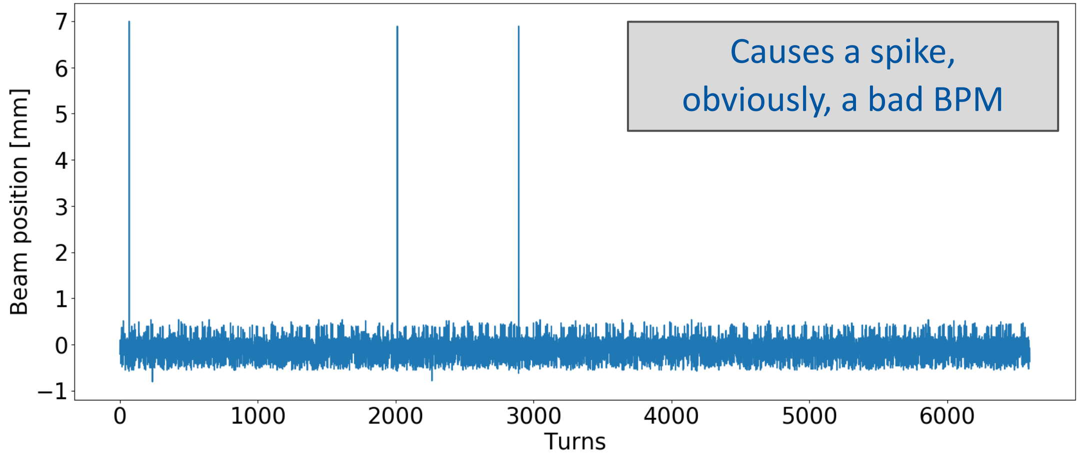
- Compute beta-beating and other optics functions

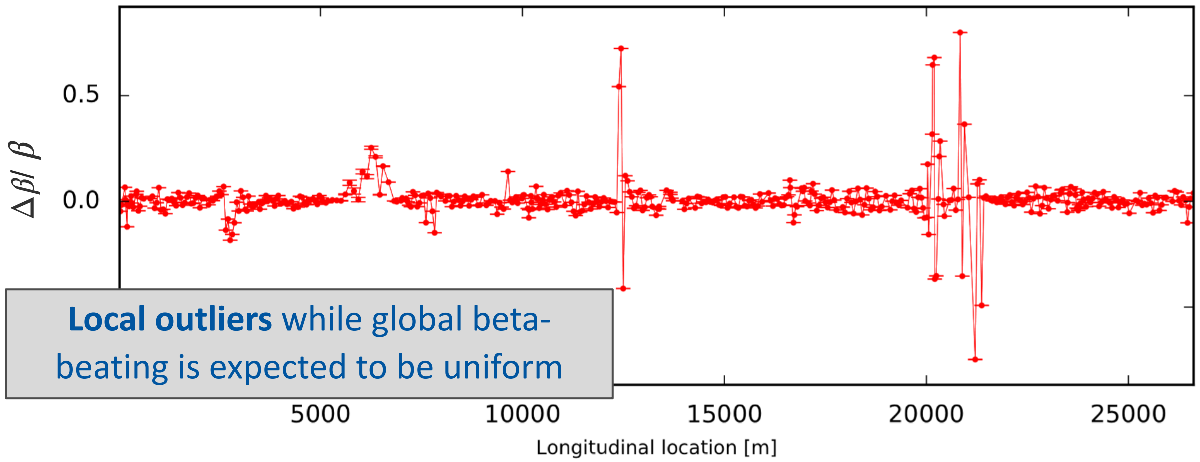
Unphysical values still can be observed

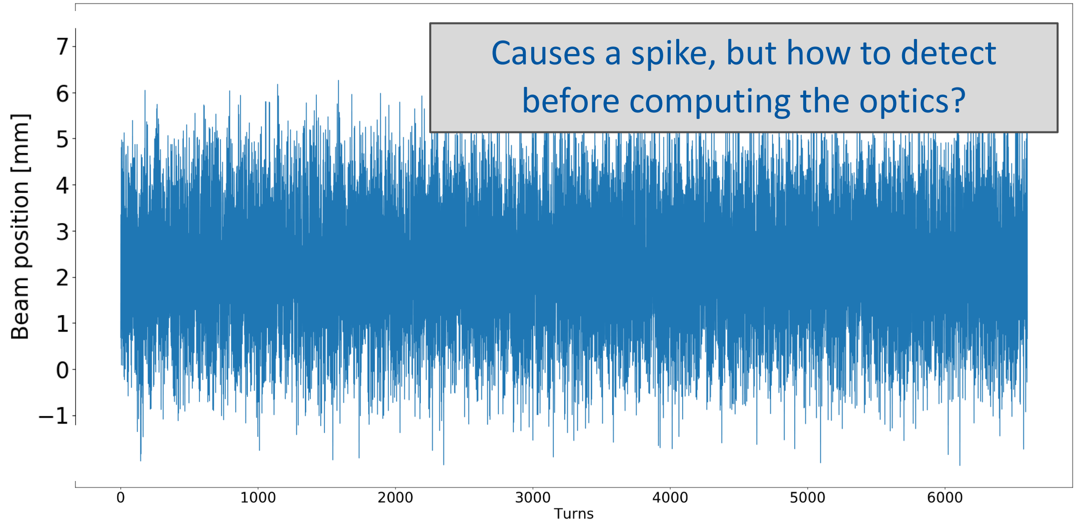
What are the limitations of traditional techniques?

Measuring the optics: challenges





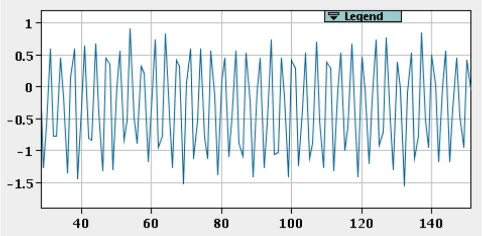


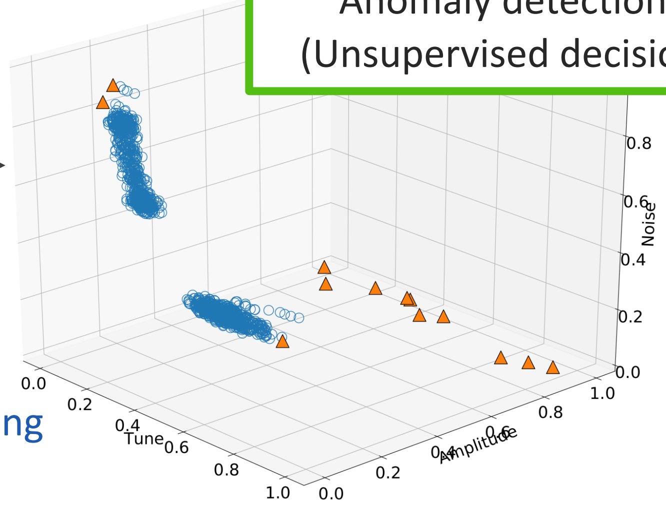


What are the limitations of traditional techniques?

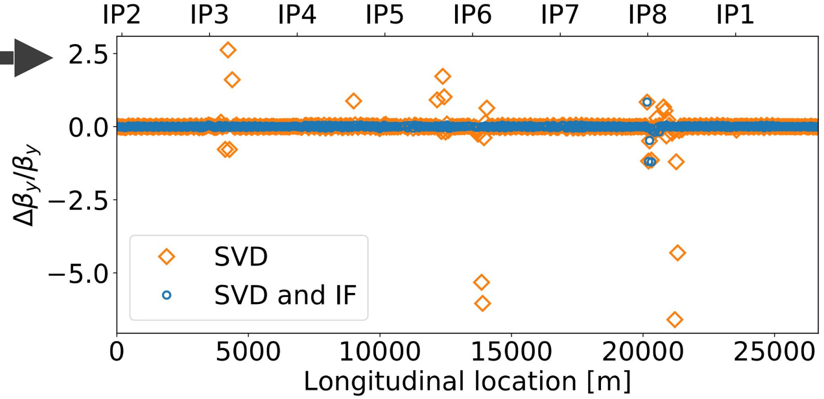
Detection and diagnostics of faulty BPMs

Harmonic analysis of all BPMs,
SVD for noise reduction





Prior to optics computation:
Anomaly detection using **Isolation Forest**
(Unsupervised decision tree-based algorithm)



- **No training, immediate results** applying IF on FFT properties of BPM signal

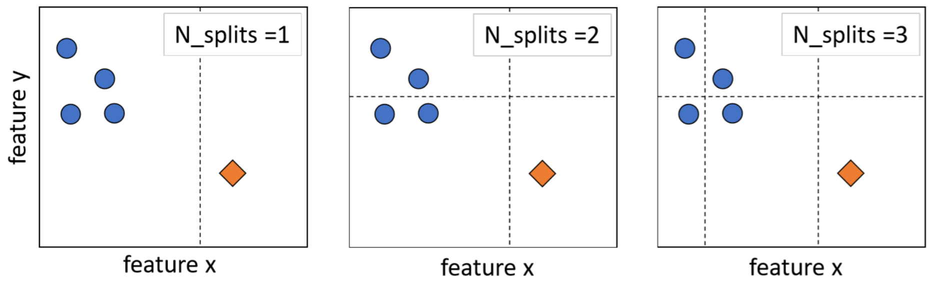
➢ **Providing information to BI experts:**

✓ Statistical analysis of data starting from 2018
✓ IF- algorithm: Identify dominant signal properties for faults classification
✓ Identified **116 critical faulty BPMs out of more than a thousand BPMs** in the LHC.

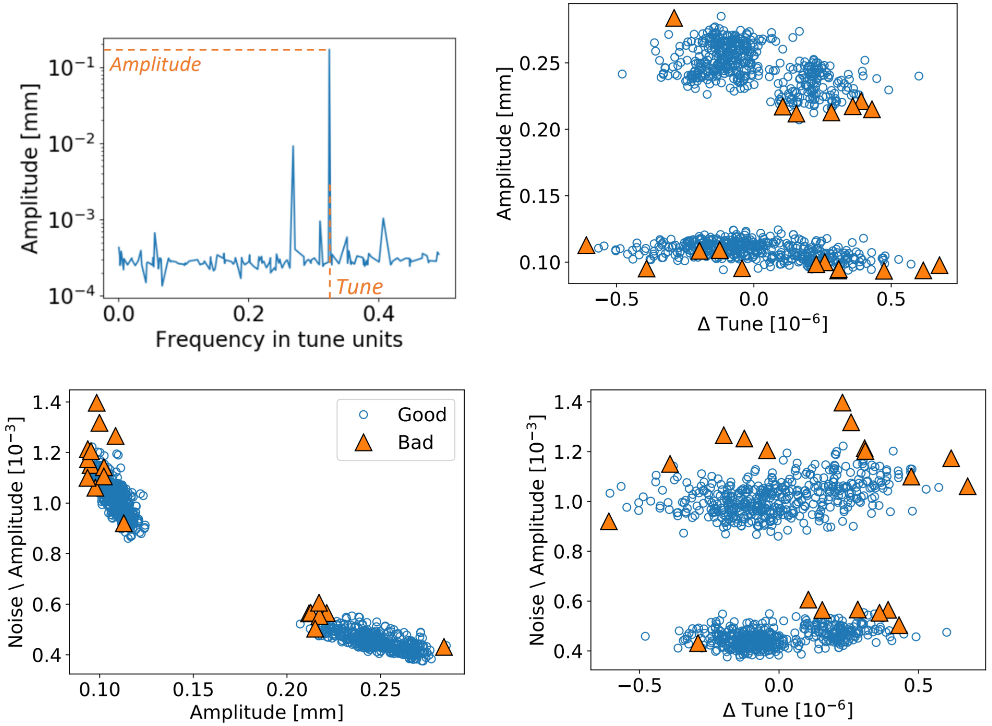
Thanks to ML: ➡ **Detection of otherwise unexplored hardware and electronics problems in BPMs**

Isolation Forest Algorithm

- Forest consists of several **decision trees***
- **Random splits aiming to “isolate” each point**
- The less splits are needed, the more “anomalous”
- **Contamination factor**: fraction of anomalies to be expected in the given data
  - → First obtained empirically from the past measurements
  - → Refined on **simulations introducing expected BPM faults**.
- **Input data: combination of several signal properties** obtained from harmonic analysis of BPM turn-by-turn measurements
  - → No additional data handling needed.
  - → No training, applied directly on the currently taken measurements data



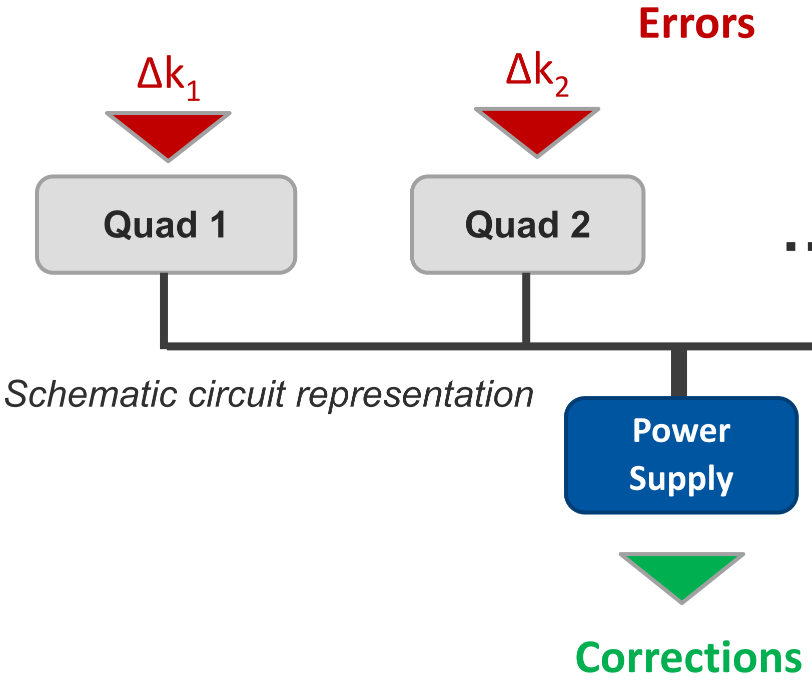
Conceptual illustration of Isolation Forest algorithm

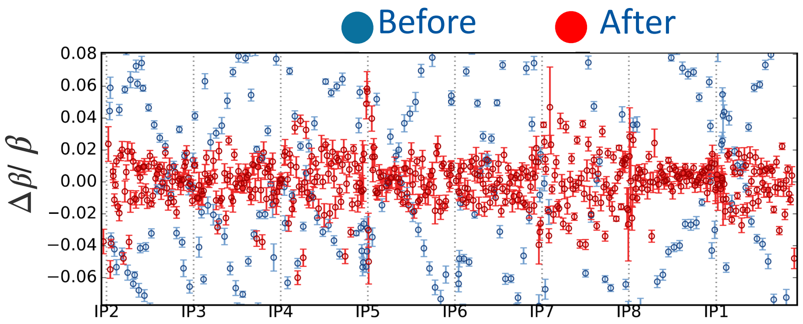


*Liu, Fei Tony, Ting, Kai Ming and Zhou, Zhi-Hua. “Isolation forest.” Data Mining, 2008. ICDM‘08.

Machine Learning for magnets errors reconstruction and beam corrections

Correcting the optics





β → Determined by **quadrupole arrangement and powering**

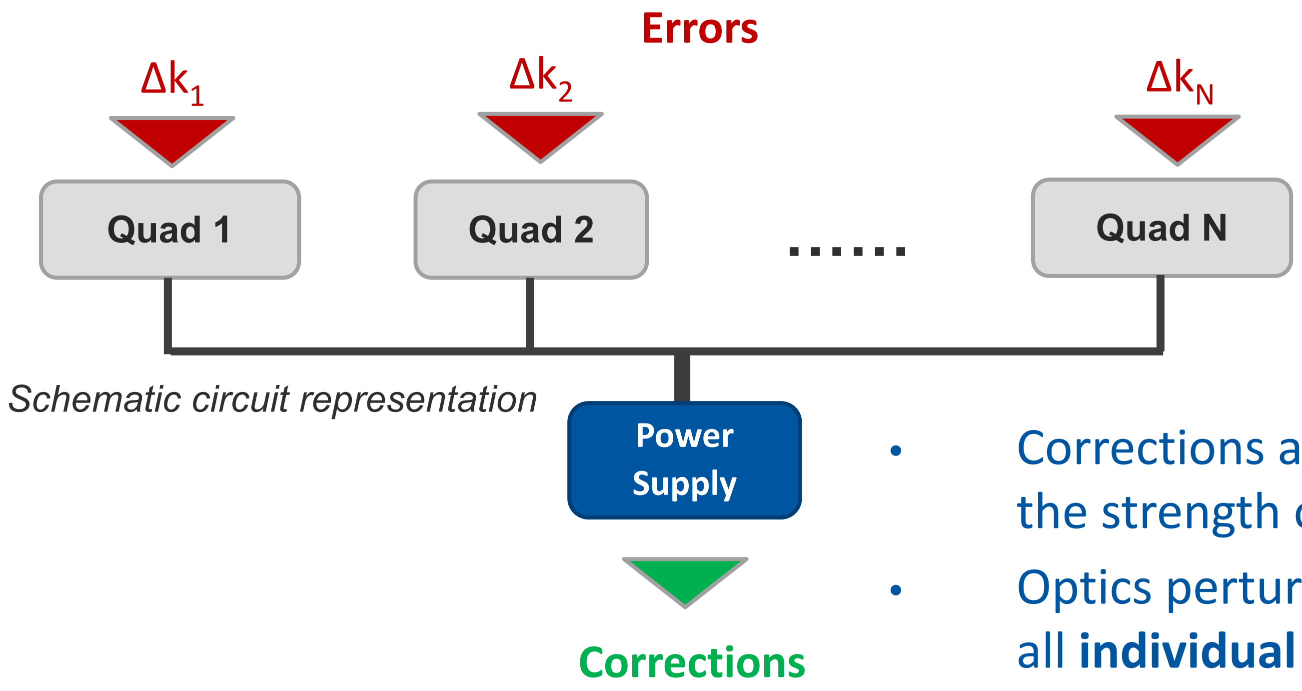
$$\frac{\Delta\beta}{\beta} = \frac{\beta_{meas} - \beta_{model}}{\beta_{model}}$$

- Access to the magnets for direct measurements is not possible during operation.

→ Beam-based measurements and corrections of lattice imperfections.

What are the limitations of traditional techniques?

Correcting the optics

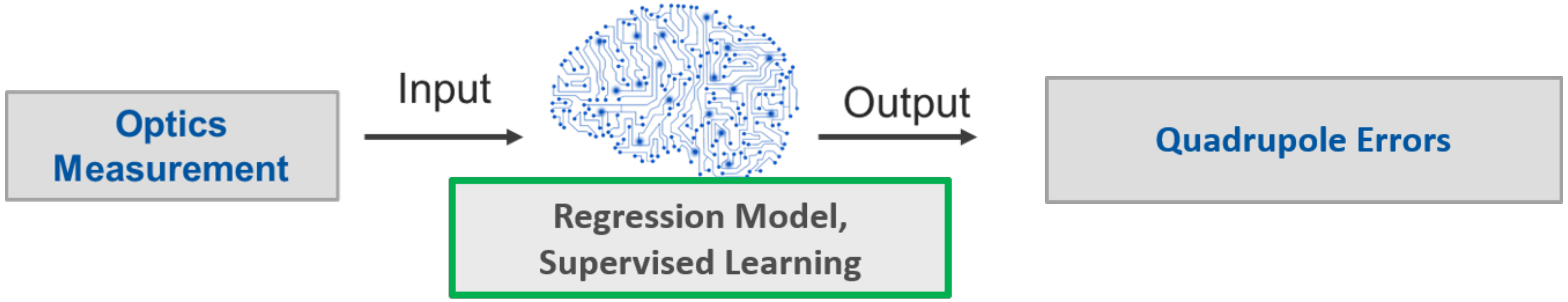


- Corrections are implemented by changing the strength of **circuits**
- Optics perturbations are caused by all **individual magnets.**

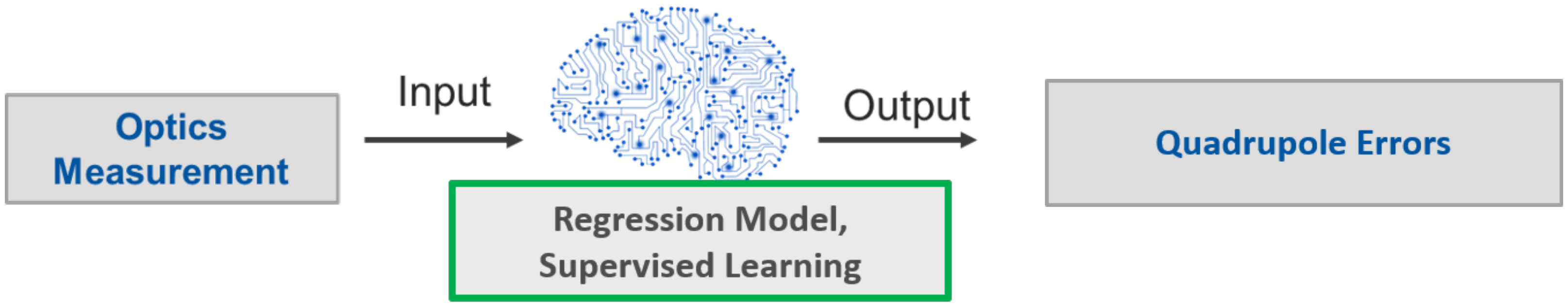
➢ What are the **actual errors of individual quadrupoles?**
➢ How to obtain the **full set of errors in one step**?

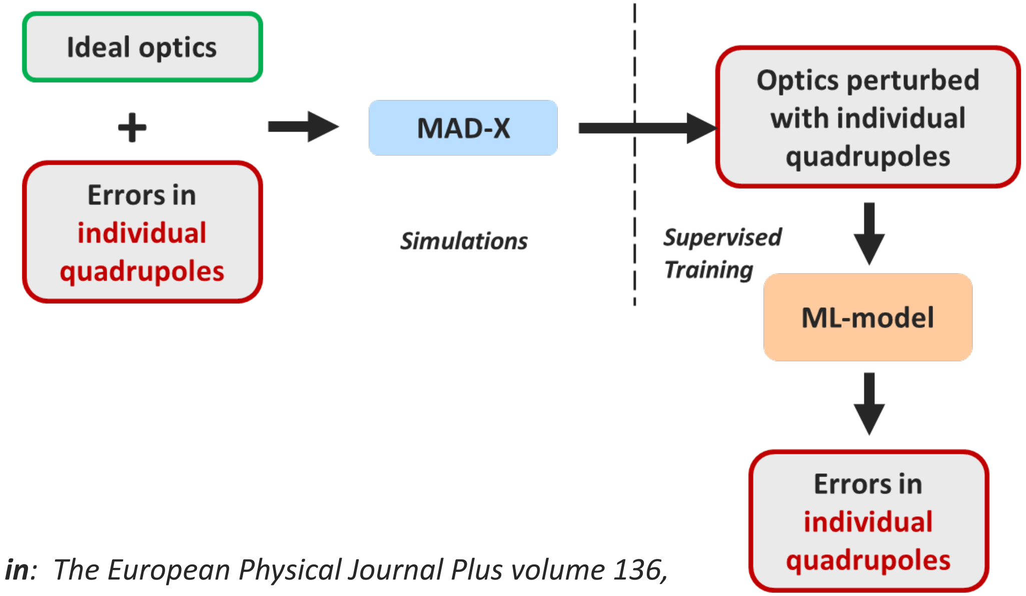
What are the limitations of traditional techniques?

Estimation of quadrupole errors



Estimation of quadrupole errors





Training ML- regression model:

- **1256 target** variables: randomly assigned field errors in quadrupoles + other error sources
- **3304 input** variables: optics functions
- Using **Linear Regression as baseline model**

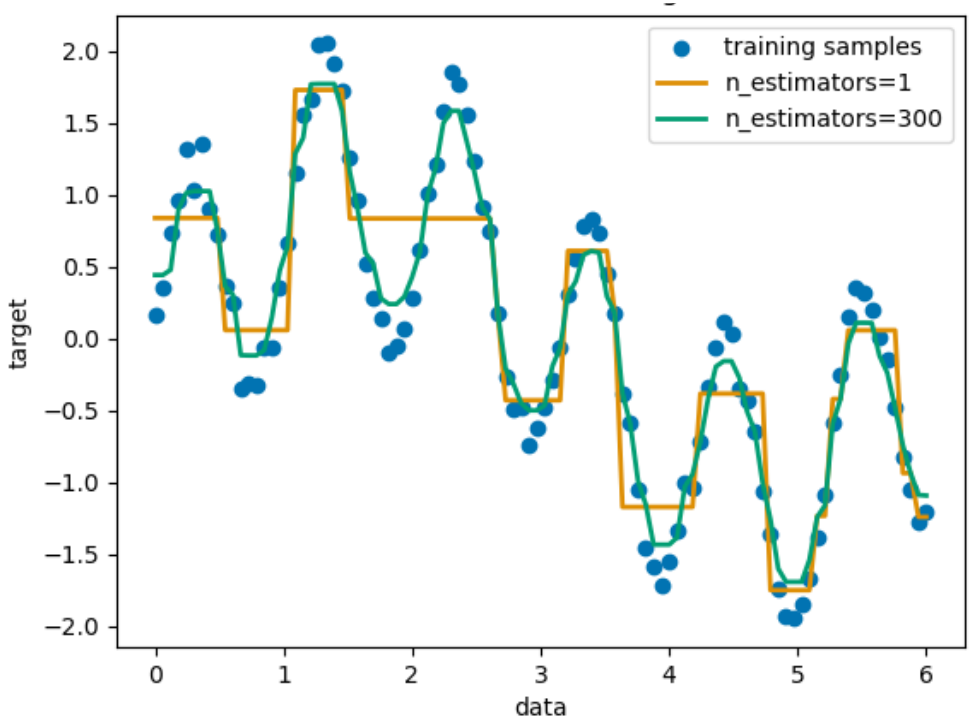
$$\min_{w} \|Xw - y\|_2^2 + \alpha \|w\|_2^2$$

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

Random Forest Regression

Supervised Learning approach:

➡ generalized model **explaining relationship between input and output** variables in **all training samples**.



Decision Trees:

- Partition data based on a sequence of thresholds
- Continuous target y, in region estimate: $c_m = \frac{1}{N_m} \sum_{i \in N_m} y_i$
- Mean Square Error: $H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y_i - c_m)^2$

Random Forest:

- Random subset of examples, train separate model on each subset
- Only random subset of features is used at each split
- Increases variance, tend not to overfit

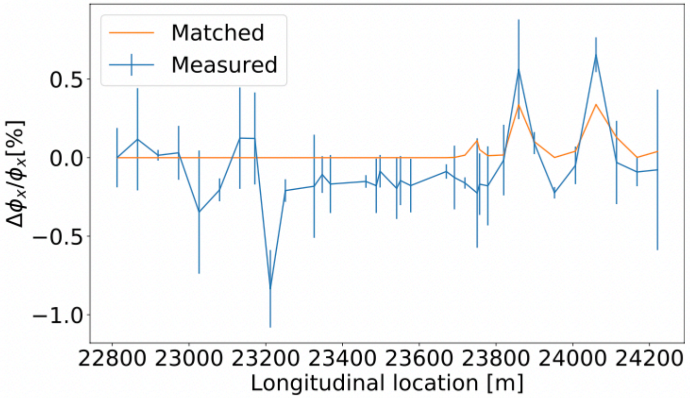
LHC commissioning 2022: beam optics corrections

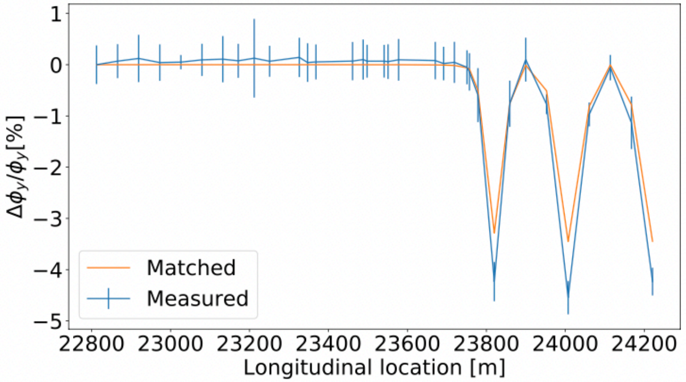
Optics Corrections path in the LHC

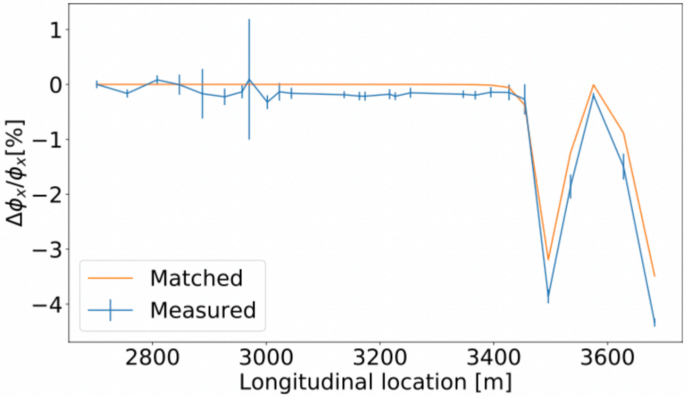
1. Beam optics measurements
2. **Reconstruct the magnet errors** ← **Predict all errors using ML-model trained on simulations!**
3. Propagate the errors within the region (simulations)
4. Compare with measurements
5. Apply the reconstructed quadrupole strength errors with opposite sign —> optics corrections!

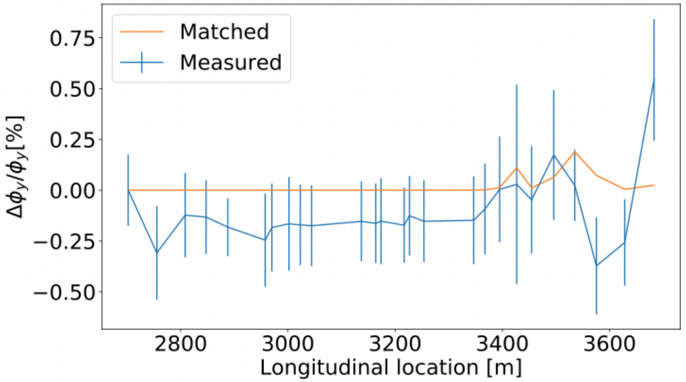
LHC commissioning 2022: beam optics corrections

Example: Corrections in Interaction Region 1, squeeze to β^* = 30 cm (challenging low beta optics)







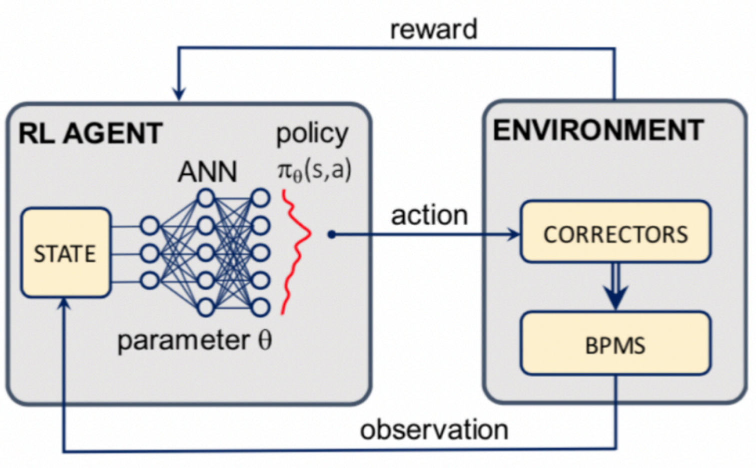


- ✓ Phase errors can be corrected applying the errors with opposite sign as correction settings
- ✓ **Simultaneous local correction in all IRs within seconds.**

→ **Potential to save operation time!**

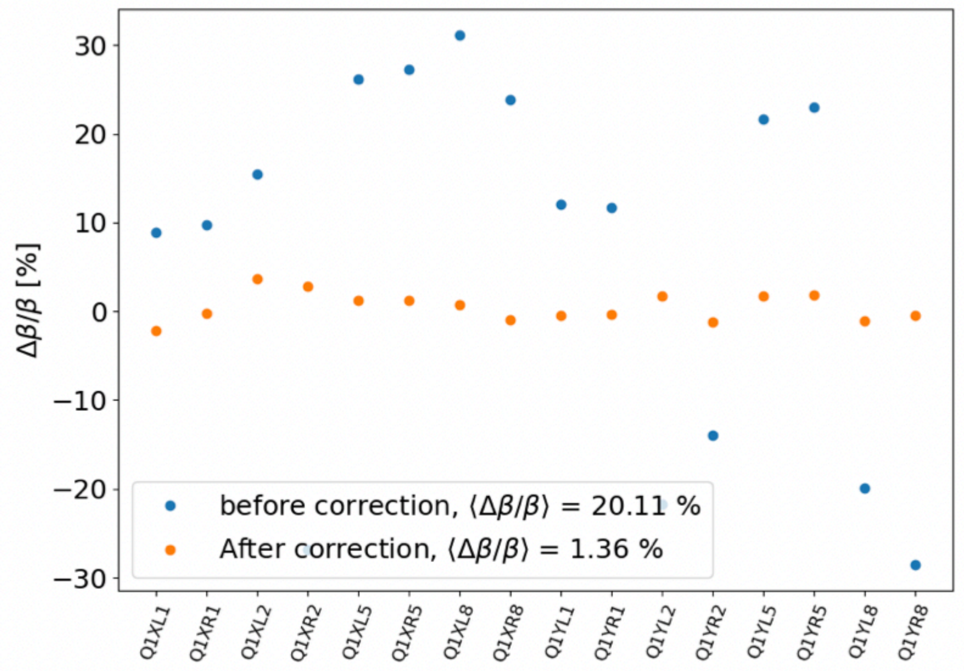
E.Fol et al.,"Experimental Demonstration of Machine Learning Application in LHC optics commissioning",IPAC'22 MOPOPT047

Look into the future: Optics control in HL-LHC



- Environment = Surrogate model of HL-LHC lattice
- Reward = Average beta-beating in IRs
- State space = Quadrupole strengths (only triplet magnets for now)
- Action space = Correctors settings

Work by Hector Garcia Morales, BE-ABP



- Uses the previously presented approach **to learn LHC model from simulated data**
- **RL algorithms** implementations based on **OpenAI**
- **PyTorch** for the training of critic networks

Results:

After the learning process, the model is **able to perform the optics correction in one single iteration** with residual β-beating of 1-2% (up to 20% initially)

Virtual Diagnostics

Virtual Optics Measurements

- ➢ How to reconstruct optics observables **without direct measurements?**

Phase advance measurements*

** Always available from turn-by-turn data at one beam excitation*

Input →

Regression Model, Supervised Learning

Output →

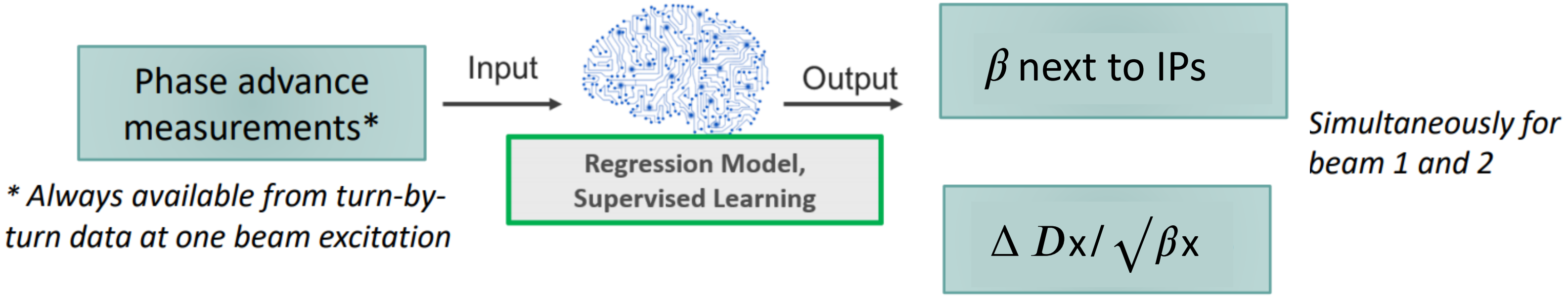
β next to IPs

$\Delta\, Dx/\sqrt{\beta x}$

Simultaneously for beam 1 and 2

Virtual Optics Measurements

➢ How to reconstruct optics observables **without direct measurements?**



Measuring beta-function in Interaction Regions:

Traditional technique: **k-modulation**

- Based on modulation of quadrupole current
- Time consuming
- Accuracy varies depending on tune measurement uncertainty, magnet errors and β^* settings.

Virtual Optics Measurements

➢ How to reconstruct optics observables **without direct measurements?**

Phase advance measurements*

Input

Output

Regression Model, Supervised Learning

β next to IPs

$\Delta\ Dx/\sqrt{\beta x}$

Simultaneously for beam 1 and 2

** Always available from turn-by-turn data at one beam excitation*

Measuring beta-function in Interaction Regions:

Traditional technique: **k-modulation**

- Based on modulation of quadrupole current
- Time consuming
- Accuracy varies depending on tune measurement uncertainty, magnet errors and β^* settings.

✓ β-functions left and right from IPs **within a few seconds vs. several minutes for k-modulation**

✓ Average accuracy: **5 % for β^* = 30 cm**.

➢ Tests during LHC commissioning 2022

β^* = 30 cm

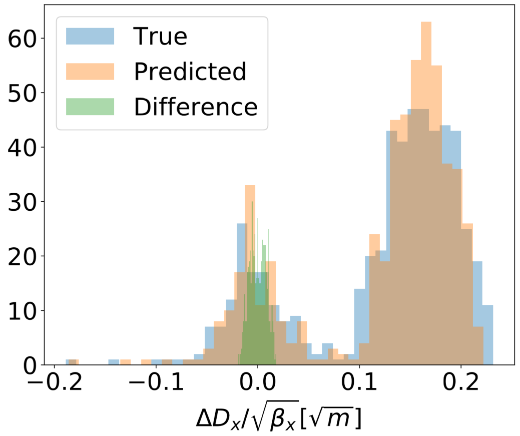
Location	K-mod β_x, β_y [m]	ML β_x, β_y [m]	$\Delta\beta/\beta_{kmod}$ x, y [%]
B1, IP1L	1262, 1074	1296, 1223	2.6, 13.8
B1, IP1R	1340, 1051	1268, 1197	5.3, 13.9
B1, IP5L	1388, 1552	1377, 1659	0.8, 6.9
B1, IP5R	1302, 1624	1369, 1642	5.2, 1.1
B2, IP1L	1406, 1773	1435, 1851	2.1, 4.4
B2, IP1R	1366, 1947	1412, 1893	3.4, 2.7
B2, IP5L	1511, 1364	1639, 1315	8.4, 3.6
B2, IP5R	1637, 1377	1632, 1303	0.3, 5.4

Virtual Optics Measurements

Horizontal Dispersion reconstruction:

- Computed by acquiring turn-by- turn data from **several beam excitations, shifting the momentum.**
- **Input**: simulated phase advance deviations given noise
- **Output**: normalized dispersion $\Delta D\mathrm{x}/\sqrt{\beta\mathrm{x}}$
- Using **linear regression model**: 10 000 samples

***Simulation** example: Beam 1*



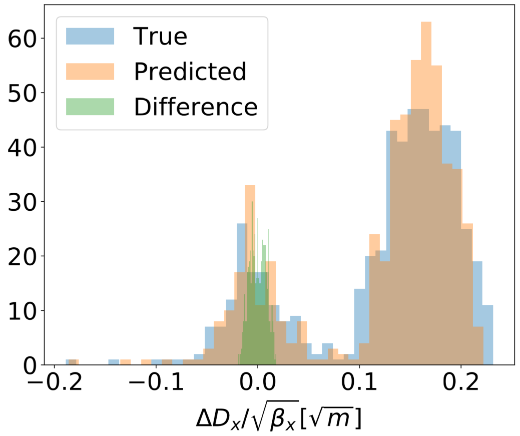
How to reconstruct optics observables without measurements?

Virtual Optics Measurements

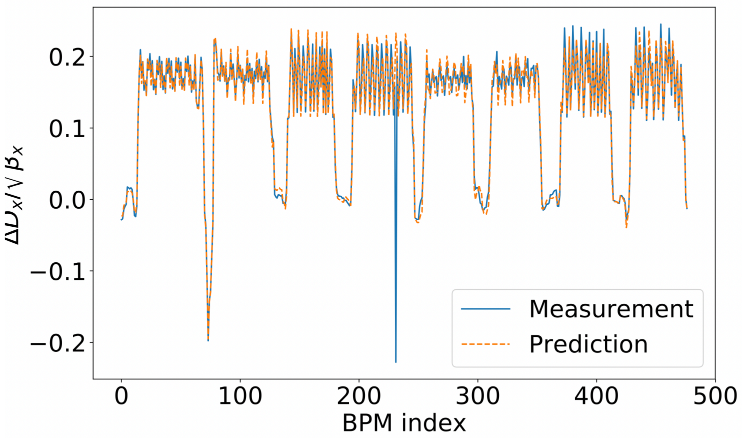
Horizontal Dispersion reconstruction:

- Computed by acquiring turn-by- turn data from **several beam excitations, shifting the momentum.**
- **Input**: simulated phase advance deviations given noise
- **Output**: normalized dispersion $\Delta D\text{x}/\sqrt{\beta\text{x}}$
- Using **linear regression model**: 10 000 samples

***Simulation** example: Beam 1*



Measurement taken during LHC commissioning**, β^* **= 30 cm



✓The relative **error of prediction is 5%** (beam 1) and **7%** (beam 2)

✓Potential speedup of machine commissioning for the same performance.

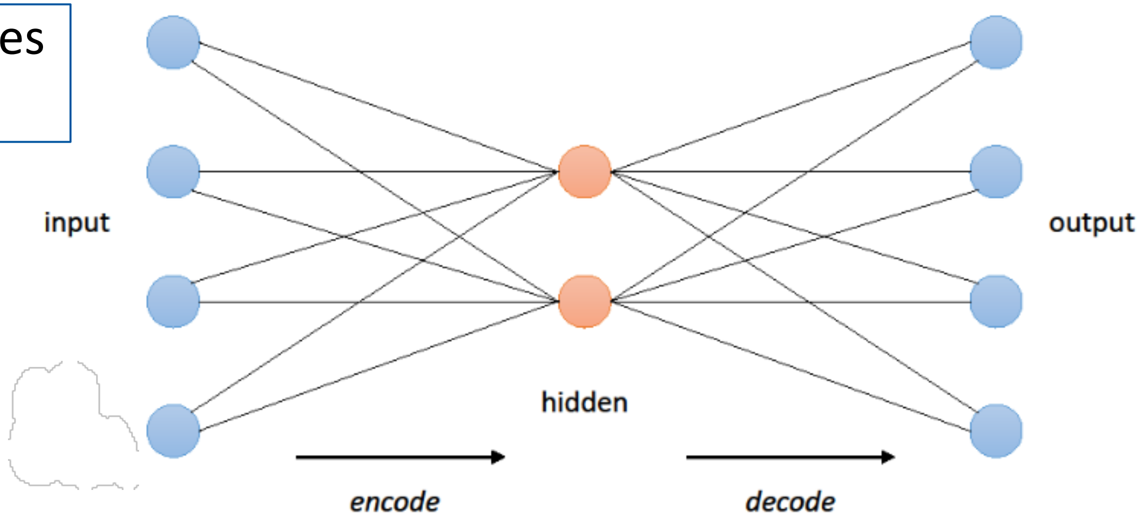
How to reconstruct optics observables without measurements?

De-noising of optics measurements

- Uncertainties in the measured optics functions → **"noise"** → Noise in the measurements degrades the performance of corrections techniques

Autoencoder Neural Network

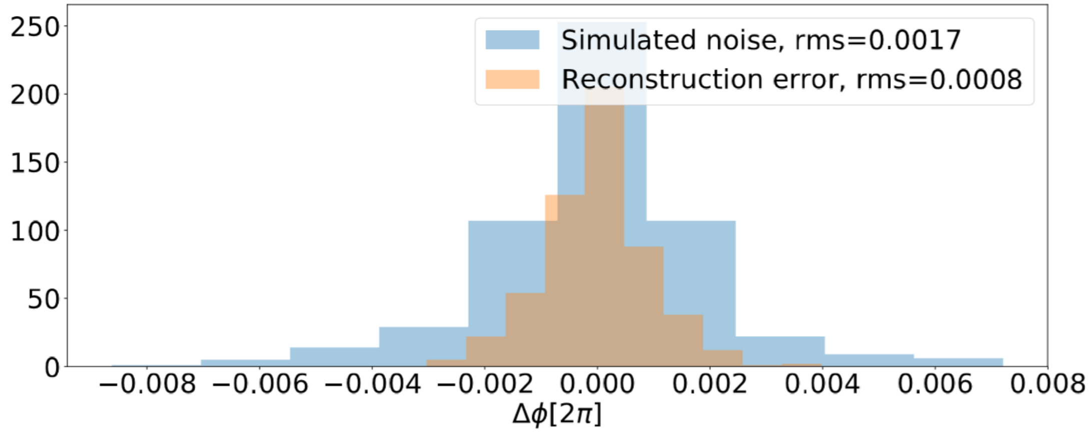
Simulated optics observables **+ noise**



Denoised optics

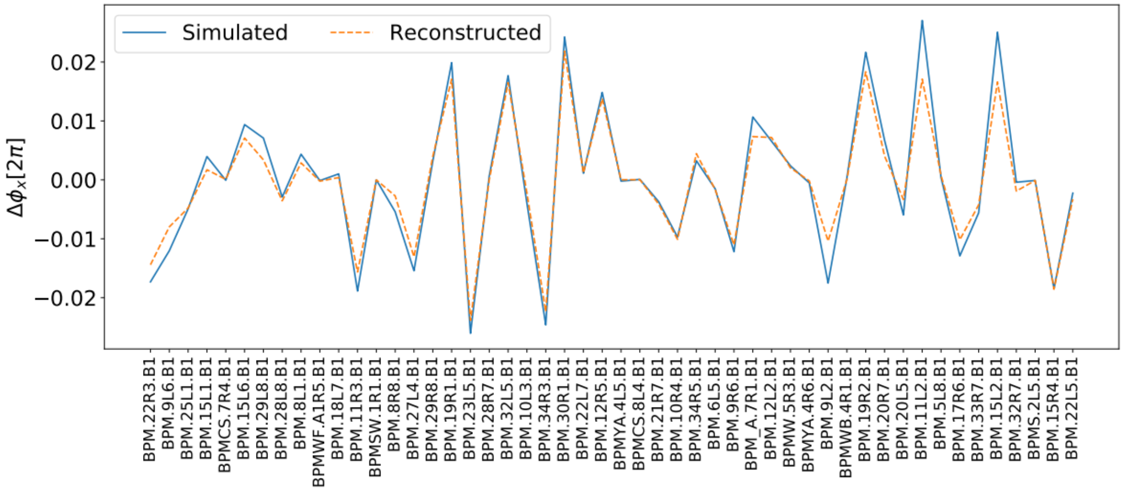
De-noising of optics measurements

Simulated data: Noise Reduction



✓ Reconstruction error is by factor 2 smaller than simulated realistic noise.

Simulated data: Reconstruction



✓ Reliable reconstruction after denoising

➢ **Potential improvement of measurements quality**
➢ **Possibility to reconstruct the phase advance at the location of faulty BPMs.**

Summary

ML in the LHC beam optics control

✓ **ML-based toolbox for beam optics analysis**

- Detection of instrumentation faults → no manual cleaning and repeated optics analysis
- Estimation of individual magnet errors → Better knowledge and control of individual optics errors
- Reconstruction of optics observables → Additional observables without dedicated measurements
- Denoising of optics measurements → Increasing the quality of the measurements

- ***More in "Machine learning for beam dynamics studies at the CERN Large Hadron Collider" https://doi.org/10.1016/j.nima.2020.164652***

ML in the LHC beam optics control

✓ **ML-based toolbox for beam optics analysis**

- Detection of instrumentation faults → no manual cleaning and repeated optics analysis
- Estimation of individual magnet errors → Better knowledge and control of individual optics errors
- Reconstruction of optics observables → Additional observables without dedicated measurements
- Denoising of optics measurements → Increasing the quality of the measurements

- ***More in "Machine learning for beam dynamics studies at the CERN Large Hadron Collider" https://doi.org/10.1016/j.nima.2020.164652***

✓ **Paving the way for new studies currently being in progress:**

- Optics corrections for High Luminosity – LHC upgrade (Reinforcement Learning)
- Exploring more complex optics error sources in the LHC: coupling corrections
- Improving Dynamic Aperture estimates using clustering
- **Optimizing the design of future colliders (Ionisation Cooling channel for a muon collider)**.

Cat!

Thank you for your attention!

ML in Muon Collider Design Studies

Muon Collider Design study [1]:

- **Reduction of transverse emittance of produced muon beams as one of the biggest challenges:**

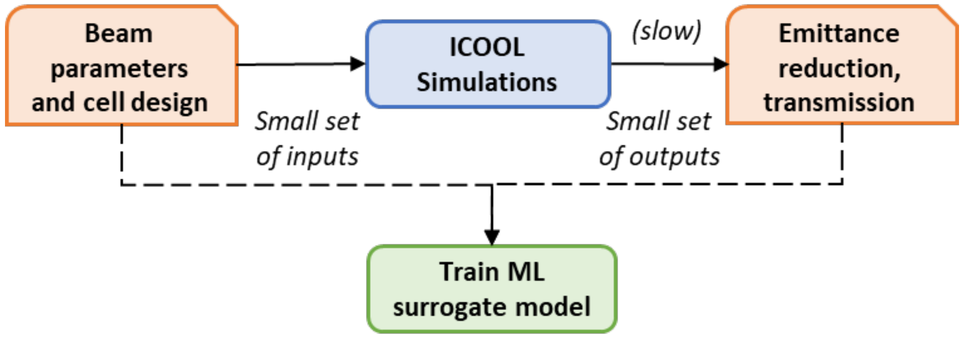
→ **Final Cooling system** with challenging design

→ High dimensional parameter space to be optimized in order to achieve low emittance, high intensity muon beams

→ Trade-off between different optimization objectives

- Extending existing simulation frameworks towards automatic, fast executing optimization.
- **Exploring application of Supervised Learning**

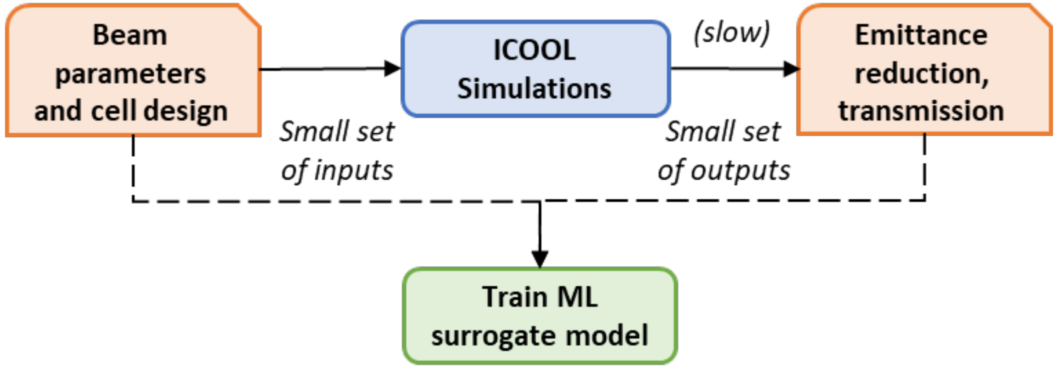
  → **surrogate models**



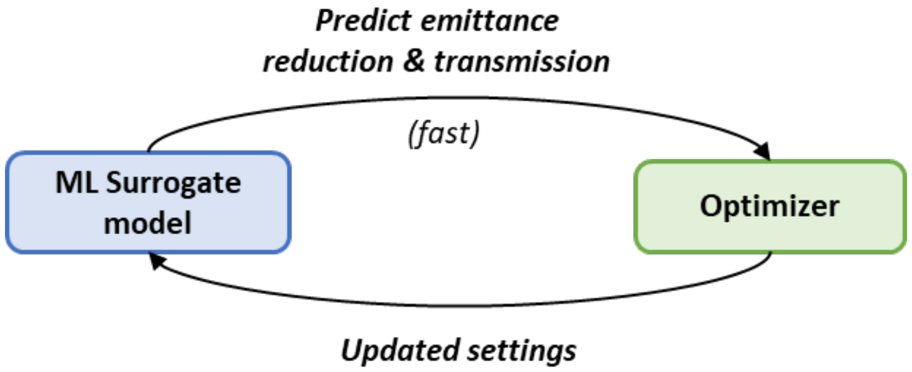
[1]: https://muoncollider.web.cern.ch

ML in Muon Collider Design Studies

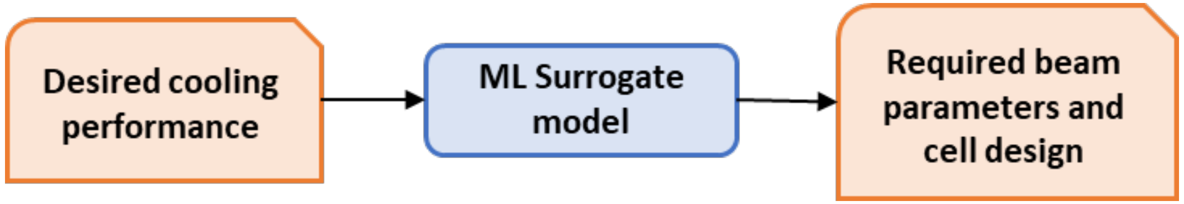
- **Exploring application of Supervised Learning → surrogate models**



1. Speeding up optimization:



2. "Backwards" design:



✓ **First results demonstrating orders of magnitude optimization speed up**

✓ **Accurate prediction of initial parameters to achieve desired cooling performance**

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

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.

Regression Models

- Linear model for *input X / output Y pairs, i* – number of pairs (training samples): $f(X, w) = w^T X$
- Squared Loss function for model optimization: $L(w) = \frac{1}{2}\sum_i \left(Y_i - f(X_i; w)\right)^2$
- Find new weights minimizing the Loss function: $w^* = \mathrm{argmin}_w L(w)$

→ **Update weights for each incoming input/output pair.**

→ **Regularization** places constraints on the model parameters (weights)

- Trading some bias to reduce model variance.
- Using L2-norm: $\Omega(w) = \sum_i w_i^2$, adding the constraint $\alpha\Omega(w)$ to the weights update rule
- The larger the value of α, the stronger the shrinkage and thus the coefficients become more robust.