Machine Learning Application in the Large Hadron Collider at CERN

Elena Fol

CERN

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

LBNL Instrumentation Colloquium, 28th of September 2022

Overview

Why do particle accelerators need ML?

Detection of instrumentation faults
(Beam Position Monitors)

Reconstruction of settings imperfections
(magnet errors)

De-noising of beam measurements

Virtual diagnostics:
obtaining beam properties without direct measurement

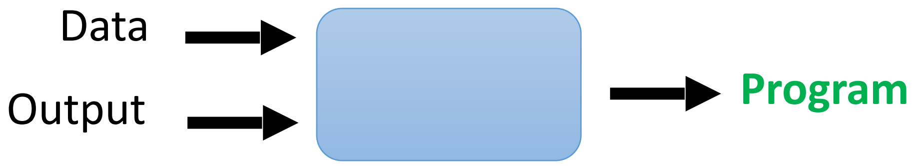
Developing Machine Learning applications

Teaching machines to learn from experience

- ## Traditional programming

creating **manually a set of commands** and rules

- ## Machine Learning approach



learn from data automatically

Learning from data: Supervised Learning

example 1
example 2
example 3
.
.
.

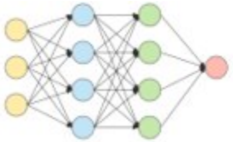
Training input data

Training output data

Learning from data: Supervised Learning

example 1
example 2
example 3
.
.
.

Training input data



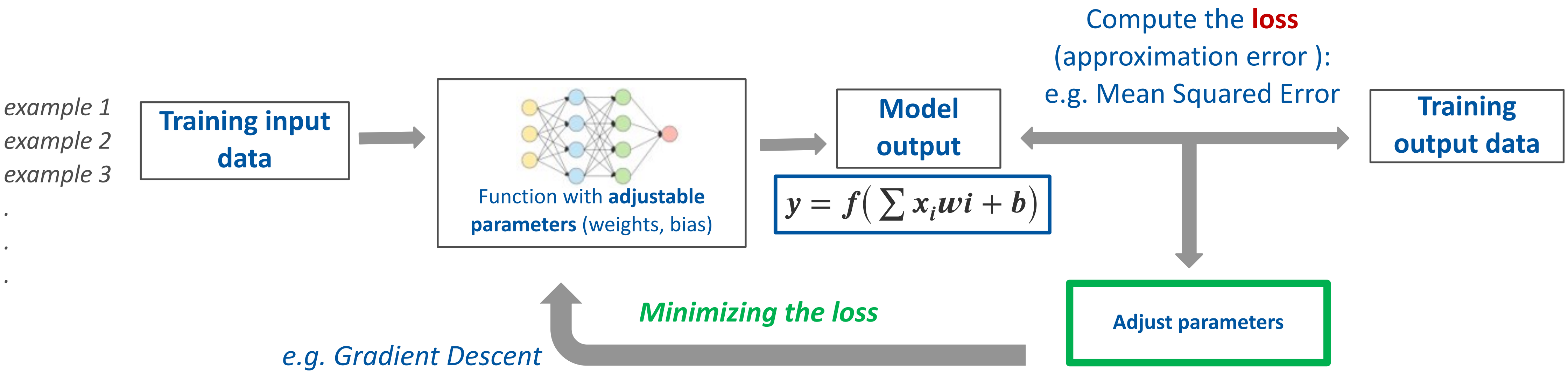
Function with **adjustable parameters** (weights, bias)

Model output

$$y = f\left(\sum x_i wi + b\right)$$

Training output data

Learning from data: Supervised Learning



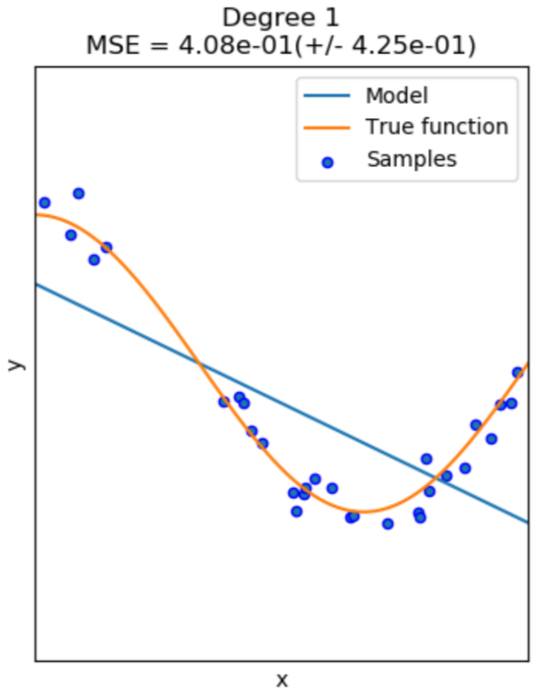
➡ **Learning from data automatically**

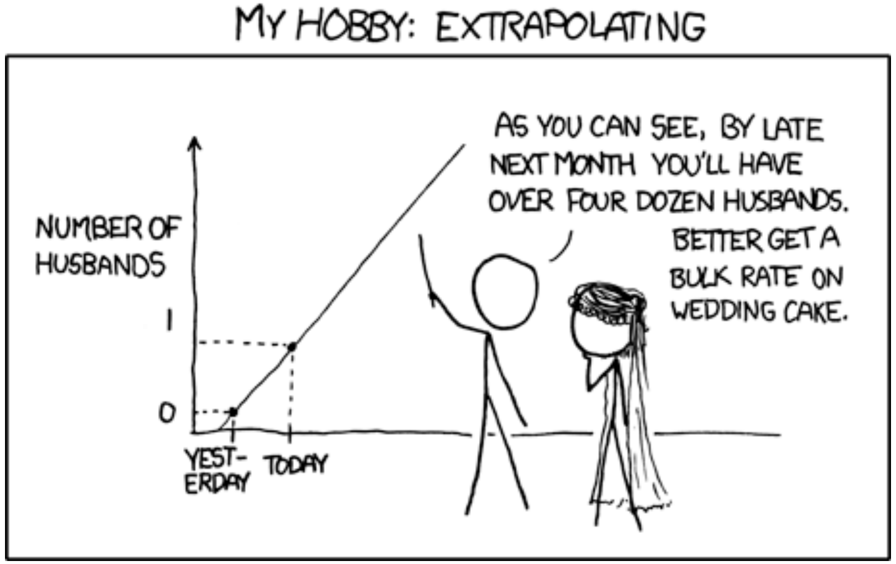
➡ **Explaining relationship between input and output variables in all training samples.**

➡ **Generalisation: the capability of explaining new cases**

- How to prove generalisation?

Training and generalization

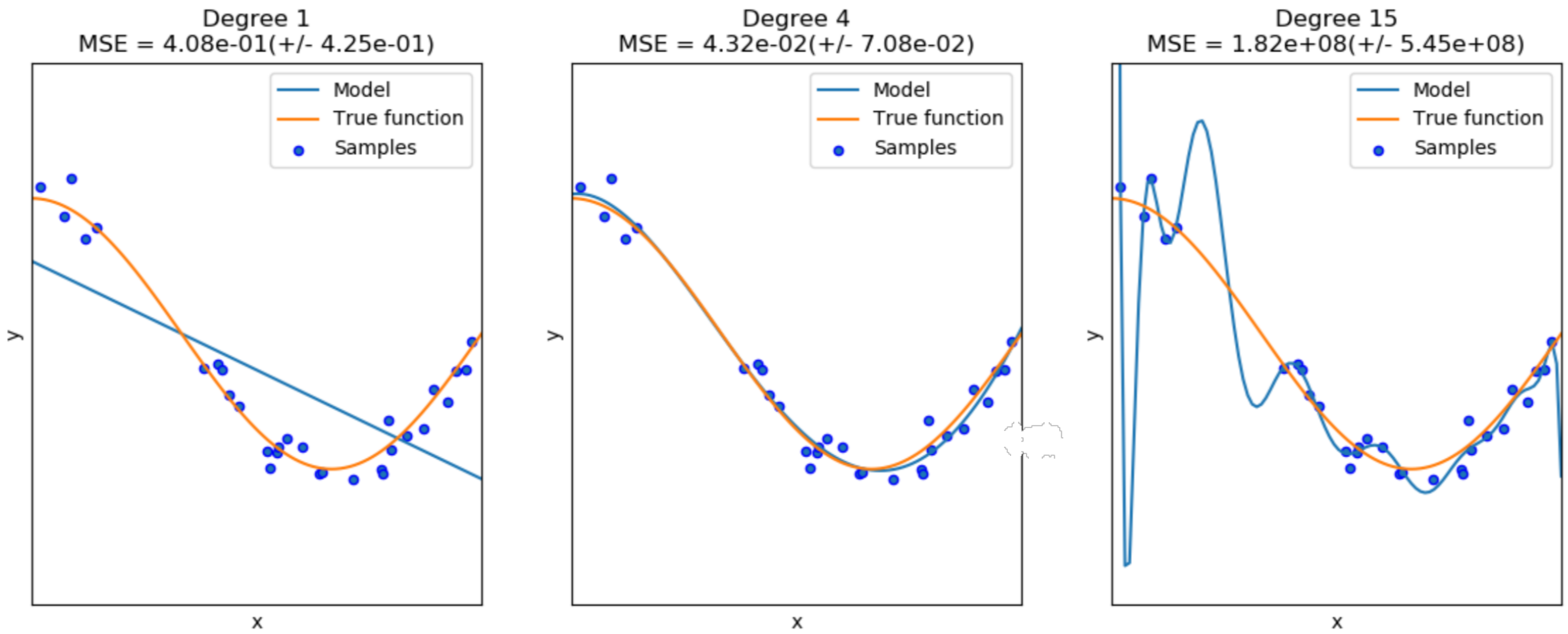




Simple models underfit

- Derivate from data (high bias)
- Do not correspond to data structure (low variance)

Training and generalization



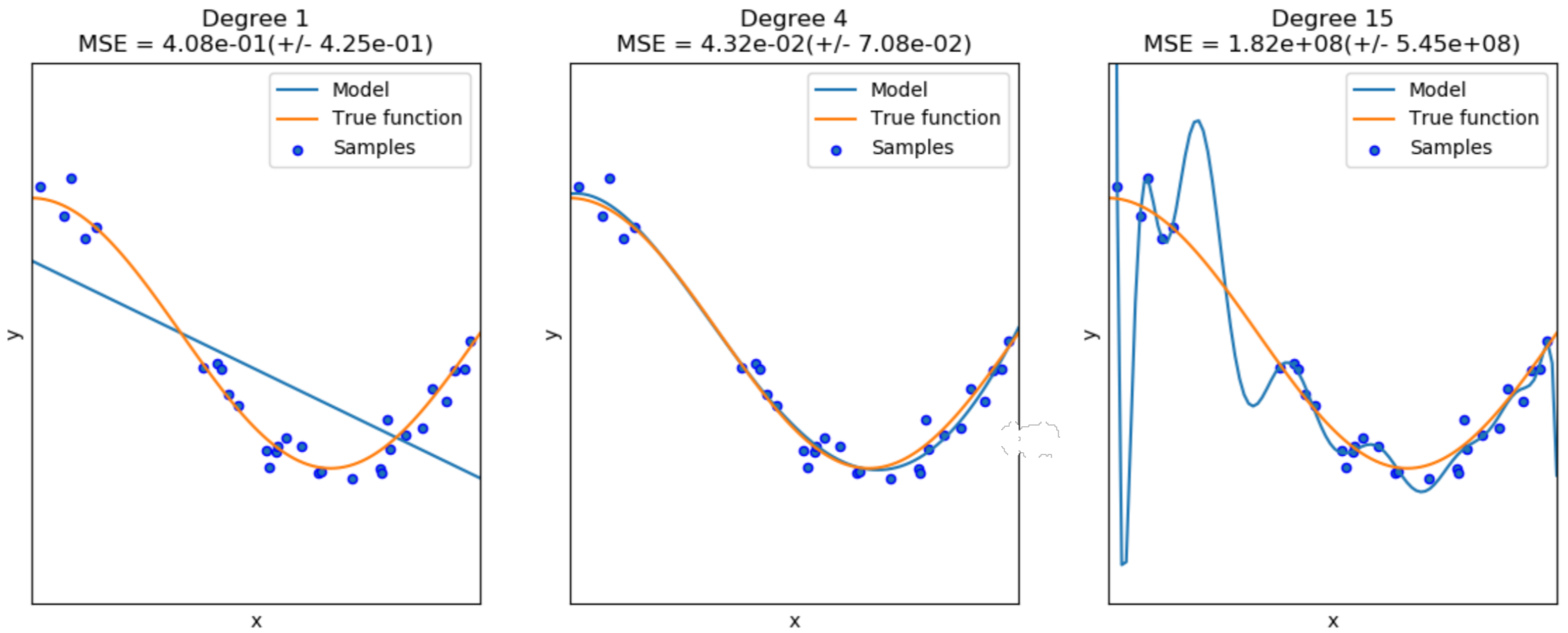
Simple models underfit

- Derivate from data (high bias)
- Do not correspond to data structure (low variance)

Complex models overfit

- Very low systematical deviation (low bias)
- Very sensitive to data (high variance)

Training and generalization



Simple models underfit ⟷ **Bias-Variance tradeoff** ⟷ Complex models overfit

Simple models underfit

- Derivate from data (high bias)
- Do not correspond to data structure (low variance)

Bias-Variance tradeoff

- **Separate data into train and test sets**
- **Find optimal model hyperparameter e.g. with cross-validation**

Complex models overfit

- Very low systematical deviation (low bias)
- Very sensitive to data (high variance)

Relevant ML concepts and definitions

Supervised Learning

- **Input/output pairs** available
- Learn a mapping function, **generalizing for all provided data**
- Predict from **unseen data**

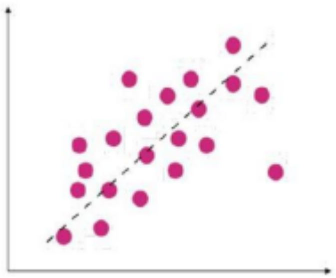
Unsupervised Learning

- **Only input** data is given
- Discover structures and patterns

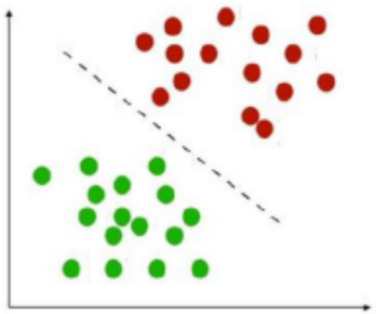
Reinforcement Learning

- No labeled dataset for training
- Interact with an environment
- Trying to learn optimal sequences of decisions

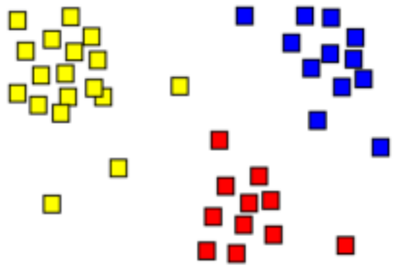
Regression

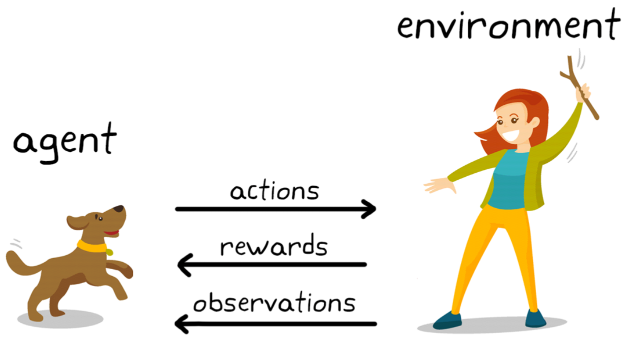


Classification



Clustering

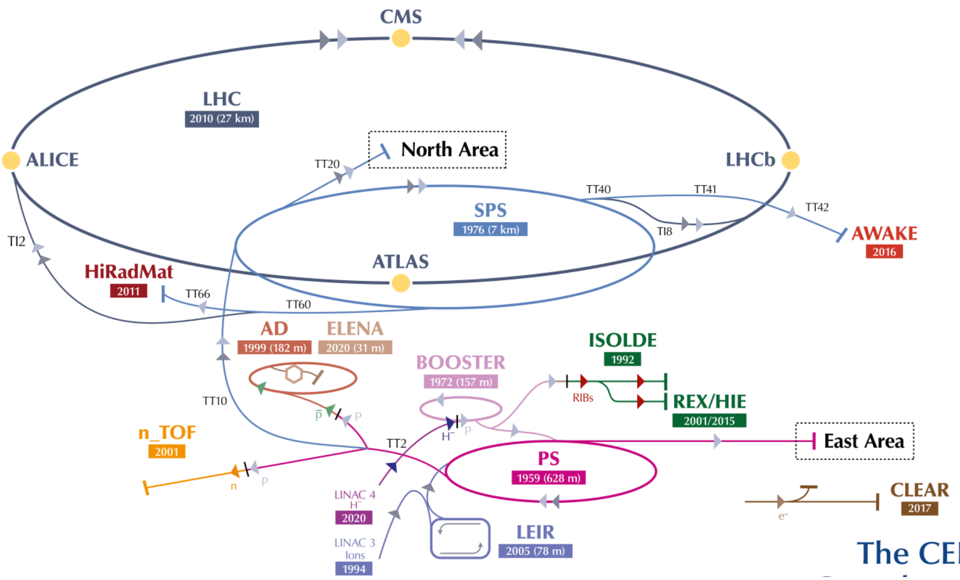




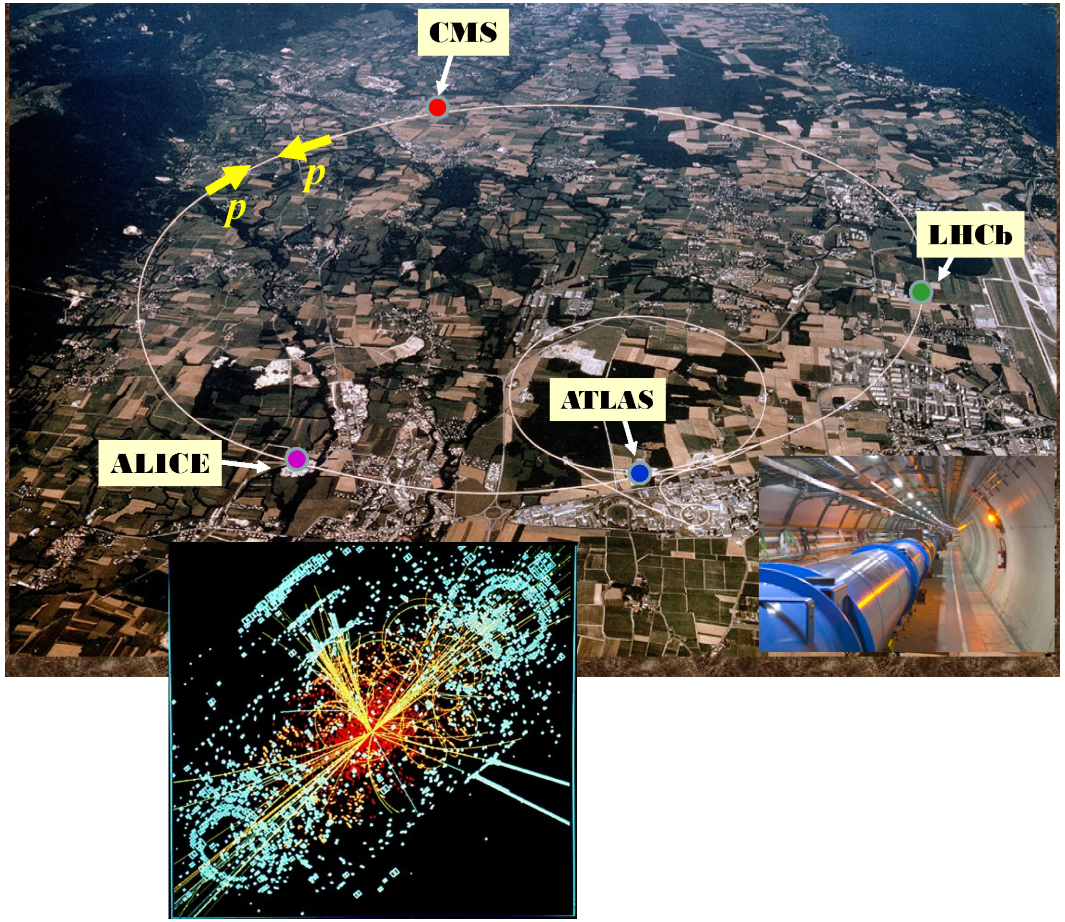
Large Hadron Collider at CERN:
Why do we need Machine Learning?

CERN and the Large Hadron Collider

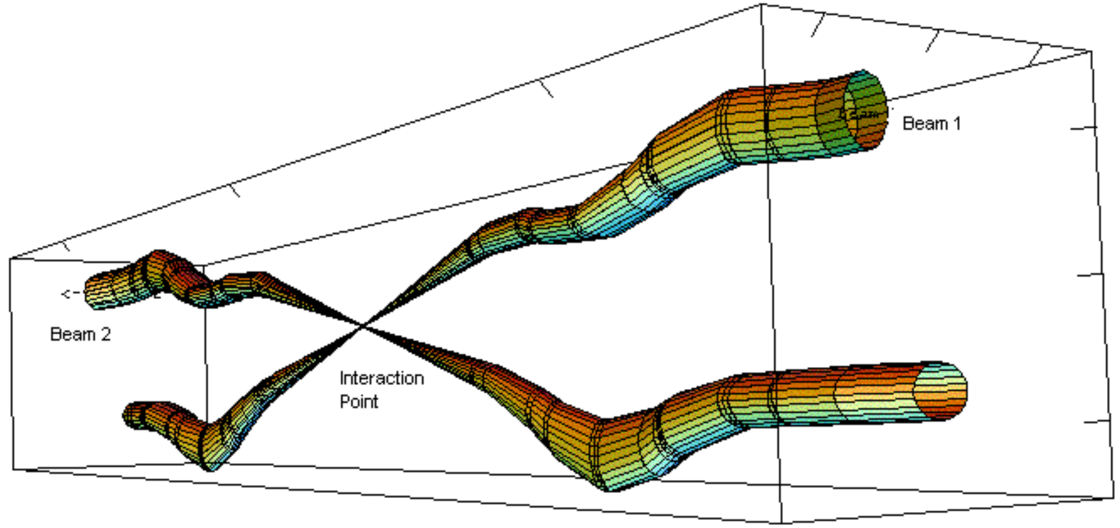
- World’s largest and highest-energy particle collider
- Powerful discovery tool: 13.6 - 14 TeV center of mass energy
- Goal: search for new elements and forces, answering open questions if fundamental research.



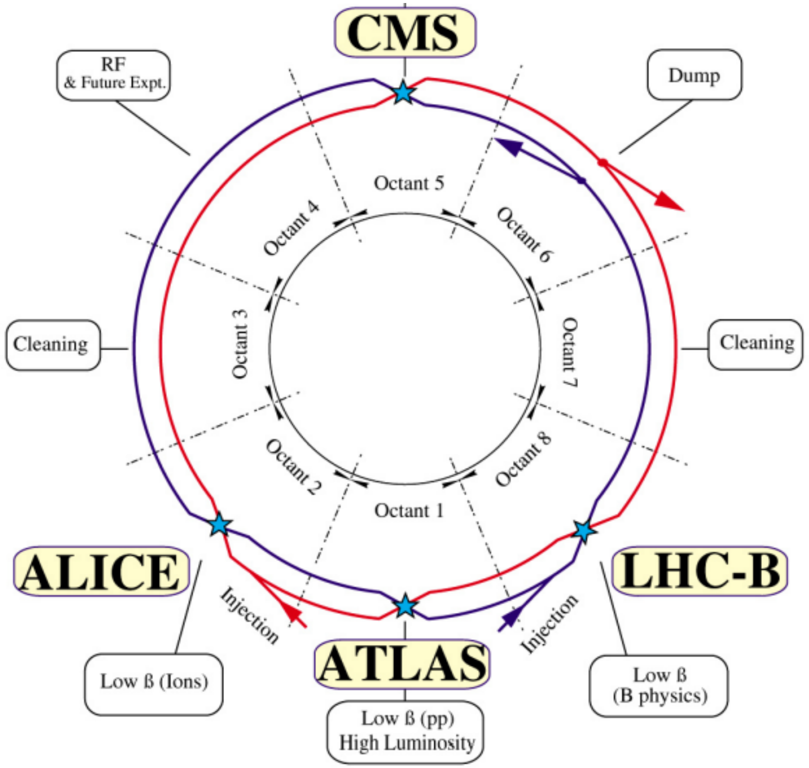
The CERN accelerator complex



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.

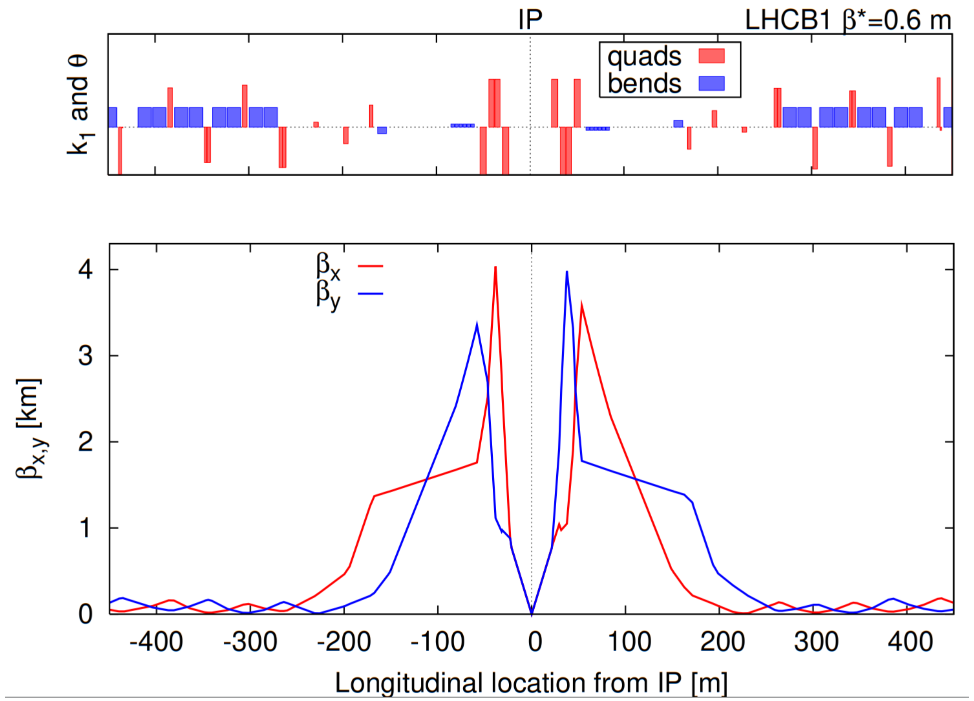
$$\mathscr{L} \sim \frac{f \cdot N^2}{4\sigma^2}$$

$$\sigma = \sqrt{\varepsilon\beta}$$

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

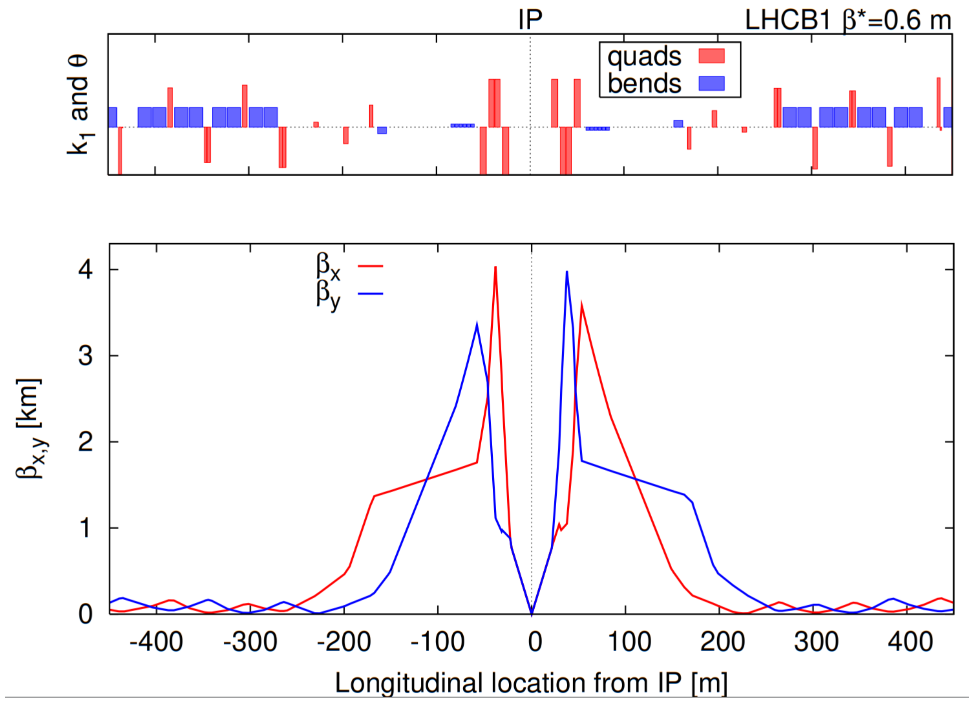
- **Luminosity:** maximize the number of collision events.

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

ε → Const

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

Optics ←



- Optics errors: **beta-beating** $\frac{\Delta\beta}{\beta} = \frac{\beta_{meas} - \beta_{model}}{\beta_{model}}$
  - Access to the magnets for direct measurements is not possible during operation.

→ Beam-based measurements and corrections of lattice imperfections.

Limitation of traditional optics control techniques

1. **Instrumentation faults** → **unreliable measurements** of beam properties and optics analysis

2. Corrections methods to compensate **measured deviations from optics design**
→what are the **actual magnet errors**?

3. Dedicated time to obtain **advanced optics observables** → how to **reduce the time** effort?

4. **Uncertainties** in the measured optics functions
→ **reduce the noise** without removing valuable information?

5. **Missing data points** due to the presence of faulty instrumentation
→ How to **reconstruct the missing data**?

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?

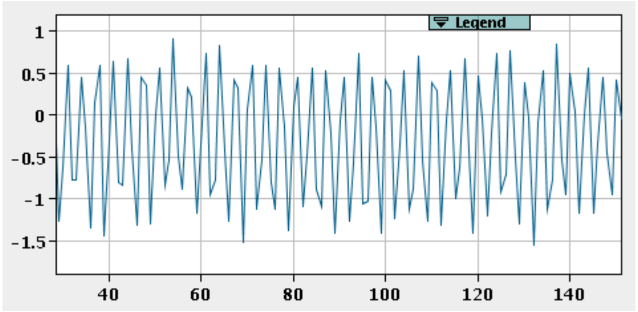
- ➢ large amount of optimization targets
- ➢ computationally expensive simulations
- ➢ direct measurements are not possible
- ➢ previously unobserved behaviour
- ➢ non-linear interacting sub-systems, rapidly changing environment.

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

Detection of instrumentation faults

Measuring the optics

Turn-by-turn beam position

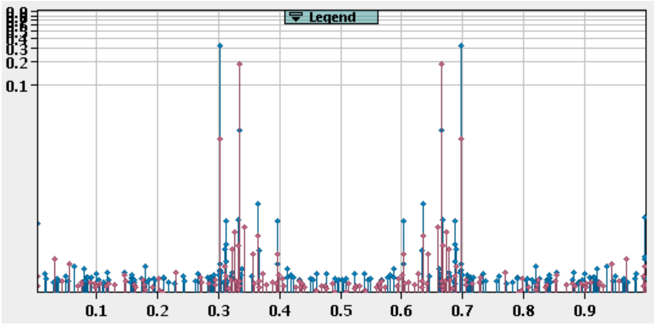


- Excite the beam to perform transverse oscillations.

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

Denoising (SVD)
Signal cuts

Spectrum



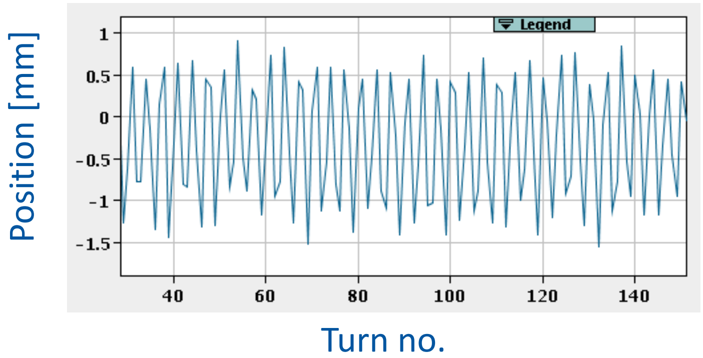
- Harmonic analysis using Fast Fourier Transform (FFT)

Semi-automatic and manual cleaning of outliers

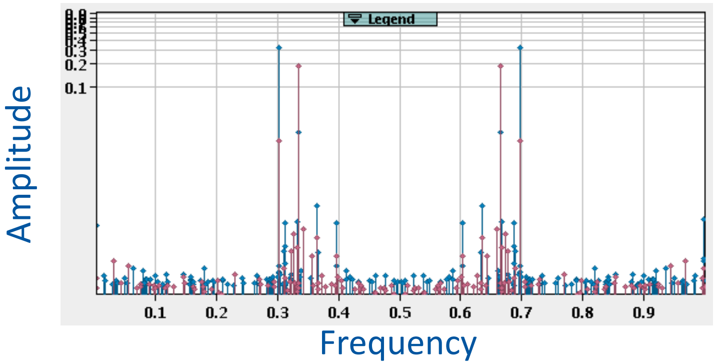
What are the limitations of traditional techniques?

Measuring the optics

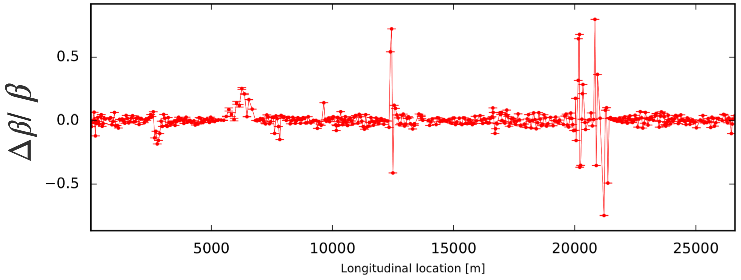
Turn-by-turn beam position



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)

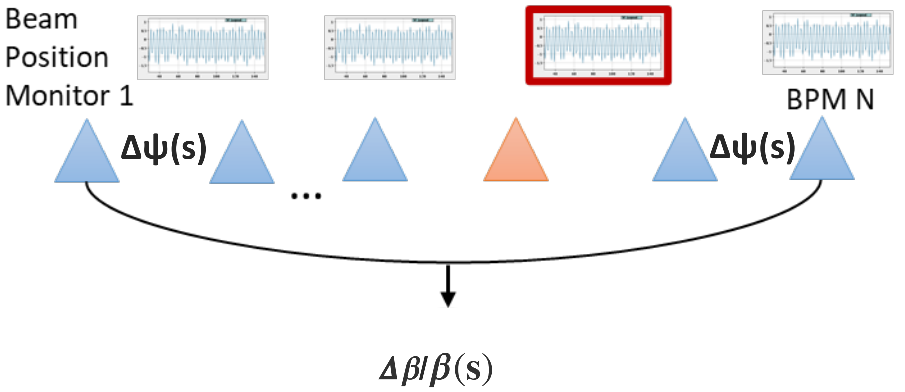
Semi-automatic and manual cleaning of outliers

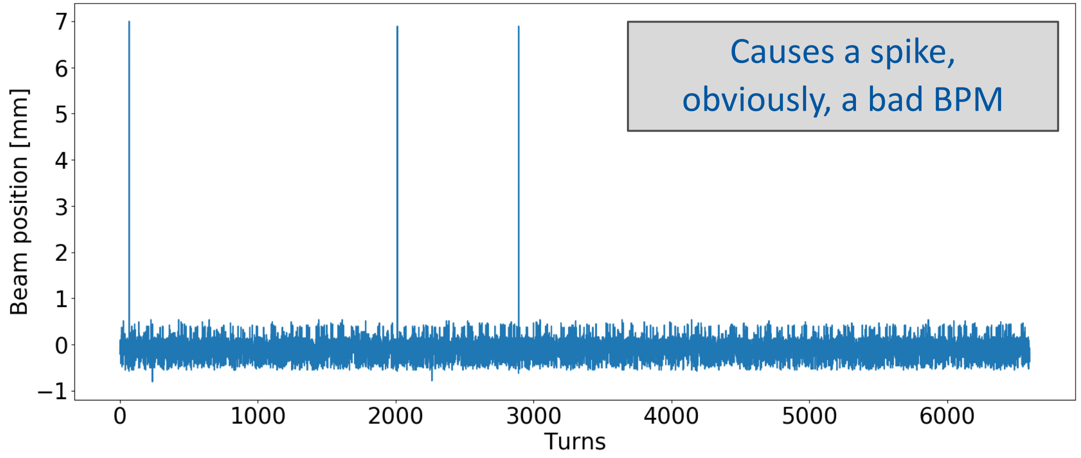
- Compute beta-beating and other optics functions

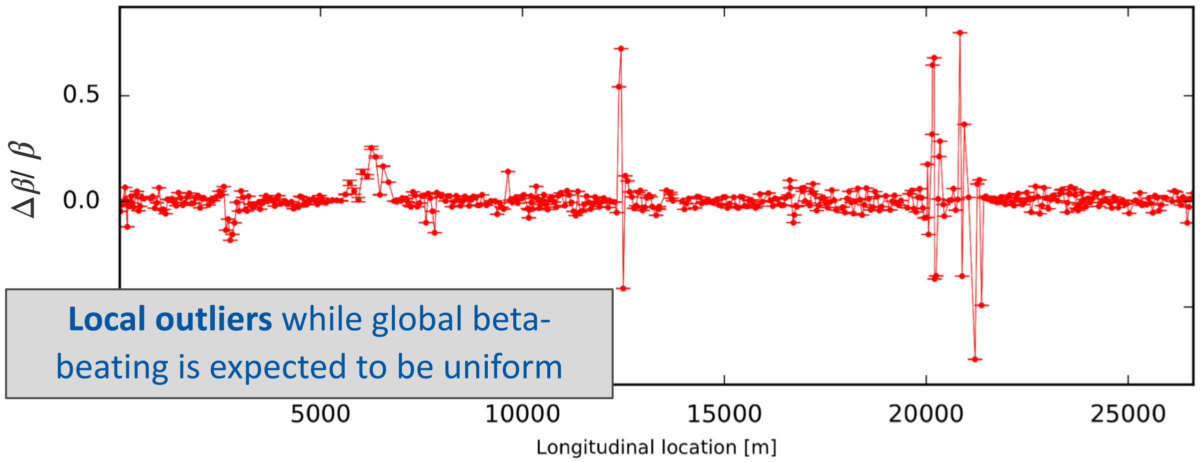
Unphysical values still can be observed

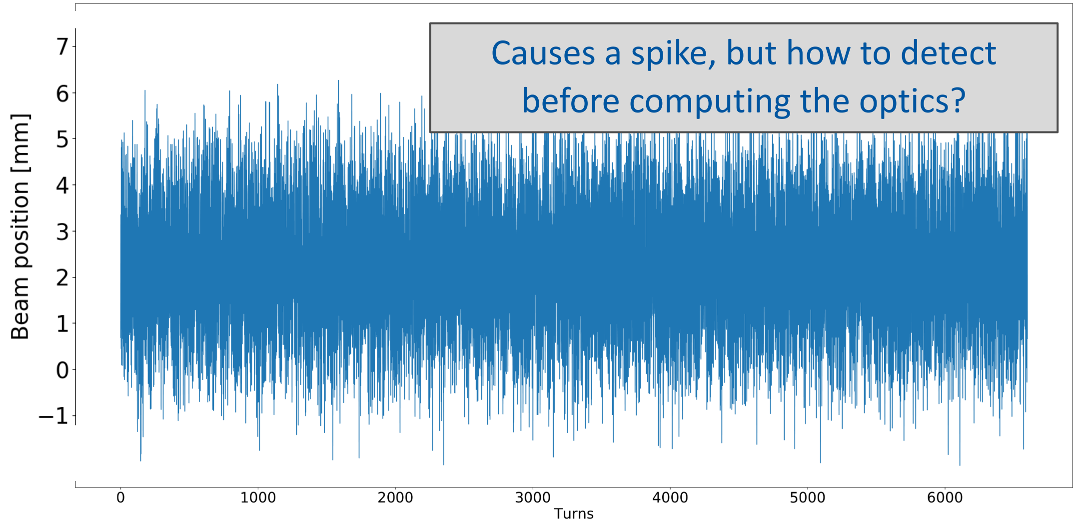
What are the limitations of traditional techniques?

Measuring the optics: challenges









Detection of faulty Beam Position Monitors

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

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

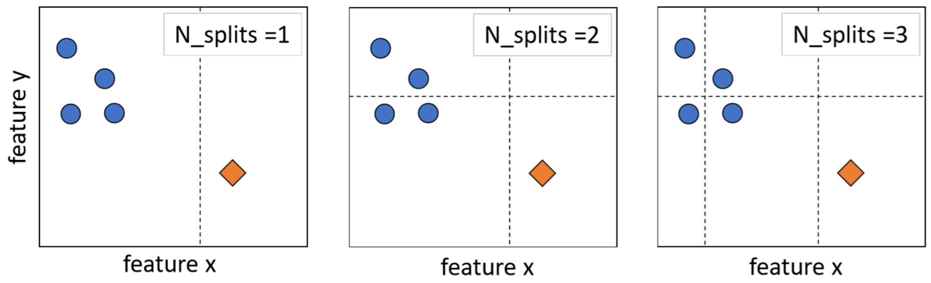
- **No ground truth → Unsupervised Learning** → ***Ability to identify anomalies without predefined thresholds or rules.***

- Applied clustering algorithms: DBSCAN, Local Outlier Factor, **Isolation Forest** (implemented with *Scikit-Learn)*

Isolation Forest Algorithm

- Forest consists of several **decision trees**
- **Random splits aiming to "isolate" each point**
- The less splits are needed, the more "anomalous"

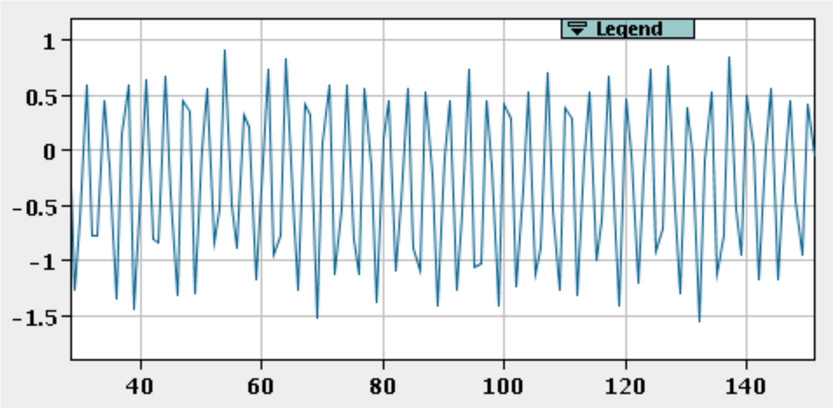
- **Contamination factor**: fraction of anomalies to be expected in the given data
  → First obtained empirically from the past measurements
  → Refined on **simulations introducing expected BPM faults**.

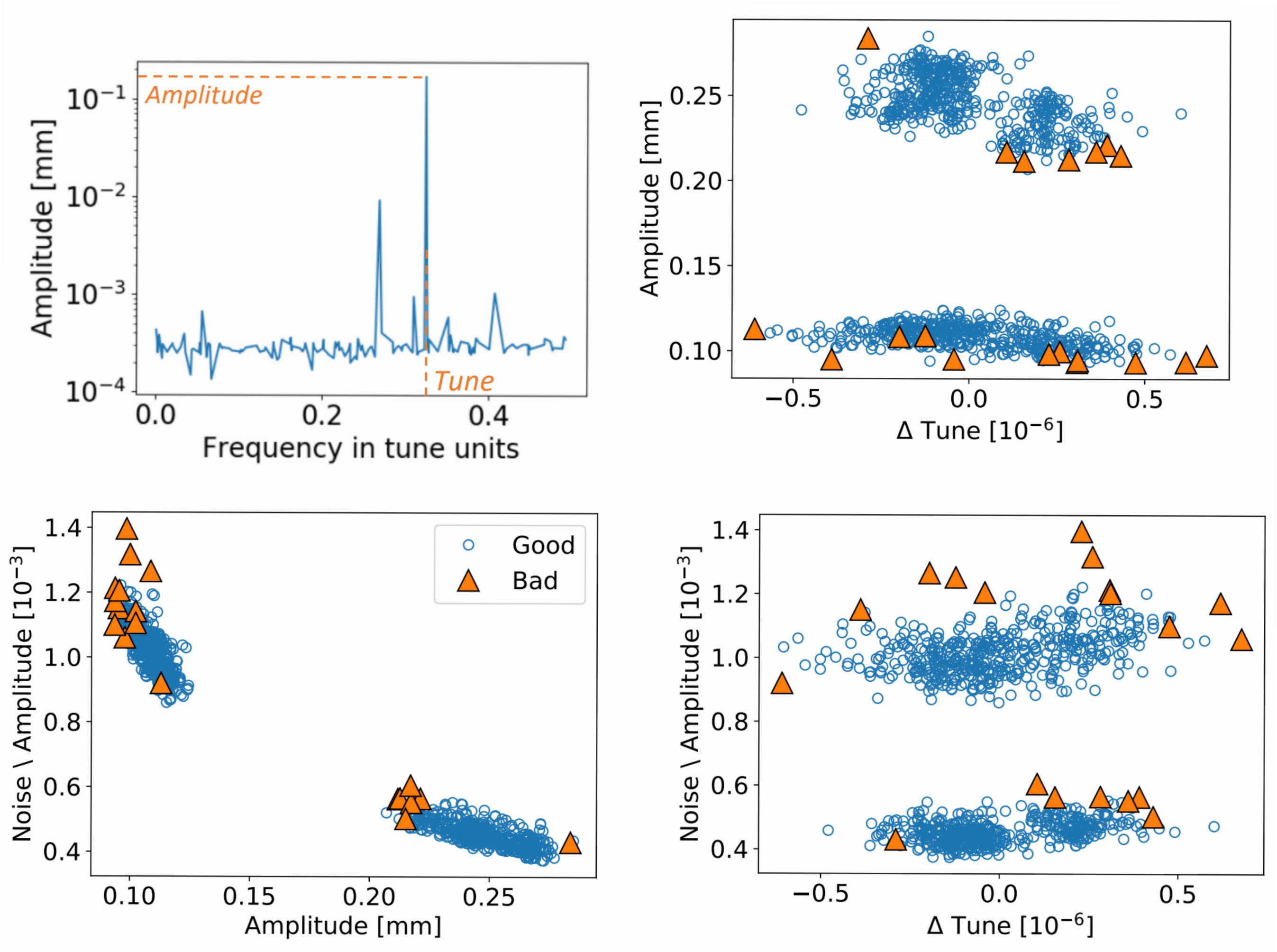


Conceptual illustration of Isolation Forest algorithm

Detection of faulty signal

Harmonic analysis of all BPMs





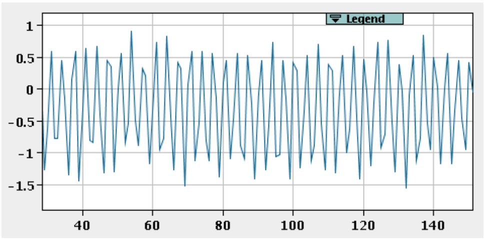
- **Input data: signal properties extracted by FFT as part of optics analysis**

  → No additional data handling needed.

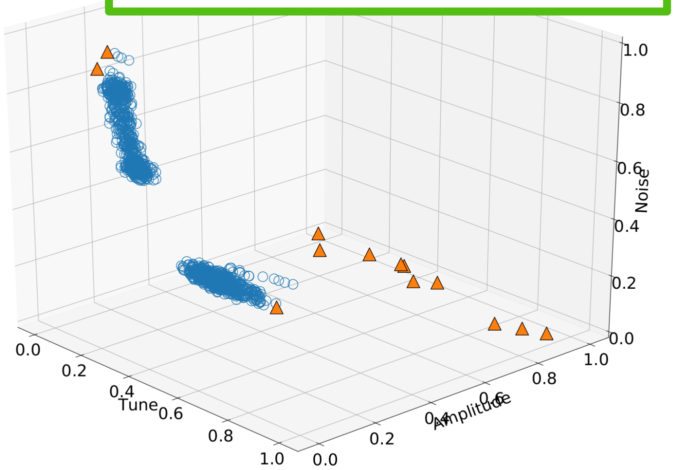
- Amplitude
- Main frequency (Tune)
- Noise (SVD analysis) ratio

Integrating ML into traditional beam optics analysis

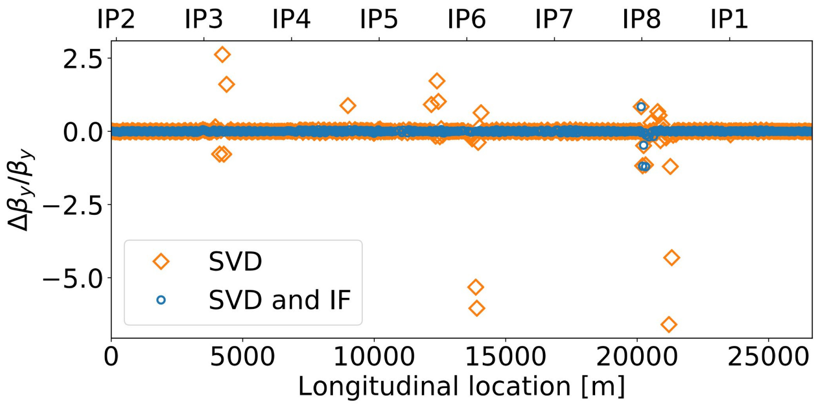
Harmonic analysis of all BPMs



Detection of faulty signal **prior to optics computation**



Avoid the appearance of erroneous optics computation



- Outlier detection based on combination of several signal properties
- Immediate results

→ no manual cleaning and repeated optics analysis

"Detection of faulty beam position monitors using unsupervised learning", Phys. Rev. Accel. Beams 23, 102805.

Are the BPMs really faulty? Beam Instrumentation checks

Advantages of IF-algorithm

- Ability to identify signal properties, which are relevant for classification between good and bad signal hints to possible instrumentation issues.

Information provided by cleaning algorithms

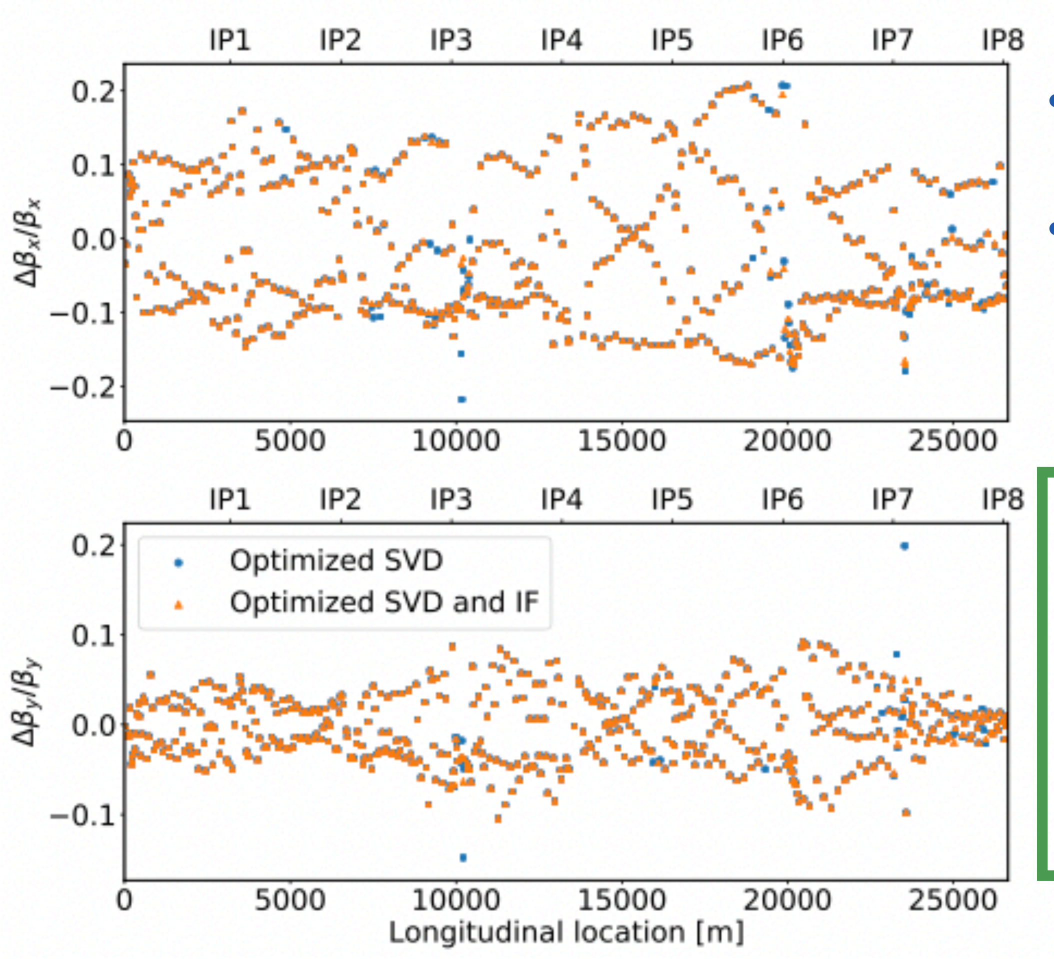
✓ Statistical analysis of data starting from 2018

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

- **Note: for the LHC with ~ 1000 BPMs, it is more critical to eliminate all faulty signal than to keep all good BPMs in the analysis data**
- Verifying false positive BPMs: removed as trade-off for detecting actual faults.

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

LHC commissioning 2022: operational results



- Instant faults detection instead of offline diagnostics.
- Full optics analysis is possible directly during dedicated measurements session instead of iterative procedure of cleaning and analysis.

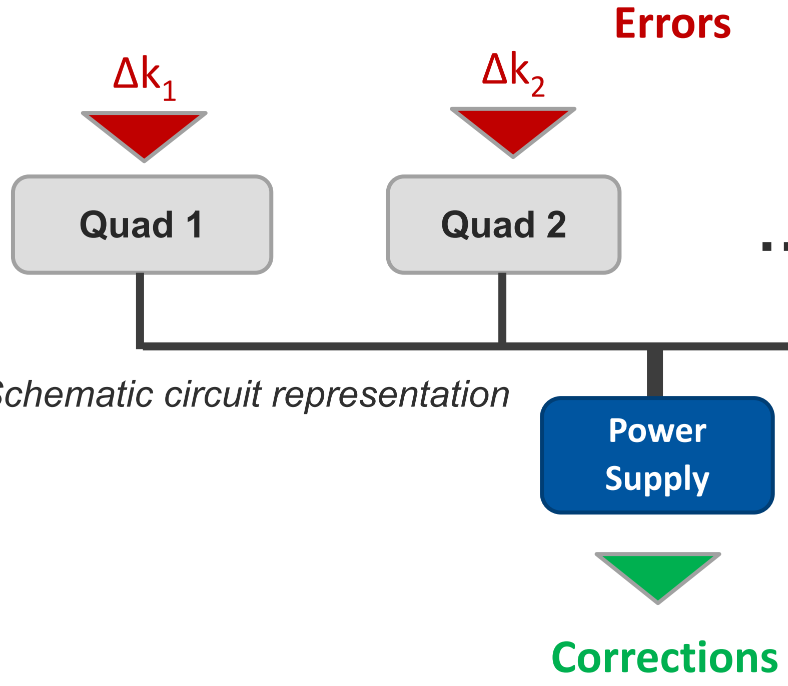
LHC Commissioning 2022:

✓ BPM upgrades considering the findings by cleaning methods.

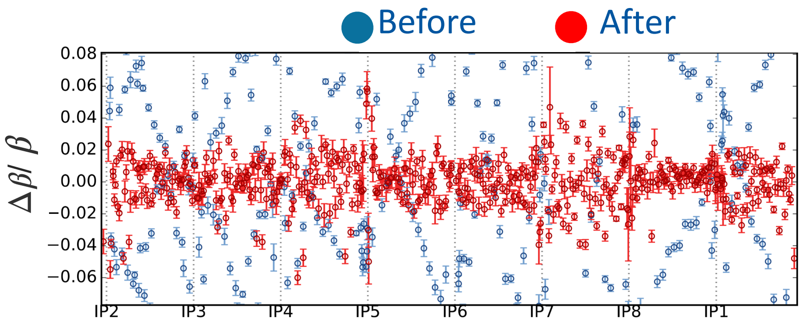
✓ Verification of **updated cleaning settings** against false positives: no negative impact on optics

Machine Learning for magnets errors reconstruction and beam corrections

Correcting the optics



Schematic circuit representation



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

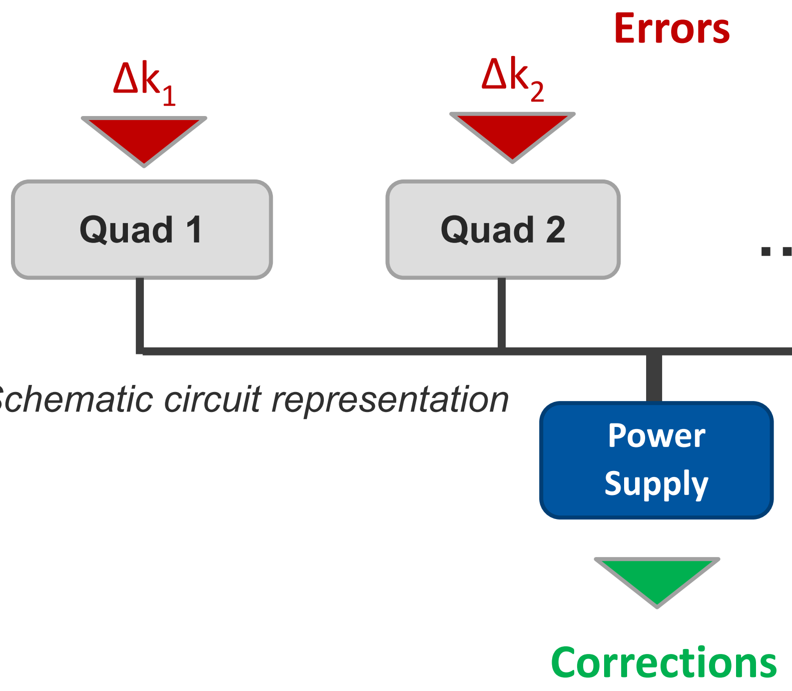
$$\frac{\Delta\beta}{\beta} = \frac{\beta_{meas} - \beta_{model}}{\beta_{model}}$$

- 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



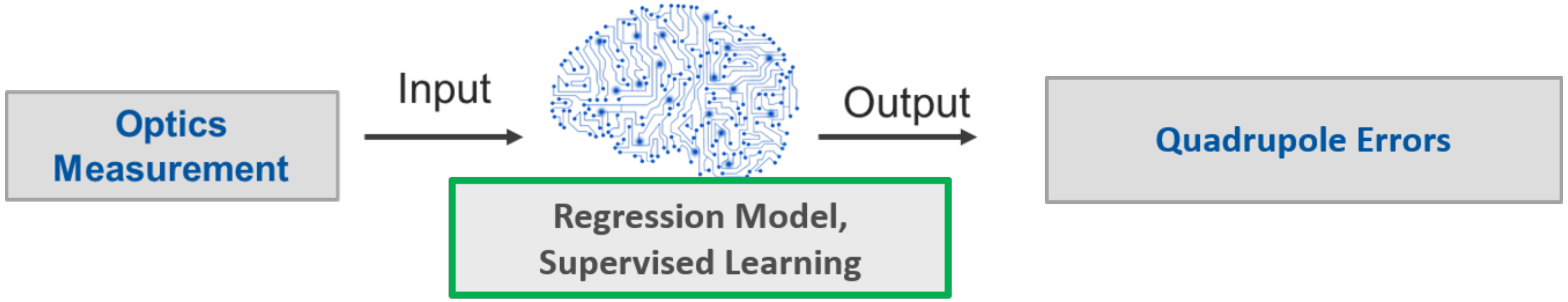
Schematic circuit representation

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

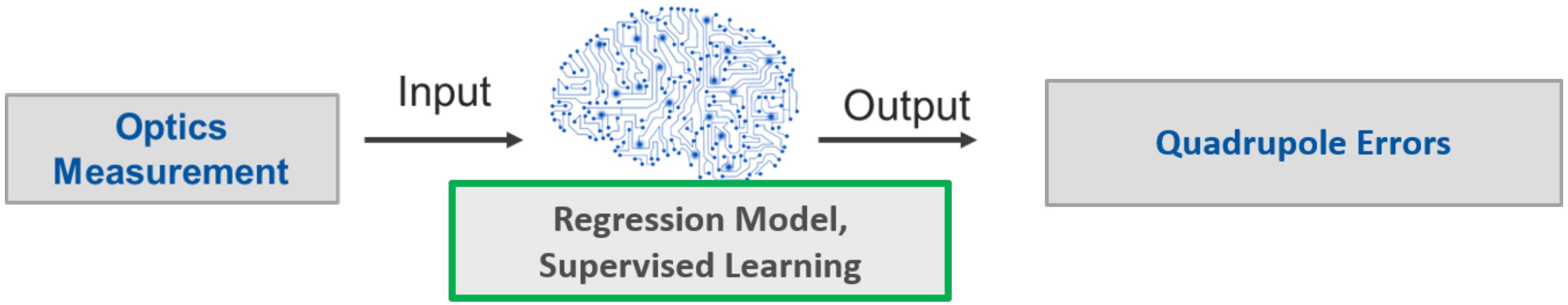
➢ What are the **actual errors of individual quadrupoles?**
➢ How to obtain the **full set of errors in one step**?

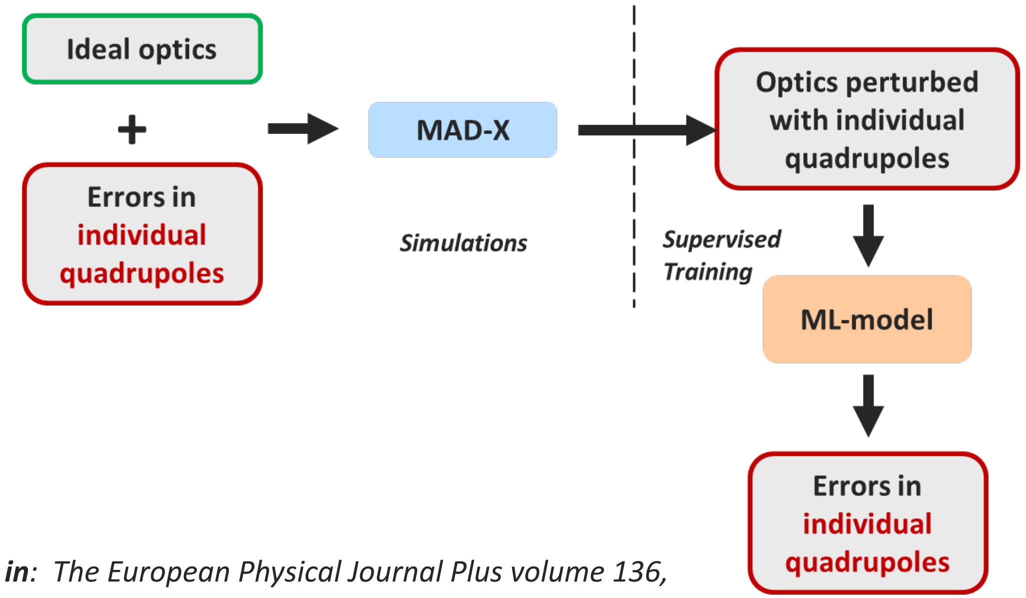
What are the limitations of traditional techniques?

Estimation of quadrupole errors



Estimation of quadrupole errors





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

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**

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

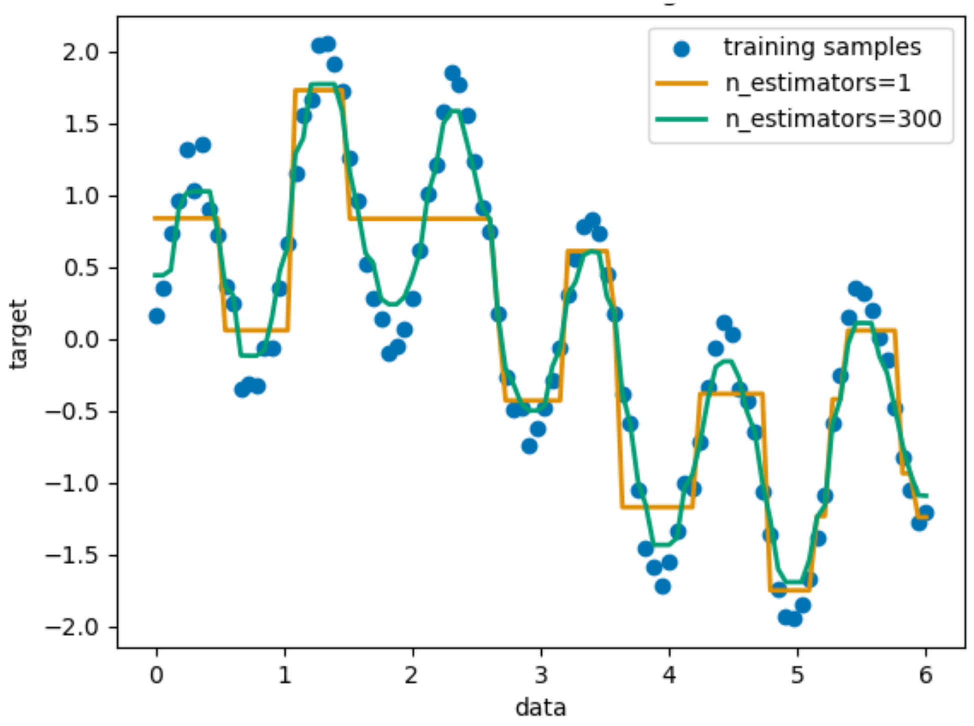
Random Forest Regression

Supervised Learning approach:

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

<u>Decision Trees:</u>

- 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: $H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y_i - c_m)^2$



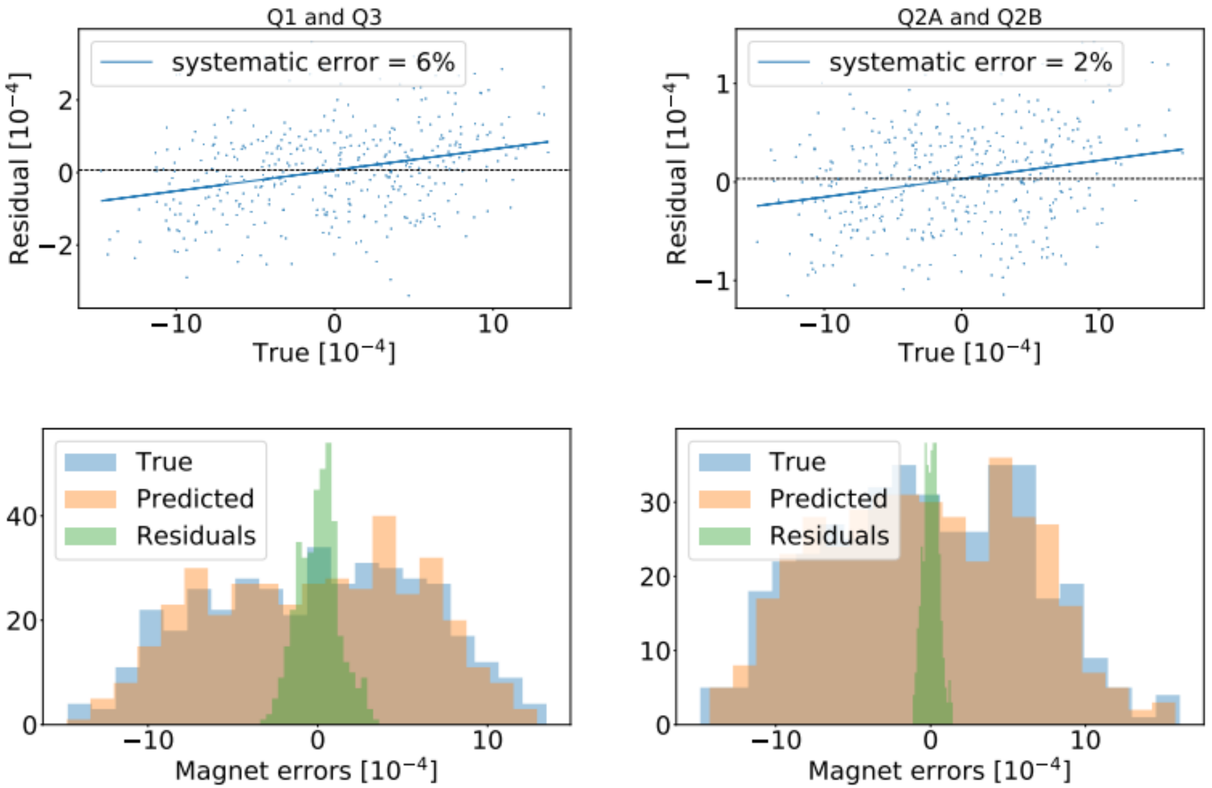
<u>Random Forest:</u>

- 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

Estimation of quadrupole errors: simulations

Simulations: **true magnet errors are known**

→ directly compare prediction to simulated data → **residual error**



Measurements: magnet errors are unknown!

How well can we correct the optics with predicted errors?

LHC commissioning 2022: beam optics corrections

Optics Corrections path in the LHC

1. Beam optics measurements

2. **Reconstruct the magnet errors**

3. Propagate the errors within the region (simulations)

4. Compare with measurements

5. Apply the reconstructed quadrupole strength errors with opposite sign —> optics corrections!

Note: using traditional techniques, this procedure is performed per Interaction Region, for Beam 1 and Beam 2

LHC commissioning 2022: beam optics corrections

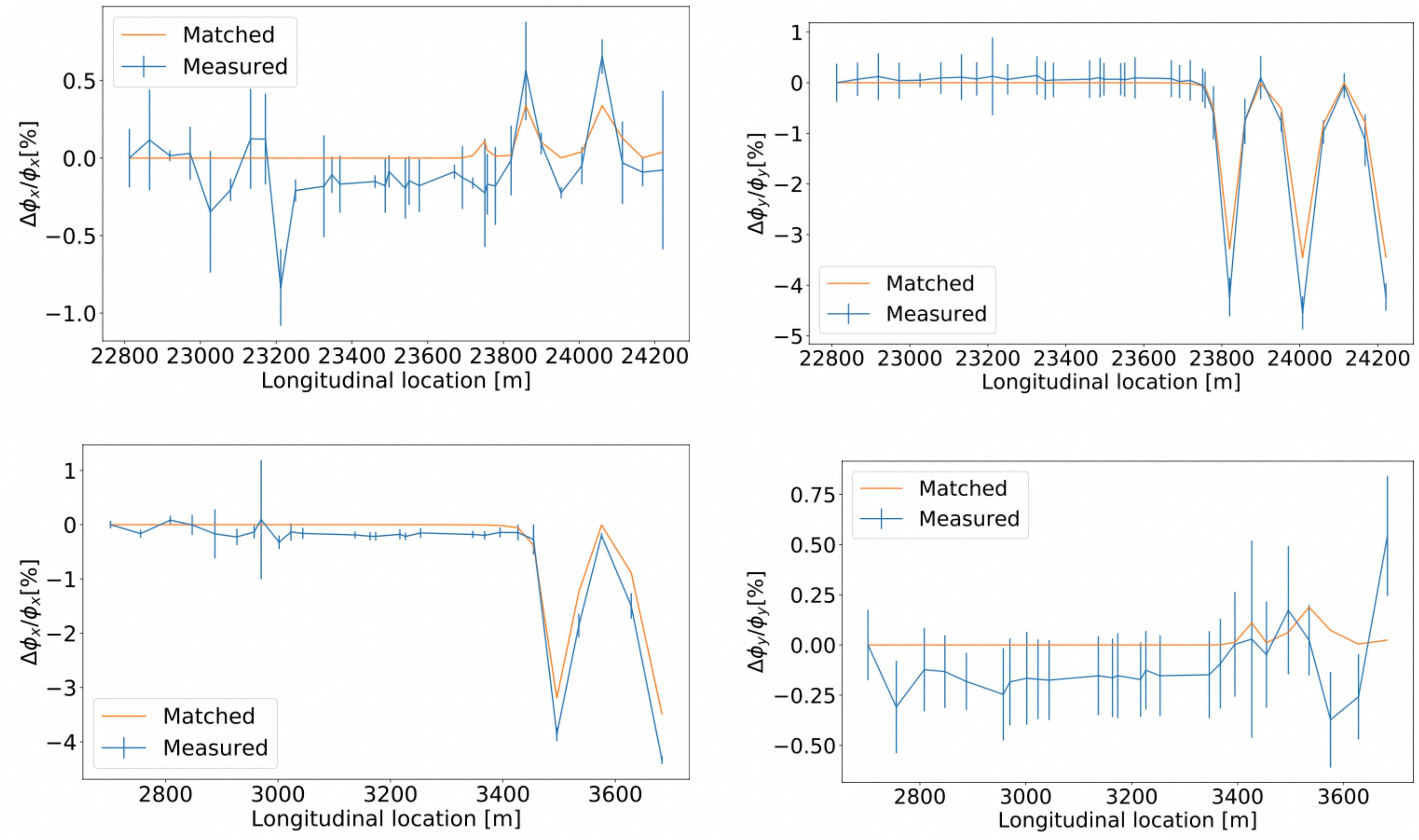
Optics Corrections path in the LHC

1. Beam optics measurements
2. **Reconstruct the magnet errors**
3. Propagate the errors within the region (simulations)
4. Compare with measurements
5. Apply the reconstructed quadrupole strength errors with opposite sign —> optics corrections!

Predict all errors using ML-model trained on simulations!

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.**

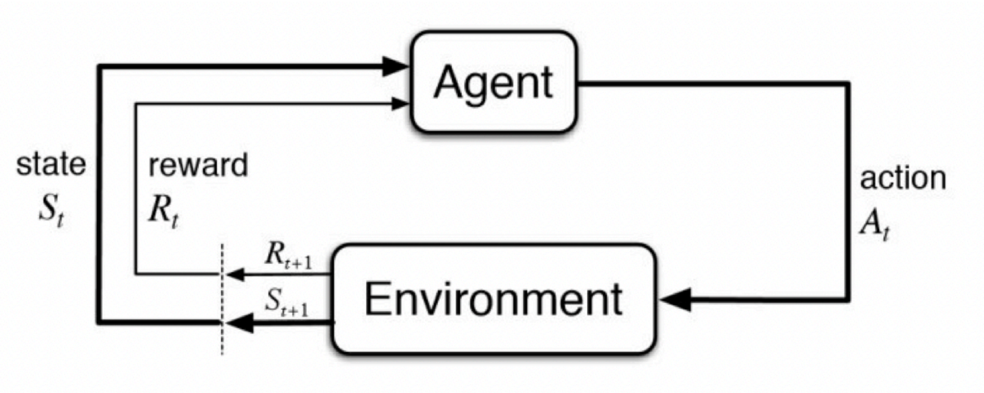
→ **Potential to save operation time!**

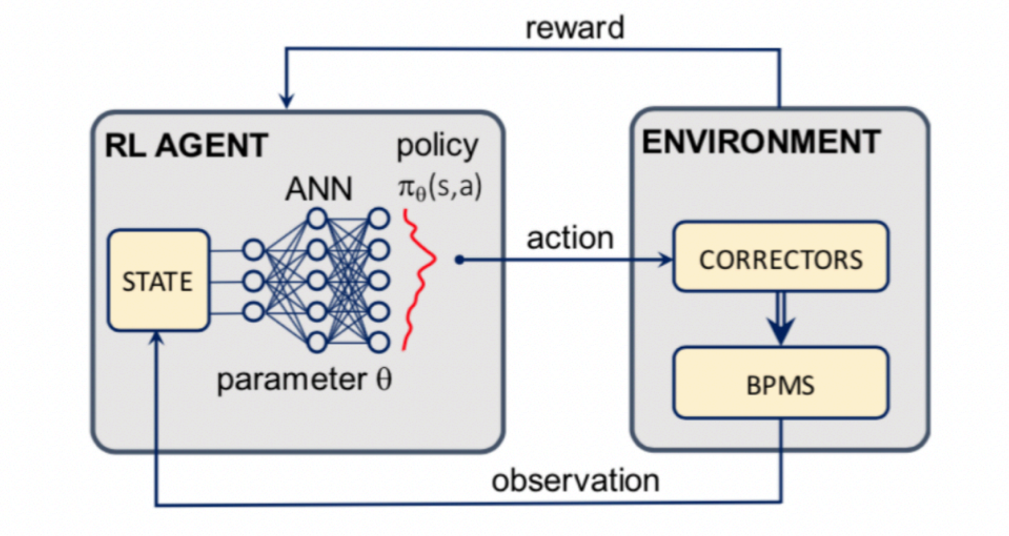
E.Fol et al.,"Experimental Demonstration of Machine Learning Application in LHC optics commissioning",IPAC'22 MOPOPT047

Look into the future: Optics control in HL-LHC

Reinforcement learning - based local corrections

- Uses the previously presented approach **to learn LHC model from simulated data**





Based on V.Kain et al., "Sample-efficient RL for CERN accelerator control"

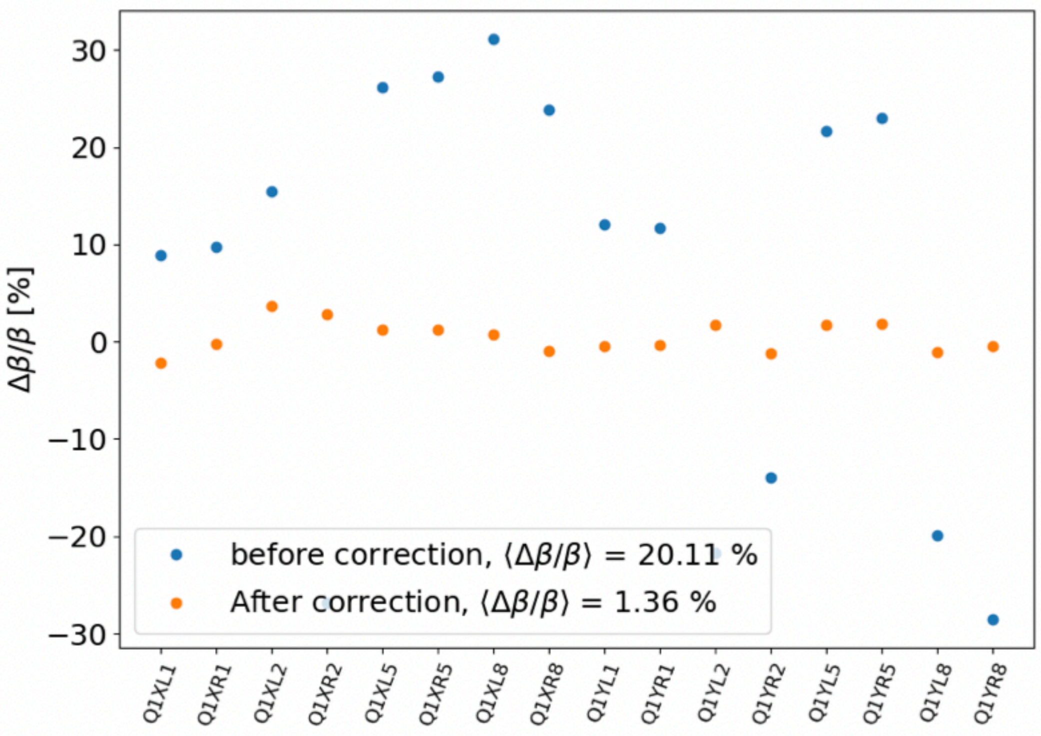
- Environment = Surrogate model of HL-LHC lattice
- Reward = Average beta-beating in IRs
- State space = Quadrupole strengths (only triplet magnets for now)
- Action space = Correctors settings

Optics control in HL-LHC studies

Implementation

- Introducing local magnetic errors in triplet magnets in one of the LHC Interaction Points
- **RL algorithms** implementations based on **OpenAI**
- **PyTorch** for the training of critic networks

Work by Hector Garcia Morales, BE-ABP



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)

De-noising of optics measurements

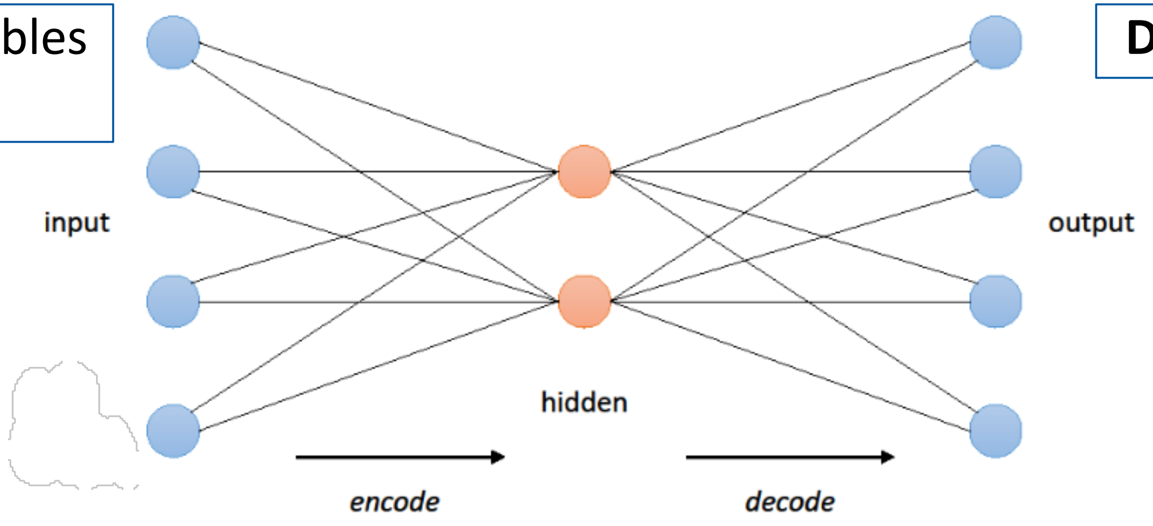
De-noising of optics measurements

- Uncertainties in the measured optics functions → **"noise"** →

Noise in the measurements degrades the performance of corrections techniques

Autoencoder Neural Network

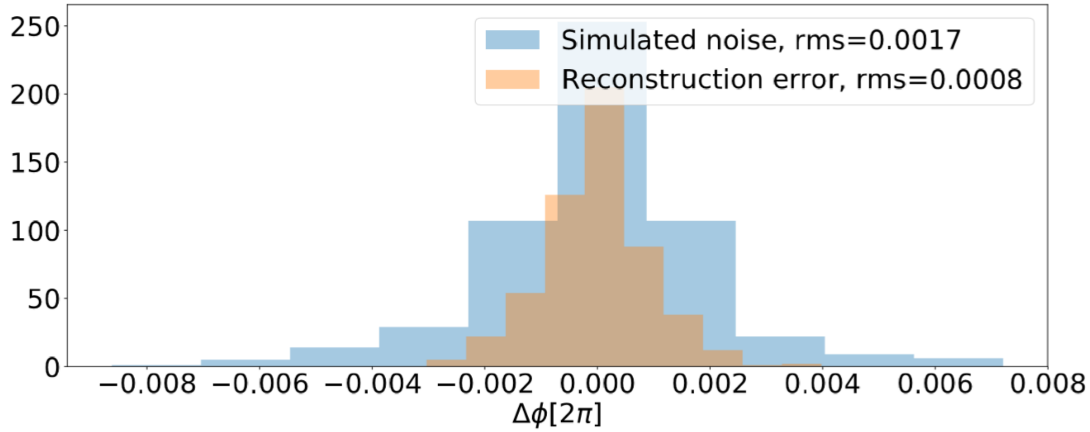
Simulated optics observables
+ noise



Denoised optics

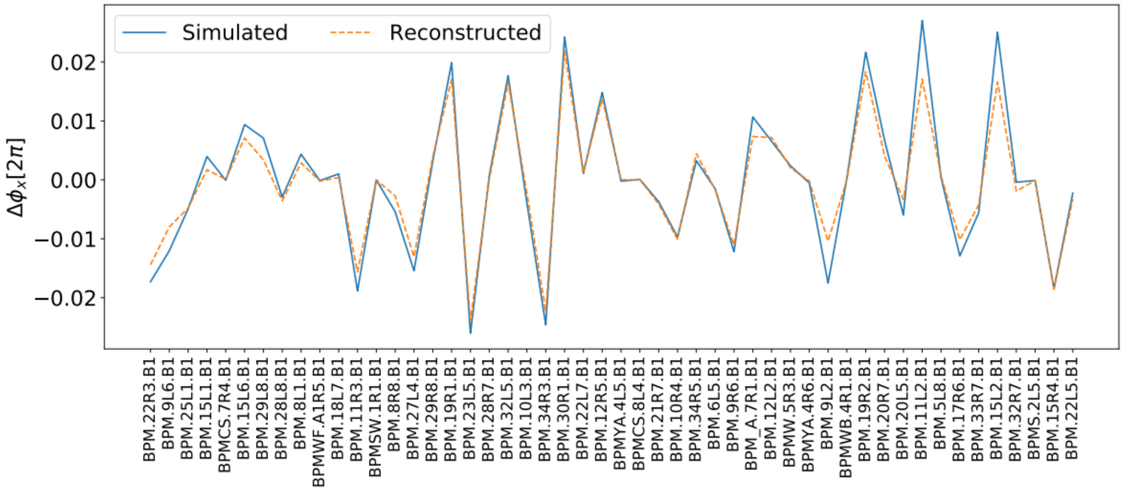
De-noising of optics measurements

Simulated data: Noise Reduction



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

Simulated data: Reconstruction



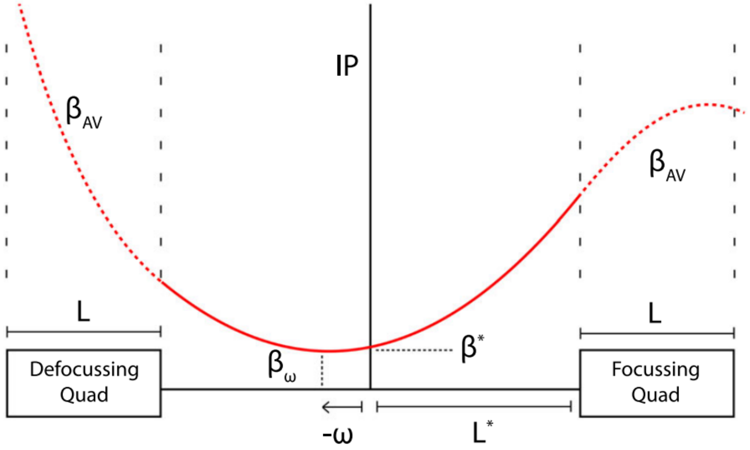
✓ Reliable reconstruction after denoising

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

Reconstruction of advanced beam optics observables

Reconstruction of β-beating in Interaction Regions

- Special technique, "k-modulation" is performed to obtain the measurements: modulation of quadrupole current **—> Time consuming!**



- β-functions at the location of the BPMs next to the IPs are needed for local correction computation
- Accuracy varies depending on tune measurement uncertainty, magnet errors and β^* settings.

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

Phase advance measurements*

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

Input

Output

Regression Model, Supervised Learning

β-function at IPs

Simultaneously for beam 1 and 2

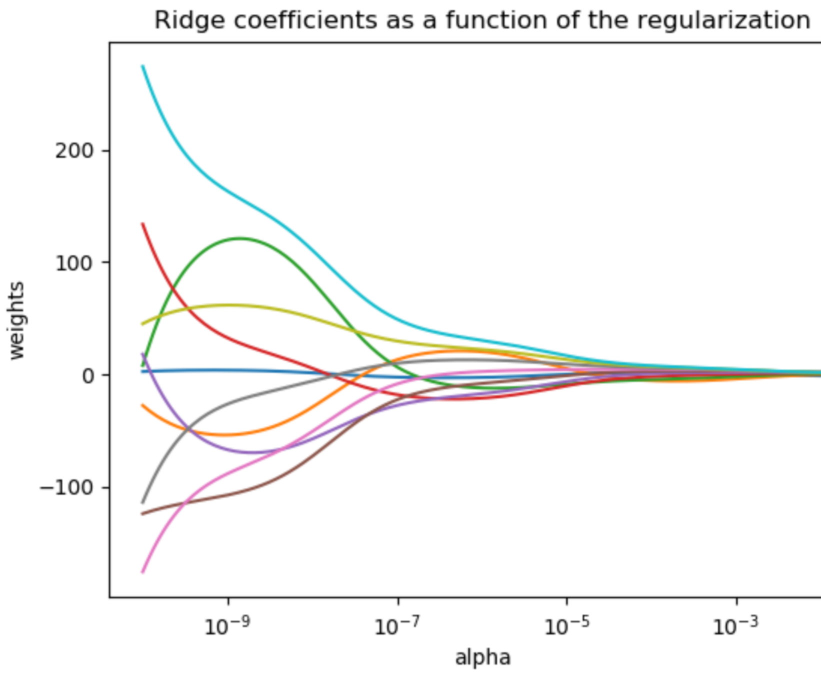
Regression Models

- 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.**

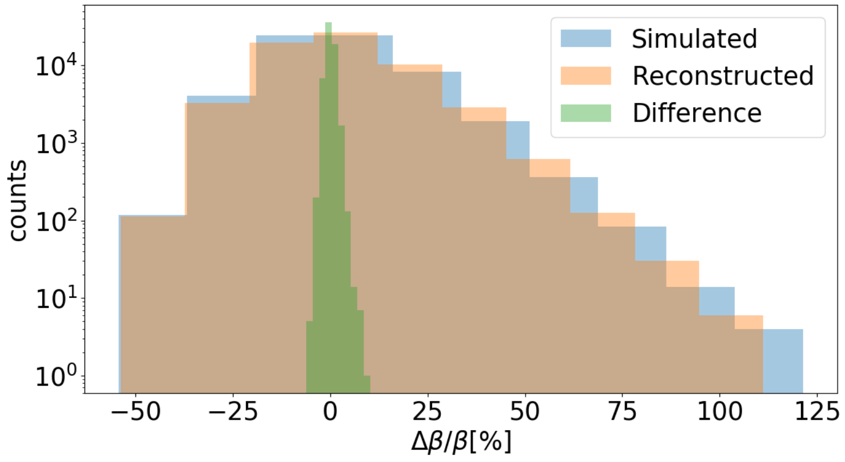
→ **Regularization** places constraints on the model parameters (weights)

- 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.



Reconstruction of β-beating in Interaction Regions

Simulations



Reconstruction error:

$$\frac{\beta_{simulated} - \beta_{reconstructed}}{\beta\_simulated} = \mathbf{1\%}$$

LHC Commissioning 2022, comparison to traditional technique

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

- β-functions next to Interaction Points **within a few seconds vs. several minutes for k-modulation**
- **Difference between prediction and measurement** : 5 % for β^* = 30 cm.

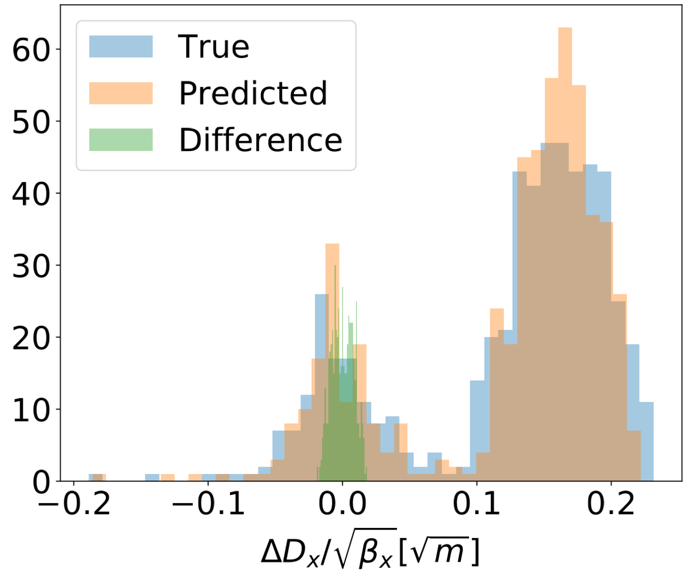
Reconstruction of horizontal dispersion

- An important optics observable (independent from BPM calibration) and is included into the computation of global optics corrections.
- Computed by acquiring turn-by- turn data from **several beam excitations, shifting the momentum.**

➡ Reconstruct **directly from phase advances on momentum,** using **ML-model trained on simulations.**

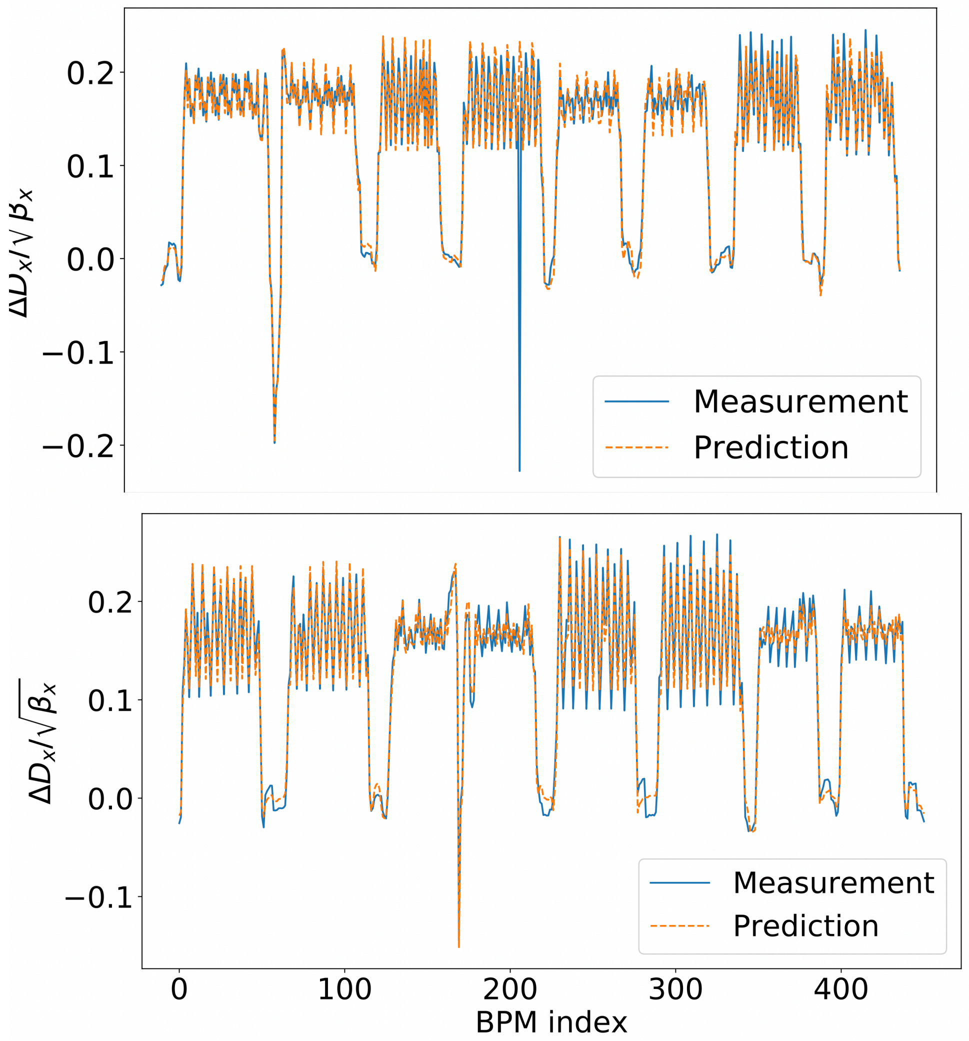
- **Input**: simulated phase advance deviations given noise
- **Output**: normalized dispersion $\Delta D_x / \sqrt{\beta_x}$
- Using **linear regression model**: 10 000 samples

***Simulation** example: Beam 1*



Operational results from LHC Commissioning 2022

- ✓ Simultaneous reconstruction of normalised dispersion in beam 1 and beam 2 requires only a few seconds.
- ✓ The relative **error of prediction is 5%** (beam 1) and **7%** (beam 2).
- **Potential speedup of machine commissioning for the same performance.**



Machine Learning as a tool for accelerator design

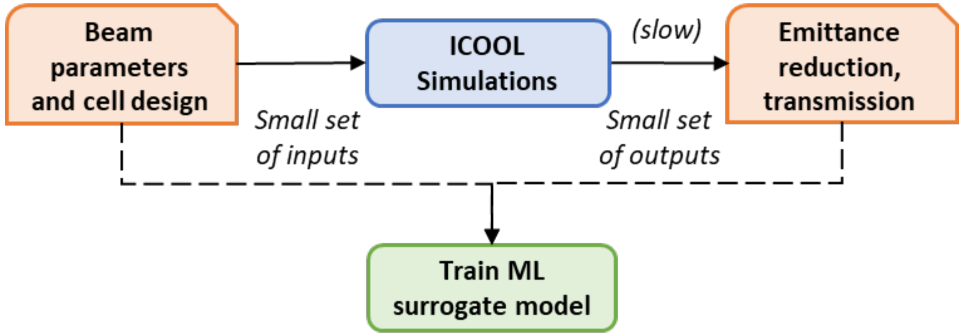
ML in Muon Collider Design Studies

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

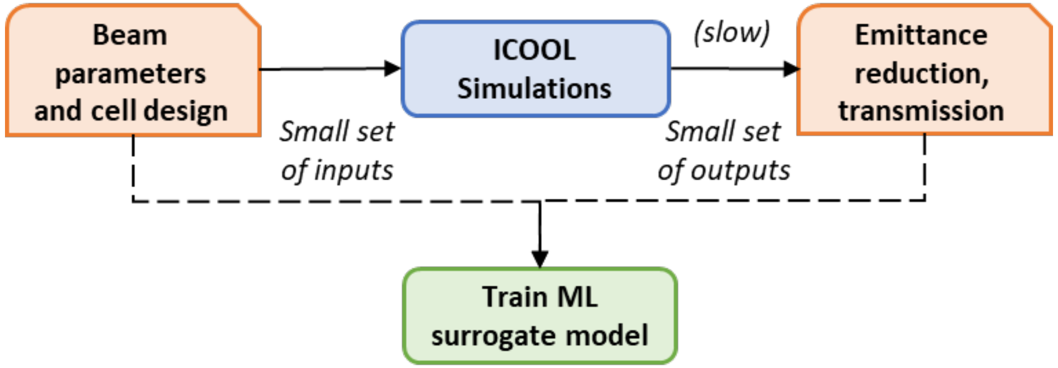
- Extending existing simulation frameworks towards automatic, fast executing optimization.
- **Exploring application of Supervised Learning**
  → **surrogate models**



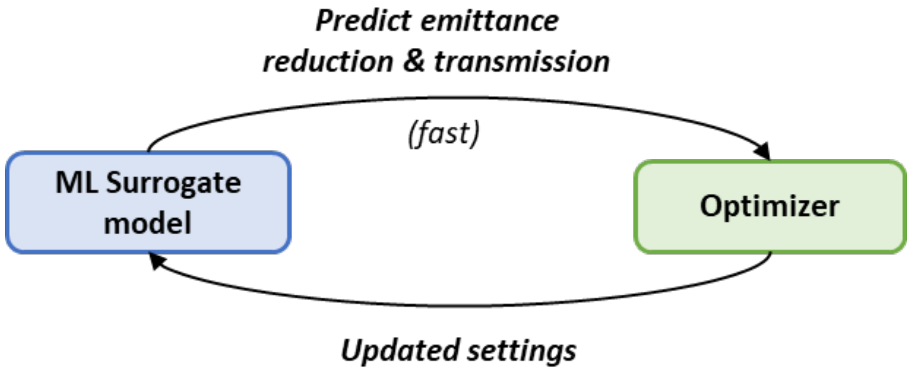
[1]: https://muoncollider.web.cern.ch

ML in Muon Collider Design Studies

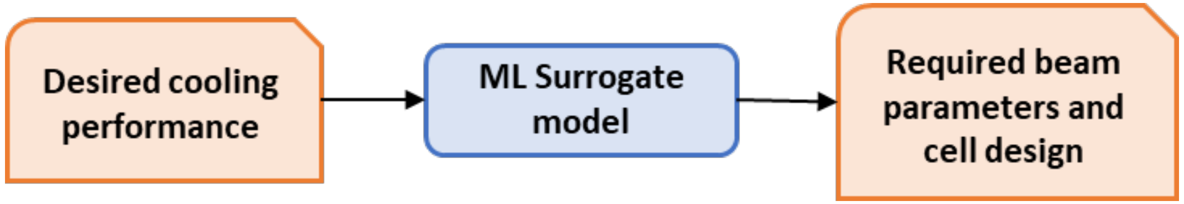
- **Exploring application of Supervised Learning**
  → surrogate models



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

Conclusions

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
- Denoising of optics measurements → Increasing the quality of the measurements
- Reconstruction of optics observables → Additional observables without dedicated 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
- Denoising of optics measurements → Increasing the quality of the measurements
- Reconstruction of optics observables → Additional observables without dedicated 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 (muon cooling).

Summary: Where can we use ML in accelerators?

Detection of instrumentation failures

Beam control and lattice imperfection corrections

Optimization and operation automation

Virtual Diagnostics

- Defining a **narrow task** (optimization of specific parameters rather than the entire machine)
- **Performance measure** of selected model (beam size, pulse energy, …)
- e.g. when no analytical solution is available, rapidly changing systems, no direct measurements are possible.

Important to identify where ML can surpass traditional methods

How much effort is needed to implement a ML solution? Is appropriate infrastructure for data acquisition available? Enough resources to perform the training?

Cat!

Thank you 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.

CERN

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.

Finding optimal (supervised) model: **Cross-Validation**

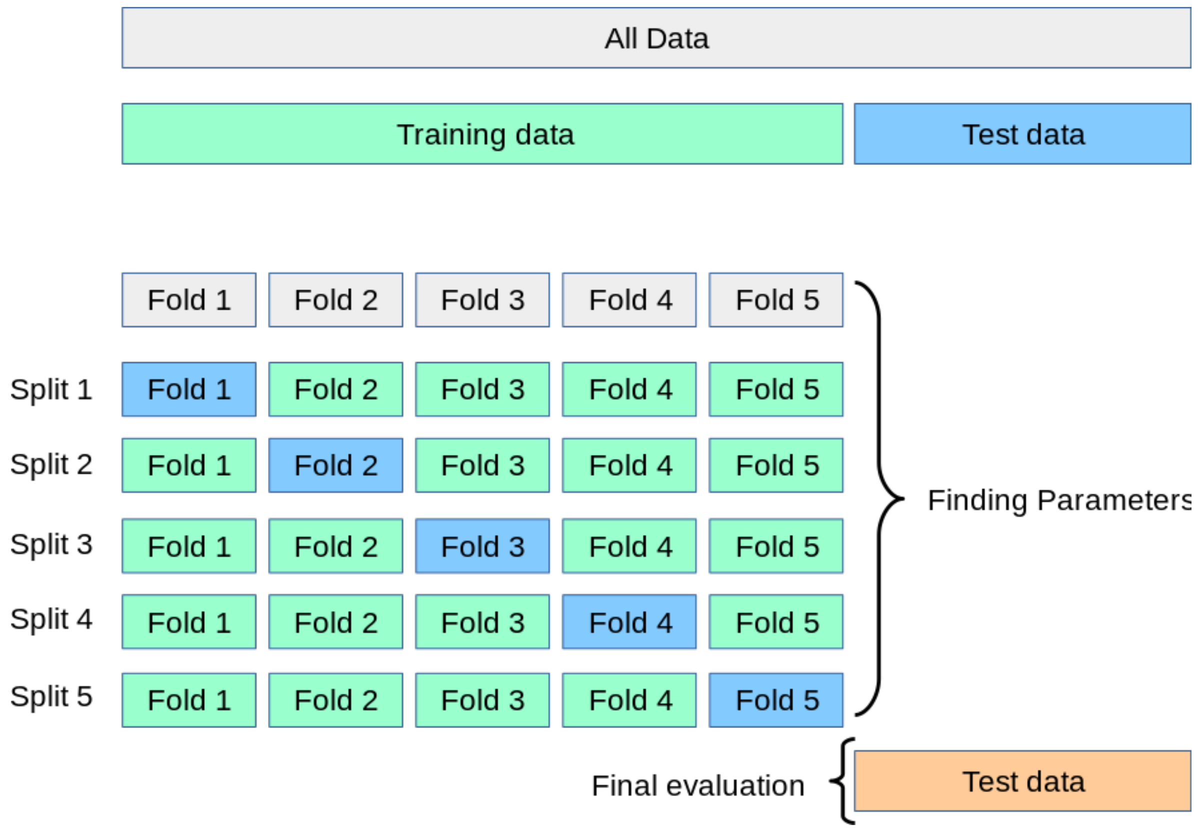
General rule: split available data

- **Train set** —> find model parameters which minimise the total loss between prediction and **known true output**
- **Test set** —> how does the model perform on **unseen data**?

How to find optimal model parameters?

Grid-search of parameters using cross-validation:

- Split train set into **k folds**
- Each fold is spitted into train and validation subsets
- In each fold, e.g. 80% of samples for training, 20% for validation
- **Repeat k times**, take the mean as model's performance score
- ➡ Robust estimation of performance for the different combination of model parameters.



https://scikit-learn.org/stable/modules/cross_validation.html