Machine Learning for Beam Optics Control

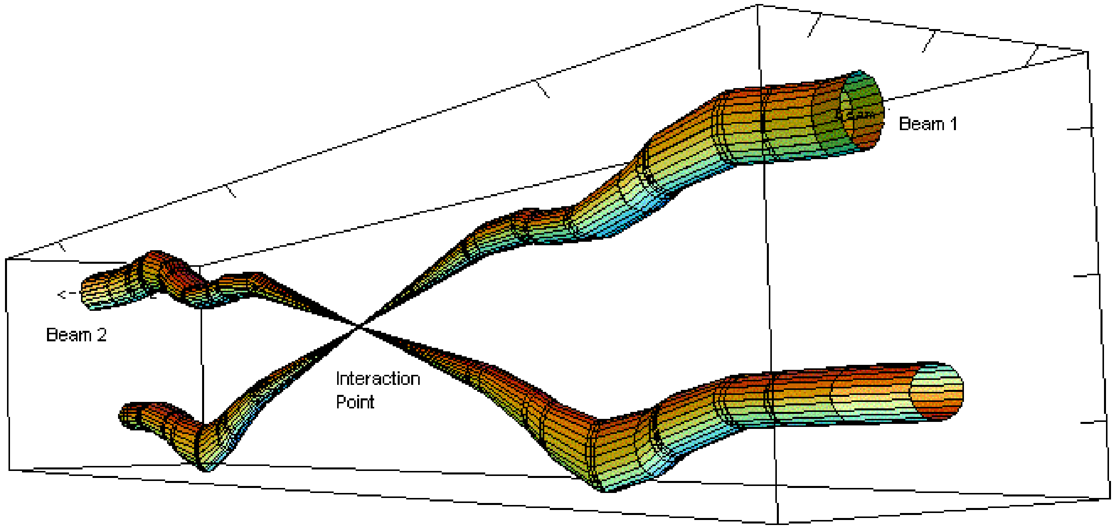
Elena Fol
BE-ABP-LAF

Many thanks to Giuliano Franchetti, Rogelio Tomás García and OMC team

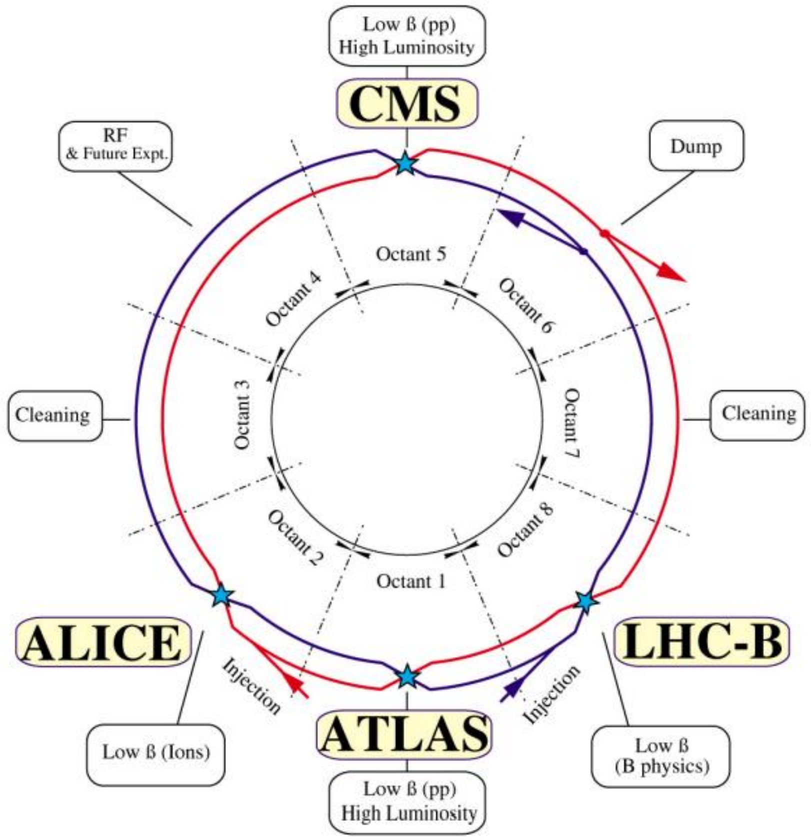
11th March, 2022

Beam Optics Control at the LHC

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.

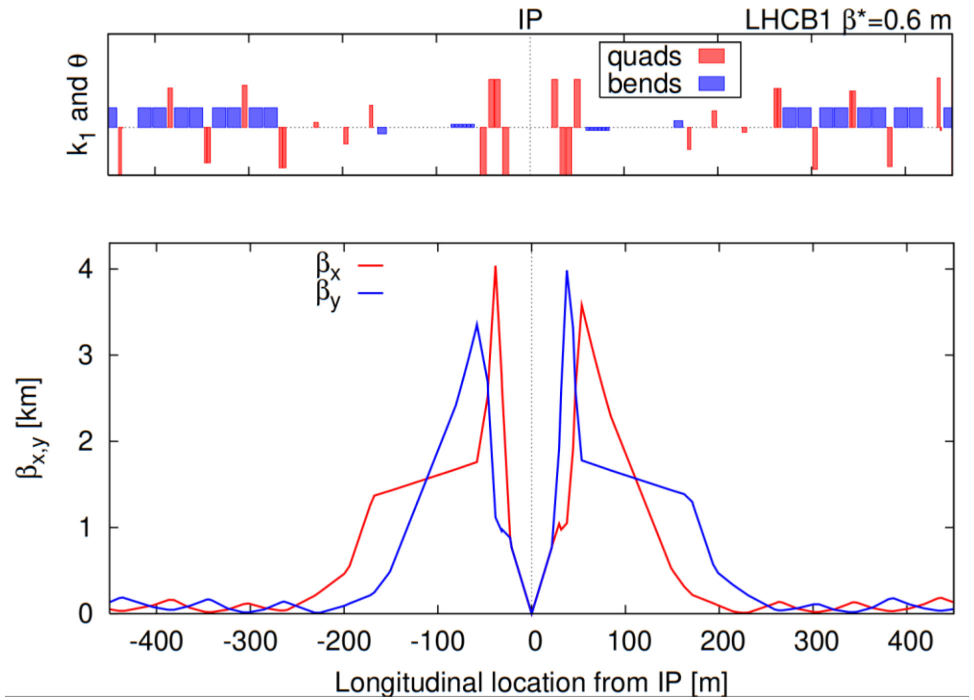
$$\mathcal{L} \sim \frac{f \cdot N^2}{4\sigma^2}$$

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

ε → Const

$\boldsymbol{\beta}$ → Determined by **quadrupole arrangement and powering**

Optics ←



Courtesy of Rogelio Tomas

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

Beam optics control at the LHC

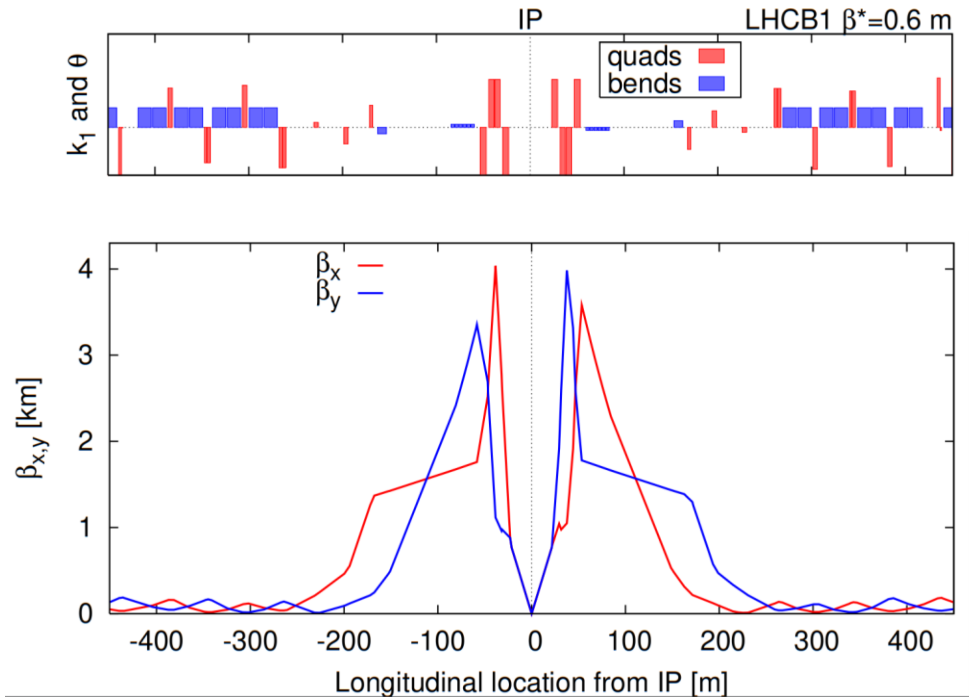
- **Luminosity:** maximize the number of collision events.

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

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

ε → Const
$\boldsymbol{\beta}$ → Determined by **quadrupole arrangement and powering**

Optics



Courtesy of Rogelio Tomas

- 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

1. **Instrumentation faults** → unreliable optics measurements

2. Corrections compensate **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 BPMs → How to **reconstruct** the missing data?

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
- direct measurements are not possible
- previously unobserved behavior

Why applying ML to optics control?

Accelerators

- Operation
- Diagnostics
- Beam Dynamics Modeling

Which limitations can be solved by ML with reasonable effort?

➢ large amount of optimization targets
➢ direct measurements are not possible
➢ previously unobserved behavior

→

Machine Learning:
✓ ***Learn arbitrary models***
✓ ***Directly from provided data***

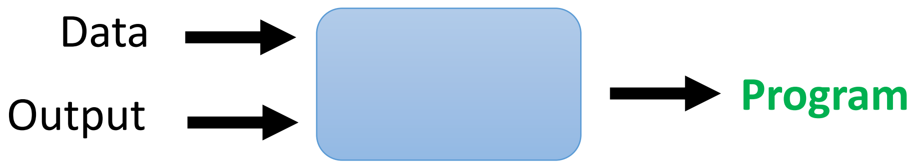
Introduction to Machine Learning

Teaching machines to learn from experience

- Traditional programming

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

- Machine Learning approach



learn from data automatically

Relevant ML concepts and definitions

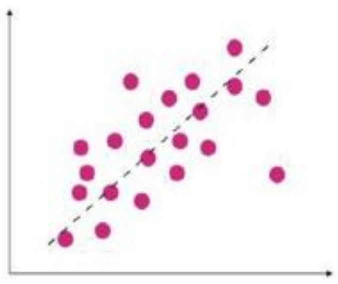
Supervised Learning

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

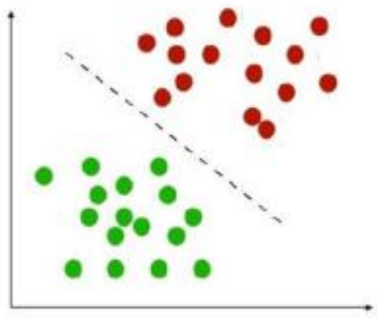
Unsupervised Learning

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

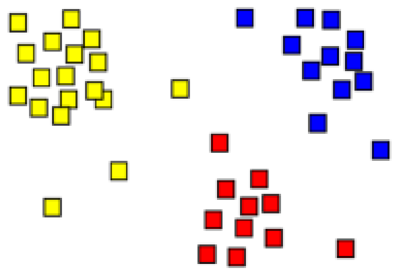
Regression



Classification



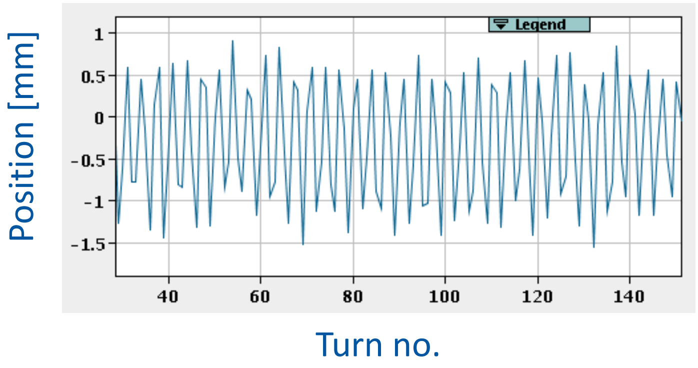
Clustering



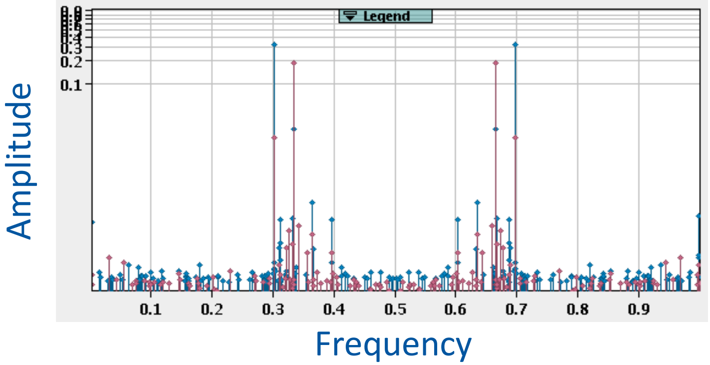
Optics Measurements at the LHC

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)

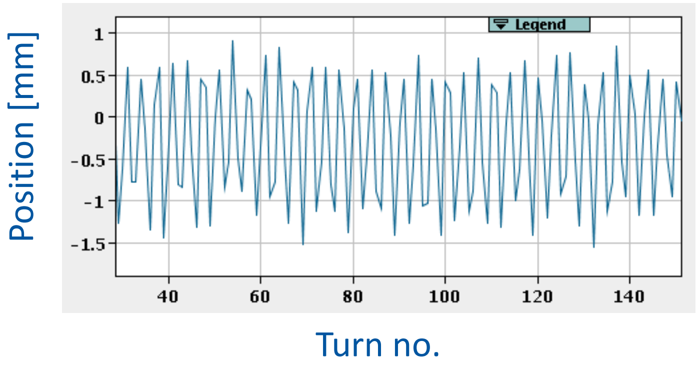
Denoising (SVD)
Signal cuts

Semi-automatic and manual cleaning of outliers

What are the limitations of traditional techniques?

Measuring the optics

Turn-by-turn beam position

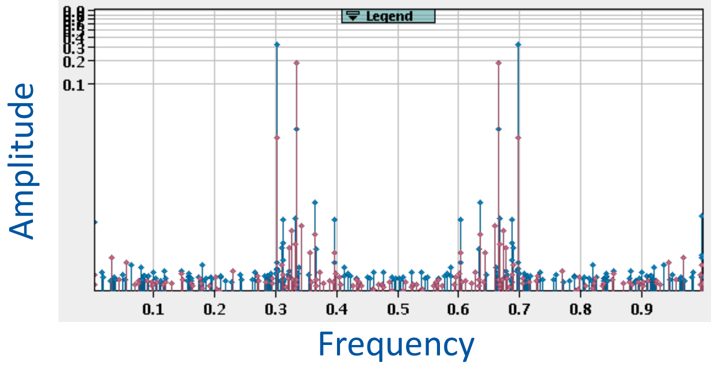


- Excite the beam to perform transverse oscillations.

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

Denoising (SVD)
Signal cuts

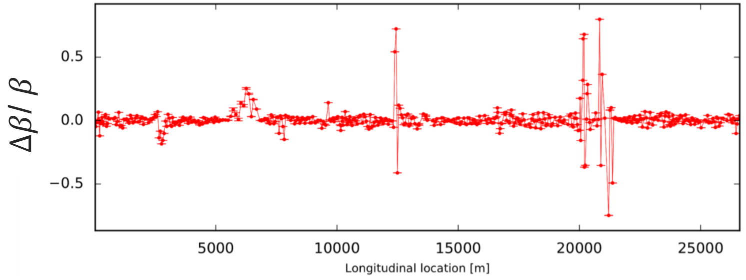
Spectrum



- Harmonic analysis using Fast Fourier Transform (FFT)

Semi-automatic and manual cleaning of outliers

Optics

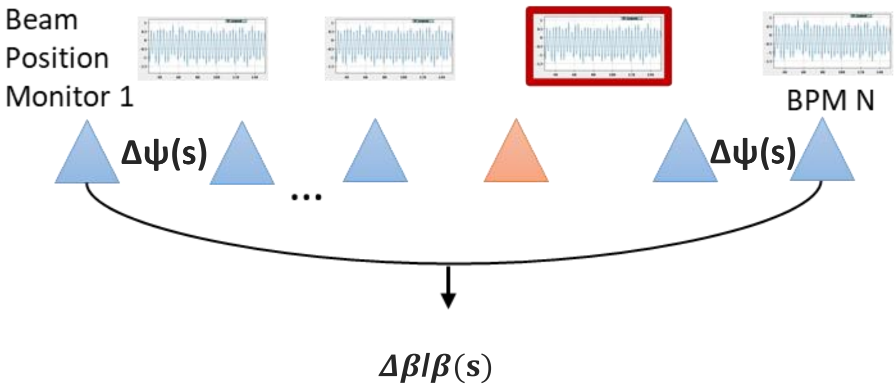


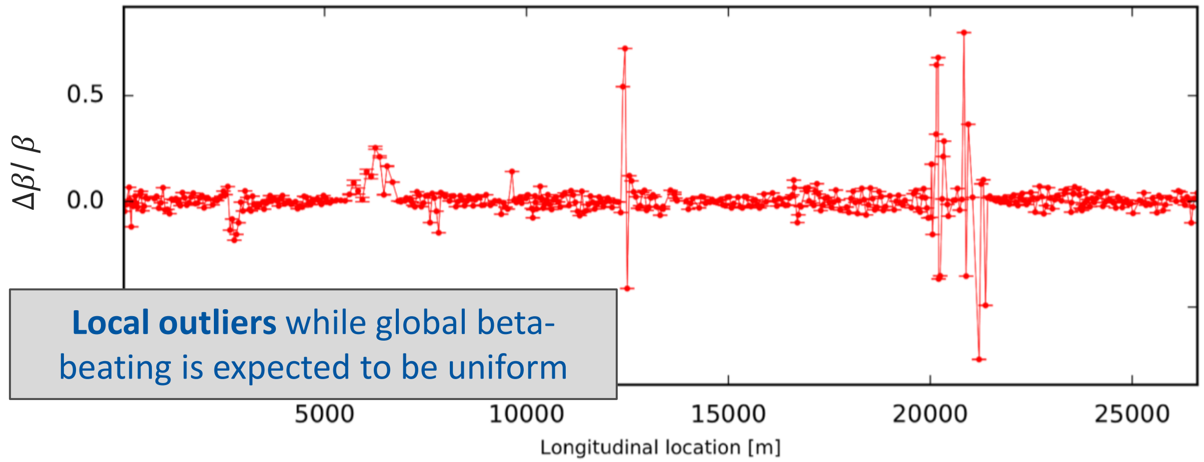
- Compute beta-beating and other optics functions

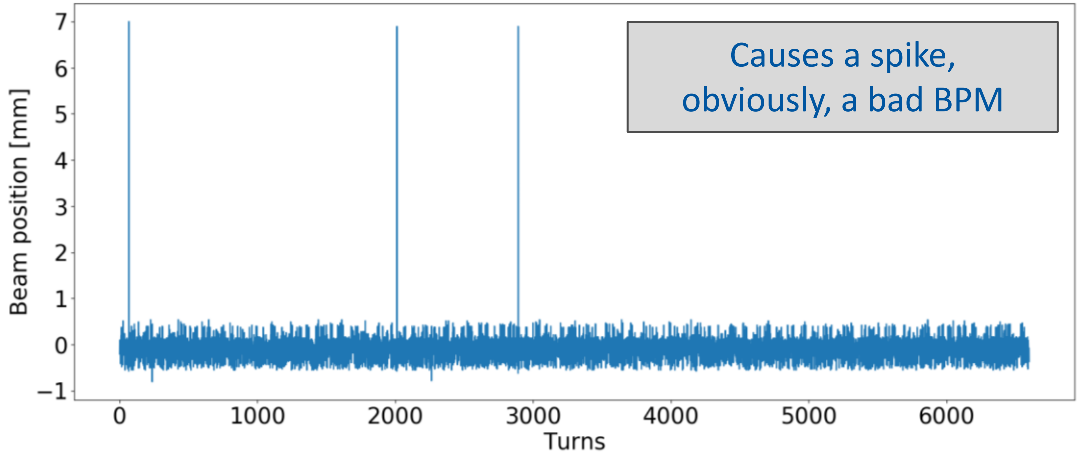
Unphysical values still can be observed

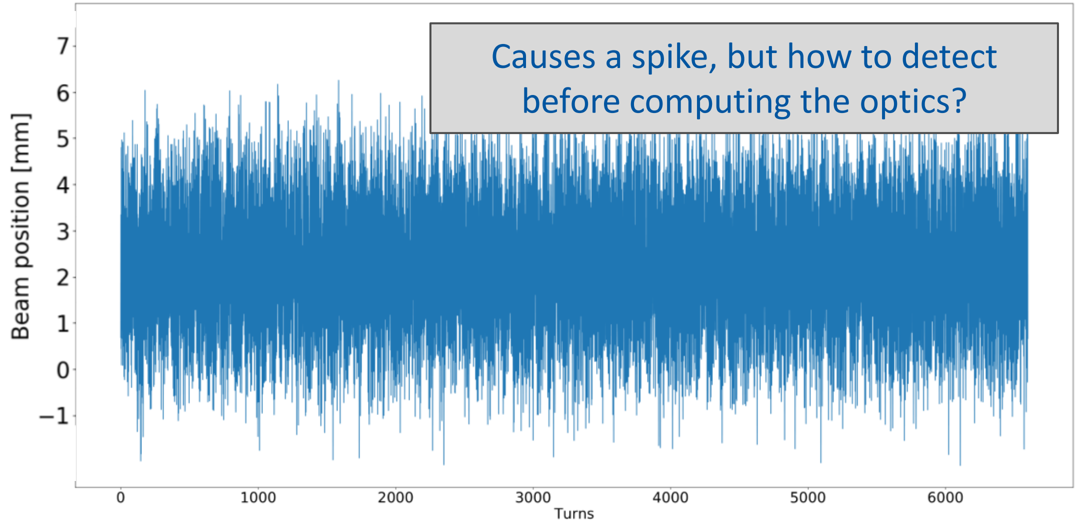
What are the limitations of traditional techniques?

Measuring the optics: challenges









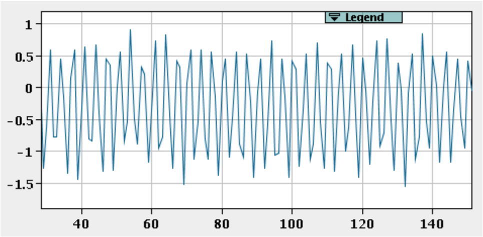
What are the limitations of traditional techniques?

Detection of faulty Beam Position Monitors

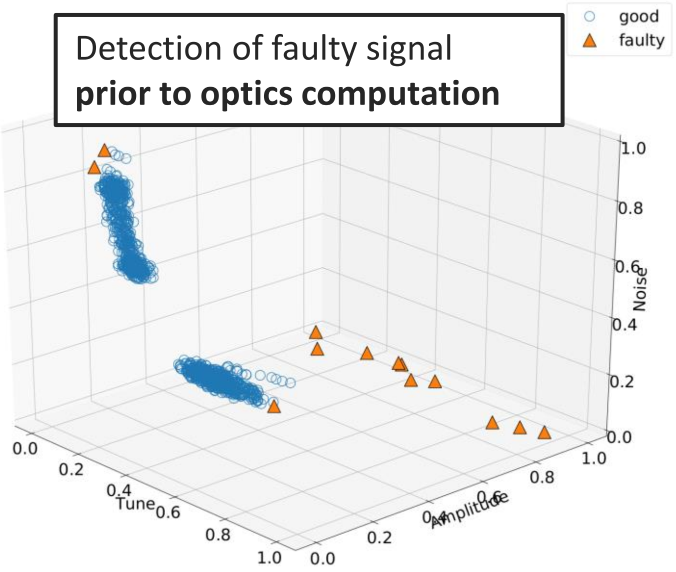
Detection of faulty Beam Position Monitors

- Faulty BPMs **are a-priori unknown**: no ground truth → **Unsupervised Learning**
- Applied clustering algorithms: DBSCAN[1], Local Outlier Factor[2], anomaly detection using **Isolation Forest**[3] implemented with *Scikit-Learn.*

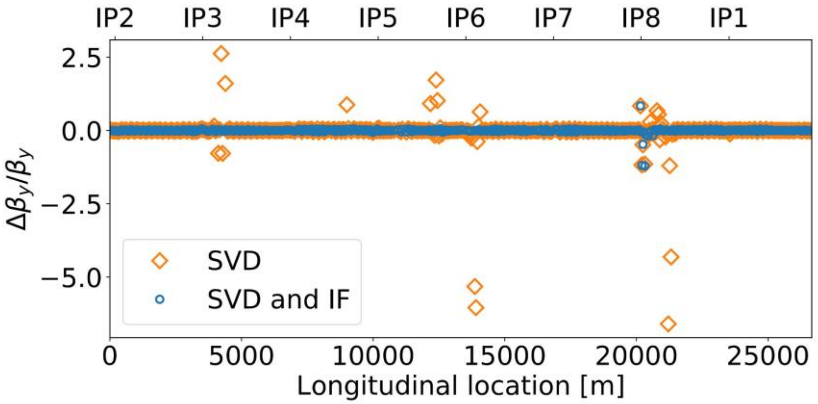
Harmonic analysis of all BPMs



Detection of faulty signal
prior to optics computation



Avoid the appearance of erroneous optics computation



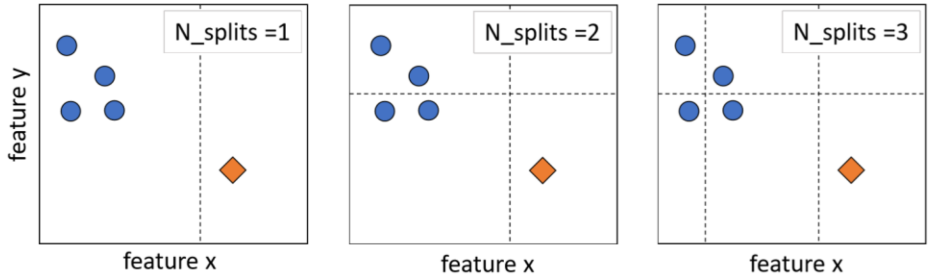
- Outlier detection based on combination of several signal properties
- Immediate results

1. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise" Ester, M., H. P. Kriegel, J. Sander

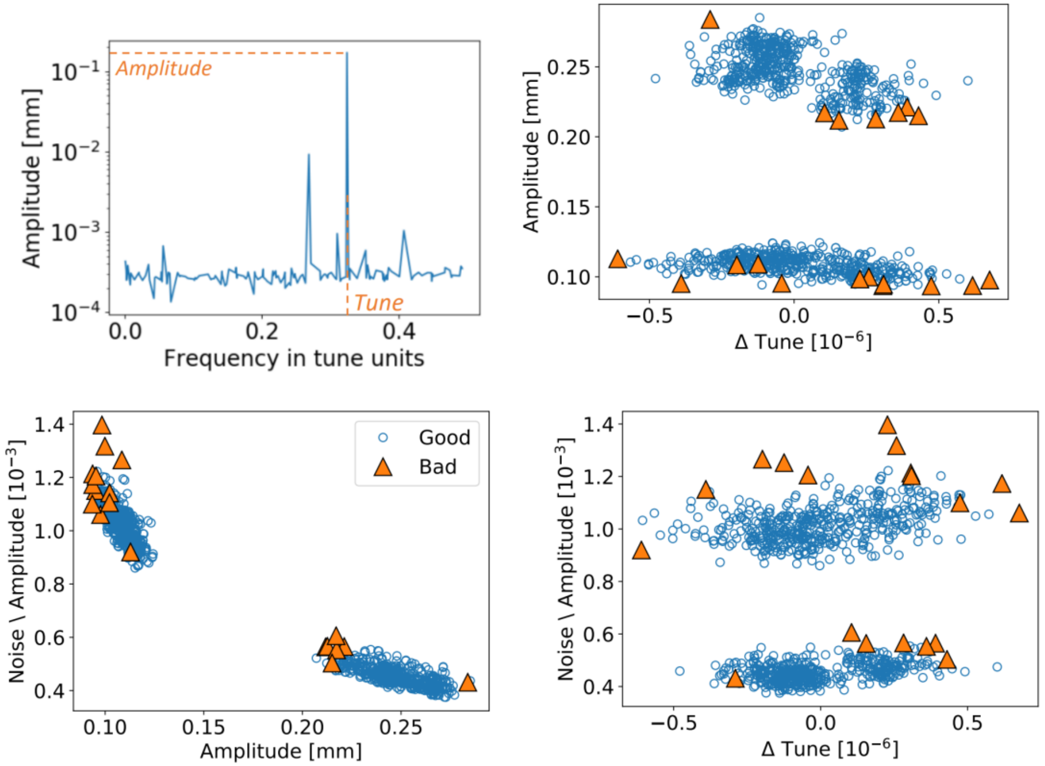
2. Breunig, M. M., Kriegel, H. P., Ng, R. T., & Sander, J. (2000, May)., LOF: identifying density-based local outliers
3. Liu, Fei Tony, Ting, Kai Ming and Zhou, Zhi-Hua. "Isolation forest." Data Mining, 2008. ICDM'08.

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**.



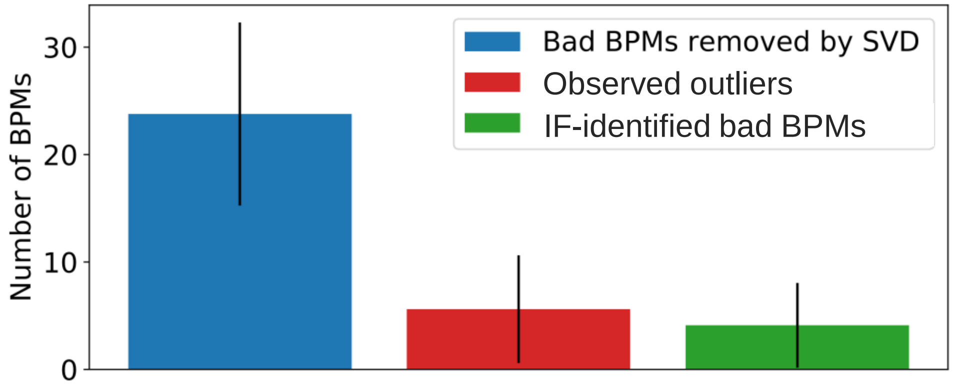
Conceptual illustration of Isolation Forest algorithm



- **Input data: combination of several signal properties** obtained from harmonic analysis of BPM turn-by-turn measurements
  → No additional data handling needed.

Detection of faulty Beam Position Monitors

Reduction of non-physical outliers in beta-beating:
Averaged cleaning results, optics measurements in 2018.



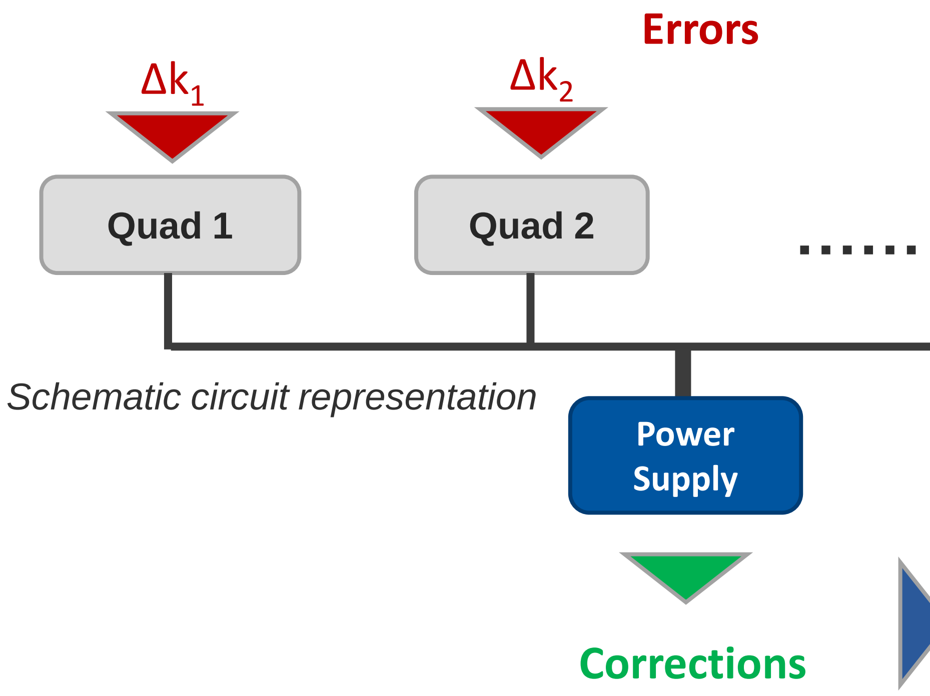
- Instant faults detection instead of offline diagnostics.
- Full optics analysis is possible directly during dedicated measurements session instead of iterative procedure of cleaning and analysis.

✓ **Fully integrated** into optics measurements at LHC
✓ **Successfully used in operation** under different optics settings.

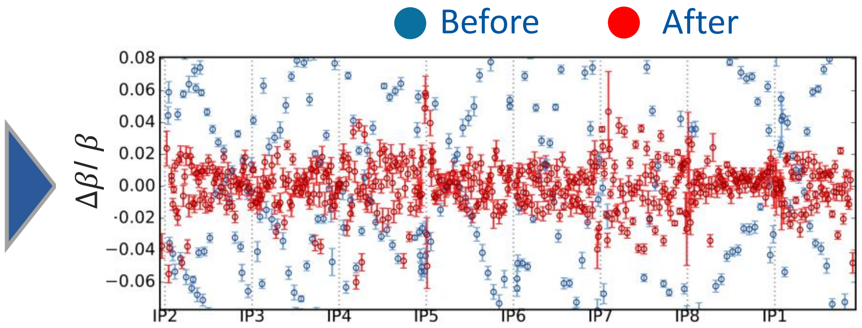
Published in: *Physical Review Accelerators and Beams:*
"Detection of faulty beam position monitors using unsupervised learning", Phys. Rev. Accel. Beams 23, 102805.

Optics Corrections: estimation of quadrupole errors

Correcting the optics



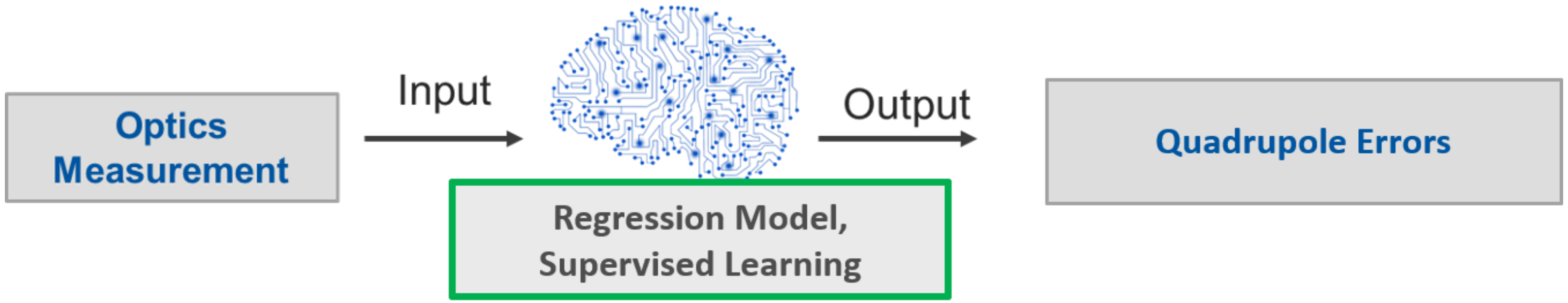
- Corrections are implemented by changing the strength of **circuits**
- Optics perturbations are caused by all **individual magnets.**



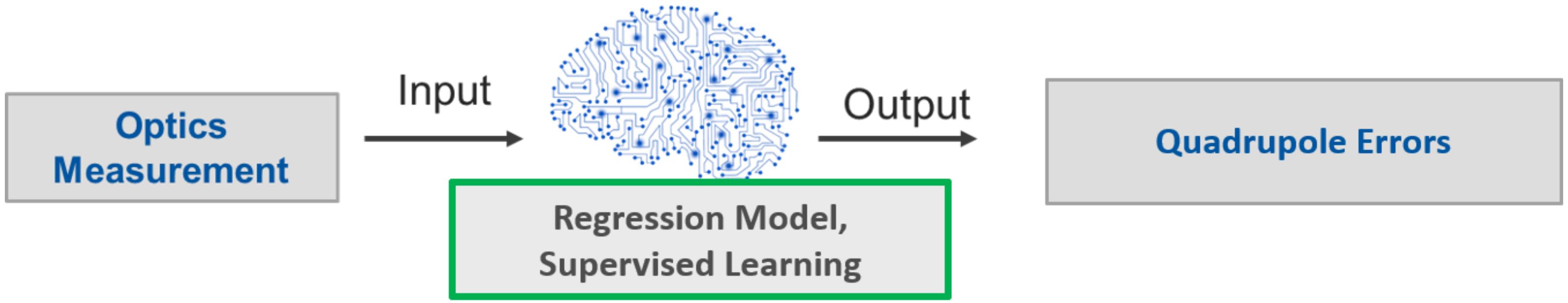
- ➢ What are the **actual errors of individual quadrupoles?**
- ➢ How to obtain the **full set of errors in one step**?

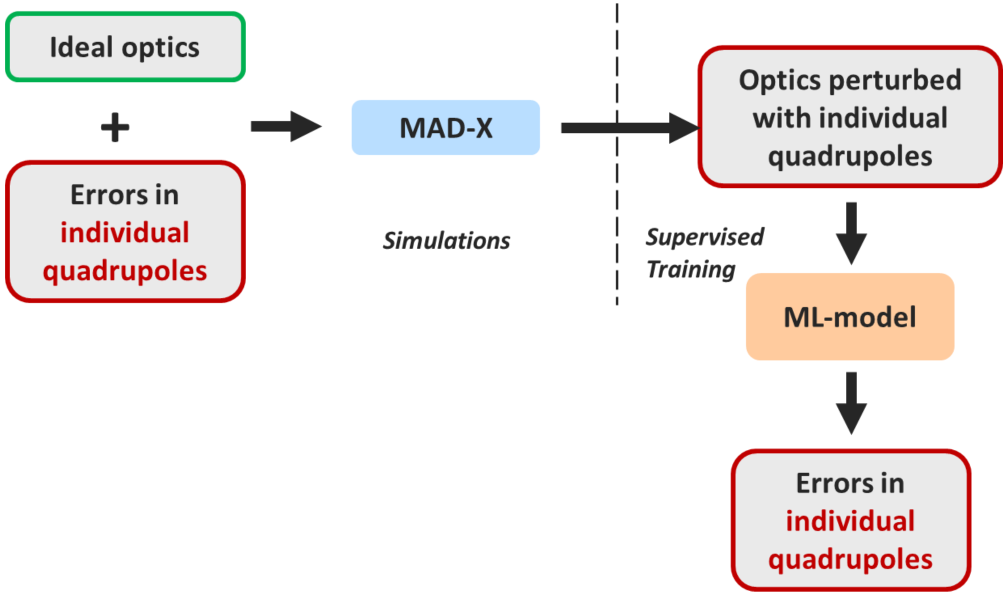
What are the limitations of traditional techniques?

Estimation of quadrupole errors



Estimation of quadrupole errors





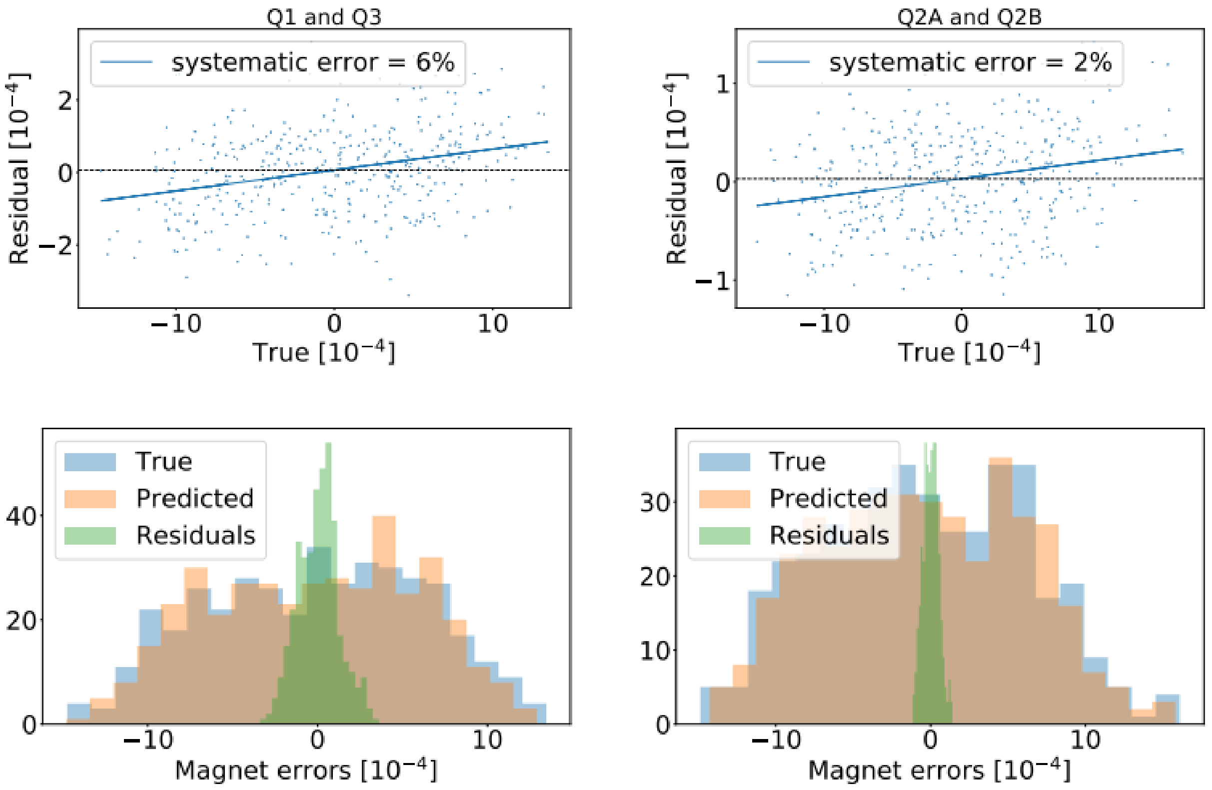
<u>Training ML- regression model:</u>

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

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

Verifying ML approach: simulations

Simulations: **true magnet errors are known**
→ directly compare prediction to simulated data → **residual error**



$$R^2(y, \hat{y}) = 1 - \frac{Var\{y - \hat{y}\}}{Var\{y\}}$$

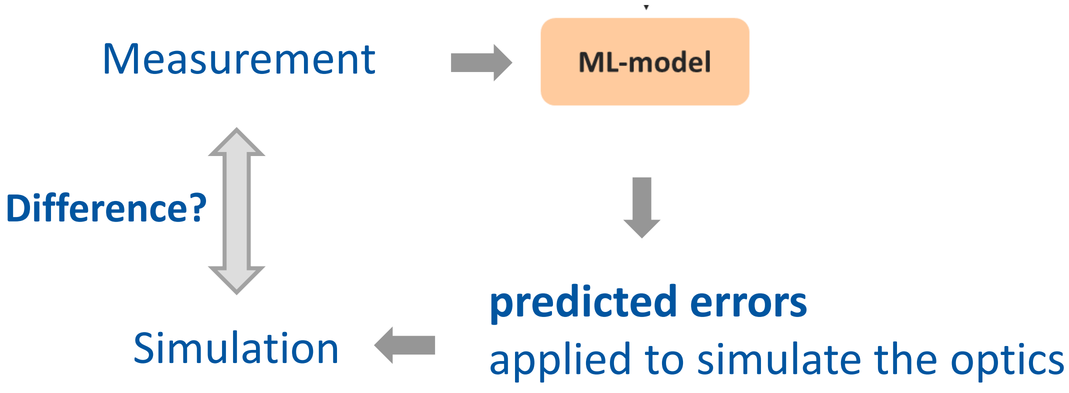
$$MAE(y, \hat{y}) = \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

How well can we correct the optics with predicted errors?

Estimation of quadrupole errors: measurements

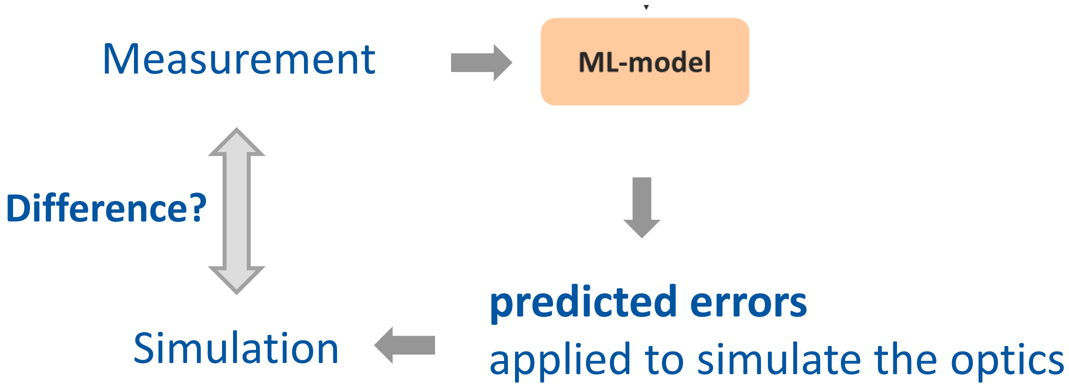
Measurements: **true magnet errors are unknown**

→ **Control beta-beating**

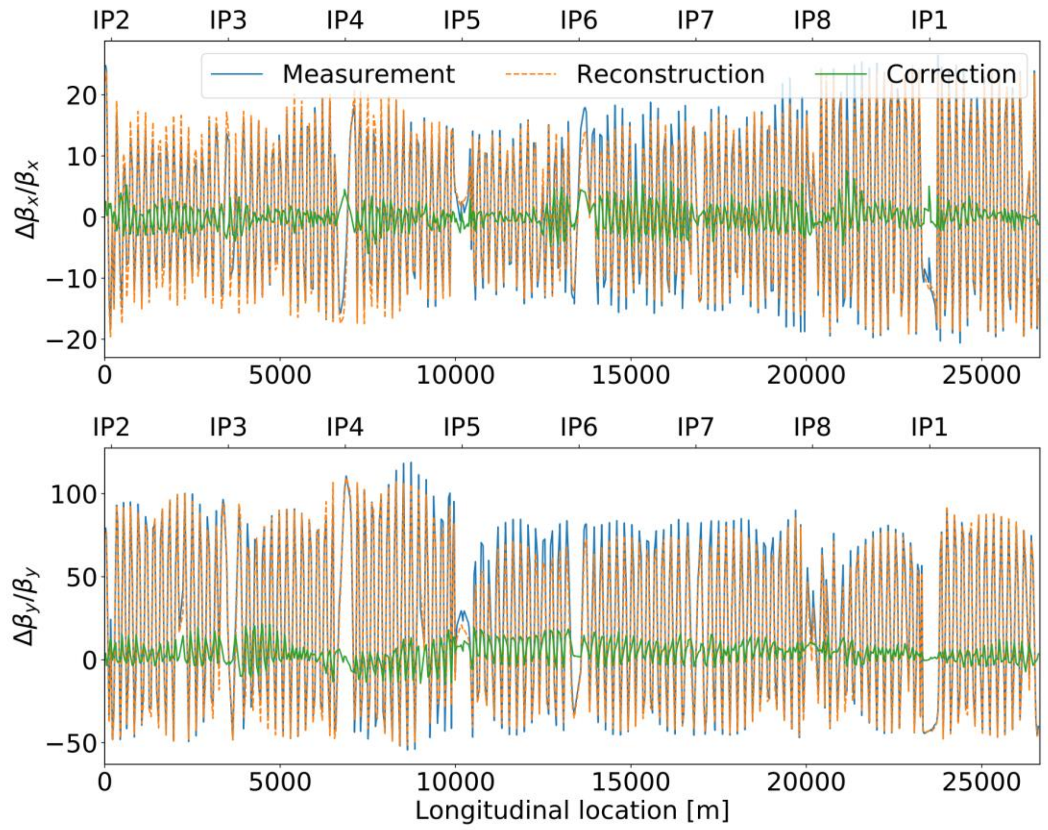


Estimation of quadrupole errors: measurements

Measurements: **true magnet errors are unknown**
→ **Control beta-beating**



*Test on **LHC optics measurements, uncorrected** machine*

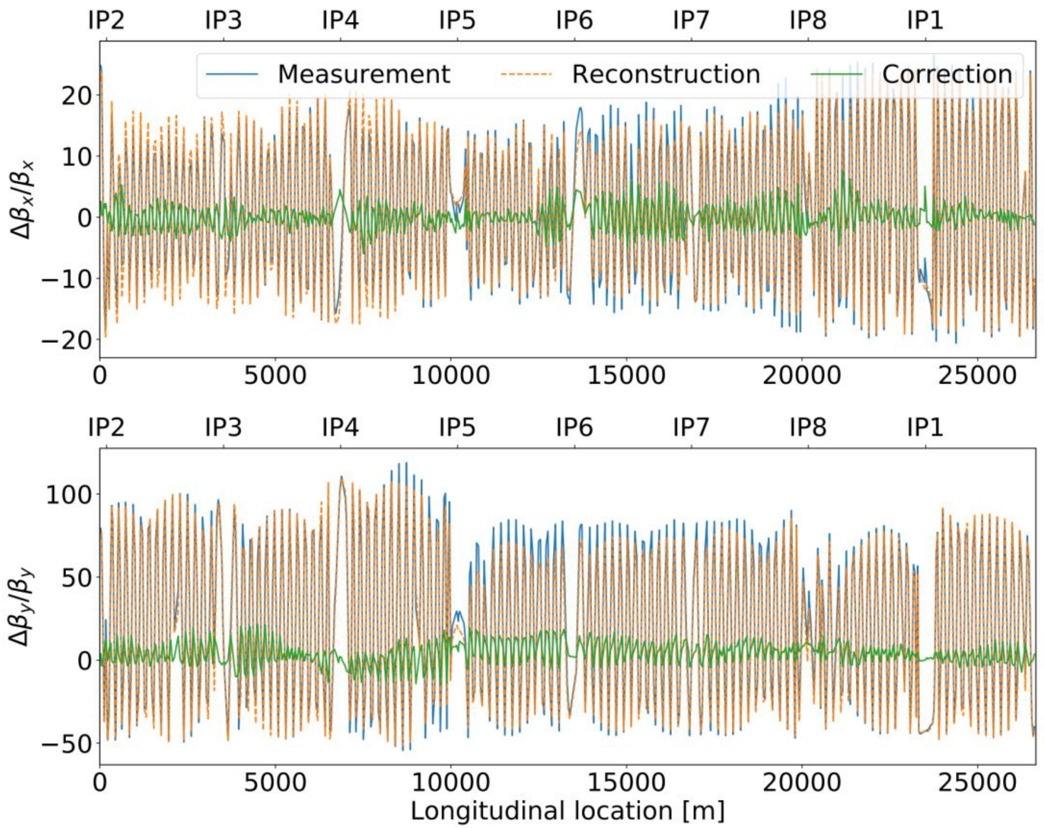


Magnet errors predicted with ML-model reproduce the measured β –beating in uncorrected machine with average rms error of 7% and below 3% at IPs.

Estimation of quadrupole errors: measurements

- ✓ New method for local optics corrections
- ✓ Improved knowledge of direct error sources
- ✓ Simultaneously obtaining quadrupole errors for both beams, at every location
  - → **Potential to save operation time**
  - → **To be tested in LHC commissioning, April 2022**

*Test on **LHC optics measurements, uncorrected** machine*



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

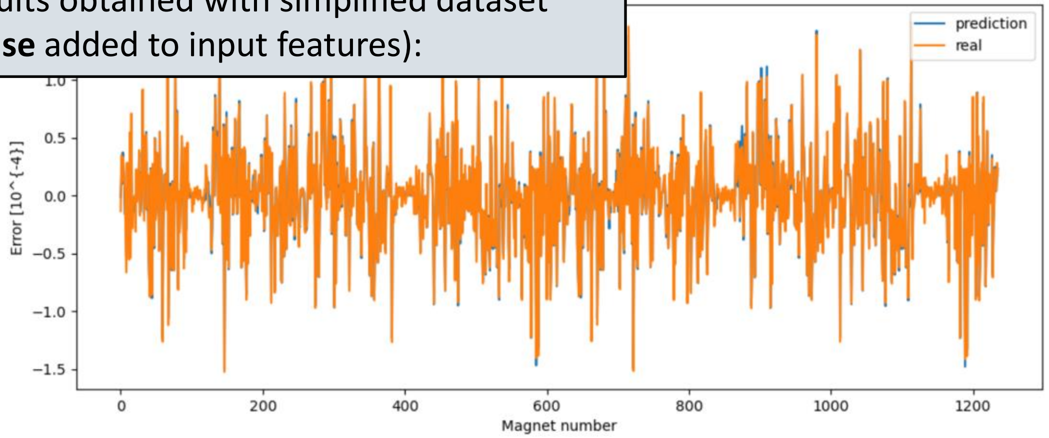
Optics control in HL-LHC studies

High Luminosity Large Hadron Collider: Upgrade of the LHC to push the performance in terms of beam size and luminosity.

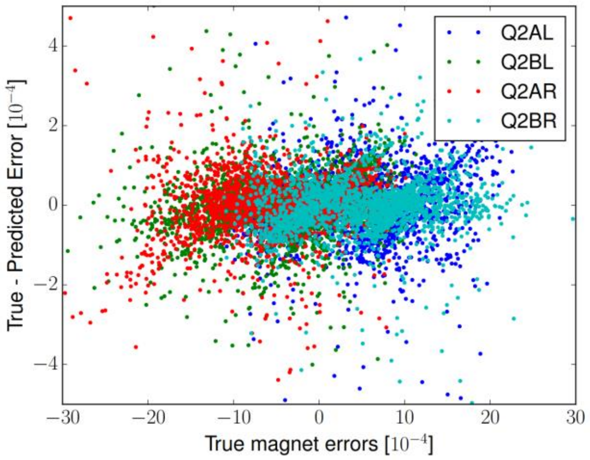
- The local linear optics correction at the IR will be essential to ensure the HL performance.
- Current LHC strategies might impose limitations → new correction strategies are needed.

Preliminary results obtained with simplified dataset (**no noise** added to input features):

Full set of quadrupoles all around the ring

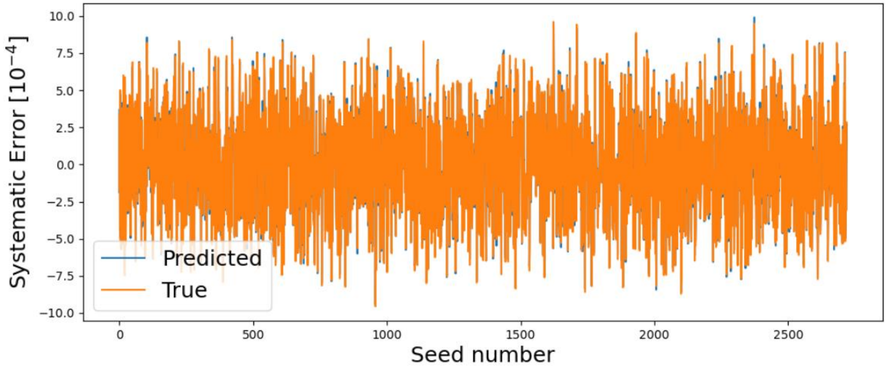


Courtesy of Hector Garcia Morales



Inner Triplet magnets in Interaction Regions

- Systematic part of the gradient error (unknown) may have a significant impact on the β-beating.



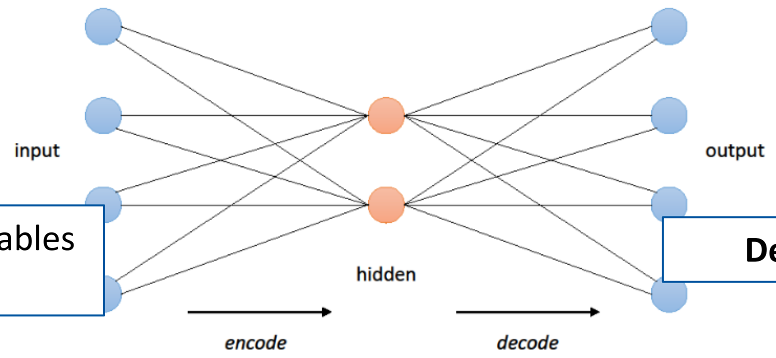
Denoising of optics measurements

Denoising of optics measurements

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

Noise in the measurements degrades the performance of corrections techniques

Autoencoder Neural Network

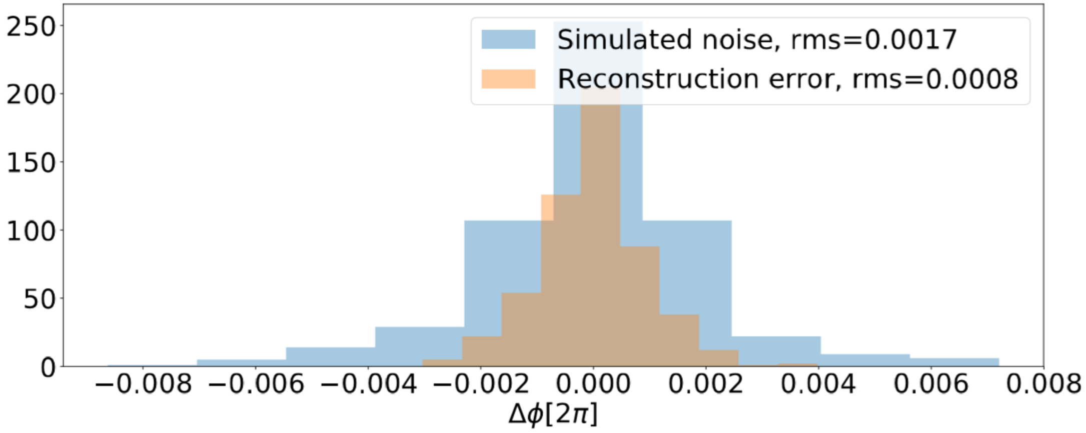


Simulated optics observables **+ noise**

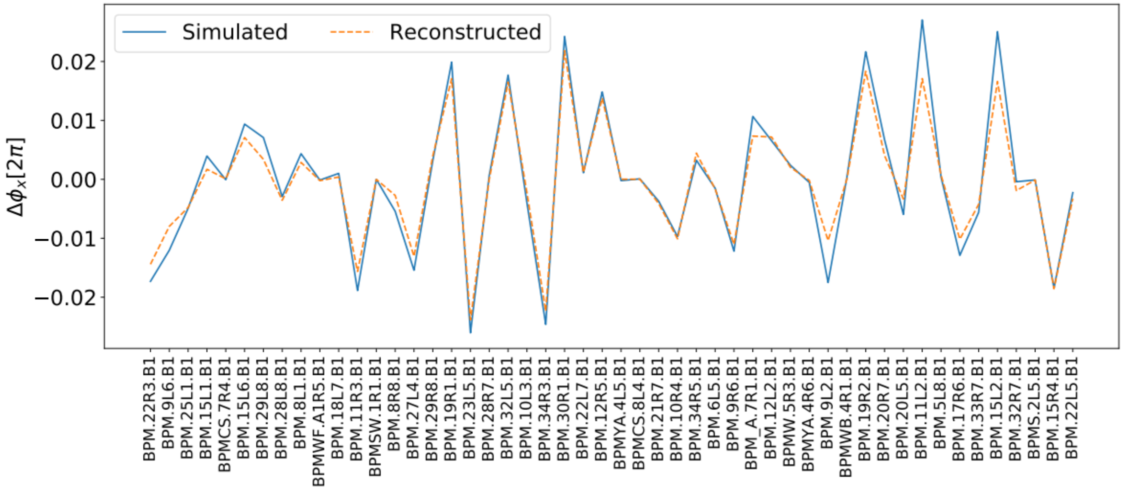
Denoised optics

Denoising of optics measurements

Simulated data: Noise Reduction



Simulated data: Reconstruction



✓ Reconstruction error is by factor 2 smaller than the noise present in the signal.

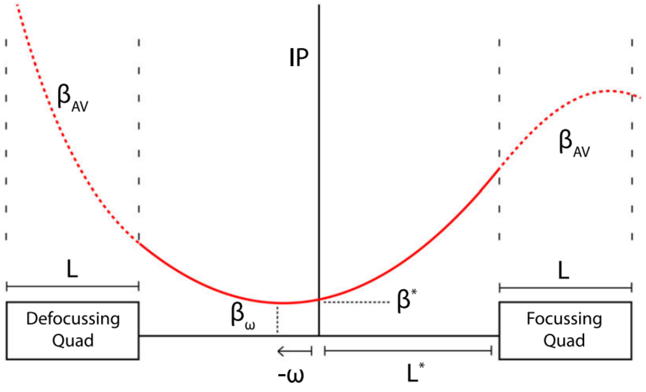
✓ Reliable reconstruction after denoising

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

Reconstruction of advanced optics observables

Reconstruction of β-beating in Interaction Regions

- Special technique to measure beta-function at IP is needed:



- Modulation of quadrupole gradient
- Computation of average beta-function
- Propogate beta-function values to IP

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

Phase advance measurements*

Input

Output

Regression Model, Supervised Learning

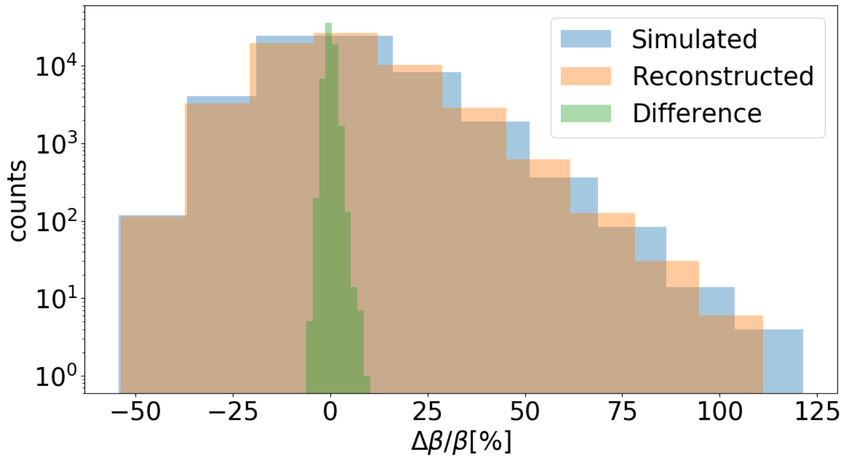
***β*-function** at IPs

Simultaneously for beam 1 and 2

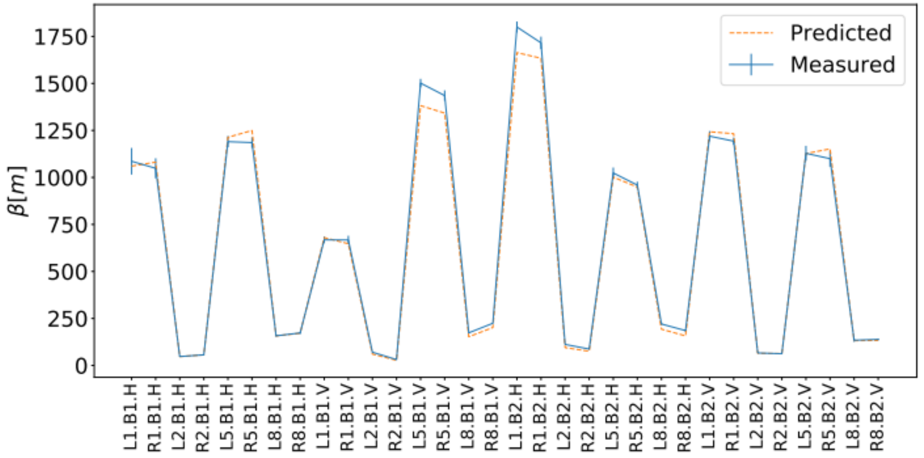
** Always available from turn-by-turn data at one beam excitation*

Reconstruction of β-beating in Interaction Regions

Simulations



LHC Measurements, BPMs left and right from Interaction Points



Reconstruction error: $\frac{\beta_{simulated} - \beta_{reconstructed}}{\beta\_simulated}$ = ***1%***

- ✓ comparable to measurement uncertainty of traditional method.

- ✓ Great potential to reduce measurements time
- ✓ Applicable to estimation of other optics observables (e.g. horizontal dispersion)

Outlook and Summary

ML applied to modeling and optimization

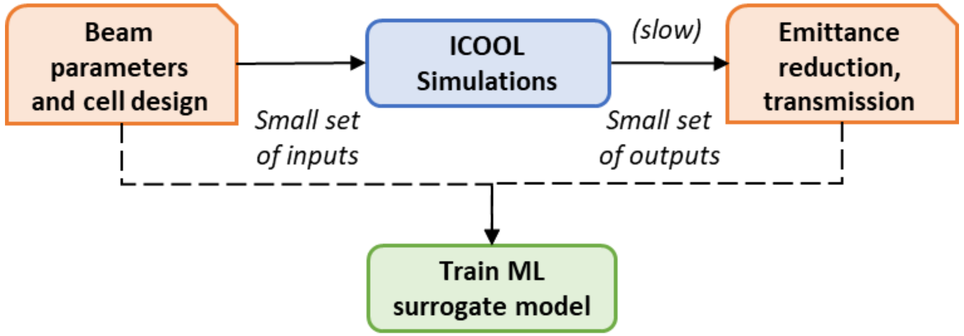
Muon Collider Design study [1]:

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

→ **Final Cooling system** with challenging design

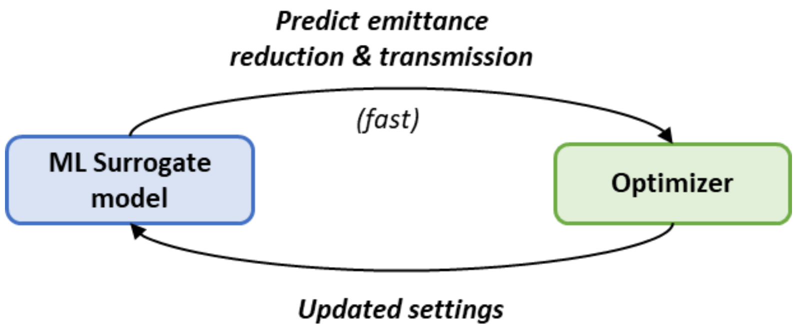
→ High dimensional parameter space to be optimized in order to achieve low emittance, high intensity muon beams

→ Trade-off between different optimization objectives

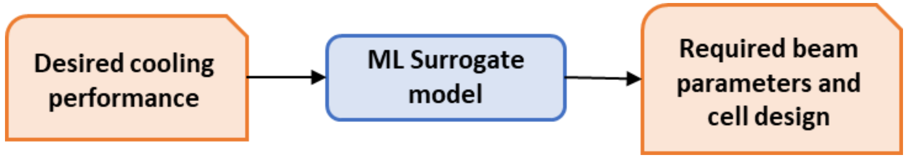


- Extending existing simulation frameworks towards automatic, fast executing optimization.
- **Exploring application of Supervised Learning**
  → surrogate models

1. 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

CERN

Achieved Results

✓ **ML-based toolbox for optics control:**

- 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

Achieved Results

✓ **ML-based toolbox for optics control:**

- 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

Outlook

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

- Optics corrections for High Luminosity – LHC upgrade:
  - local correction
  - exploring Reinforcement Learning for determining correctors settings.
- Exploring more complex optics error sources: coupling corrections
- Optimizing the design of future colliders.

Further References

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

- **Opportunities in Machine Learning for Particle Accelerators**
  https://arxiv.org/abs/1811.03172

- **Optimization and Machine Learning for Accelerators (USPAS course)**
  https://slaclab.github.io/USPAS_ML/

Thank you for your attention!

Back-up slides

Learning from data

example 1
example 2
example 3
.
.
.

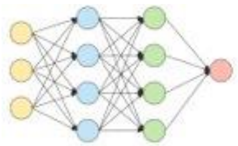
Training input data

Training output data

Learning from data

example 1
example 2
example 3
.
.
.

Training input data



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

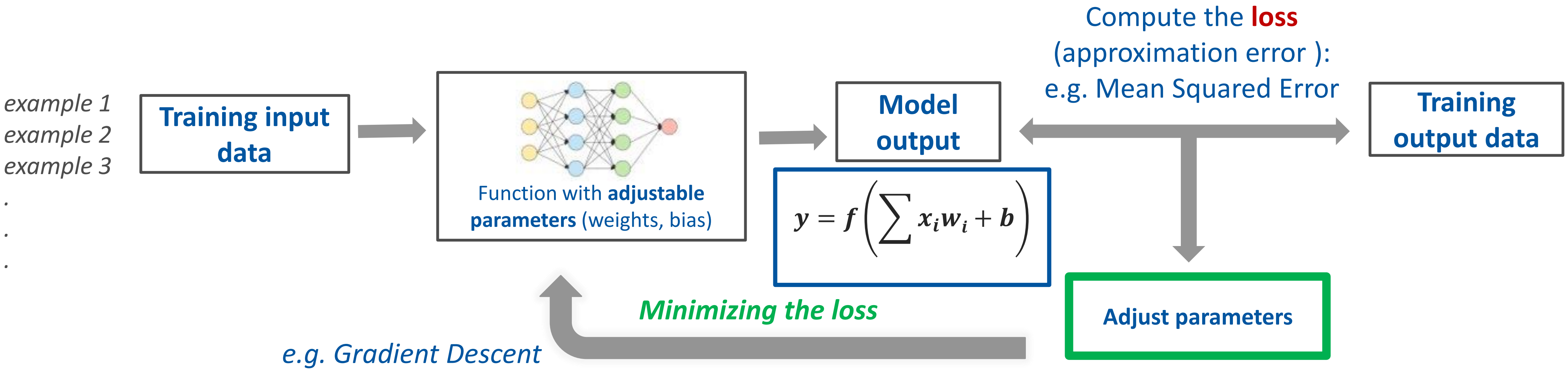
Model output

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

Training output data

What is "Learning"?

Learning from data



➢ Generalized model **explaining relationship between input and output** variables in **all training samples**.

Regression Models

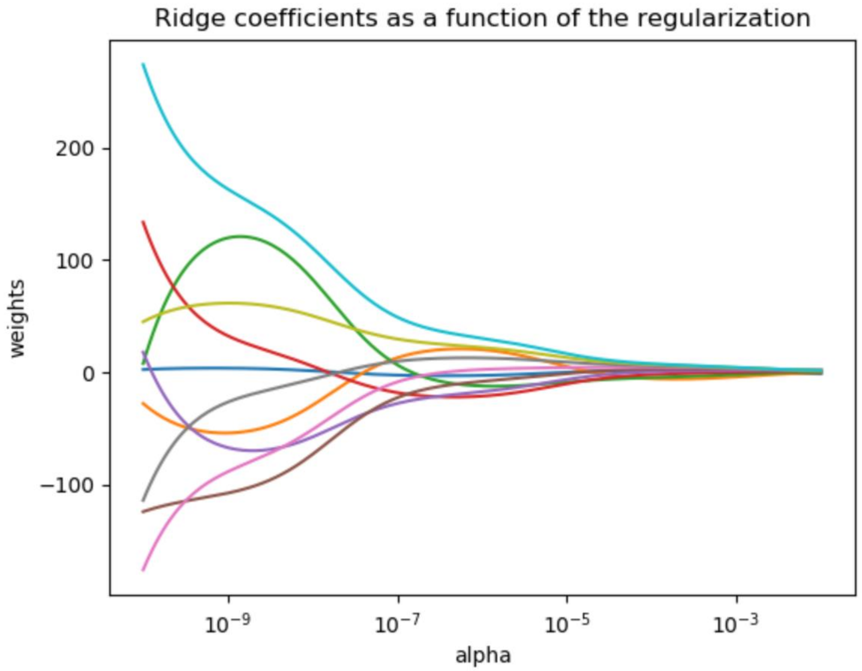
- Linear model for *input X / output Y pairs, i* – number of pairs (training samples): $\boldsymbol{f(X,w) = w^T X}$
- Squared Loss function for model optimization: $\boldsymbol{L(w) = \frac{1}{2}\sum_i \left(Y_i - f(X_i;w)\right)^2}$
- Find new weights minimizing the Loss function: $\boldsymbol{w^*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

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

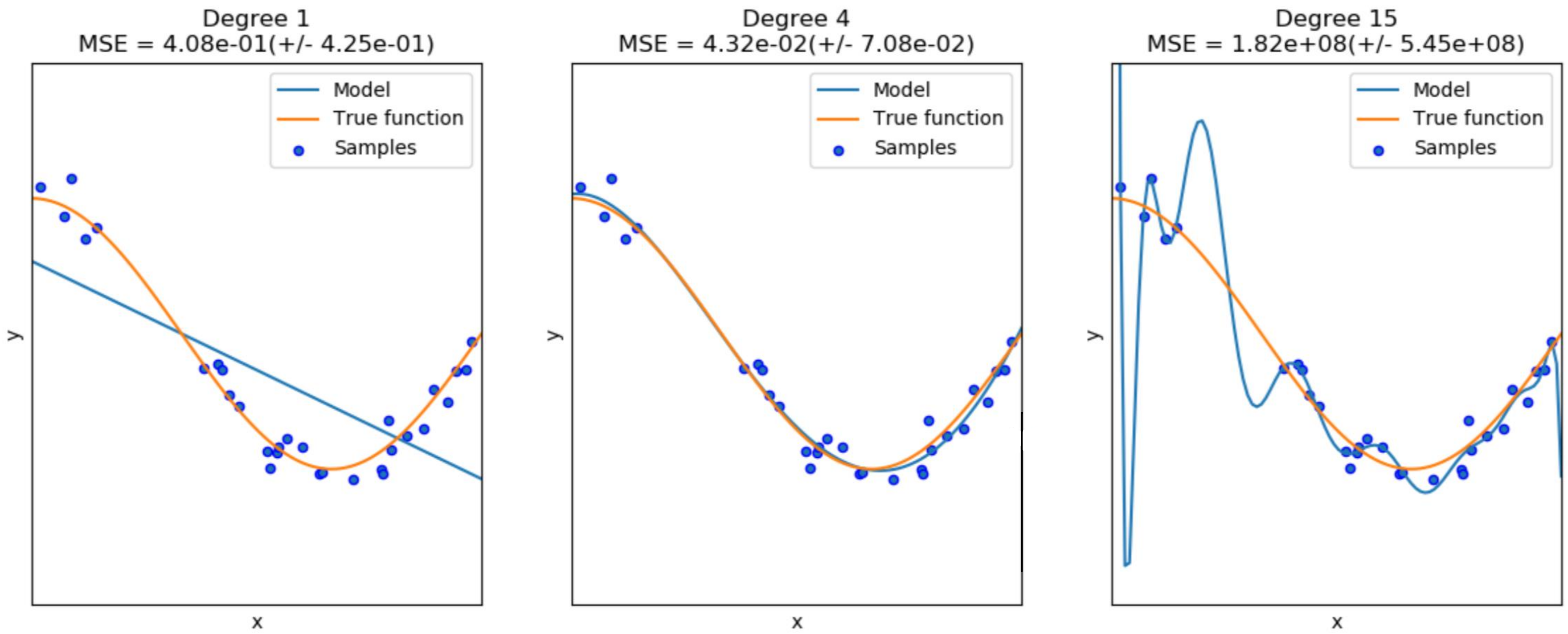
Too much "flexibility" in weights update can lead to *overfitting*

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

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



Training and generalization: no perfect model needed!



Simple models underfit

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

We don‘t want „look up tables“

We don‘t want unreliable prediction

→ Bias-Variance tradeoff

Complex models overfit

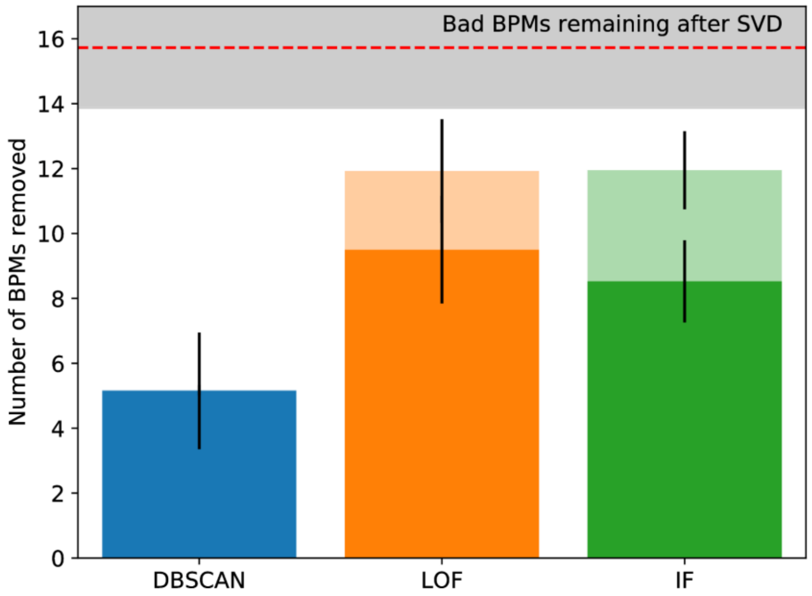
- Very low systematical deviation (low bias)
- Very sensitive to data (high variance)

Faulty BPMs detection: simulation study

- Comparing different suitable techniques:

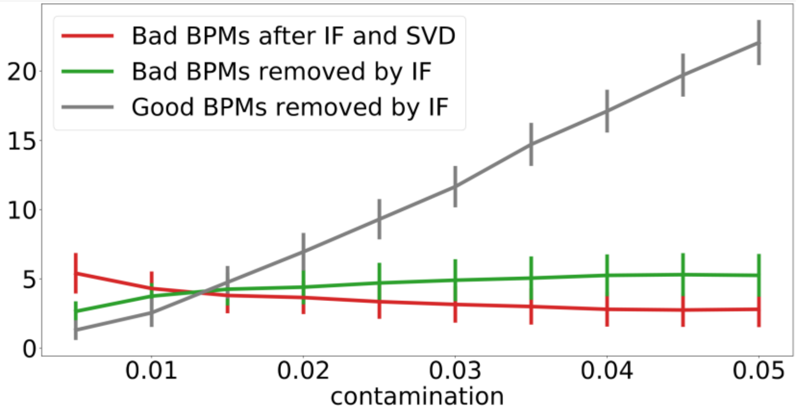
The presence of a single faulty BPM has more significant negative impact on the optics computation than the absence of a good BPM

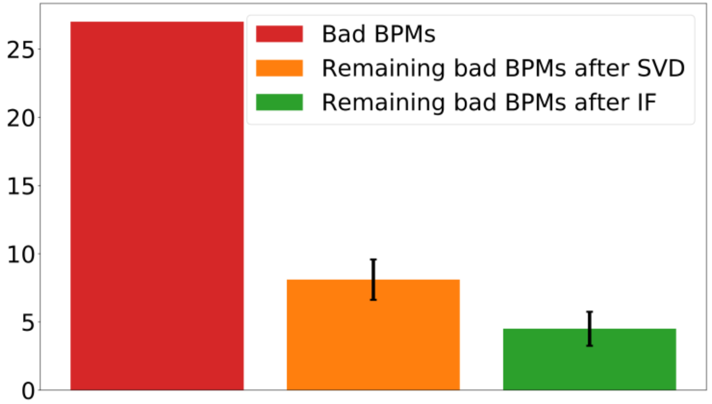
→ IF is preferred method for the LHC.



- Tuning of IF-algorithm **after finding optimal settings for SVD-cleaning**:

→ Trade-off between eliminating bad BPMs and removing good BPMs as side effect by setting the expected contamination rate



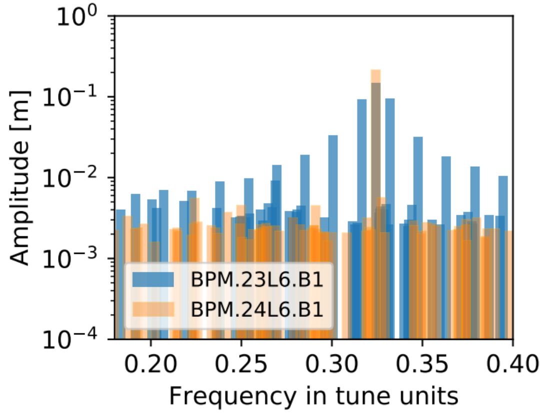


- Averaged results over 100 simulations

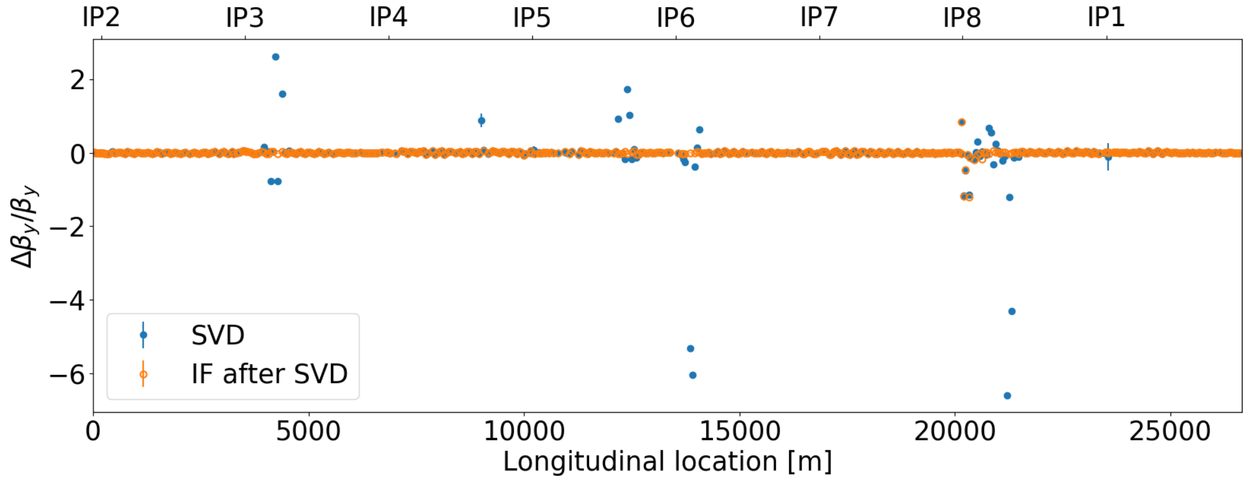
IF in the LHC operation: detecting unknown failures

- Some artifacts in the signal are known to be related to BPM failures (manual cleaning would time consuming, but potentially possible).
- **How to deal with unknown failure modes?**

Several BPMs with unusual pattern in the spectra indicating a new failure mode



First observed in: "Analysis of tune modulations in the LHC", D.W. Wolf
Related to BPM failure: L. Malina, "Noise and stabilities", https://indico.cern.ch/event/859128/



Since IF is based on the structures in given data
➢ ***Ability to identify previously unknown failures***