Supervised Learning-based optics corrections in circular accelerators

Elena Fol

R. Tomás, G. Franchetti

Optics Measurements and Corrections Team,
BE-ABP, CERN

Johann-Wolfgang Goethe University

Intelligent Process Control Seminar
20.04.2021

Outline

- **Introduction**
- **Estimation of magnetic errors from optics measurements**

  - General concept
  - Results on simulations
  - Results on experimental data

- **Denoising and reconstruction of optics functions**

  - Autoencoder
  - Linear models
  - Results

- **Conclusions**

I. Introduction

Applying Machine Learning to Beam Optics

PhD project: Application of Machine Learning to Accelerator Optimization with the focus on beam optics.

- Why and how is the beam optics controlled in the LHC?
- Where are the limitations of traditional techniques?
- Which ML concepts and algorithms can be applied?
- Achieved results?

Applying Machine Learning to Beam Optics

PhD project: Application of Machine Learning to Accelerator Optimization with the focus on beam optics.

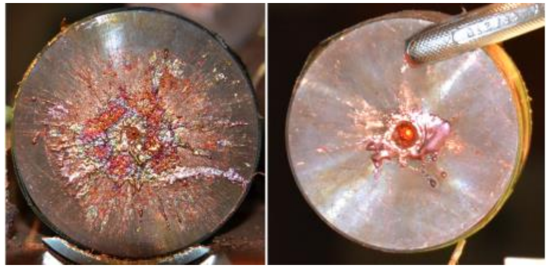
- **Why and how is the beam optics controlled in the LHC?**
- Where are the limitations of traditional techniques?
- Which ML concepts and algorithms can be applied?
- Achieved results?

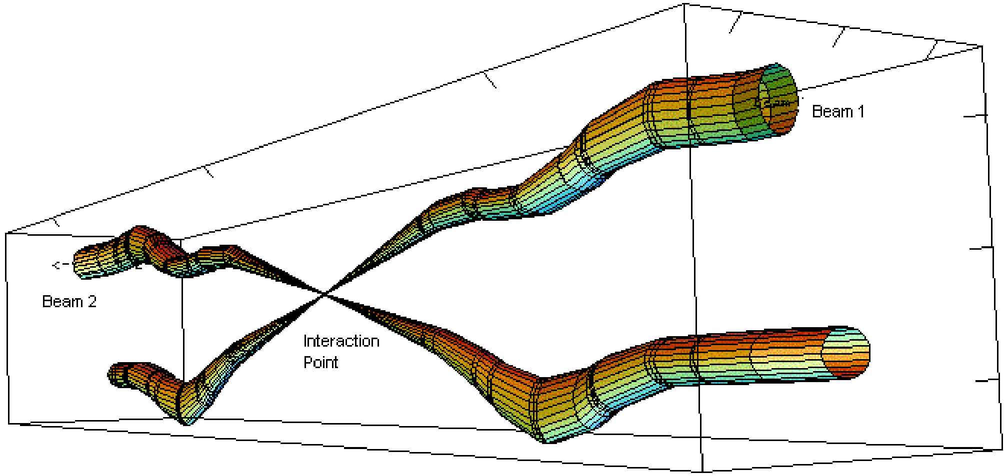
Beam optics control:

- Magnetic errors and misalignments change **beam size** - optics
- Adjust **magnetic strengths** – optics corrections
- Reliable and precise **measurements of optics functions** are needed to compute corrections.

Importance of beam optics control:

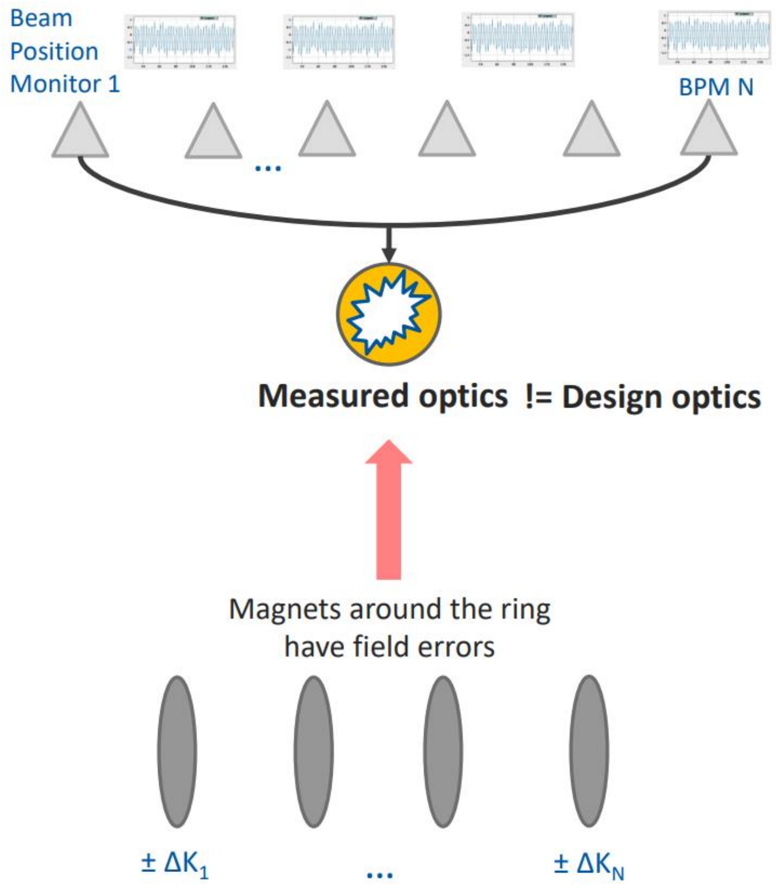
- **Collision rate** depends on the beam size
- Beam optics imperfections can lead to **machine safety** issues.





Relative beam sizes around IP1 (Atlas) in collision

Limitations of traditional techniques for optics corrections?



- Optics corrections algorithms aim to **compensate the measured optics deviations** from design
  - → What are the actual currently present **magnetic errors**?

- Advanced techniques for computation of optics functions require **additional measurements and operational time**
  - → How to obtain advanced analysis **from available measurements**?

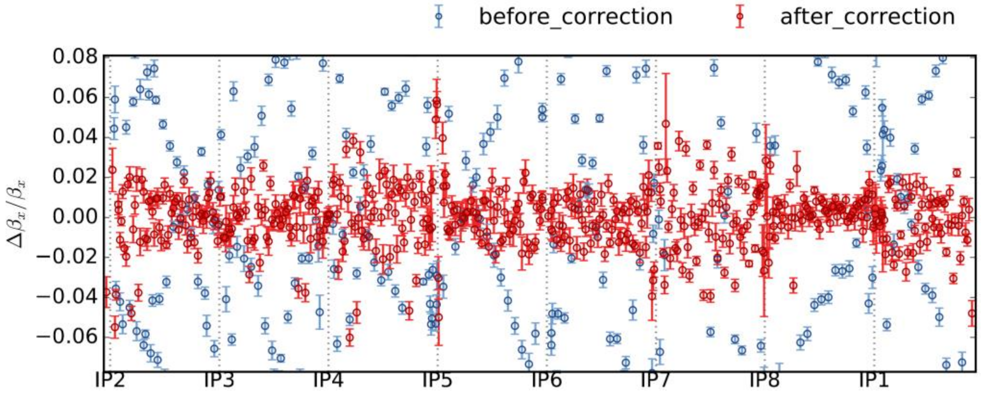
- **Noise** in the measured optics functions
  - → How to **reduce the noise** without removing valuable information?

- **Missing data points** due to the presence of faulty BPMs
  - → How to **reconstruct** the missing data?

I. Estimation of magnetic errors

Optics corrections at the LHC

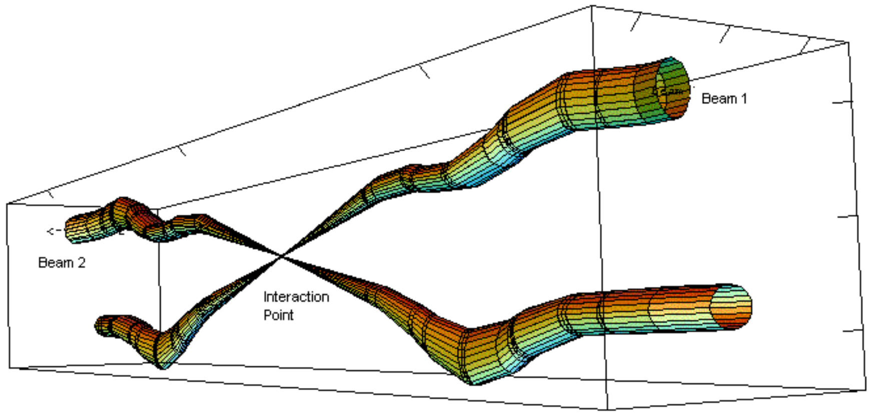
- Corrections aim to **minimize the difference between the measured and design optics** by changing the strength of corrector magnets – single quadrupoles and **quadrupoles powered in circuits**.



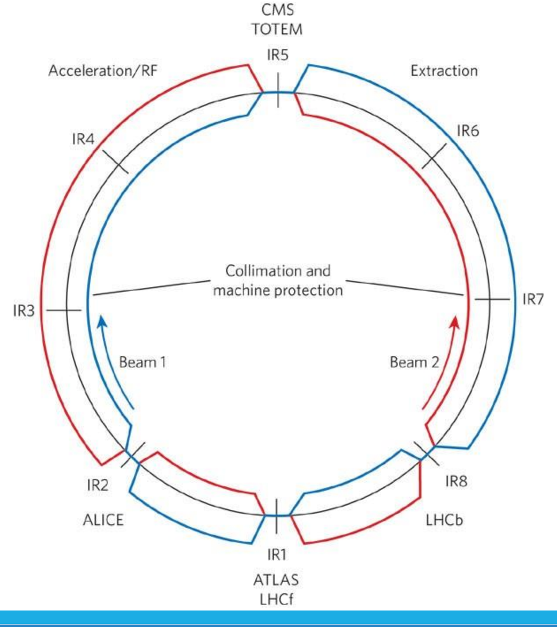
<u>Optics corrections in the LHC are currently based on</u>:
– **Local** corrections around Interaction Points
(e. g. Segment-by-Segment method)

– **Global** corrections using a *Response Matrix* between available correctors and optics observables.

– For each beam **separately.**

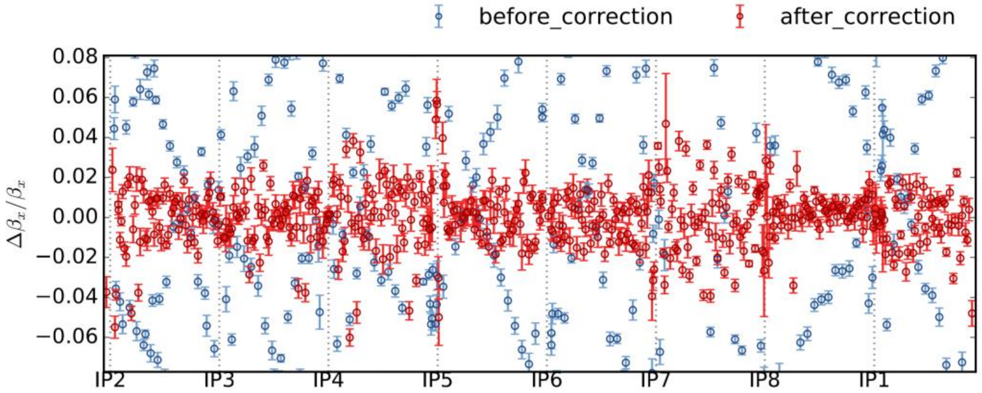


Relative beam sizes around IP1 (Atlas) in collision



Optics corrections at the LHC

- Corrections aim to **minimize the difference between the measured and design optics** by changing the strength of corrector magnets – single quadrupoles and **quadrupoles powered in circuits**.



- What is the actual error of each individual magnet?

Optics corrections in the LHC are currently based on:
– **Local** corrections *around Interaction Points*
(e. g. Segment-by-Segment method)

– **Global** corrections using a *Response Matrix* between available correctors and optics observables.

- Appropriate weights of observables in the response matrix are adjusted manually**.**

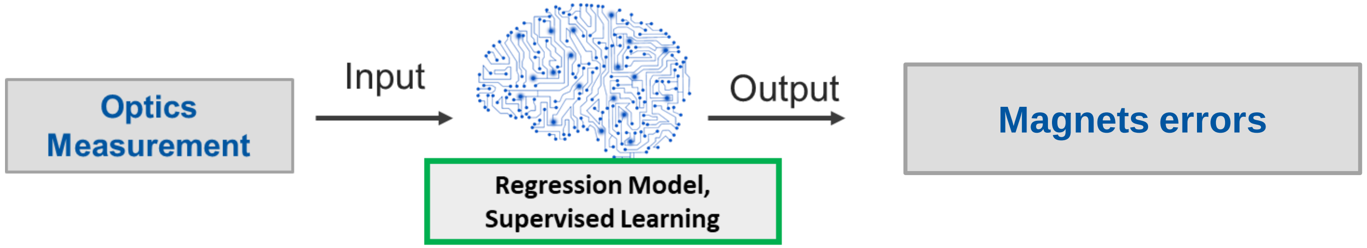
– For each beam **separately.**

- How to determine the whole set of errors for both beams simultaneously?

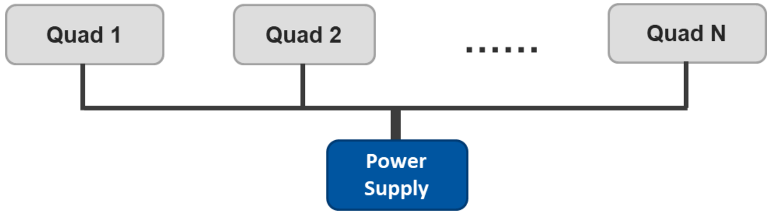
Supervised Learning & multivariate regression

General concept

- Train **supervised regression model** to predict magnet errors from optics perturbations caused by these errors.



- Corrections are implemented by changing the strength of **circuits** – magnets powered in series



Schematic circuit representation

- Optics perturbations are caused by **single magnets** all around the ring
- Training data has to consist of pairs: *"input – <u>correlated to</u> – **known** target values"*

Predict **single quadrupole errors** directly correlated with the optics perturbations.

- **Correlations** between magnetic errors and optics deviations from design can be **learned by ML-model**.
- Large dataset is needed in order to train a regression model: **simulations!**

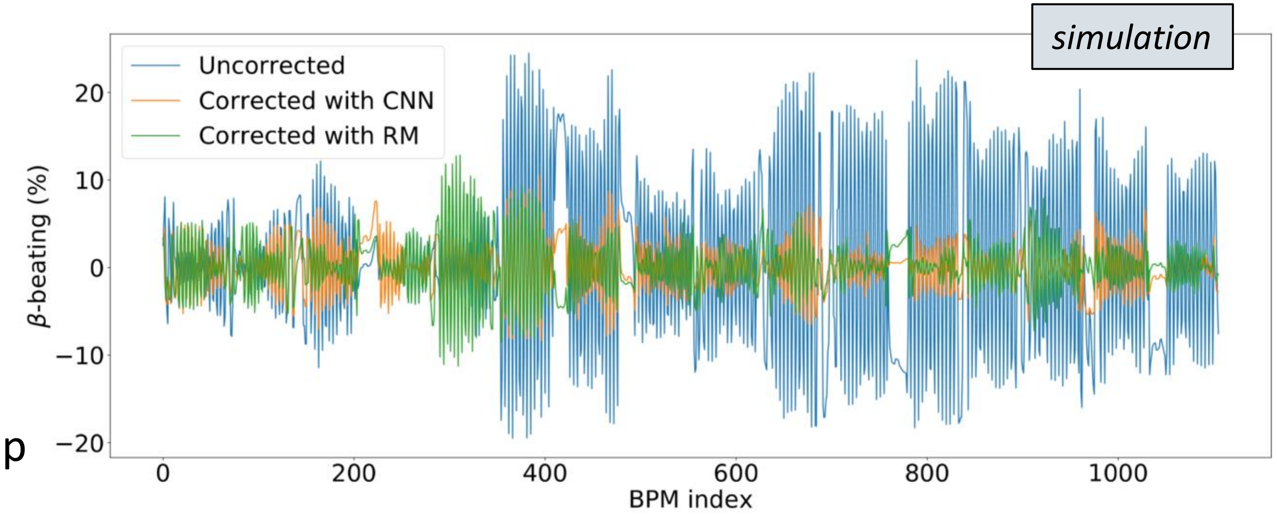
Simplified studies: optics deviations caused by circuits errors

- **Training** data: perturb the optics by changing the strength in the **circuits (quadrupoles powered in series)**
- **Validation**: simulations perturbed with errors in **individual quadrupoles**

Different algorithms are compared:
Orthogonal Matching Pursuit, Random Forest, Convolutional Neural Network:

- Similar results
- Linear Regression as baseline model:
  - easier to interpret,
  - faster to train,
  - mostly linear effects are present in simulations.

- Increasing the complexity of simulations step by step by adding additional error sources, exploring limitations of regression models.



→ *Correction results using Convolutional Neural Network are similar to Response Matrix.*

Linear Regression model as predictor

Linear model for *input X, output Y - pairs*, *i* – number of pairs (training samples), with *weights w:*

$$f(X, w) = w^T X$$

Residual sum of squares as **loss function** for model optimization:

$$L(w) = \sum_i \left(Y_i - f(X_i; w)\right)^2$$

Find **new weights** minimizing the Loss function:

$$w^* = \arg min_w L(w)$$

Update weights for each incoming input/output pair

- Generalized model explaining relationship between input and output variables in **all training samples**.
- Test the model on unseen validation data.

→ How to improve the predictive power of the model?

Weights update regularization & bagging

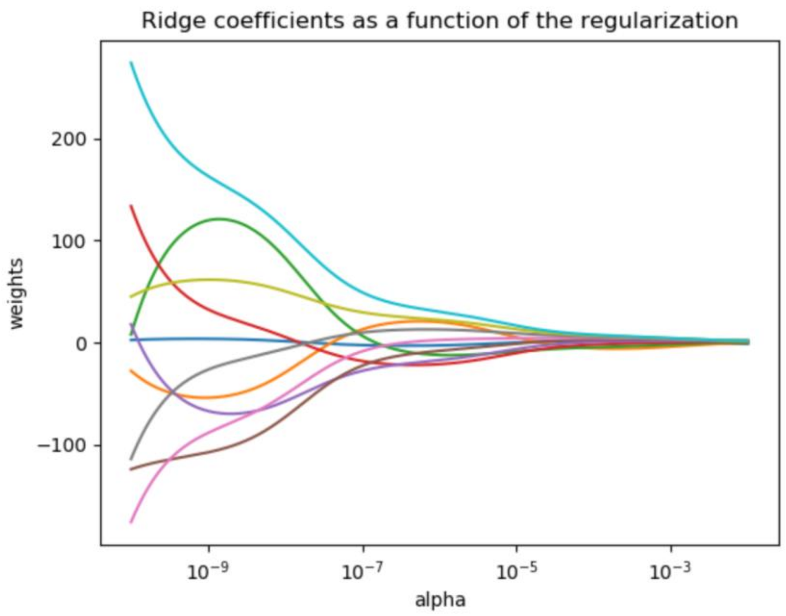
Too much "flexibility" in weights update can lead to *overfitting*

→ **Regularization** places constraints on the model parameters

- Trading some bias to reduce model variance

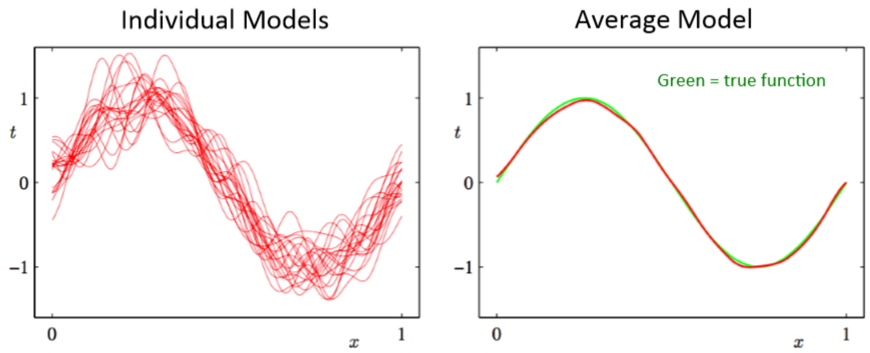
- Using **L2-norm**: $\boldsymbol{\Omega(w) = \sum_i w_i^2}$, adding the constraint $\boldsymbol{\alpha\Omega(w)}$ to the weights update rule: **Ridge Regression**

- The larger the value of $\boldsymbol{\alpha}$, the stronger the shrinkage and thus the coefficients become more robust.



→ **Bagging**: Bootstrap Aggregating: reduce variance of the model, without increasing systematic error of prediction:

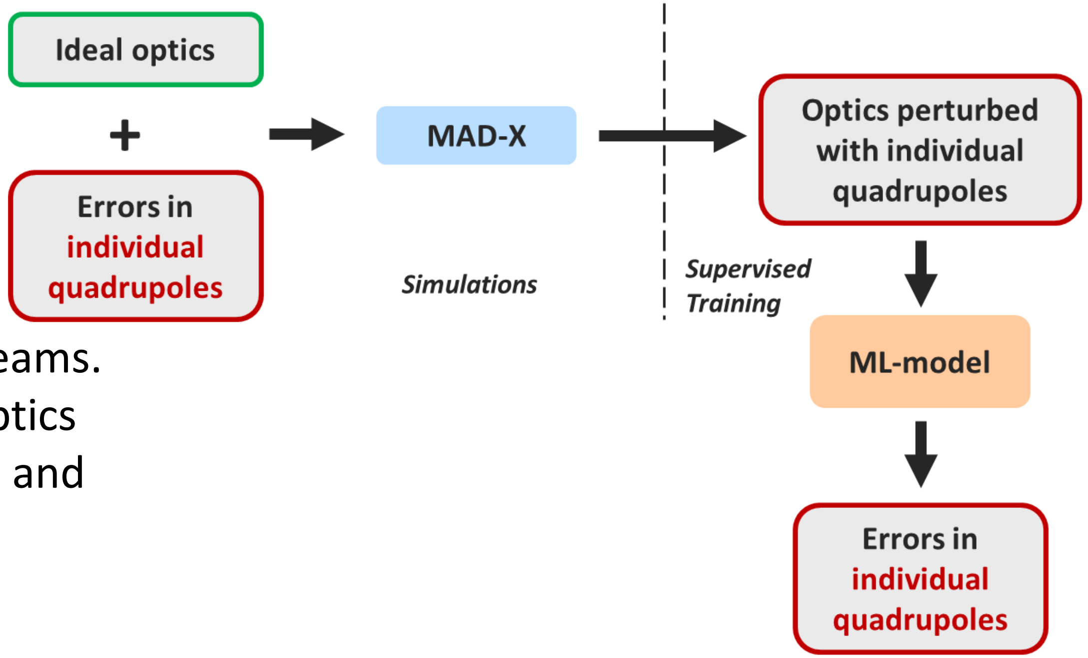
- Ensemble of slightly different models
- Train a separate model on a subset of training data
- Average output of each predictor for the final output.



[Bishop, "Pattern Recognition"]

Data generation for LHC and model training

- Optics for β*=40 cm, collision mode, 6.5 GeV
→ to test the model on available measurements of uncorrected machine (LHC commissioning in 2016)



Data generation for training and test on simulations

- **1256 target** variables
  - assigned gradient errors in the **all** quadrupoles, **both** beams.
- **3304 input** variables: simulated deviations from the design optics in betatron phase advance, normalized dispersion at all BPMs and β at BPMs next to Interaction Points.
- Adding realistic **noise** estimated from the measurements.

➡ **Realistic training data to make adequate prediction from measurements.**

<u>Selected model:</u>

- Scikit-Learn implementation of **Ridge Regression** (regularization parameter α=0.001)
- **Bagging-estimator** (combining 10 Ridge Regression – models)
- 80000 training samples (divided into training and test sets)

How to evaluate trained models?

- "**ML point of view**": compare predicted magnet errors with corresponding true values.

  Figures of merit: $MAE(y,\hat{y}) = \sum_{i=1}^{n} |y_i - \hat{y}_i|$ $\quad R^2(y,\hat{y}) = 1 - \frac{Var\{y-\hat{y}\}}{Var\{y\}}$

- **In terms of optics**:
  ML-model input: **optics** perturbed with magnet errors to be predicted

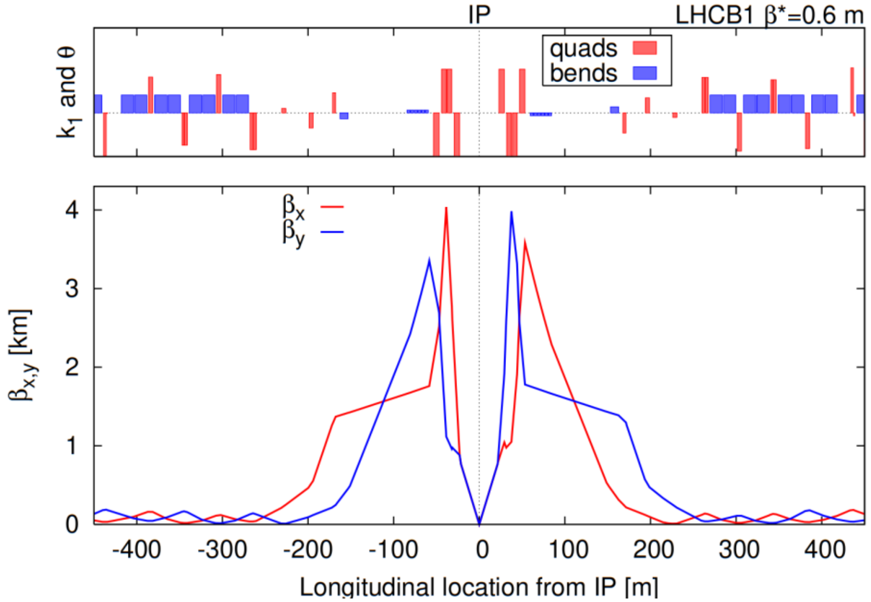
  ML- model output: **magnet errors** estimated from optics perturbations

How well can be reconstructed?

- Quadrupole **magnets close to the IPs** produce the **largest optics perturbations.**
- **Triplet**: assembly of quadrupole magnets used for a reduction of β-function at the IPs.
  → Important to verify if ML-model can produce reliable reconstruction of these errors.

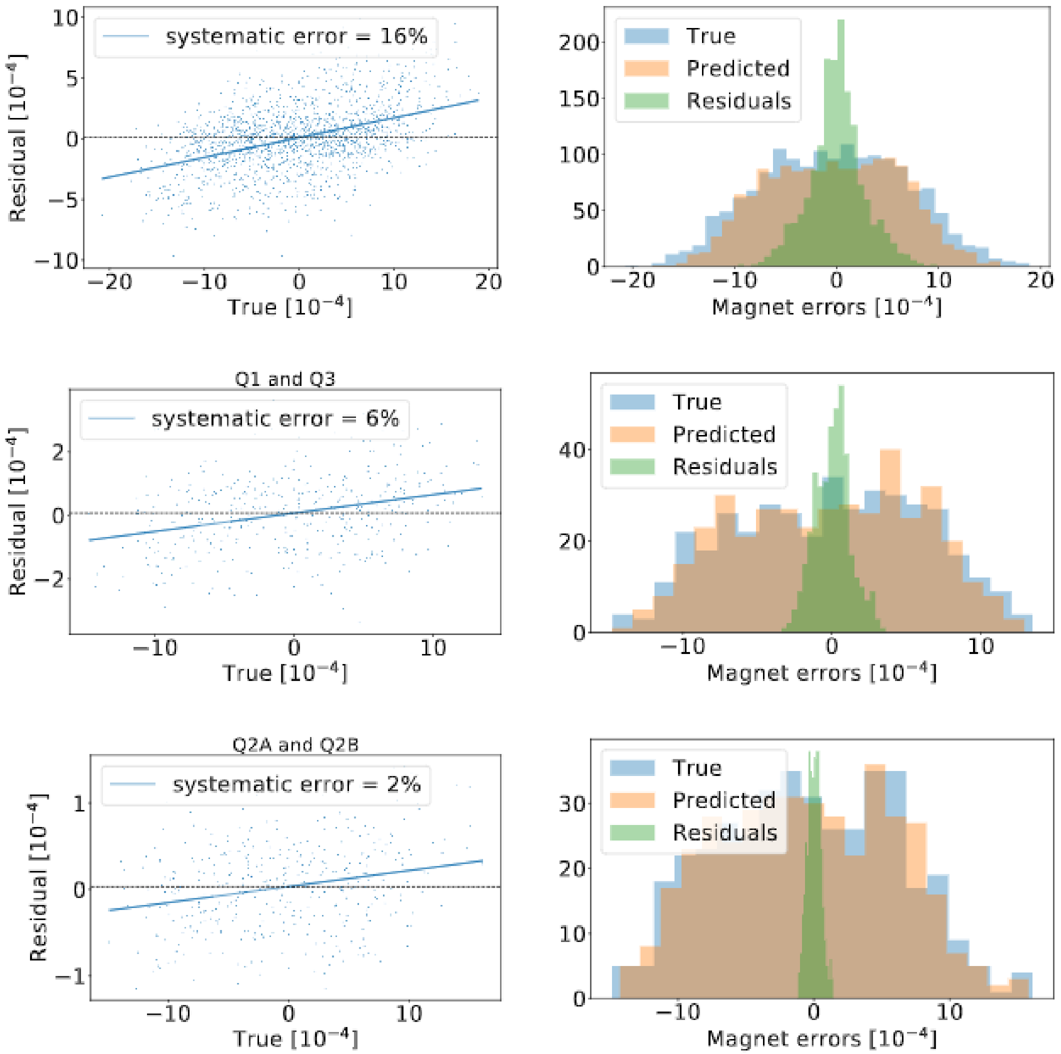
ML-model is trained to **predict all quadrupoles** in the machine.

→ **Evaluating** triplets and arcs magnet errors prediction separately



Results on simulations: errors of prediction

- Comparison between true simulated and predicted errors:



Field errors in individual magnets around Interaction Points

systematic prediction error (bias) → 16**%,**
random error ~ 25%.

Combining individual quadrupole errors according to the powering scheme in the LHC

6% systematic prediction error

2% systematic prediction error

Reconstructing optics with predicted magnet errors

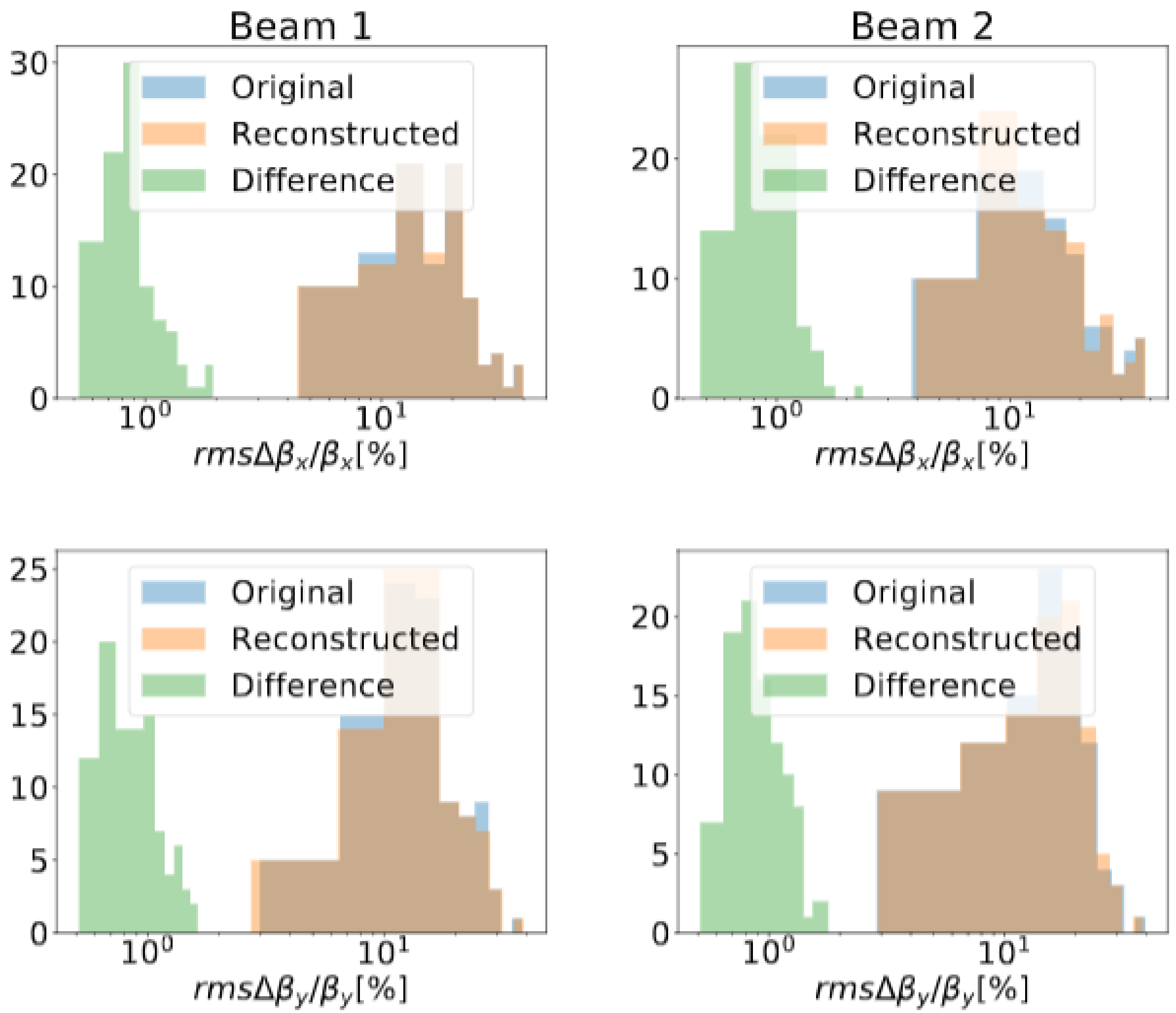
Ideal optics +
simulated errors = perturbed optics

Difference
$\Delta\beta/\beta_{mdl}$?

Ideal optics +
predicted errors = reconstructed optics

→ Very good agreement between the optics simulated **with true magnetic** errors and simulations generated **with the errors predicted by the model.**

Predicted magnet errors agree with the errors that introduced the original beta-beating
→ beta-beating can be corrected.

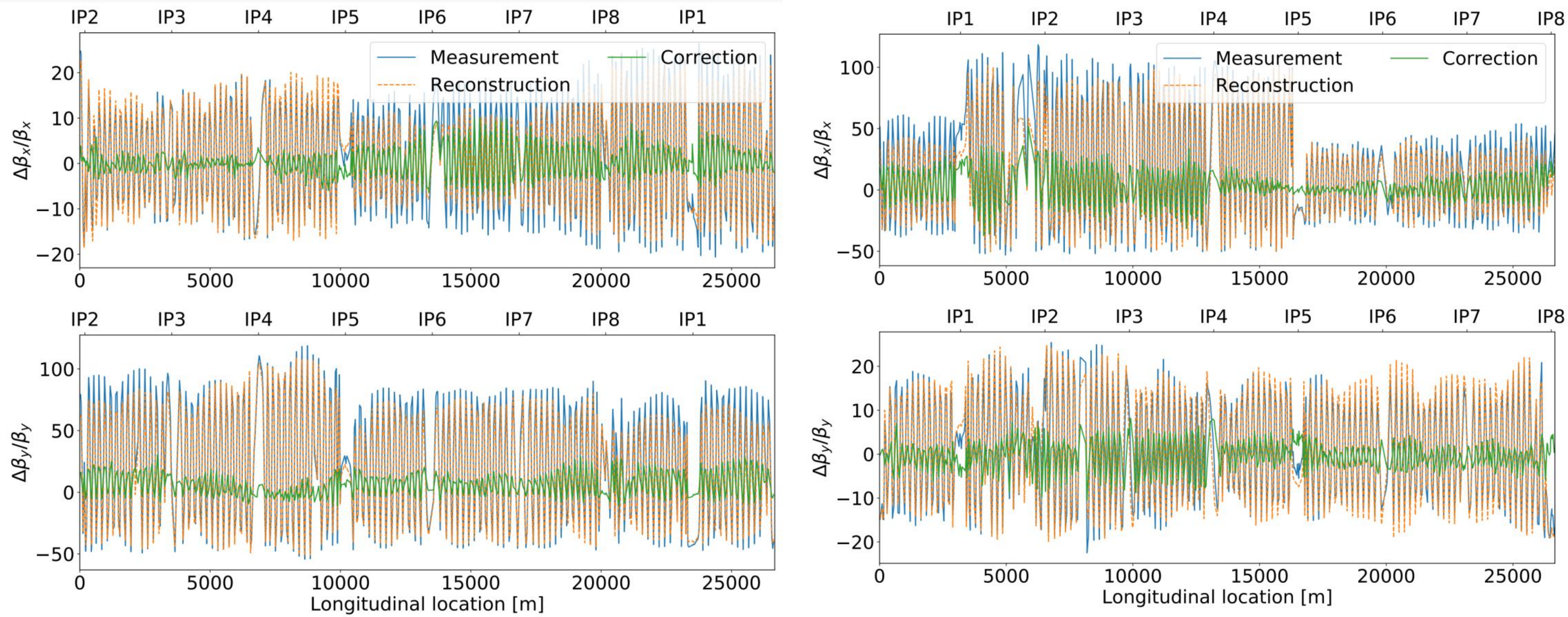


Results on experimental data: 2016 LHC commissioning

“Ground-truth” of magnet errors is unknown unlike simulations.

1. Use predicted magnet errors to simulate optics perturbation
2. Compare produced simulation to actual measurement

→ **Residual error of measured optics reconstruction** (≈ potential correction)

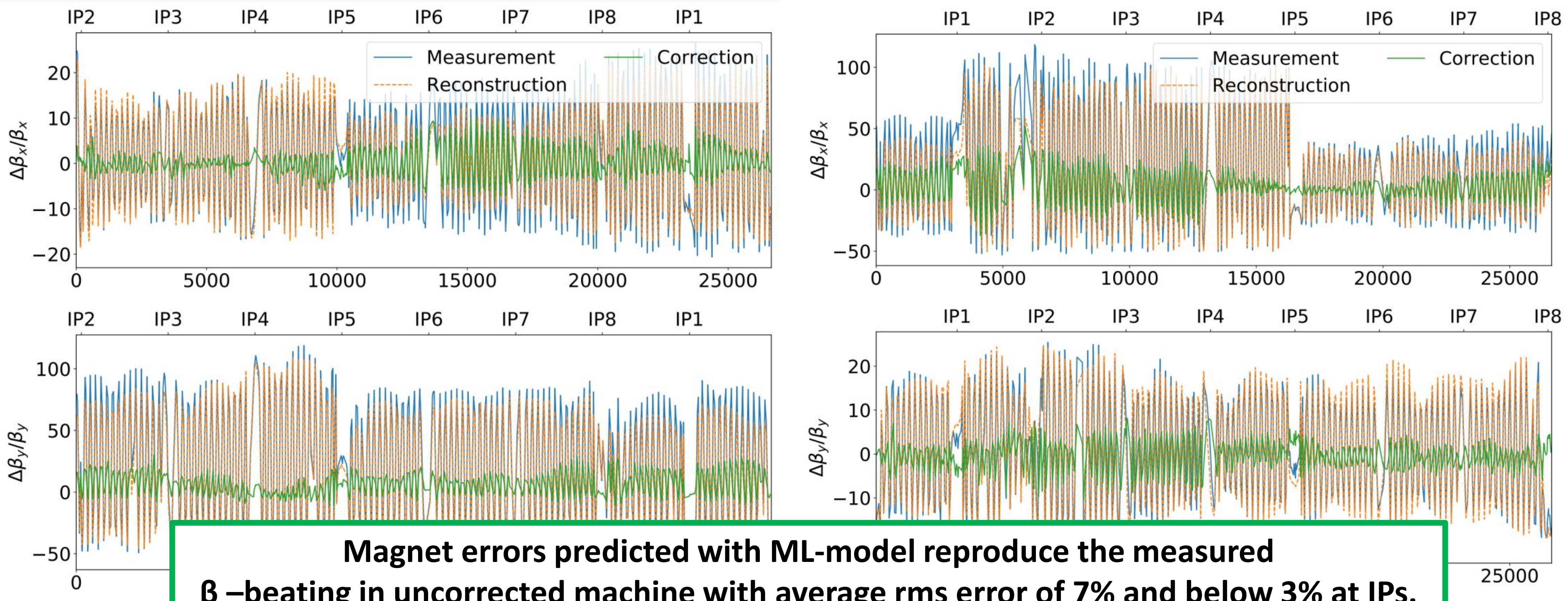


Results on experimental data: 2016 LHC commissioning

"Ground-truth" of magnet errors is unknown unlike simulations.

1. Use predicted magnet errors to simulate optics perturbation
2. Compare produced simulation to actual measurement

→ **Residual error of measured optics reconstruction** (≈ potential correction)



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.

HL-LHC studies

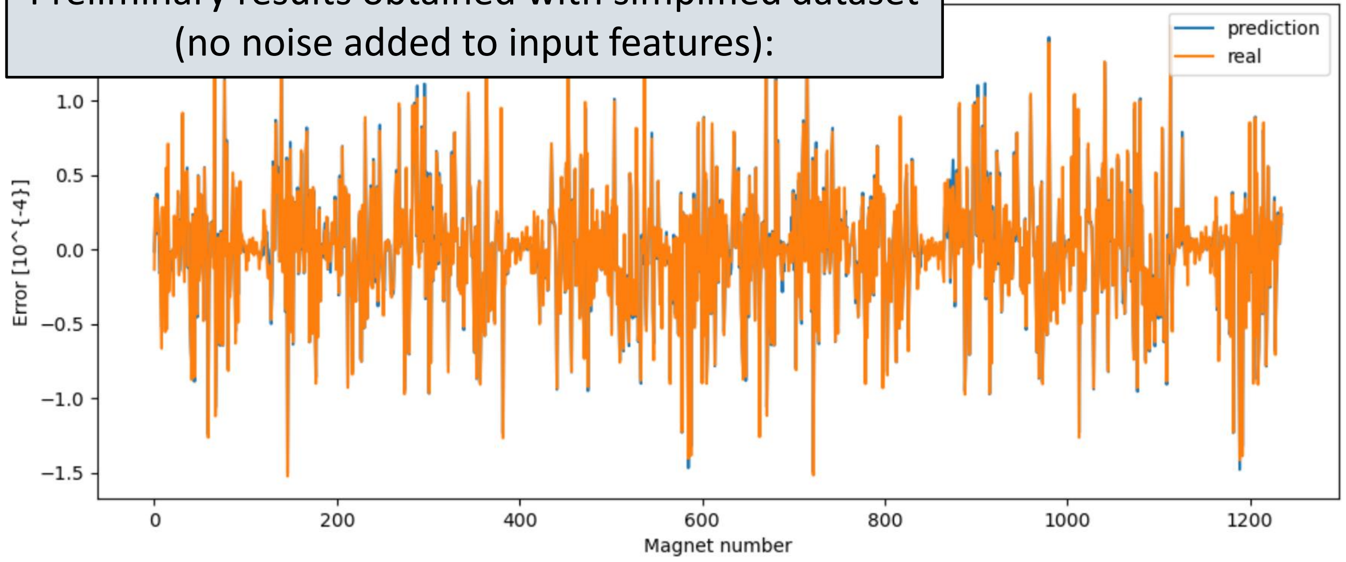
High Luminosity Large Hadron Collider: Upgrade of the LHC to push the performance in terms of beam size and luminosity.

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

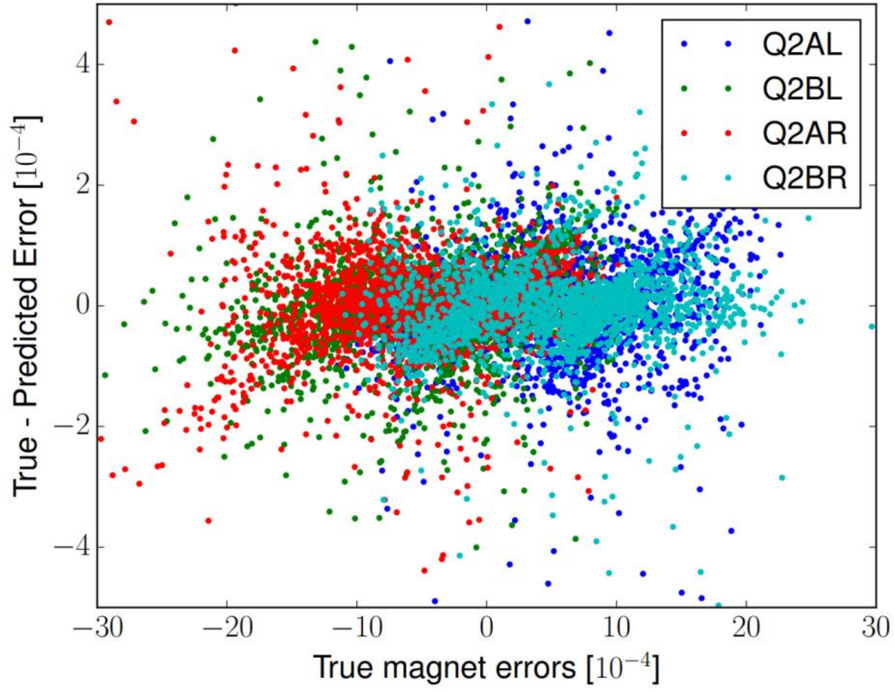
→ new correction strategies are needed.

Courtesy of Hector Garcia Morales

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



Full set of quadrupoles all around the ring



Inner Triplet magnets in IRs

HL-LHC studies

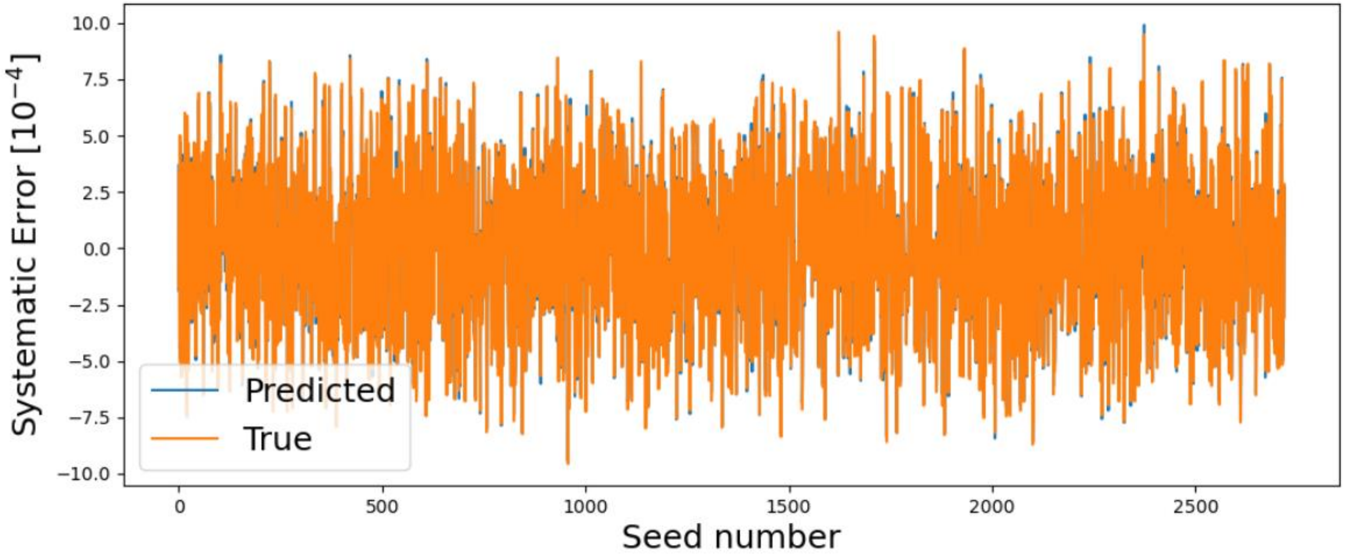
Quadrupole error prediction: Predicting the systematic error

One of the major concerns with the errors in the triplet magnets:

- systematic part of the gradient error (unknown) may have a significant impact on the β-beating.
- The systematic error can be estimated by averaging the error of the different triplet magnets:

$$S \approx \langle \Delta K / K \rangle$$

Courtesy of Hector Garcia Morales



Samples for training/test: 32000 [0.80/0.20]

Model: Ridge.
Regularization parameter: $\alpha = 10^{-4}$

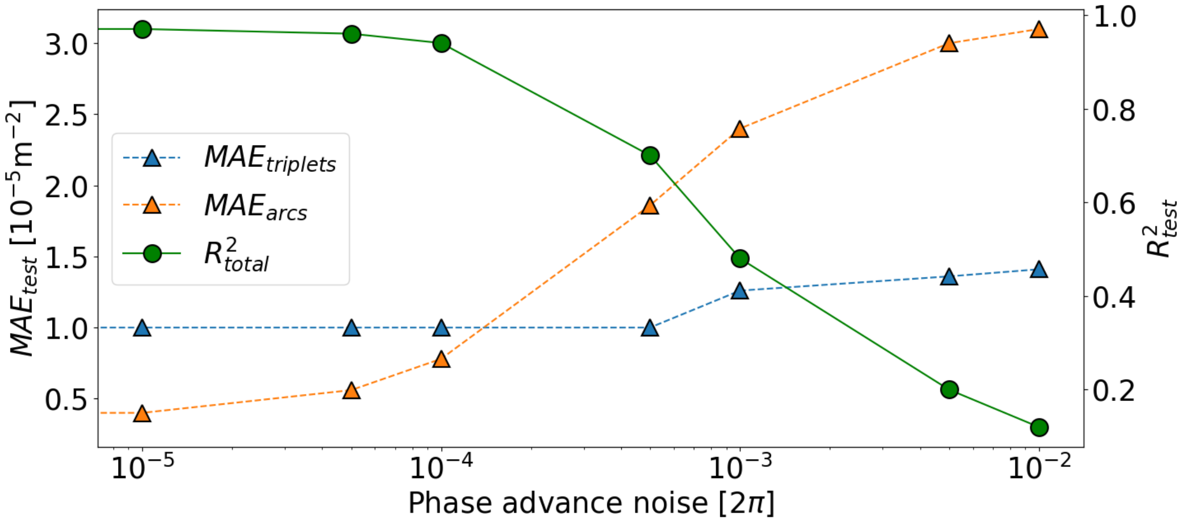
Model Scores

- R2 = 0.89/0.86
- MAE = 3.3/3.8 [10^{-6}]

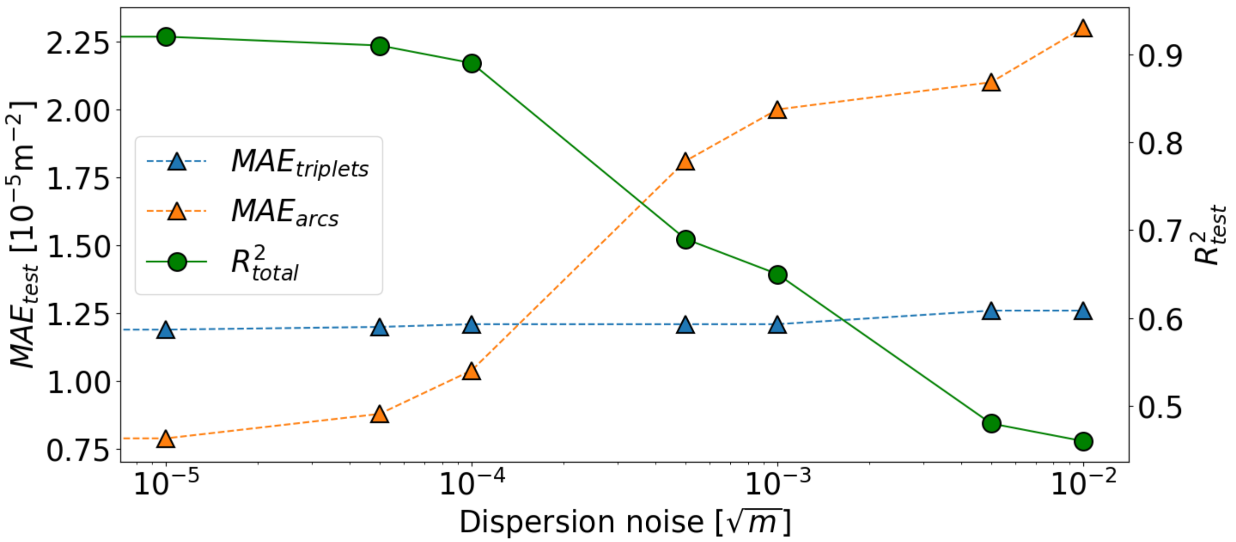
III. Denoising and reconstruction of optics functions

Effect of the noise

Model scores depending on the **phase advance noise** (other input features are not used)



Model scores depending on the **dispersion noise,** phase advance noise is unchanged



- Prediction of **magnetic errors in the arcs** sections suffers from the presence of noise
- Simulations in the **absence of noise** → **very high ML-model scores**
- Increasing prediction quality possible with **more precise measurements** of optics functions used as regression model input.

Experimental data: possible issues

- Training models on **simulations** data: **full set of input features** is always available
- Issues with using measurements as input to make new predictions:
  - General: faulty BPMs → **missing values** at the location of cleaned BPMs
  - Normalized dispersion and β at BPMs next to IPs: special measurements techniques are needed

    → **Features are not always available** e. g. depending on the measurement procedure.
- **Noise in the input data** affects the prediction of the regression models significantly.

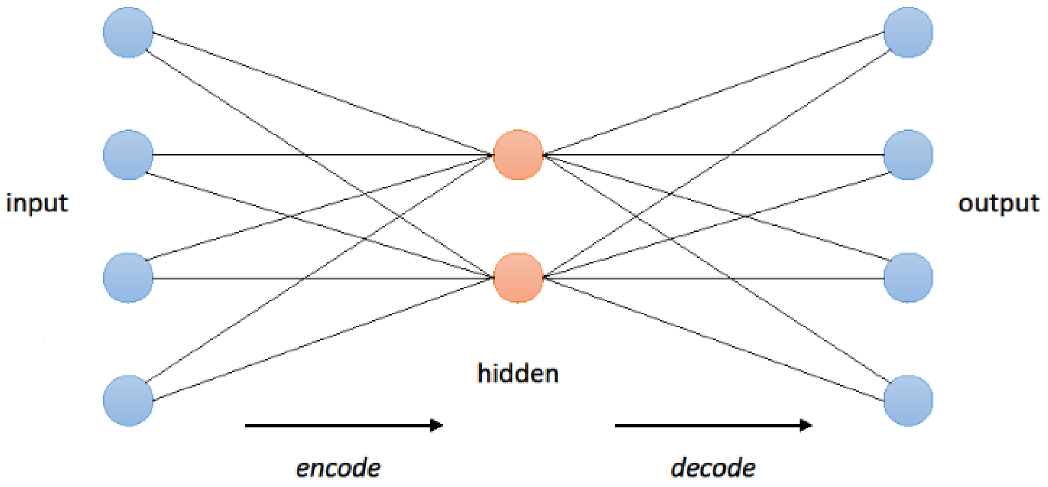
How to deal with missing and noisy data?

Experimental data: possible issues

- Training models on **simulations** data: **full set of input features** is always available
- Issues with using measurements as input to make new predictions:
  - General: faulty BPMs → **missing values** at the location of cleaned BPMs
  - Normalized dispersion and β at BPMs next to IPs: special measurements techniques are needed

  → **Features are not always available** e. g. depending on the measurement procedure.
- **Noise in the input data** affects the prediction of the regression models significantly.

How to deal with missing and noisy data? ⇒ **Denoising Autoencoder**



Encoder: **compressing** the input data to lower dimensions
Decoder: **reconstructing** the data into original input.

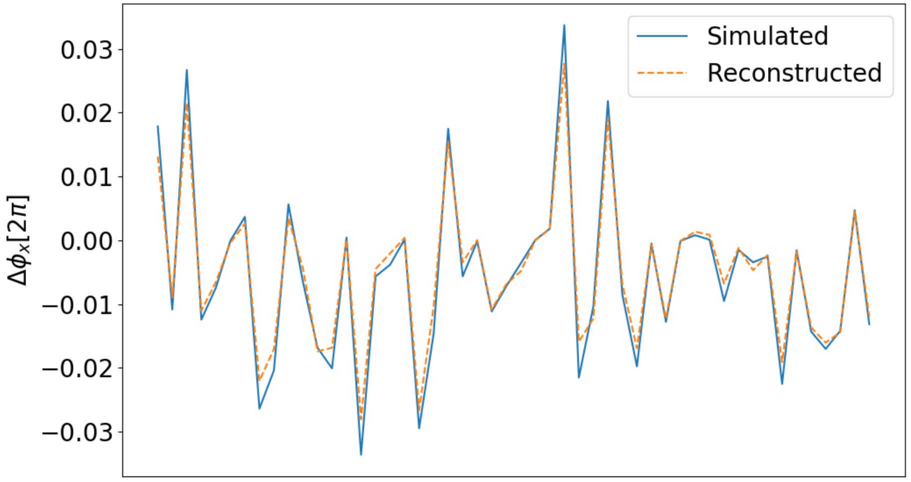
- A special neural network designed to reproduce given input as output of the network

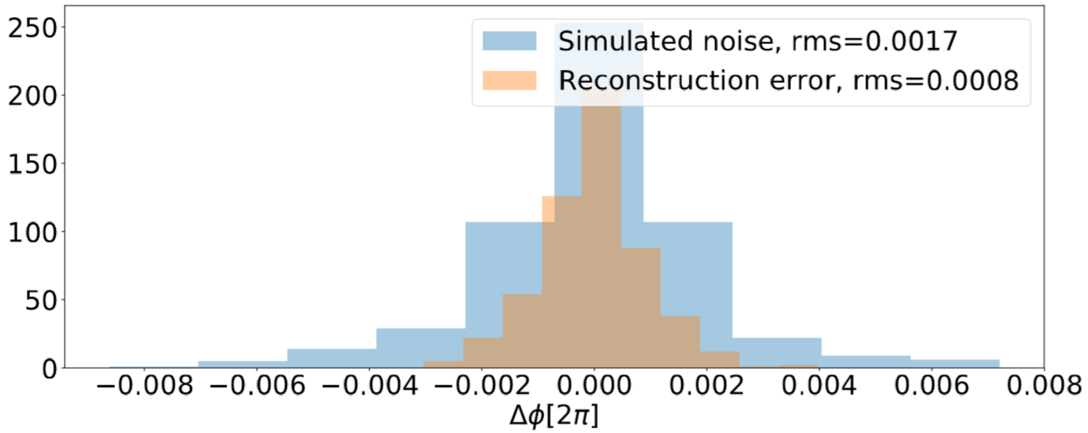
Applications:
- Denoising of data
- Dimensionality reduction
- Generative modeling
- Supervised and unsupervised learning

Reconstruction and denoising of phase advance deviations

- Input: simulated phase advance deviations given noise and replacing 10% of values with 0 (faulty BPMs)
- Output: original simulated phase advance deviations
- Autoencoder with 4 hidden layers, 10000 samples

Reconstruction of missing values in a validation sample



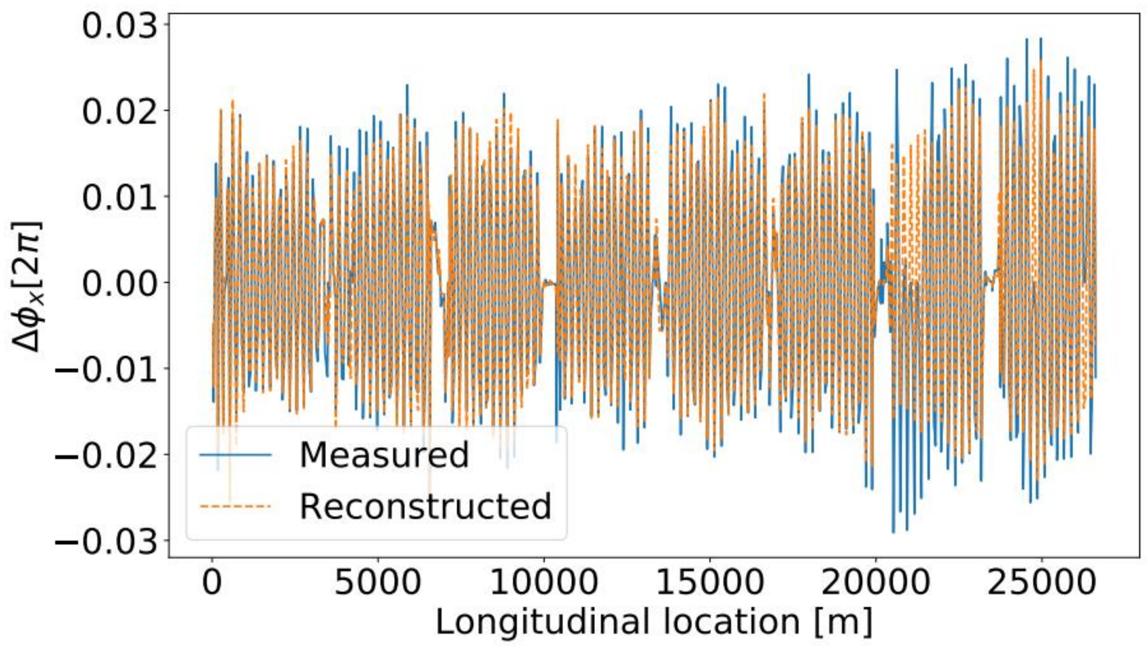


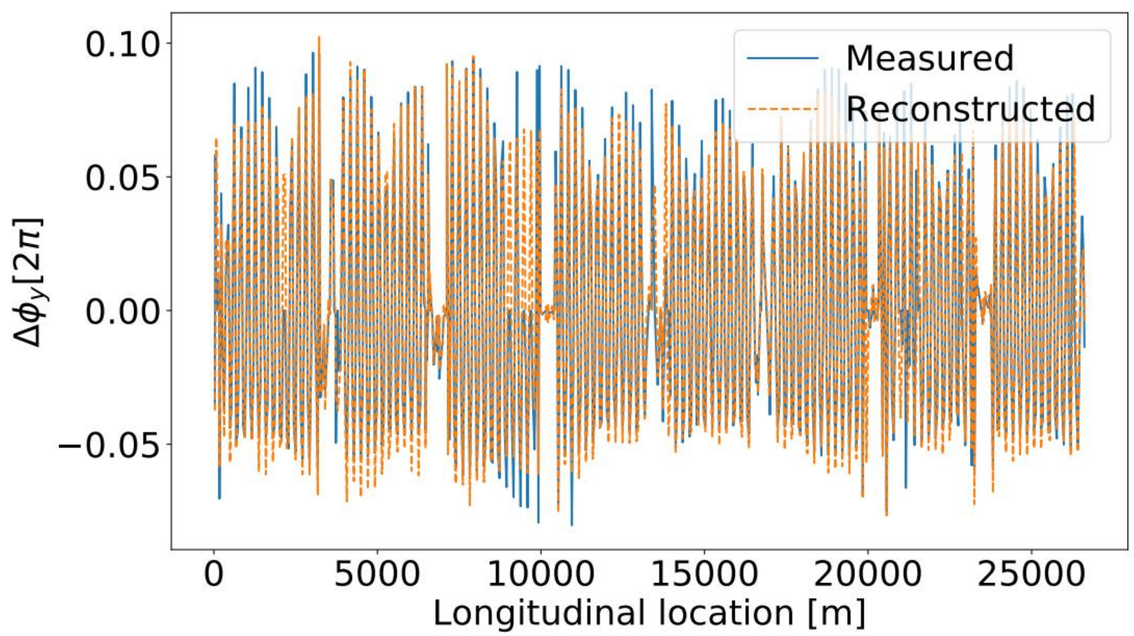
✓ **Missing BPMs**: possibility to obtain reliable estimation of the phase advance deviations at the location of faulty BPMs.

✓ **Full set of phase advance deviations**: reconstruction error is by factor 2 smaller than simulated realistic noise.

Reconstruction of phase advance: **experimental data**

Measurements: β*=40 cm, LHC commissioning 2016, Beam 1, horizontal and vertical planes





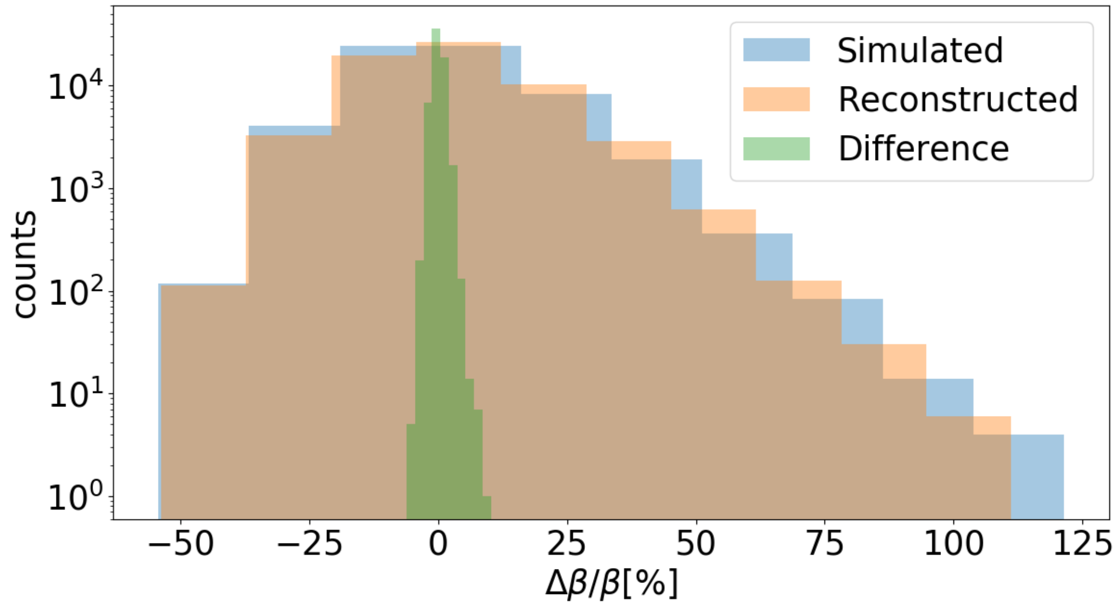
✓ Prediction of phase advance deviation from the model **agrees well** with the measured values **at all available BPMs**

→ **Reliable reconstruction** of the values **at the location of cleaned BPMs** signal.

Reconstruction of $\boldsymbol{\beta}$ - function

- β-function at IPs and at the location of the triplet quadrupoles is computed by performing k-modulation technique
- β-function around IPs provides important information for the estimation of triplet errors, but data is not always available (e. g. due to the measurements procedure in the past)

***Simulation:** summary of 1000 seeds*



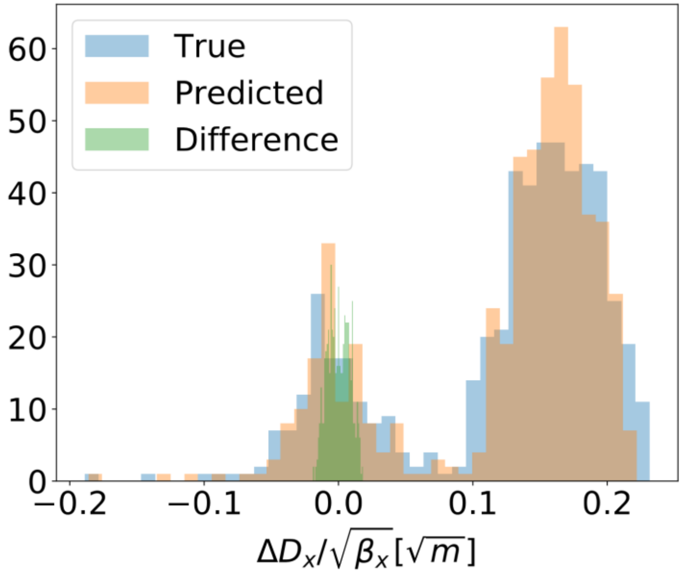
- **Input**: simulated phase advance deviations given noise (beam 1 and 2, horizontal and vertical planes)
- **Output**: $\Delta\beta$ errors at 2 BPMs left and right from IPs 1, 2, 5 and 8 (32 variables in total)
- Ridge Regression, 10 000 training samples

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

Reconstruction of **normalized dispersion**

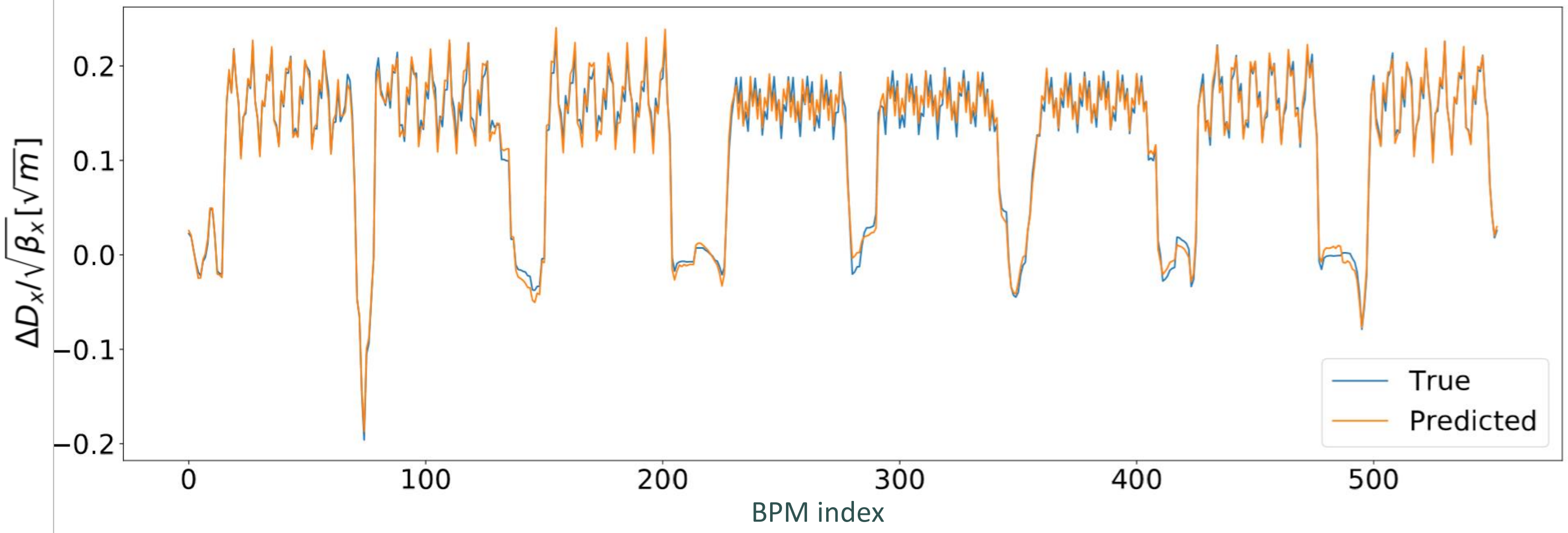
- **Input**: simulated phase advance deviations given noise
- **Output**: normalized dispersion $\Delta D_x/\sqrt{\beta_x}$
- Using **linear regression model**: Ridge Regression, 10 000 samples

***Simulation** example: Beam 1*



Simulated rms $\Delta D_x/\sqrt{\beta_x}$: 0.0802 $\sqrt{m}$

RMS-error between simulation and reconstruction: 0.007 $\sqrt{m}$



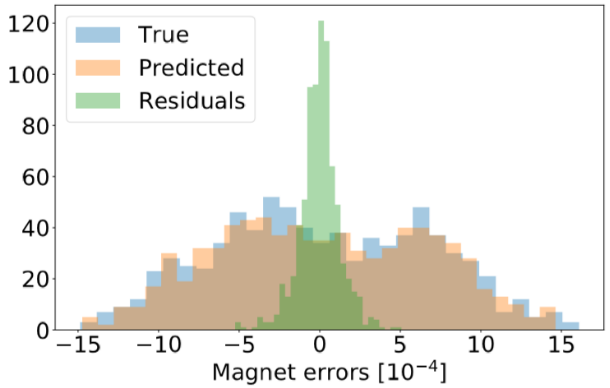
Conclusion and outlook

Optics corrections based on Supervised Regression models:

- Optics corrections today are done in two steps (local and global).
- ✓ ML-models allow to **predict all quadrupole errors** for both beams simultaneously, local and global errors in one step
- ✓ Promising results on simulations and experimental data, especially for **optics corrections in Interaction Regions** (2 - 6% systematic error)
- ✓ Tested on **different optics settings** (“ballistic” optics, triplets switched off)

Current limitations:

- Only linear error sources in training simulations
- Prediction of arc magnet errors highly depends on the noise in the measured optics observables.

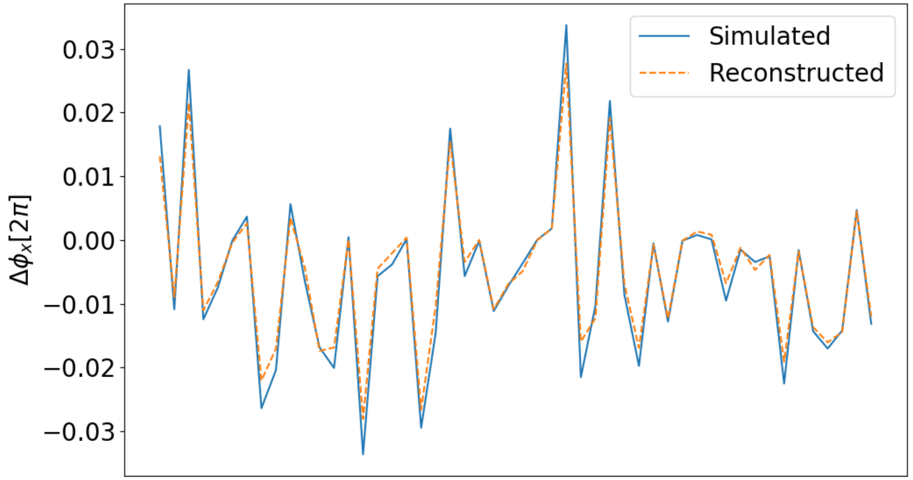


Residual error for a group of triplet quadrupoles

Conclusion and outlook

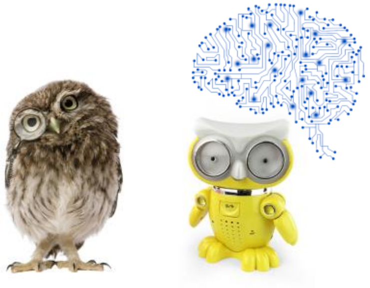
Denoising and reconstruction of missing data:

- ✓ Successfully demonstrated on simulations the possibility to **reduce noise** in phase advance measurements **using autoencoder.**
- ✓ **Reconstruction of missing features** for the magnet errors prediction → tested on measurements data.
- ✓ Providing **estimates of optics functions**, when time costly measurements techniques cannot be performed.



Outlook:

- ○ **Correctors settings (circuits strengths)** from predicted individual errors
- ○ Integration into operational LHC software infrastructure.



Thank you very much for your attention!