TECHNISCHE UNIVERSITÄT DARMSTADT

PHYSICS-INFORMED BAYESIAN OPTIMIZATION FOR CLOSED ORBIT CORRECTION

Victoria Isensee, Adrian Oeftiger, Oliver Boine-Frankenheim

09.10.2024

MOTIVATION

We want to learn something about the machine → Gain better models of the accelerator:

- Misalignments of dipoles, quadrupoles, sextupoles ...
- Nonlinearities
- Noise

Synchrotron SIS18 at GSI:

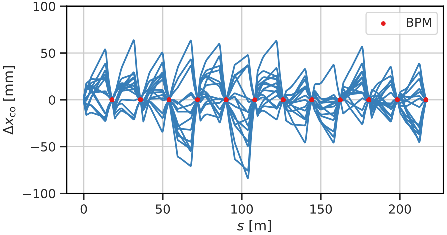
Source: A. Oeftiger

Utilizing the learned model of the accelerator for closed orbit correction

TECHNISCHE UNIVERSITÄT DARMSTADT

MOTIVATION

Simulations: Synchrotron SIS18 closed orbit correction



Remaining closed orbits with RMS=0 at BPMs

Conventional closed orbit correction methods (e.g. SVD-based [1] or FFT-based [6]) are widely adopted:

- The concept is to minimize the deviation **at the BPMs** (Beam Position Monitors)
- Remaining multitude of residual closed orbit errors
- Objective: Achieve a minimal discrepancy between the closed orbit and the target orbit **throughout the entire ring**

→ A model for in between BPMs is needed

TECHNISCHE UNIVERSITÄT DARMSTADT

MOTIVATION FOR PHYSICS-INFORMED BO

A model for in between BPMs:

- Lattice fitting methods
  (e.g. LOCO/NOECO (nonlinear optics from off-energy closed orbits)) [4][2]
  $\rightarrow$ Necessitates (repeated) measurement of the orbit response matrix (ORM)
- Black-box Bayesian Regression method
  $\rightarrow$ Does not include beam dynamics

Physics-informed Bayesian Optimization (BO):

Combines a probabilistic modeling perspective with full beam dynamics around the machine in order to infer pattern between the BPMs in an effective manner

GOALS

Phase 1: Physics-informed Bayesian Regression

1. Development of a physics-informed surrogate model with uncertainty quantification not only at BPMs but also for *in between* BPMs
2. Construction of the model in the most effective manner with the minimum amount of measurement data required

Phase 2: Physics-informed Bayesian Optimization

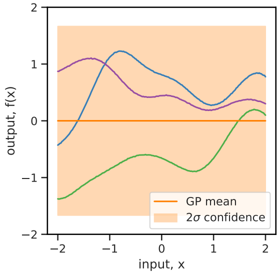
→ Closed-orbit correction in between BPMs taking into account a model for distortion
→ Achieve minimal deviation at *specific location in between BPMs* (e.g. at septum)
→ Extract lattice functions (e.g. for dispersion correction)

CONTENTS

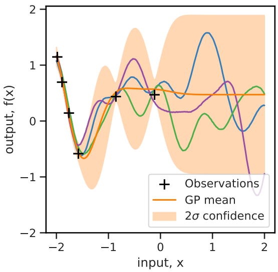
BASICS: GAUSSIAN PROCESS

- A probability distribution (Gaussian Process, GP) acts as a surrogate model
- The GP is fully described by mean function and the covariance function or kernel
- Realizations $Y_m(x)$ conditioned on observations (+)

prior



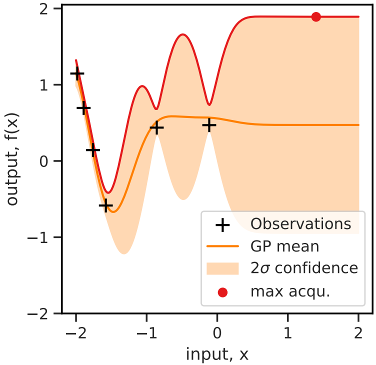
posterior



(based on: “Gaussian processes for machine learning.” by C.Rasmussen and C. Williams.)[3]

BASICS: BAYESIAN OPTIMIZATION

Searches for optimum in efficient way and controls trade-off between exploration and exploitation



- Posterior with acquisition function "Upper Confidence Bound" (UCB)
- Maximum of acquisition function is next most promising evaluation point
- With the next evaluation point the GP is updated
- Hyperparameters are the parameters of the kernel and acquisition function

TECHNISCHE UNIVERSITÄT DARMSTADT

APPROACH: KEY INGREDIENTS

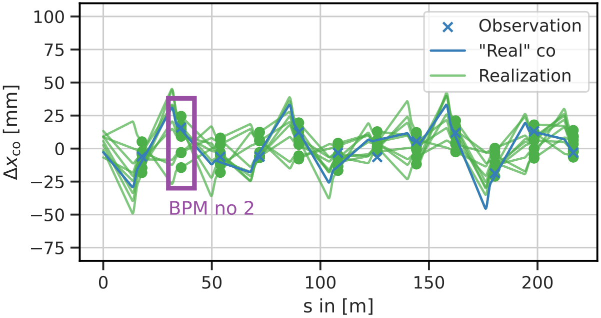
Objective: To build the physics-informed BO model for the regions in between the BPMs

- A GP for every BPM with the input vector of corrector angles $\vec{\theta} = (\theta_{\text{steerer1}}, \ldots, \theta_{\text{steerer12}})$ and the output deviation Δx_{atBPMi}
- Hyperparameters are the **distributions** $(\mu_{\text{quad}}, \sigma_{\text{quad}})$ of the misalignments of the quadrupoles (dipoles)
- Realizations are simulated using the Monte Carlo (MC) method with MAD-X (simulation tool) to include full beam dynamics around the machine
- Kernel (and mean function) is **estimated** utilizing the realizations [5]

(Paper: "Physics-informed kriging: A physics-informed Gaussian process regression method for data-model convergence." by X. Yang, G. Tartakovsky, and A. Tartakovsky)

APPROACH

Simulation example for the closed orbit of the SIS18:



- Fixed vector $\vec{\theta}$ (steerer angles)
- Observations from “real” accelerator including BPM noise
- One realization corresponds to one specific quadrupole setting

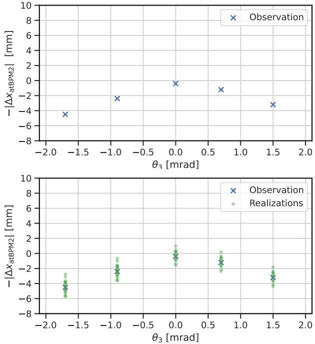
APPROACH

A GP for every BPM with input **vector** $\vec{\theta} = (\theta_{\text{steerer1}}, \dots, \theta_{\text{steerer12}})$ and output Δx_{atBPMi}

How to create the GPs:

1. "Measure" a set of observations of the "real" machine
2. Set initial values for hyperparameters $\mu_{\text{quad}}, \sigma_{\text{quad}}$ (quadrupole misalignment)
3. Simulate m realizations at "observation location" (with Monte Carlo method and simulation tool MAD-X)

Example for BPM no 2:
One component, θ_3, of $\vec{\theta}$ varies, while the remaining components are fixed



TECHNISCHE UNIVERSITÄT DARMSTADT

APPROACH

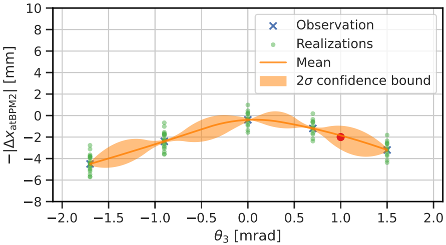
4. Estimate mean and covariance function of the realizations

$$\mu_{\mathrm{MC}}(\vec{\theta}) = \frac{1}{M}\sum_{m=1}^{M} Y^m(\vec{\theta})$$

$$k_{\mathrm{MC}}(\vec{\theta},\vec{\theta}') = \frac{1}{M-1}\sum_{m=1}^{M}(Y^m(\vec{\theta}) - \mu_{\mathrm{MC}}(\vec{\theta}))(Y^m(\vec{\theta}') - \mu_{\mathrm{MC}}(\vec{\theta}'))$$

5. Construct the GPs and condition them to the observed data

Example for BPM no 2:
One component, θ_3, of $\vec{\theta}$ varies, while the remaining components are fixed



For test value θ_3^* a prediction can be made

APPROACH

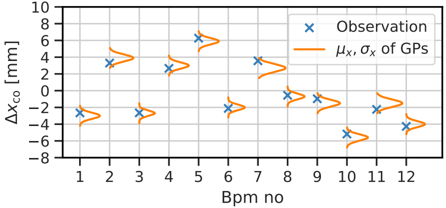
4. Estimate mean and covariance function of the realizations

$$\mu_{\mathrm{MC}}(\vec{\theta}) = \frac{1}{M}\sum_{m=1}^{M} Y^m(\vec{\theta})$$

$$k_{\mathrm{MC}}(\vec{\theta},\vec{\theta}') = \frac{1}{M-1}\sum_{m=1}^{M}(Y^m(\vec{\theta}) - \mu_{\mathrm{MC}}(\vec{\theta}))(Y^m(\vec{\theta}') - \mu_{\mathrm{MC}}(\vec{\theta}'))$$

5. Construct the GPs and condition them to the observed data

For test value θ_3^* a prediction can be made:



TECHNISCHE UNIVERSITÄT DARMSTADT

APPROACH

6. Learning and refinement of the model:
   - Adapt hyperparameters to data = **Learning** the model
     $\rightarrow$ By maximizing the Log Marginal Likelihood (LML)
     $\rightarrow$ Find $\mu_{\text{quad}}, \sigma_{\text{quad}}$ for every quadrupole and find noise term
   - Maximize the acquisition function Upper Confidence Bound (UCB) with focus on exploration to find next $\vec{\theta}$

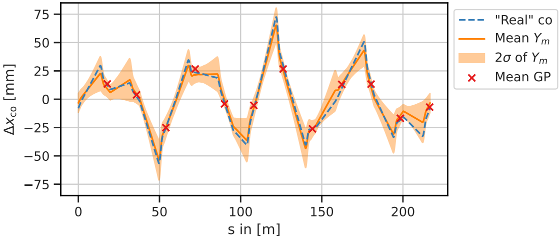
7. Mean and standard deviation of the closed orbit from realizations

8. Use physics-informed surrogate model for optimization

FIRST RESULTS

A **learned distribution** of quadrupole misalignments results in a model of the closed orbit with uncertainty quantification

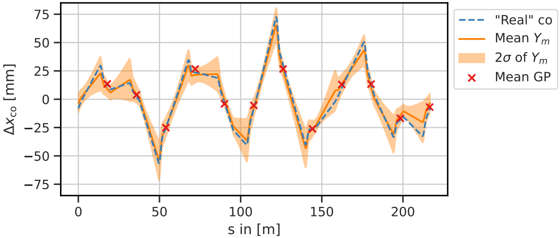
This is what we wanted:



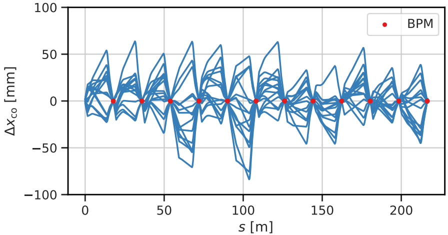
FIRST RESULTS

A **learned distribution** of quadrupole misalignments results in a model of the closed orbit with uncertainty quantification

This is what we wanted:



Closed orbit correction with SVD method:



TECHNISCHE UNIVERSITÄT DARMSTADT

CONCLUSION

A physics-informed model based on Gaussian processes is developed:

- Represents not only at BPMs but also for **in between** BPMs
- Includes **uncertainty quantification**
- Includes **noise handling**
- Constructed in an **effective manner** with the minimal amount of measurements

OUTLOOK

1. Analyse, if the adaption of the hyperparameters is sufficient

2. Examine, how robust the method is and which values can be established for the noise term

3. Use the model for **optimization** and achieve minimal deviation:
    - At the BPMs
    - Throughout the entire ring
    - At specific location in between BPMs

4. Apply to simulations of the more challenging lattice model of the synchrotron SIS100

5. Test the implemented optimization methods during beam time

[1] Y Chung, G Decker, and K Evans. "Closed orbit correction using singular value decomposition of the response matrix". In: *Proceedings of International Conference on Particle Accelerators*. IEEE. 1993, pp. 2263–2265.

[2] David K Olsson, Åke Andersson, and Magnus Sjöström. "Nonlinear optics from off-energy closed orbits". In: *Physical Review Accelerators and Beams* 23.10 (2020), p. 102803.

[3] Carl Edward Rasmussen and Christopher K.I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. Cambridge, Mass. [u.a.], 2006. ISBN: 026218253X.

[4] James Safranek. "Experimental determination of storage ring optics using orbit response measurements". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 388.1-2 (1997), pp. 27–36.

[5] Xiu Yang, Guzel Tartakovsky, and Alexandre Tartakovsky. "Physics-informed kriging: A physics-informed Gaussian process regression method for data-model convergence". In: *arXiv preprint arXiv:1809.03461* (2018).

[6] L.H. Yu et al. “Real time harmonic closed orbit correction”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 284.2 (1989), pp. 268–285. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(89)90292-1. URL: https://www.sciencedirect.com/science/article/pii/0168900289902921.

TECHNISCHE UNIVERSITÄT DARMSTADT

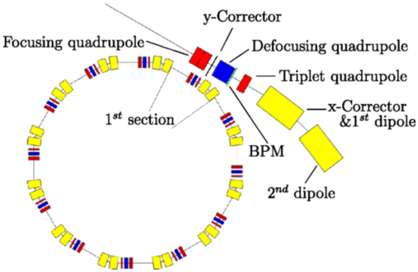
Thank you for your attention!
Questions?

BACK UP

TECHNISCHE UNIVERSITÄT DARMSTADT

SYNCHROTRON SIS18

Schematic of the lattice of the synchrotron SIS18:



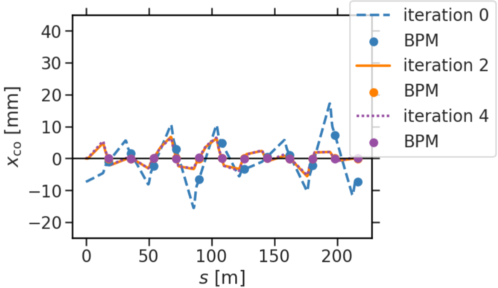
- Focusing, defocusing and triplet quadrupole in each section
- 12 horizontal corrector magnets and 12 vertical corrector magnets (correction angles: $\vec{\theta}$)
- 12 BPM to measure the position of the beam in the horizontal (Δx) and 12 BPM in the vertical plane (Δy)

Source: Mirza, S. & Singh, R. & Forck, P. & Klingbeil, H. (2019). Closed orbit

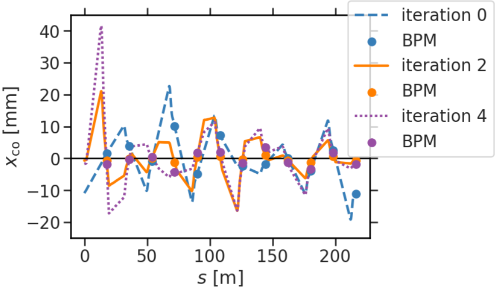
SIMULATION: FAILING OF THE SVD CORRECTION METHOD

Simulationof the conventional SVD correction method applied to broken-symmetry high-transition-energy SIS18 optics (sigma optics):

Standard optics:



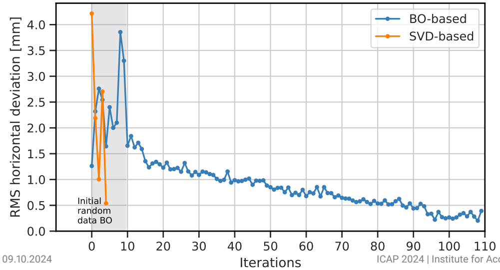
Sigma optics:



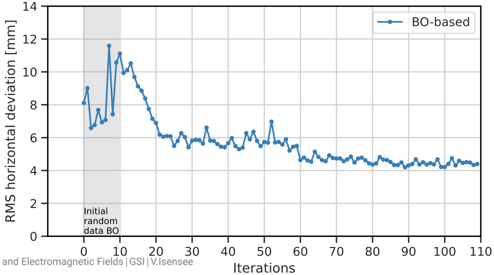
EXPERIMENTAL RESULTS: AUTOMATIC BO-BASED CORRECTION

Automatic BO-based correction of closed orbit using the standard and the broken-symmetry high-transition-energy SIS18 optics. Each evaluation of the objective function requires three acceleration cycles. The gray area marks the initialization phase of the algorithm

Standard optics:



Sigma optics:

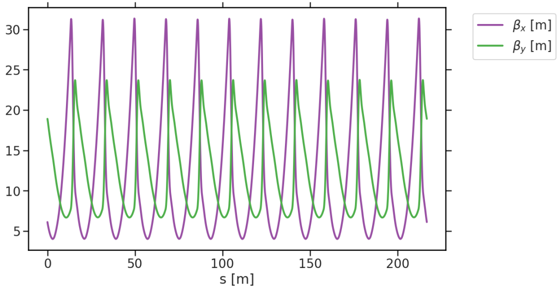


TECHNISCHE UNIVERSITÄT DARMSTADT

SIGMA OPTICS

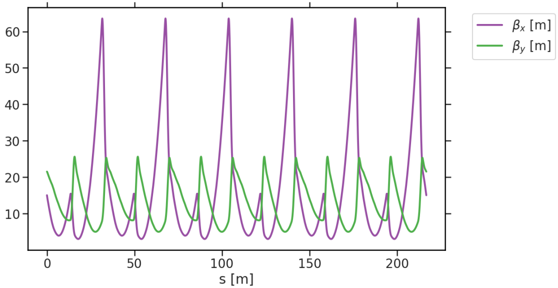
Extraction optics: *Standard optics*

- Doublet optics
- For comparison



Challenging optics: *Sigma optics*

- Shifting γ_t by splitting up even and uneven sector quadrupole families
- Asymmetric setting

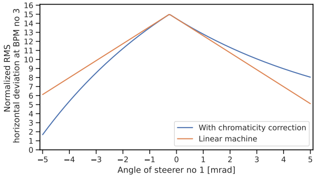


MOTIVATION FOR PHYSICS-INFORMED BO

- Through nonlinearities
- Closed orbit distortion:

$$x_{\mathrm{BPM3}}(s) = \theta_1 \cdot \sqrt{\beta_x(s_0) \cdot \beta_x(s)} \cdot \frac{\cos(|\Delta\psi_x(s)| - \pi Q_x)}{2\sin(\pi Q_x)}$$

Simulations SIS18:



TECHNISCHE UNIVERSITÄT DARMSTADT

CONSTRUCTION OF THE GAUSSIAN PROCESS

The empirical mean of the realizations:

$$\mu_{\mathrm{MC}}(x) = \frac{1}{M}\sum_{m=1}^{M} Y^m(x)$$

The empirical covariance:

$$k_{\mathrm{MC}}(x, x') = \frac{1}{M-1}\sum_{m=1}^{M}(Y^m(x) - \mu_{\mathrm{MC}}(x))(Y^m(x') - \mu_{\mathrm{MC}}(x'))$$

Find the posterior distribution at test array x^*:
$Y(x^*)|X, y, \alpha \sim \mathcal{N}(m(x^*), \sigma^2(x^*))$

$$m(x^*) = \mu_{\mathrm{MC}}(x^*) + c_{\mathrm{MC}}^T(C_{\mathrm{MC}} + \alpha E)^{-1}(y - \mu_{\mathrm{MC}})$$
$$\sigma^2(x^*) = s_{\mathrm{MC}}^2(x^*) + c_{\mathrm{MC}}^T(C_{\mathrm{MC}} + \alpha E)^{-1}c_{\mathrm{MC}}$$

With $s_{\mathrm{MC}}^2 = k_{\mathrm{MC}}(x^*, x^*)$ and $c_{\mathrm{MC}} = k_{\mathrm{MC}}(x^{(1)}, ..., x^{(N)})$

TECHNISCHE UNIVERSITÄT DARMSTADT

OUTLOOK

1. Analysis if the nonlinearity is sufficient

2. Examination which values can be established for noise term α

3. To predict $\Delta x(s_1)$, evaluate with dipole shift

   $$x_{\mathrm{COD}}(s) = \theta \cdot \sqrt{\beta_x(s_0) \cdot \beta_x(s)} \cdot \frac{\cos(|\Delta\psi_x(s)| - \pi Q_x)}{2\sin(\pi Q_x)}$$

   for thin and linear lattice

4. Use distribution of β_x, D_x and evaluate $x_{\mathrm{COD}}(s)$ to predict $\Delta x(s_1)$