Differentiable Preisach Modeling for Particle Accelerator Systems with Hysteresis

Ryan Roussell
rroussel@slac.stanford.edu
Machine Learning Based Accelerator Control

Goals:
• Automate routine tasks + improve performance
• Enable new capabilities

Challenges:
• Practical constraints and complexities of realistic accelerators
• Incorporating prior knowledge
• Scaling
Controlling Beam Optics with Electromagnets

Electron beam

Quadrupole (focusing) magnets

Insertable diagnostic screens
Modeling Hysteresis Effects in Accelerators

We generally only have control over applied currents to magnetic elements

- Current optimization schemes ignore hysteresis effects
- Need to develop a fast-modeling strategy to improve online control without measuring the hysteresis curve directly
Hysteresis Effects in Electromagnets

Applied magnetic field
\( \mathbf{H}_{0:t} = \{ H_0, H_1, \ldots, H_t \} \)

Magnetization
\( x_t = M(\mathbf{H}_{0:t}) \)

Hysteresis model

Minor loop

Major loop
Realistic Accelerator Magnets

$1.26Q3.5(4034)$: set on the way UP

$p_0 = 8.467995e-02 \pm 3.243181e-03$
$p_1 = 5.803860e-01 \pm 8.669393e-04$
$p_2 = -9.392291e-04 \pm 6.034721e-05$
$p_3 = 1.155373e-04 \pm 1.118039e-05$

RMS = 0.00641 A
NDF = 9

$b_2/b_1 \approx 0.62\% < 39\%$

$R = 20 \text{ mm}$

hardat.ru1

http://www-group.slac.stanford.edu/met/MagMeas/MAGDATA/FACET_II/Quad/1.26Q3.5-159974-034/strdat.ru1
The Preisach Model for Hysteresis

Model magnetic domains as individual “hysterons”

Determine the hysteron density

Determine hysteron states via history of applied fields

\[ M(H_{0:t}) = \int_{S^+} \mu(\beta, \alpha) d\beta d\alpha - \int_{S^-} \mu(\beta, \alpha) d\beta d\alpha \]
Parametric Preisach Modeling

\[ \mu(\alpha, \beta) = f(\alpha, \beta; \theta) \]

Sutor et al. (2010)
Non-parametric Preisach Modeling

\[ \mu(\alpha, \beta) = f(\alpha, \beta; \theta) \rightarrow \mu_i = \mu(\alpha_i, \beta_i) \]

However, this greatly increases the number of free parameters. How to solve?
Differentiable Non-parametric Preisach Modeling

Keep track of derivative information during every calculation step.

Enables gradient based optimization of model error with respect to all free parameters using the chain rule.

Easily optimize models with ~10k free parameters.

Image credit: Ayoosh Kathuria
Discretize the space into \( n \times n \) grid and treat the density at each grid point as a free parameter

\[
\bm{\mu} = \{\mu_1, \mu_2, ..., \mu_N\}
\]

\[
M(H_{0:t}; \bm{\mu}) = \sum_{i=1}^{N} \mu_i \gamma_i(H_{0:t})
\]

\[
\text{loss}(\bm{\mu}) = \text{MSE}(Y, M(H; \bm{\mu}))
\]

\[
\bm{\mu}^* = \arg\min_{\bm{\mu}} \text{loss}(\bm{\mu})
\]
Modeling SLAC Quadrupole Magnets

Measured integrated gradient

Normalized Deviation

Polynomial + Perturbation = Total
Modeling SLAC Quadrupoles

Polynomial fit error: 0.23%  Train error: 0.015%  Test error: 0.051%

Roussel R. et al., Accepted by PRL
Can we extend this process to modeling magnets in the beamline already?

Combine the hysteresis model with a flexible model for the beam response: a **Gaussian Process**, train both models simultaneously

Allows us to measure hysteresis characteristics using **beam-based** measurements
Gaussian Process Surrogate Models

Prior belief

Posterior belief

Use Gaussian Processes to represent noisy beam attributes as a function of magnetic fields.

Gaussian Processes - Hyperparameters

Example: APS Injector
Full Stack Modeling

Modeling w/o Hysteresis

Modeling w/ Hysteresis

Roussel R. et al., Accepted by PRL
Full Stack Optimization

Now use the joint model to optimize accelerator inputs to achieve beam objectives.

Applied magnetic field
\( H_{0,t} = \{H_0, H_1, \ldots, H_t\} \)

Hysteresis model

Magnetization
\( x_t = M(H_{0,t}) \)

Gaussian process model

Beam measurement
\( Y_t = f(x_t) + \epsilon \)
Benefits:

- Specify trade-off between exploration and exploitation
- Inherently improves model accuracy in regions of interest
- Enables serial or parallelized optimization strategies
Beamline Optimization with Hysteresis
Beamline Optimization with Hysteresis

\[ \beta^* = 2.0 \]

\[ \beta^* = 0.1 \]

\[ \sqrt{\Delta x \Delta y} \] (mm)

0 50 100 150 200

Iteration

GP, \( H_\varepsilon = 0 \)
GP, \( H_\varepsilon = 0.1 \)
GP, \( H_\varepsilon = 0.4 \)
H-GP, \( H_\varepsilon = 0.1 \)
H-GP, \( H_\varepsilon = 0.4 \)
Conclusions

We can create flexible, high-fidelity models of magnetic elements exhibiting hysteresis.

We can combine the **hysteresis** and **beam response** models to create a complete model of the beam response as a function of controllable parameters.