Differentiable Preisach Modeling for Particle Accelerator Systems with Hysteresis

Ryan Roussel

rroussel@slac.stanford.edu

U.S. DEPARTMENT OF ENERGY | Stanford University

SLAC NATIONAL ACCELERATOR LABORATORY

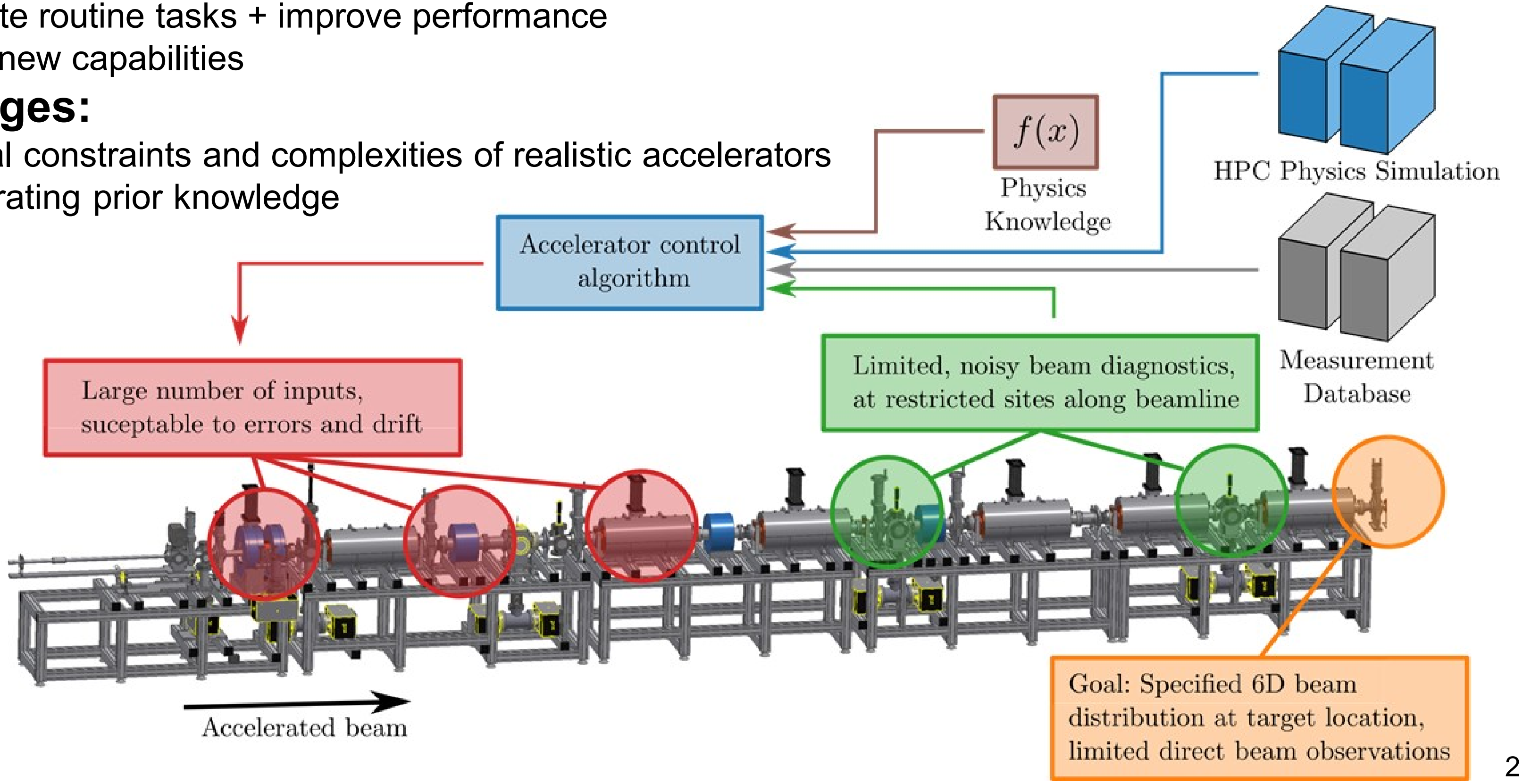
Machine Learning Based Accelerator Control

Goals:

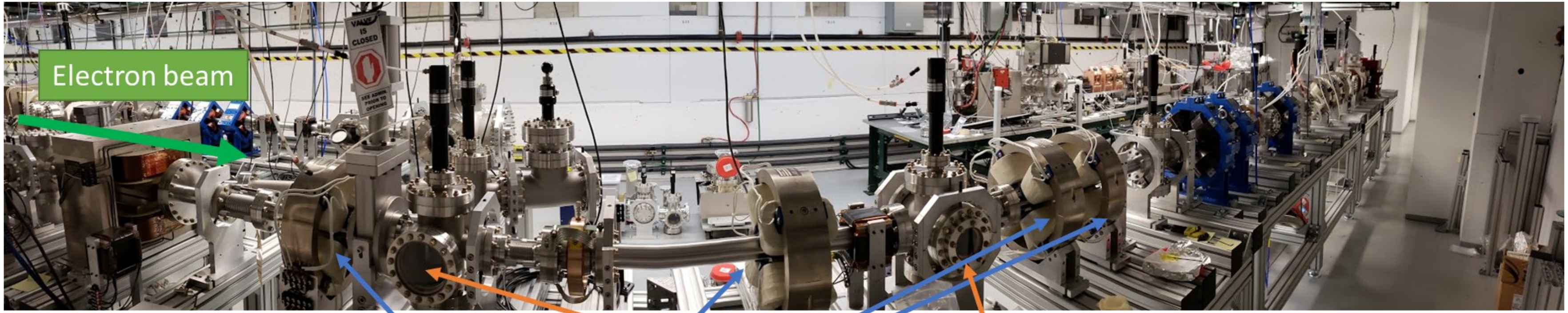
- Automate routine tasks + improve performance
- Enable new capabilities

Challenges:

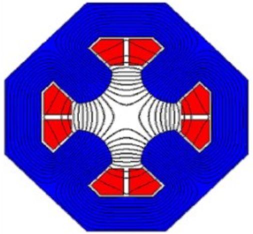
- Practical constraints and complexities of realistic accelerators
- Incorporating prior knowledge
- Scaling



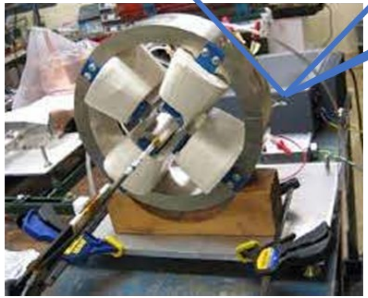
Controlling Beam Optics with Electromagnets



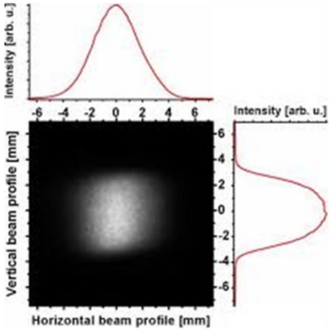
Quadrupole (focusing) magnets

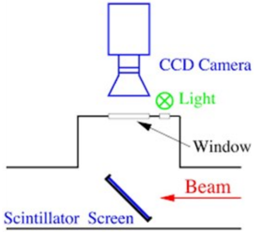


standard quadrupole



Insertable diagnostic screens

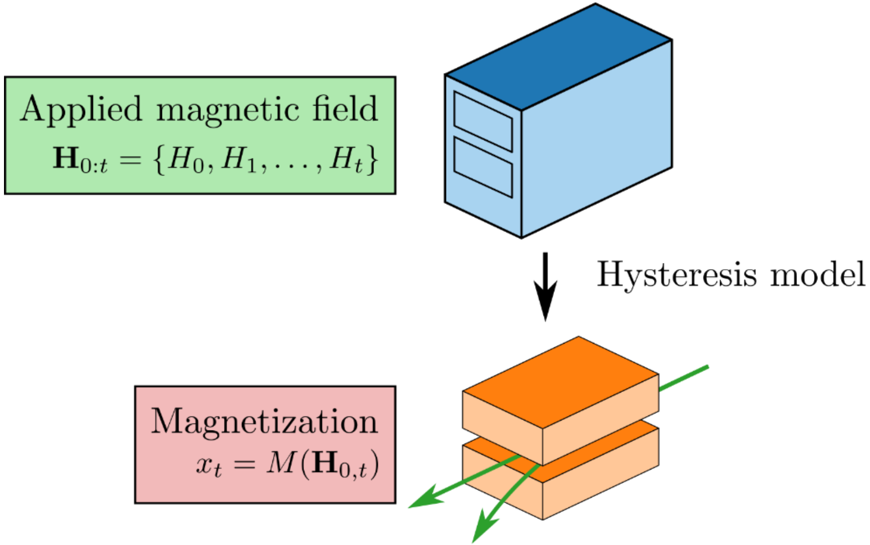




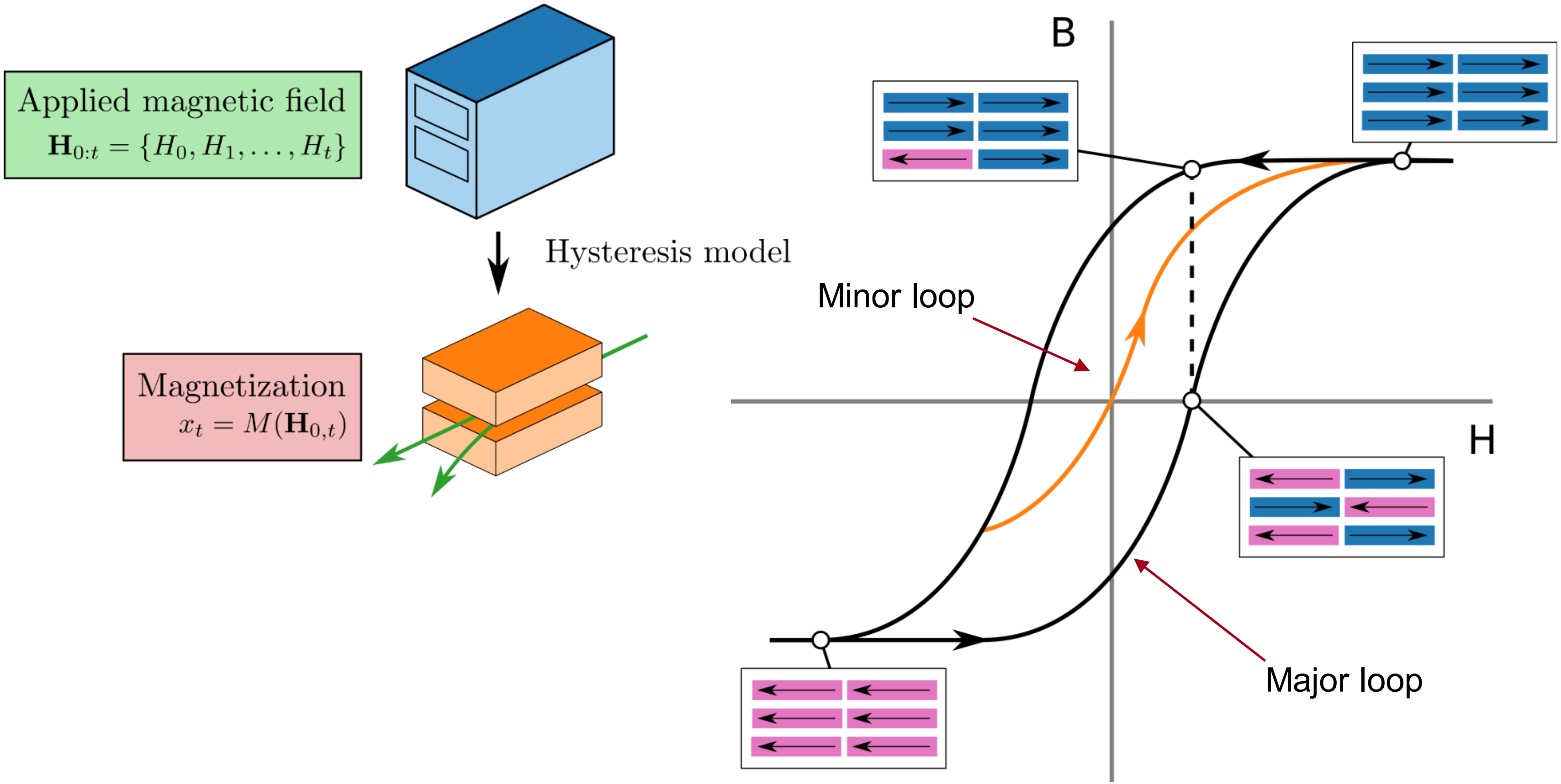
Modeling Hysteresis Effects in Accelerators

We generally only have control over applied currents to magnetic elements

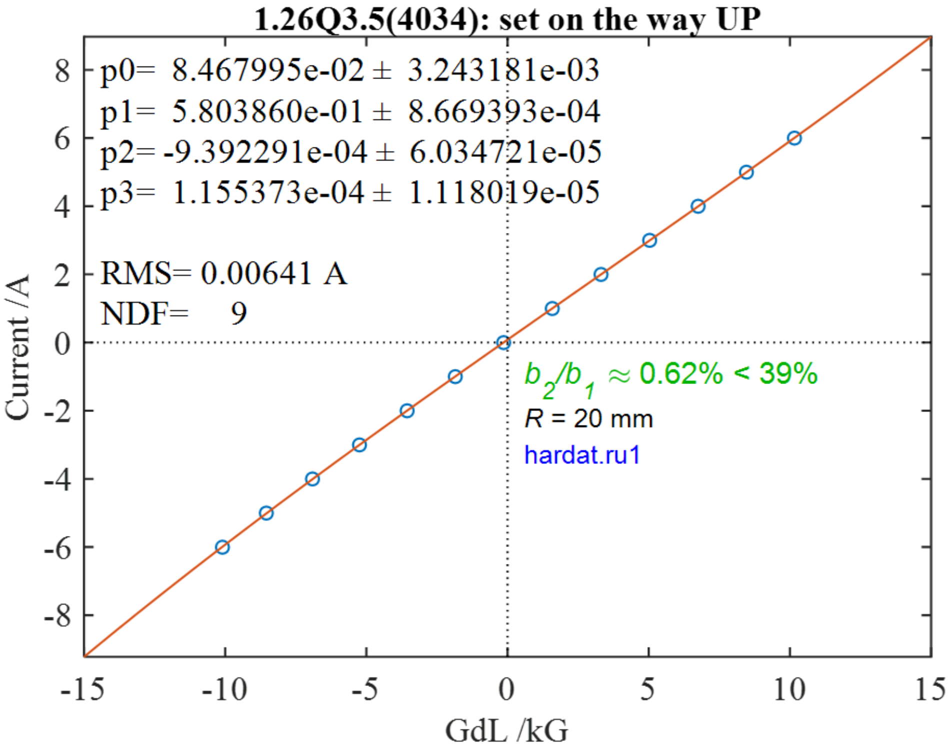
- Current optimization schemes ignore hysteresis effects
- Need to develop a fast-modeling strategy to improve online control without measuring the hysteresis curve directly



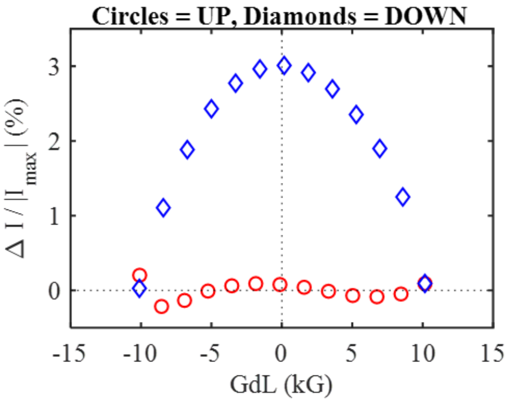
Hysteresis Effects in Electromagnets



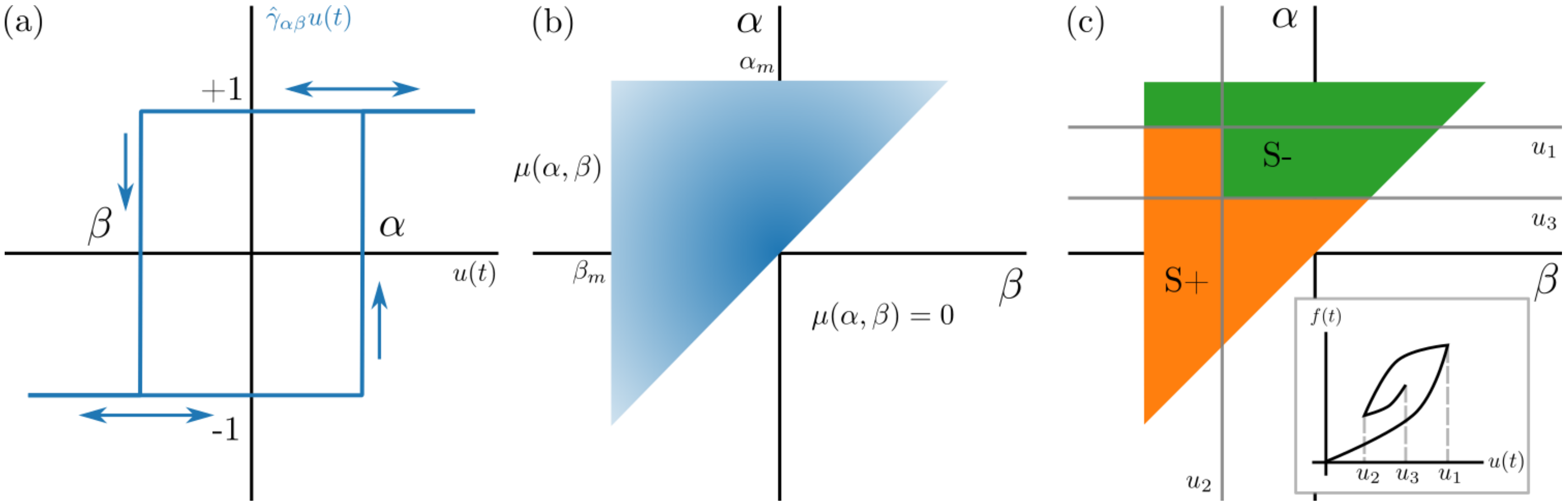
Realistic Accelerator Magnets



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



The Preisach Model for Hysteresis



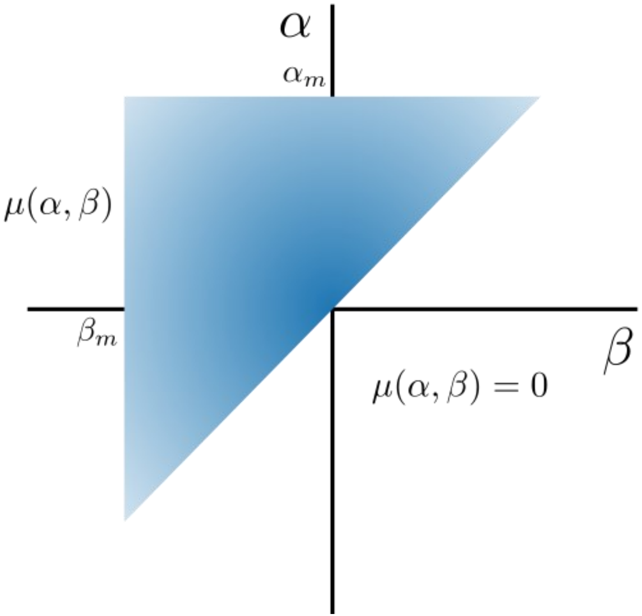
Model magnetic domains as individual "hysterons"

Determine the hysteron density

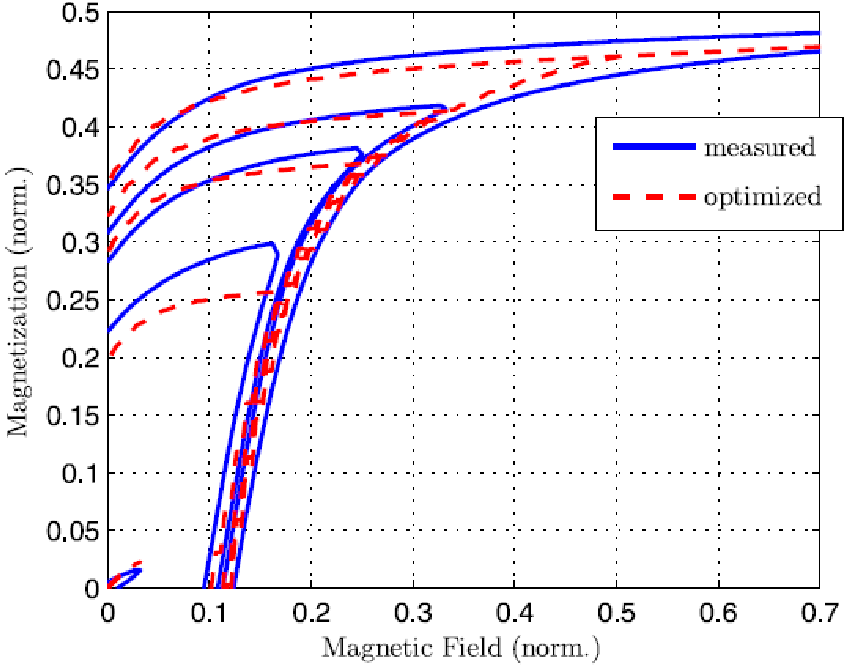
Determine hysteron states via history of applied fields

$$M(\mathbf{H}_{0:t}) = \int_{S+} \mu(\beta,\alpha)d\beta d\alpha - \int_{S-} \mu(\beta,\alpha)d\beta d\alpha$$

Parametric Preisach Modeling



$$\mu(\alpha, \beta) = f(\alpha, \beta; \theta)$$

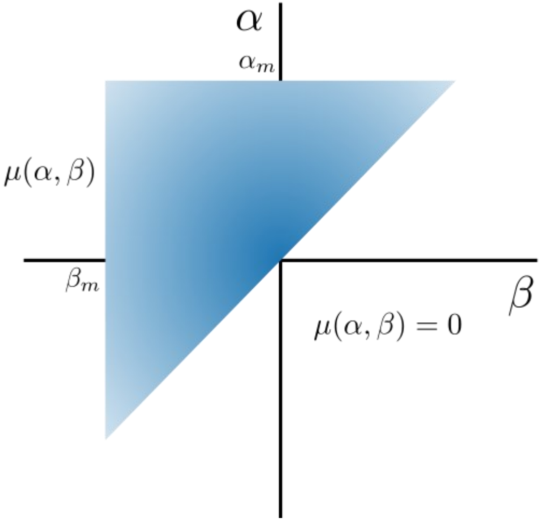


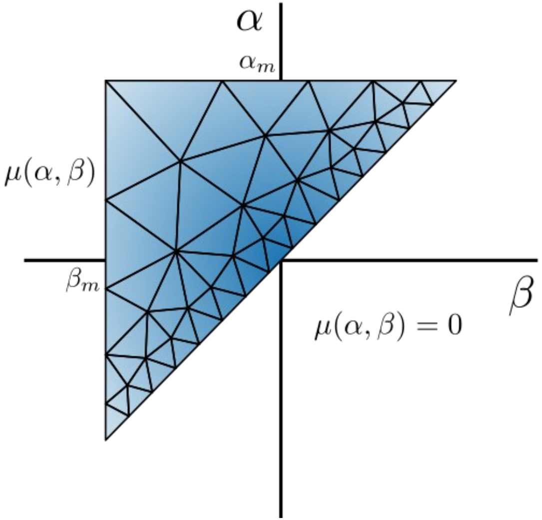
$$\mu_{\mathrm{DAT}}(\alpha, \beta) = \frac{A}{1 + \{[(\alpha + \beta)\sigma]^2 + [(\alpha - \beta - h)\sigma]^2\}^{\eta}}$$

Sutor et al. (2010)

Non-parametric Preisach Modeling

Limited by model selection





$\mu(\alpha,\beta) = f(\alpha,\beta;\theta)$ → $\mu_i = \mu(\alpha_i,\beta_i)$

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.

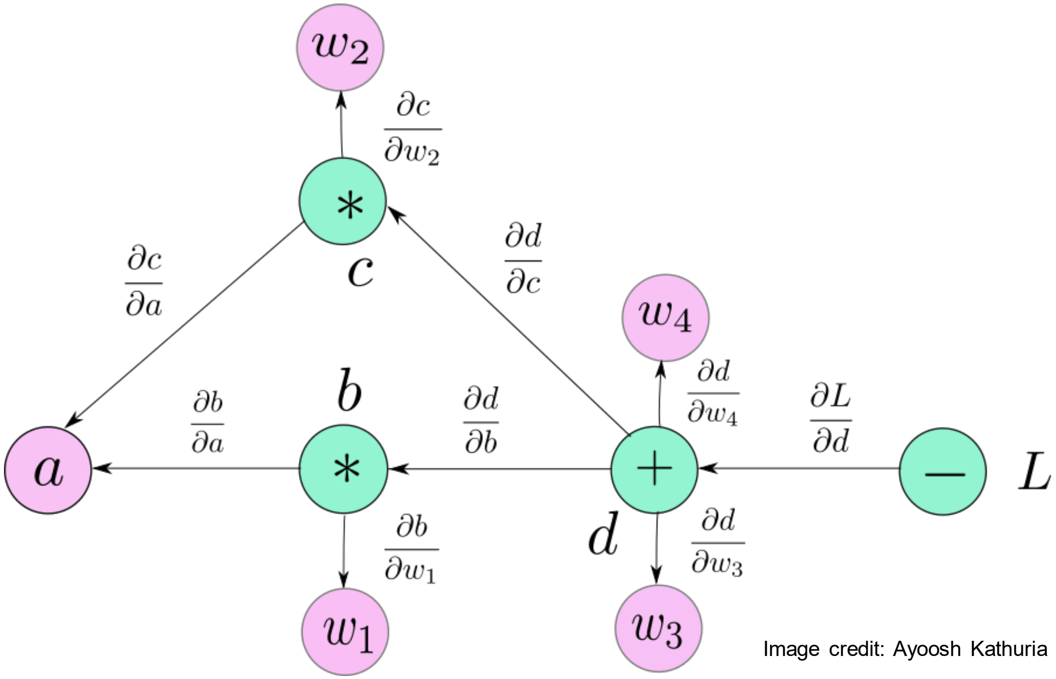


Image credit: Ayoosh Kathuria

Simple Example

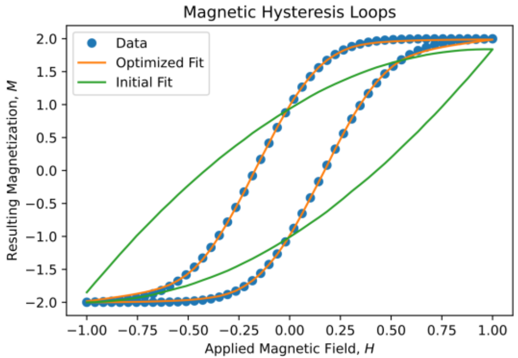
Discretize the space into n x n grid and treat the density at each grid point as a free parameter

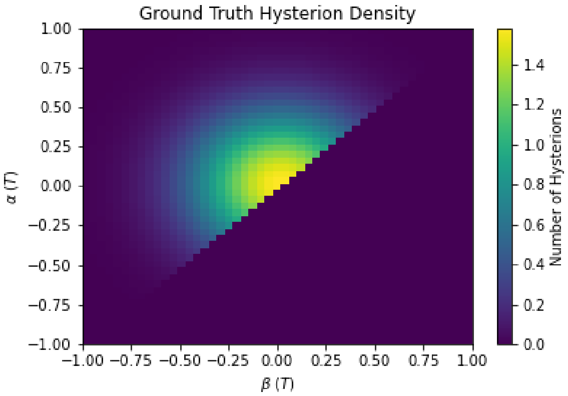
$$\boldsymbol{\mu} = \{\mu_1, \mu_2, \dots, \mu_N\}$$

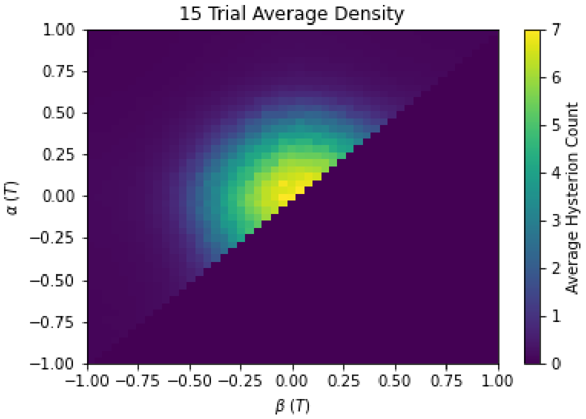
$$M(H_{0:t}; \boldsymbol{\mu}) = \sum_{i=1}^{N} \mu_i \gamma_i(H_{0:t})$$

$$loss(\boldsymbol{\mu}) = \mathrm{MSE}(\boldsymbol{Y}, \boldsymbol{M}(\boldsymbol{H}; \boldsymbol{\mu}))$$

$$\boldsymbol{\mu}^* = \mathrm{argmin}_{\boldsymbol{\mu}}\, loss(\boldsymbol{\mu})$$





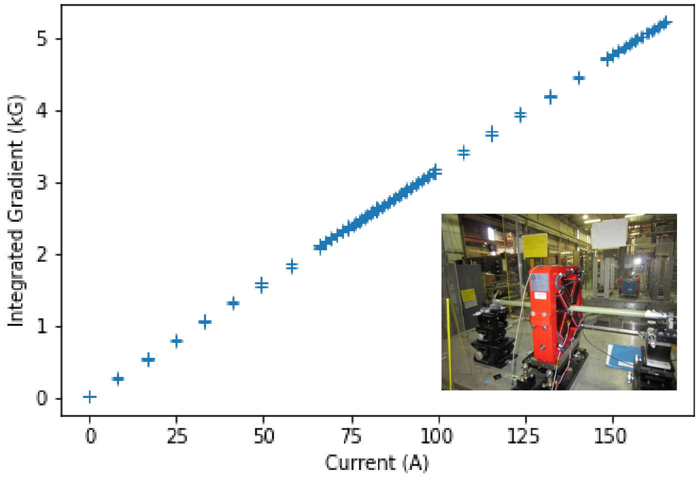


PyTorch

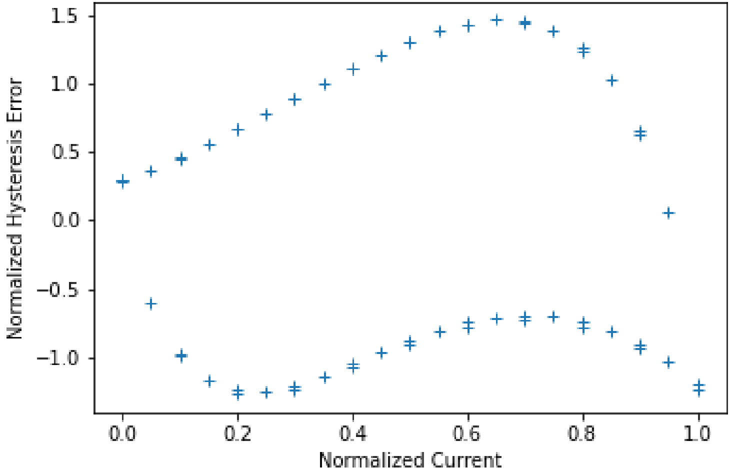
Modeling SLAC Quadrupole Magnets

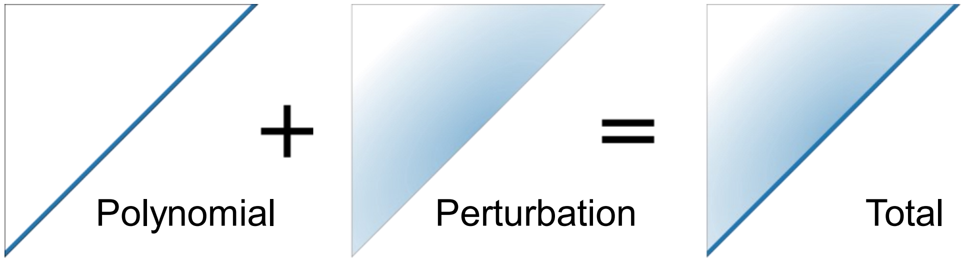
SLAC

Measured integrated gradient

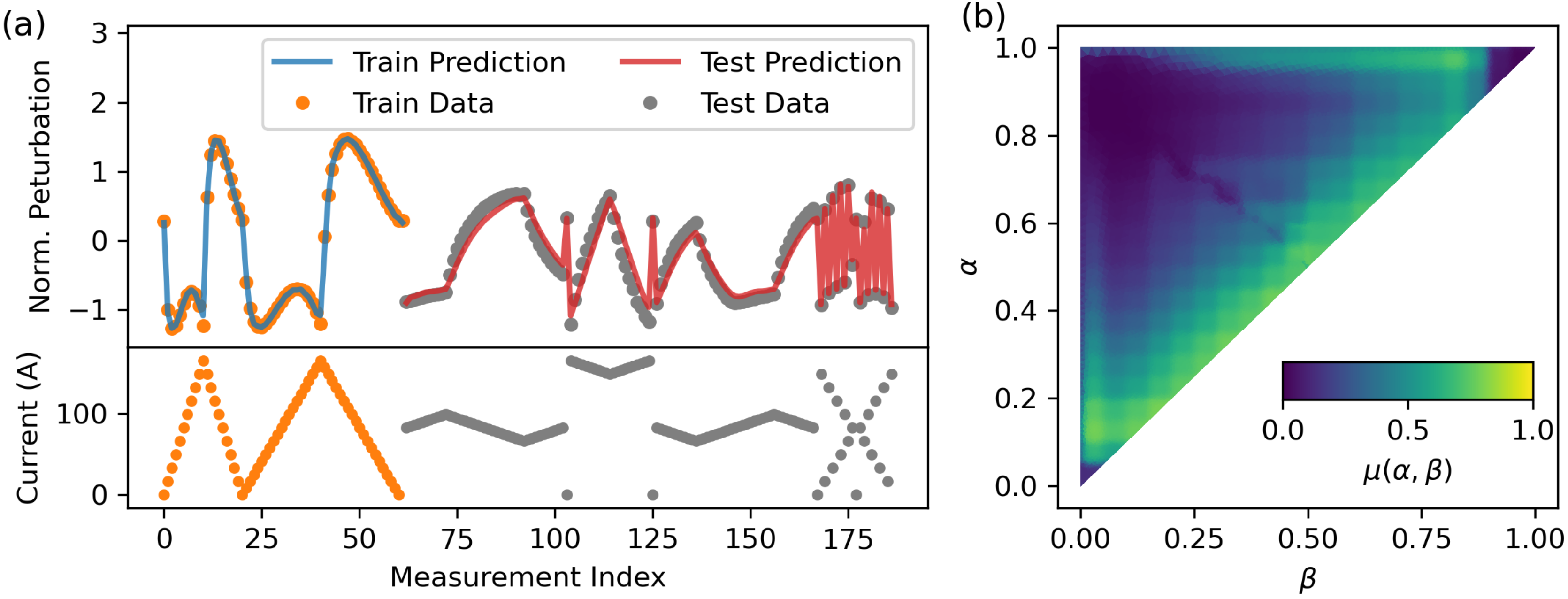


Normalized Deviation





Modeling SLAC Quadrupoles



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

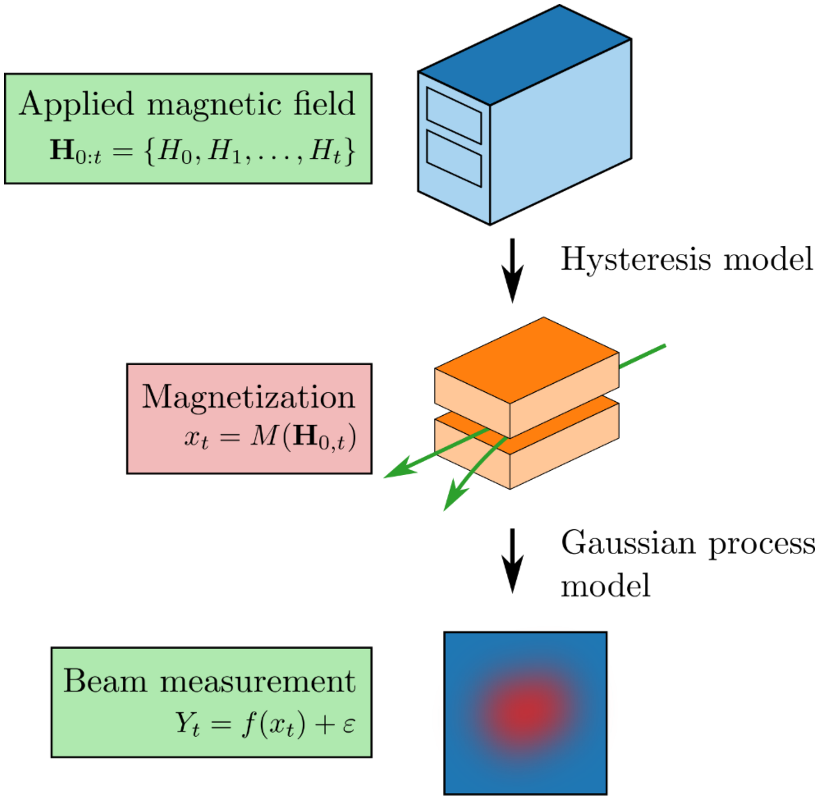
Roussel R. et al., Accepted by PRL

Full Stack Modeling

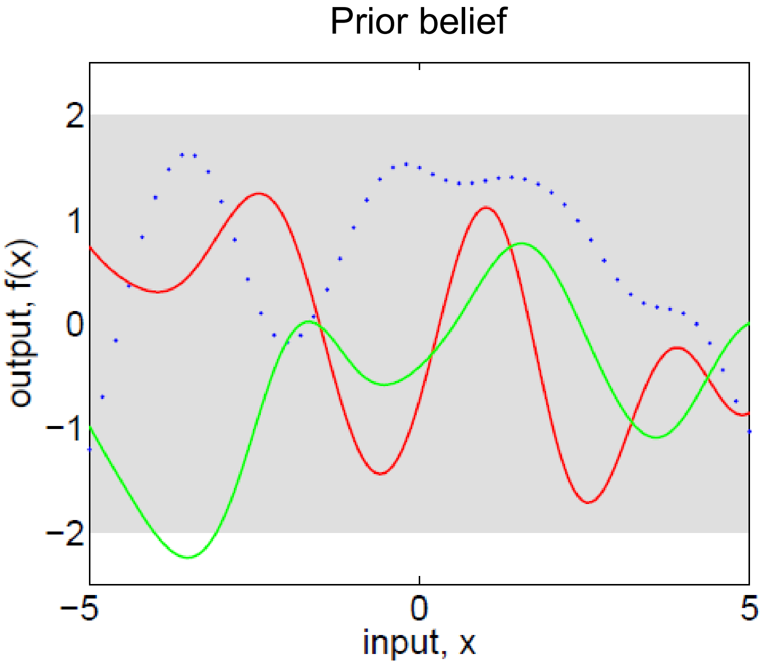
Can we extend this process to modeling magnets **in the beamline already?**

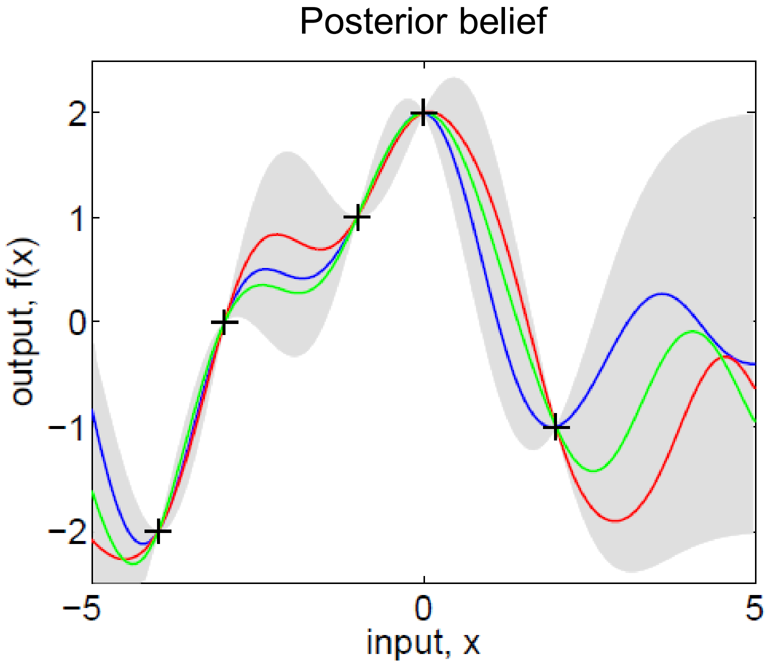
Combine the hysteresis model with a flexible model for the beam response: a **Gaussian Process,** train both models **simultaneously**

Allows us to measure hysteresis characteristics using **beam-based** measurements



Gaussian Process Surrogate Models

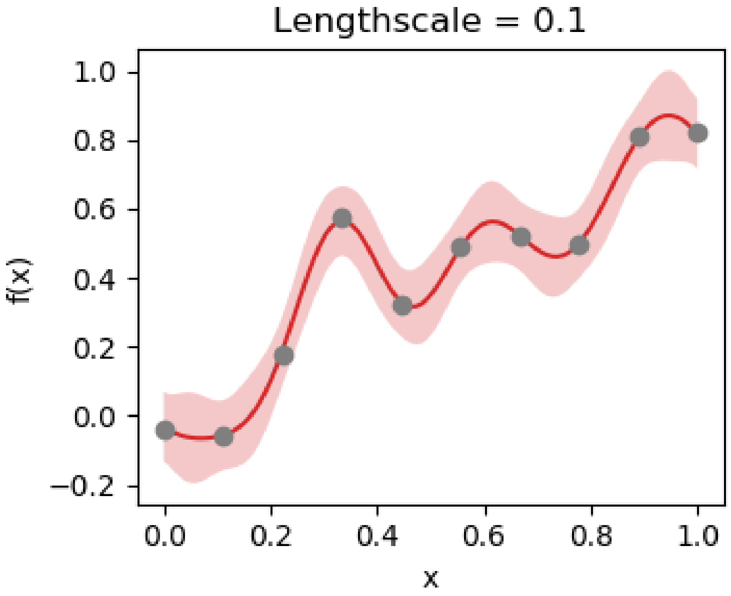


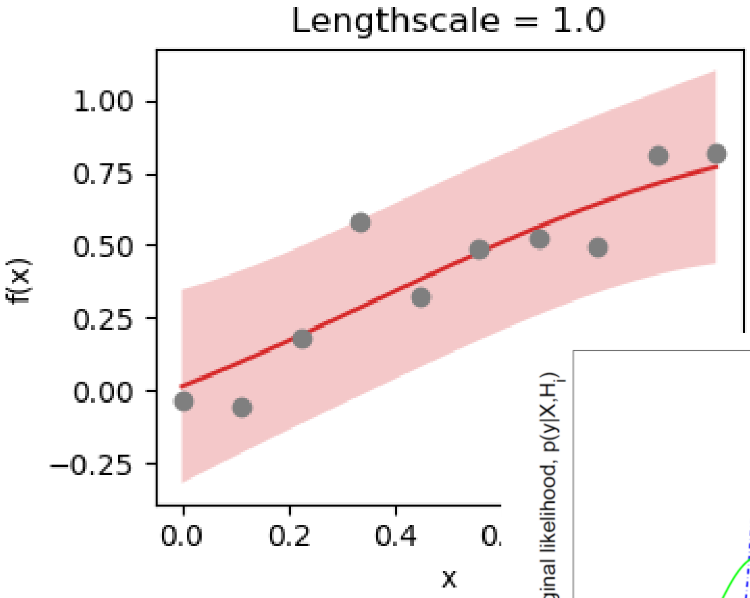


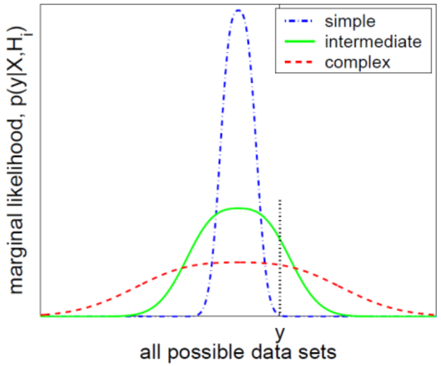
Use Gaussian Processes to represent noisy beam attributes as a function of magnetic fields.

Rasmussen, C.E.; Williams, C.K.I (2006).

Gaussian Processes - Hyperparameters



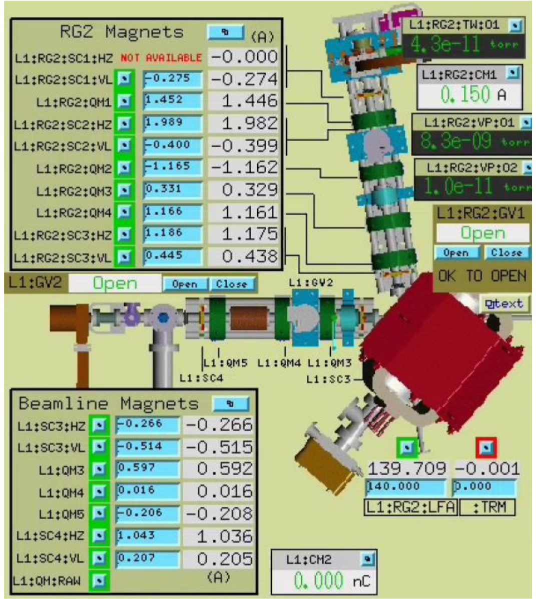


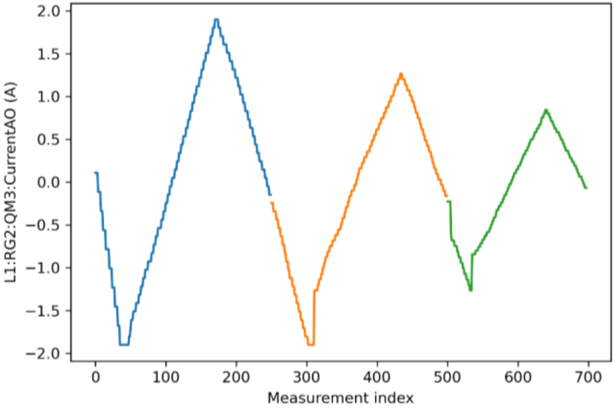


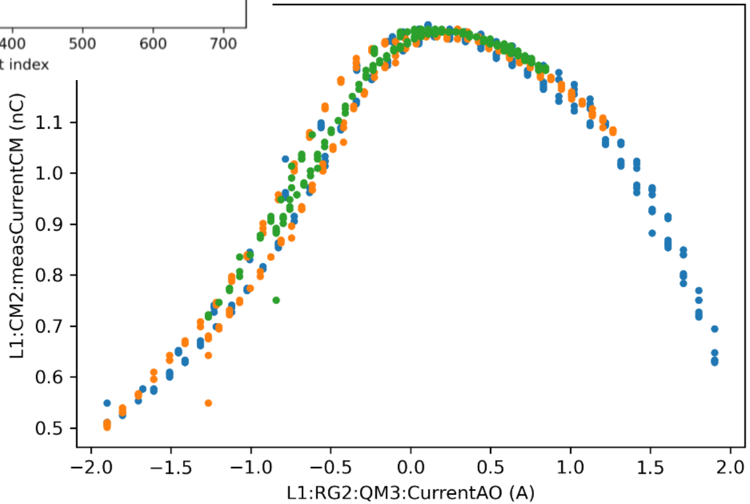
Rasmussen, C.E.; Williams, C.K.I (2006).

Example: APS Injector

Argonne NATIONAL LABORATORY

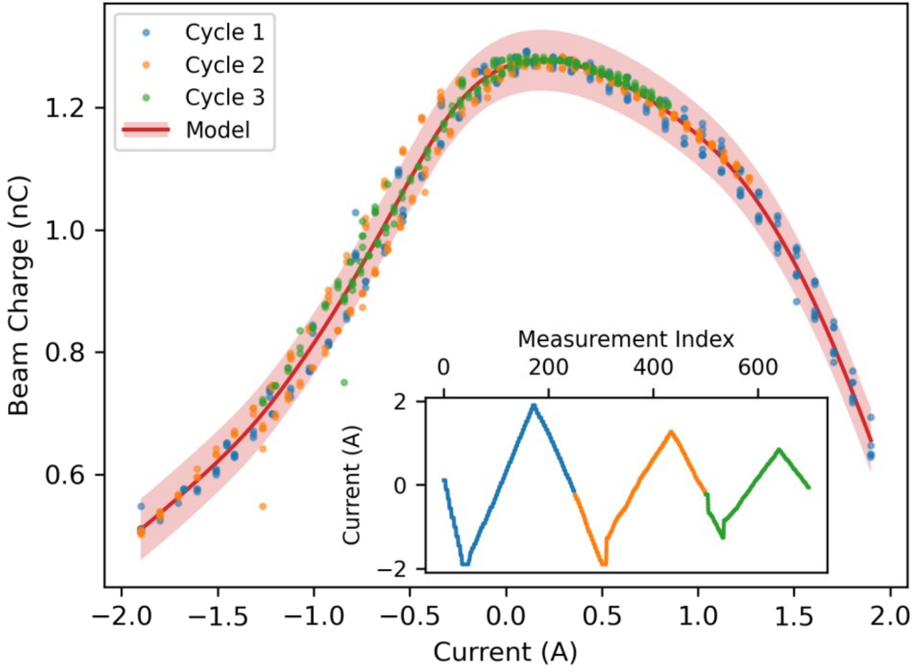




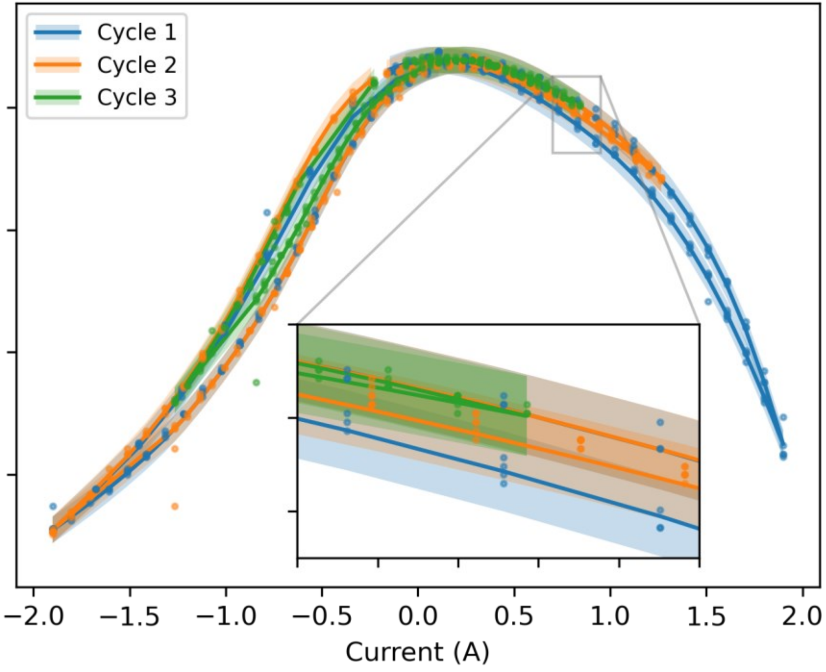


Full Stack Modeling

Modeling w/o Hysteresis



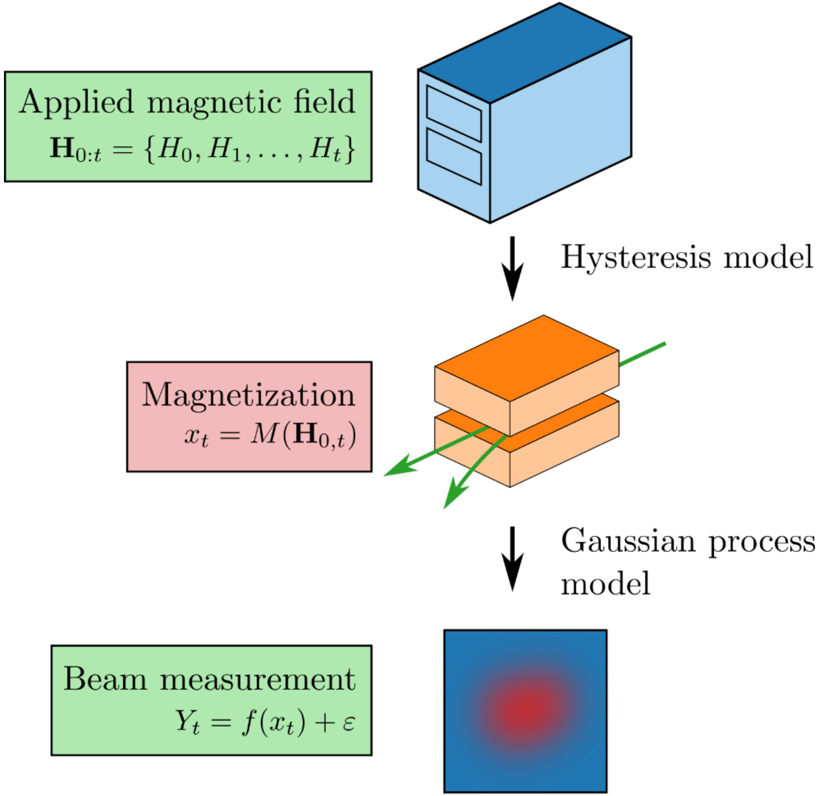
Modeling w/ Hysteresis



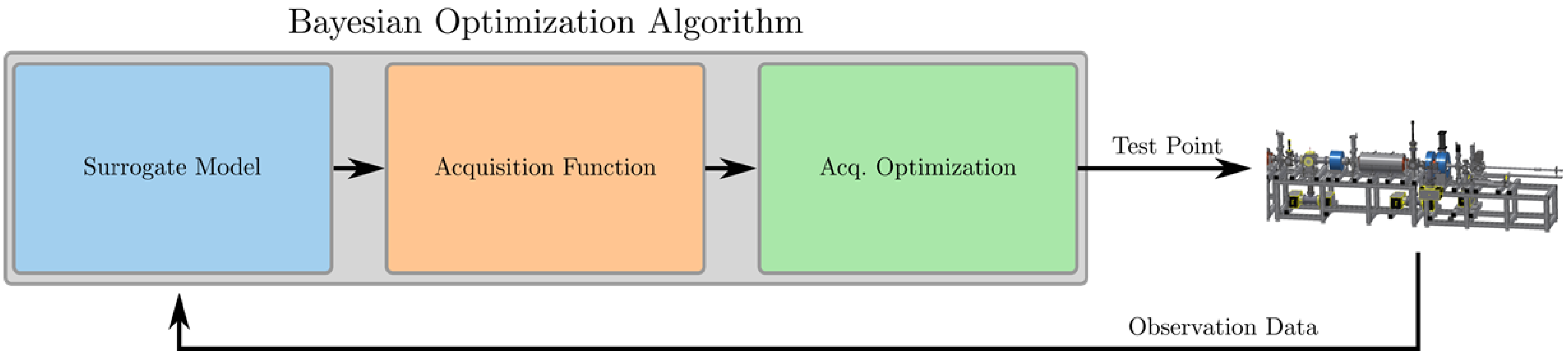
Roussel R. et al., Accepted by PRL

Full Stack Optimization

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

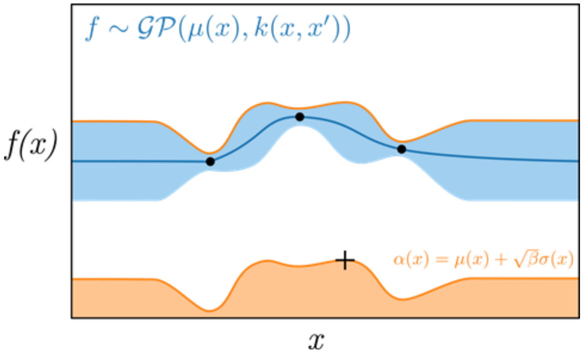


Machine Learning Based Accelerator Control

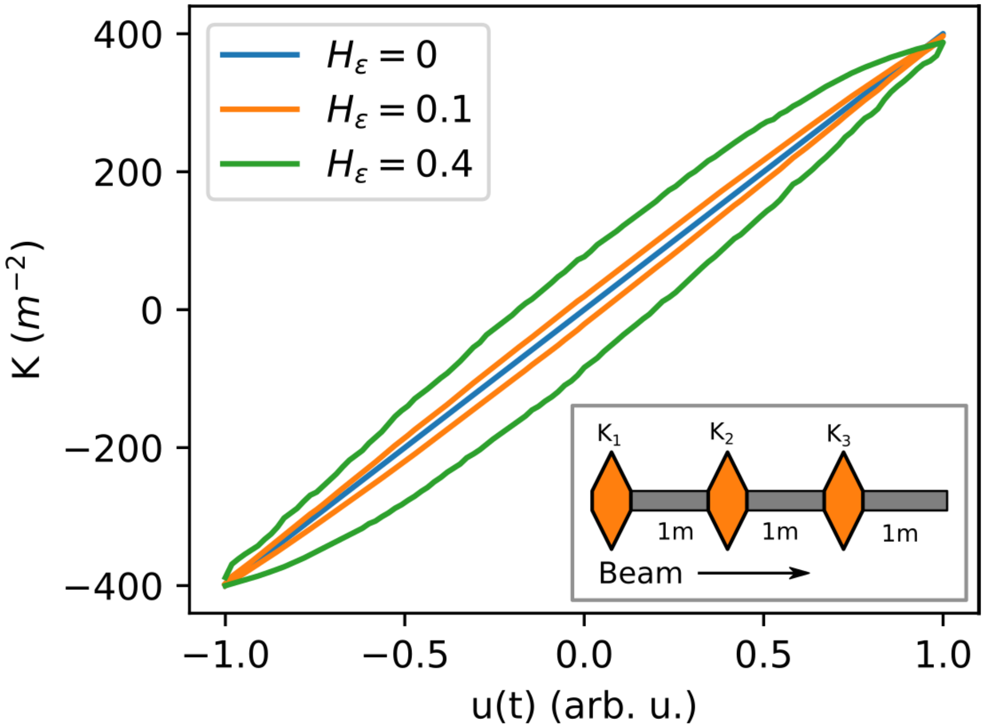


Benefits:

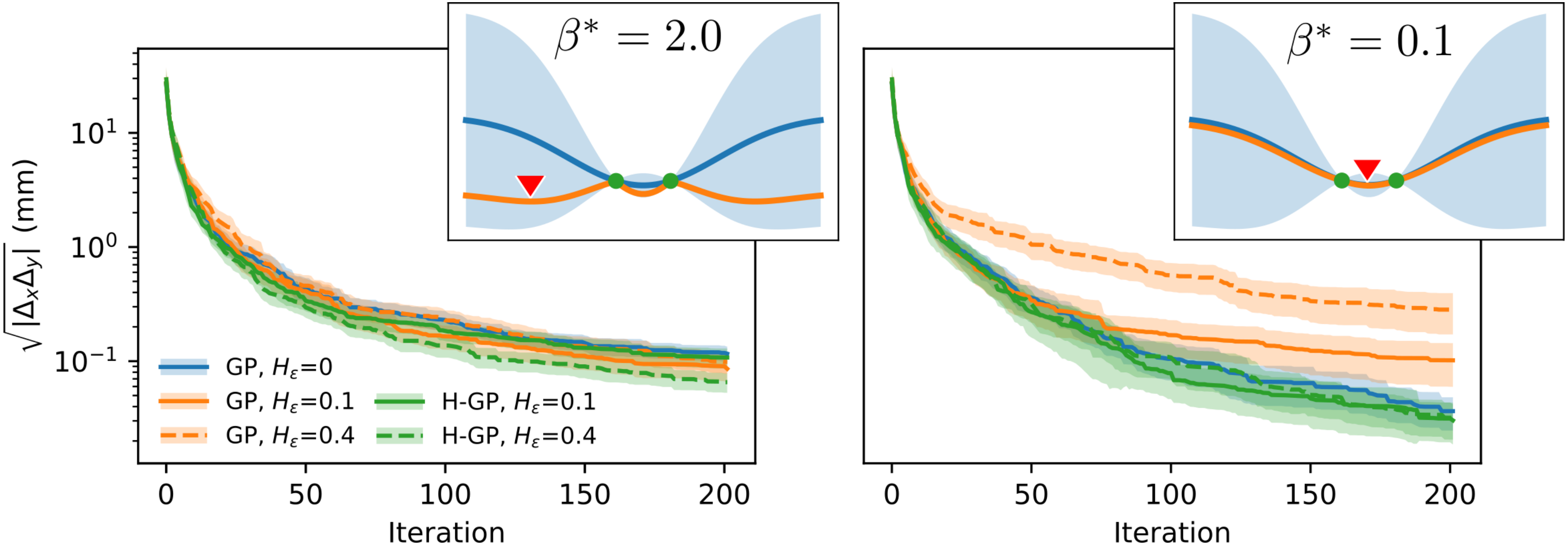
- Specify trade-off between exploration and exploitation
- Inherently improves model accuracy in regions of interest
- Enables serial or parallelized optimization strategies



Beamline Optimization with Hysteresis



Beamline Optimization with Hysteresis



Roussel R. et al., Accepted by PRL

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**.