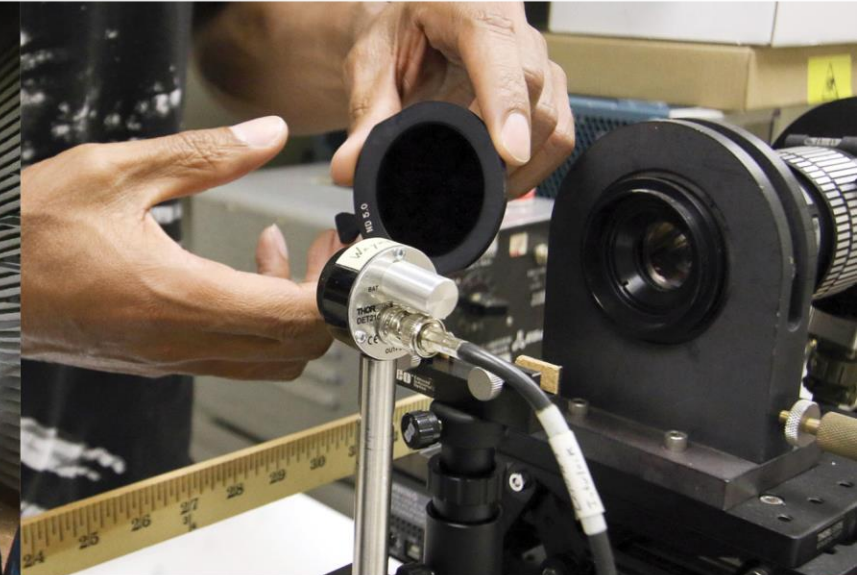


**Presented by Paul Stanik III**

5<sup>th</sup> ICFA Beam Dynamics Mini-Workshop on Machine Learning for Particle Accelerators

This work was done by Mission Support and Test Services, LLC, under Contract No. DE-NA0003624 with the U.S. Department of Energy and the National Nuclear Security Administration's Office of Defense Programs, and supported by the Site-Directed Research and Development Program. DOE/NV/03624--2138.



# Improving Fast Beam Transport Simulations Using Transfer Learning

# Authors

- ▶ Paul Stanik III <sup>1, 2</sup> – stanikp@nv.doe.gov
- ▶ Brendan T. Morris <sup>1</sup>
- ▶ Evan R. Scott <sup>2</sup>
- ▶ Trevor J. Burris-Mog <sup>2</sup>
- ▶ Piotr Wiewior <sup>2</sup>

1. University of Nevada-Las Vegas, Las Vegas, NV 89154
2. Nevada National Security Sites, North Las Vegas, NV 89030

# Outline

- ▶ Introduction
- ▶ Objective
- ▶ Data
  - Low Fidelity Simulations
  - High Fidelity Simulations
- ▶ Machine Learning
  - Results
  - Pruning
- ▶ Future Works

- ▶ Fast simulations of intense relativistic electron beams are sufficient for tuning an accelerator's magnetic transport field.
  - Can't capture all relevant beam physics.
  - Not sufficient for optimization problems due to model limitations.
  
- ▶ Simulations that capture all relevant beam physics are significantly more computationally expensive.
  - Less useful for optimization problems.
  
- ▶ Transfer learning for optimization problems using both types of simulations.

- ▶ **Replace a high-fidelity simulation with a machine learning model using limited data**
- ▶ Low-fidelity simulations → large amount
- ▶ High-fidelity simulations → very little
  - Experimental “surrogate” data
- ▶ Transfer learning from low-fidelity to high-fidelity
- ▶ Simulated data to be used → **Beam envelope**
- ▶ Why beam envelope?
  - Importance for linear induction accelerators (e.g., Scorpius)
  - Precursor for magnetic transport lattice optimizations

# Data



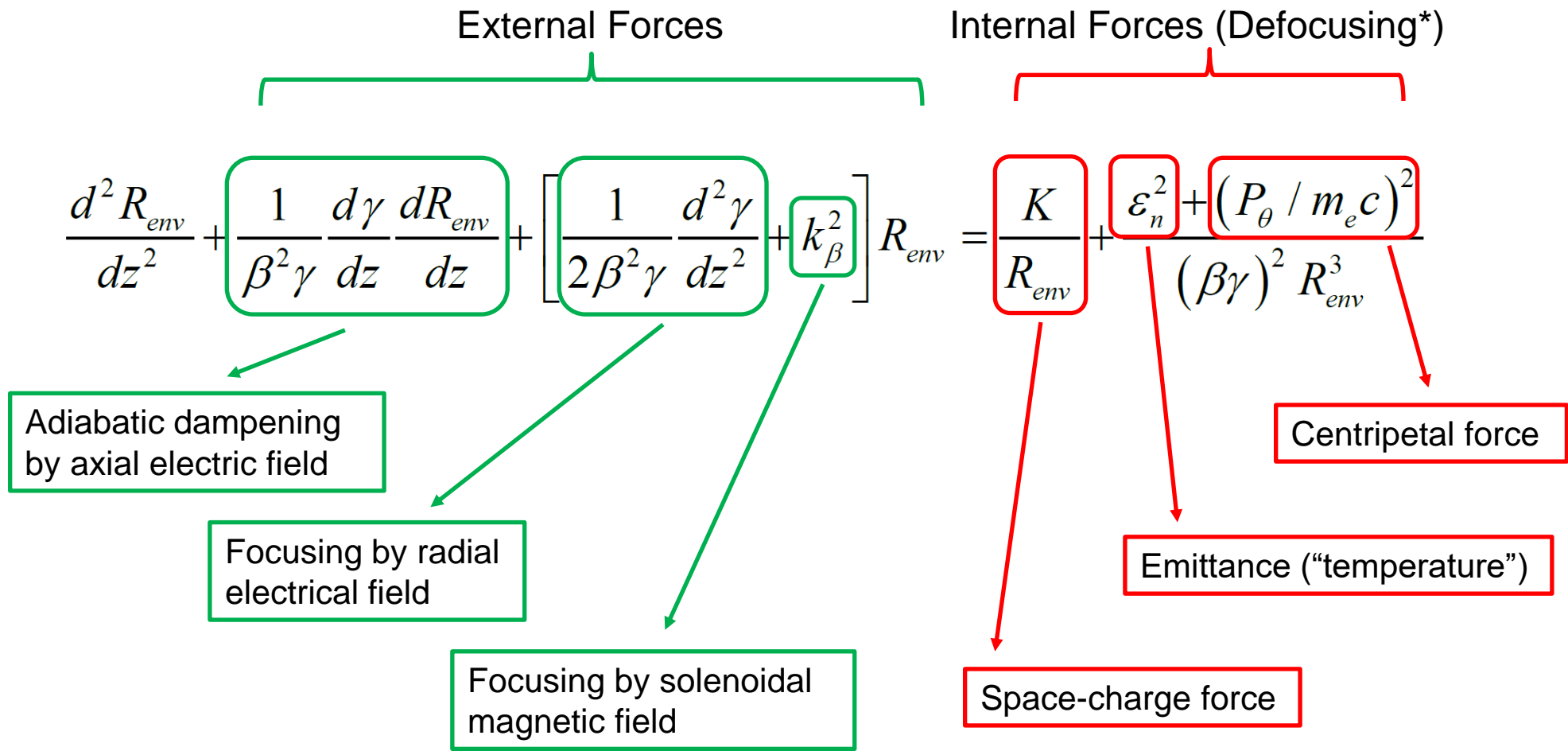
# Low-Fidelity Simulations

- ▶ Radius using beam envelope equation<sup>1</sup>
- ▶ **Runtime:** 0.0171 seconds

$$\frac{d^2 R_{env}}{dz^2} + \frac{1}{\beta^2 \gamma} \frac{d\gamma}{dz} \frac{dR_{env}}{dz} + \left[ \frac{1}{2\beta^2 \gamma} \frac{d^2 \gamma}{dz^2} + k_\beta^2 \right] R_{env} = \frac{K}{R_{env}} + \frac{\varepsilon_n^2 + (P_\theta / m_e c)^2}{(\beta\gamma)^2 R_{env}^3}$$

1. Ekdahl Jr., C. A., "The beam envelope equation and practical applications", LANL, 2019-08-2021

# Beam Envelope Equation



\*Defocusing in LIA vacuum

# High Fidelity Simulations

## ▶ Large Scale Plasma (LSP)

- Particle-in-cell code
- Calculates forces on particles per timestep
- Includes more relevant physics
- **Runtime:** ~30 minutes

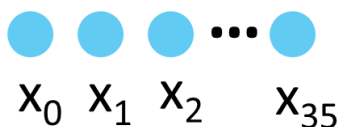
## ▶ “Surrogate” experimental data

## ▶ Dataset sizes (# of shots)

- Low-fidelity: 73,773
- High-fidelity: 971

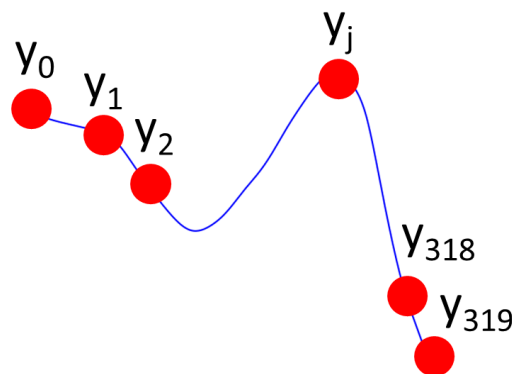
## ▶ Data shape:

- Input: 36 (randomized currents)



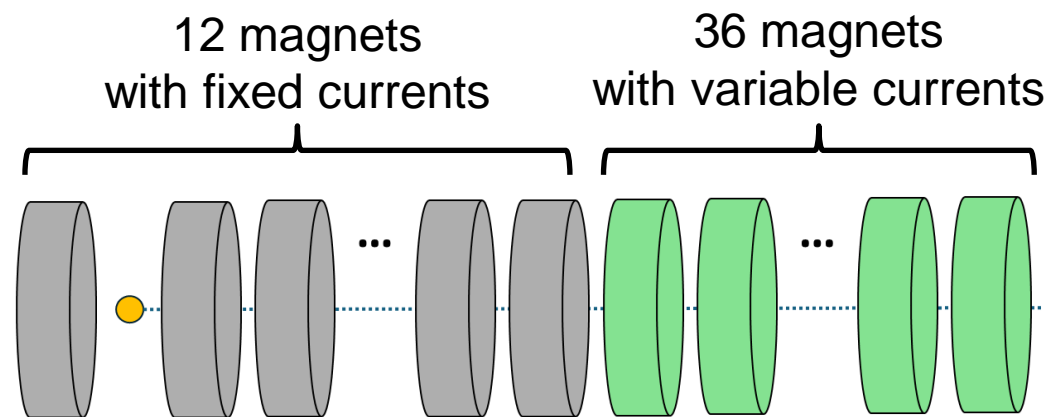
- Output: 320 (control points of beam envelope)

- Step length where  $dz \approx 2\text{cm}$



## ▶ Data splits

- Test = 30%
- Train + Validation = 70%
  - Train = 70%
  - Validation = 30%



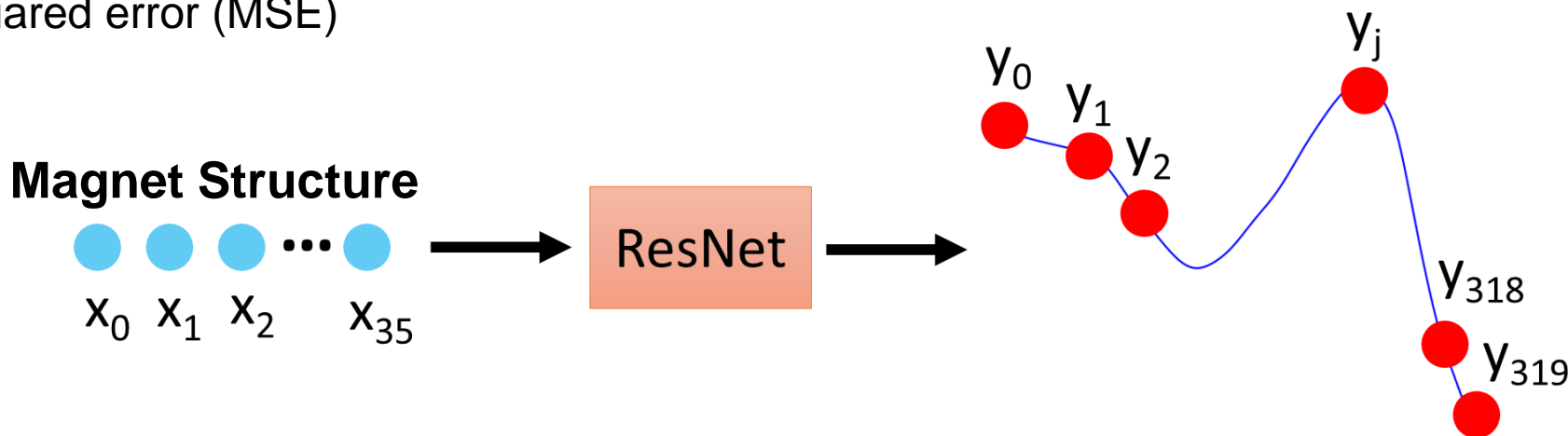
# Machine Learning



# Machine Learning

- ▶ PyTorch
- ▶ Model architecture
  - Residual network<sup>2</sup> using dense layers
  - Optimizer
    - AdamW with L2 regularization
- ▶ Evaluation metrics
  - Mean squared error (MSE)

- ▶ Input: magnet currents
- ▶ Output: beam envelope
- ▶ Train: Low-Fidelity
- ▶ Transfer learn: High Fidelity



# Quantitative Results

	Train on Lo-Fid	Train on Lo-Fid & Test on Hi-Fid	Train on Lo-Fid & Transfer on Hi-Fid	Hi-Fid MSE % Decrease
MSE (Test)	0.00263	3.14	0.00906	99.7%

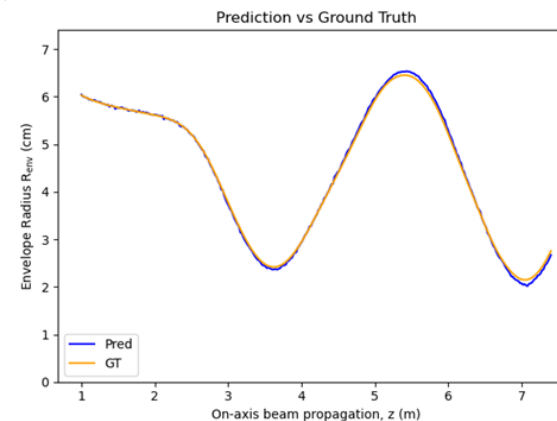
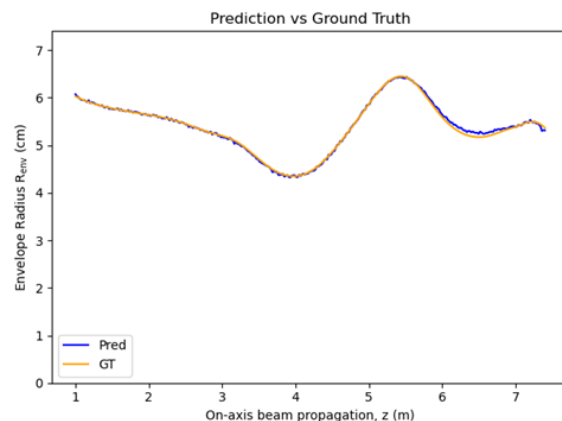
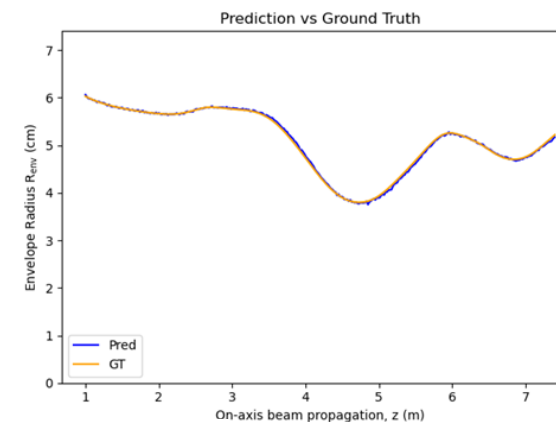
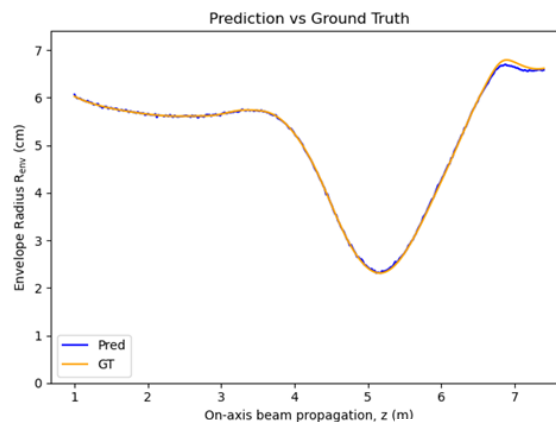
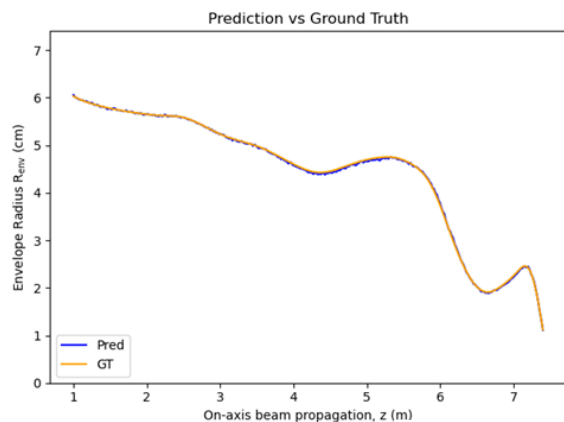
	Lo-Fid	Hi-Fid	ML Model	ML Speed Increase vs Hi-Fid
Speed	0.0171 s	~30 minutes	0.0793 s	99.9%

Using INTEL(R) XEON(R) GOLD 5520+ Processor

Lo-Fid = Low-Fidelity  
Hi-Fid = High-Fidelity  
ML = Machine Learning

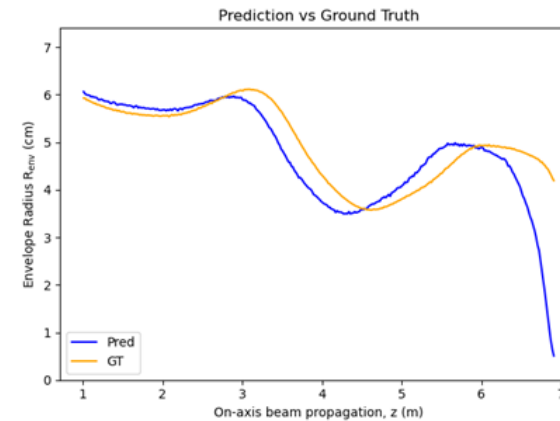
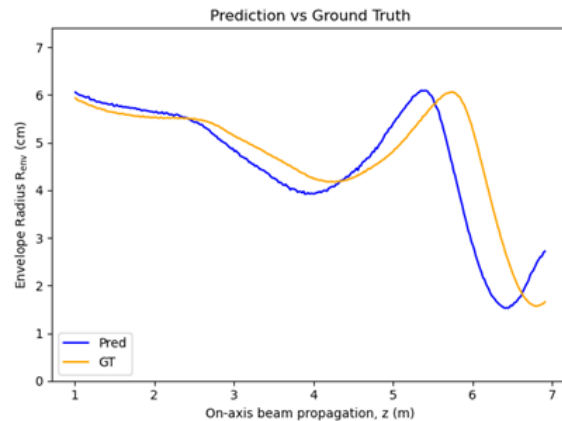
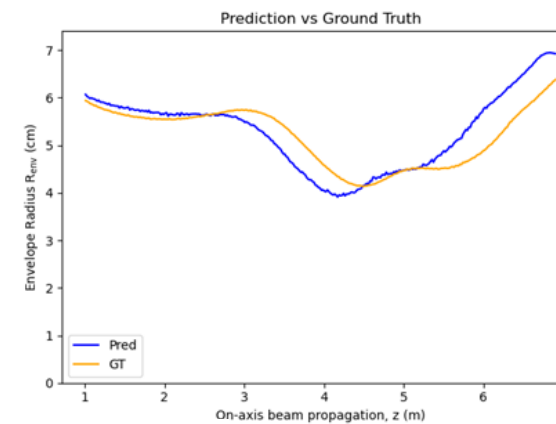
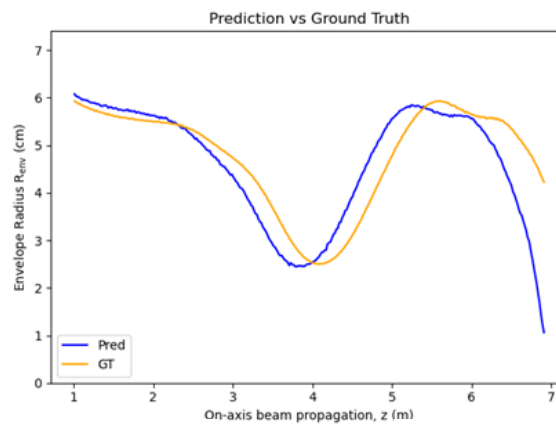
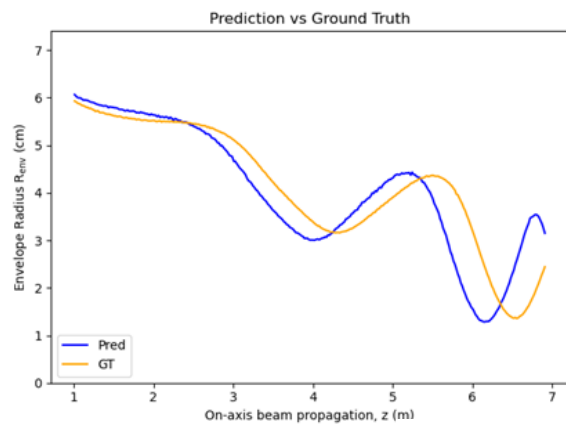
# Train on Lo-Fid

- ▶ **Test MSE: 0.00263**
- ▶ Strong agreement between model output and ground truth



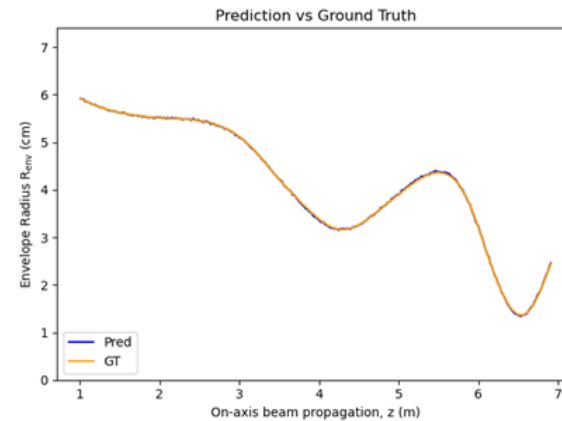
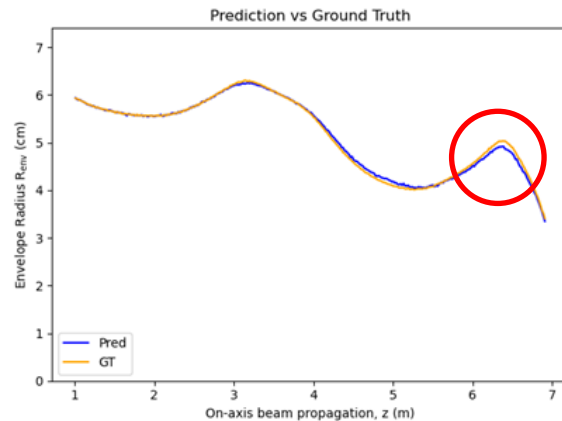
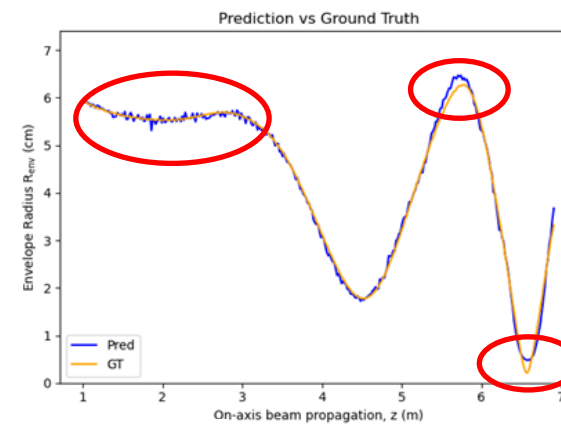
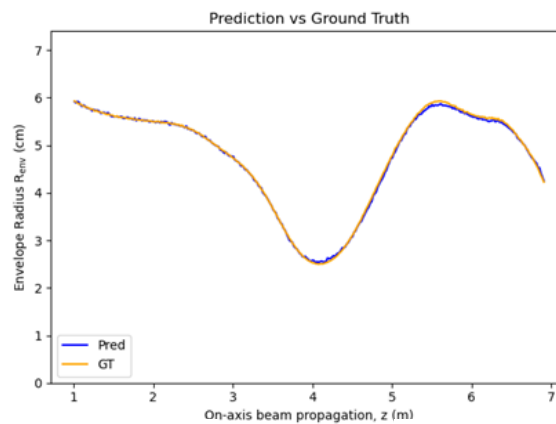
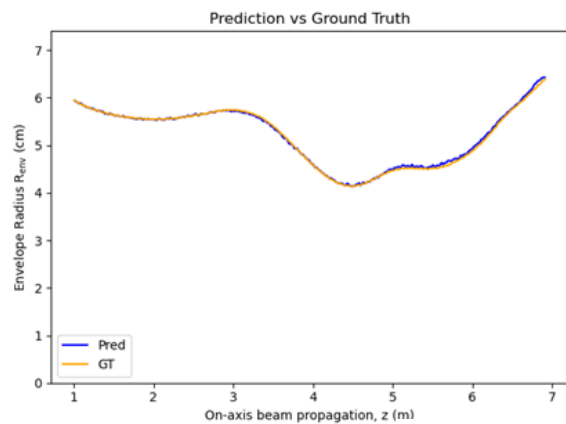
# Train on Lo-Fid & Test on Hi-Fid

- ▶ Forward pass of high-fidelity data on model pretrained on low-fidelity data
- ▶ **Test MSE: 3.14**
- ▶ Shape generally same but shifted



# Train on Lo-fid & Transfer on Hi-Fid

- ▶ **Test MSE: 0.00906**
- ▶ Acceptable results despite noise and gaps



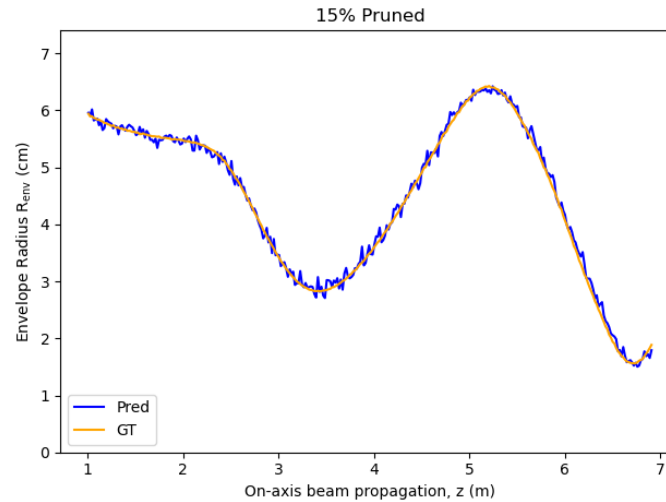
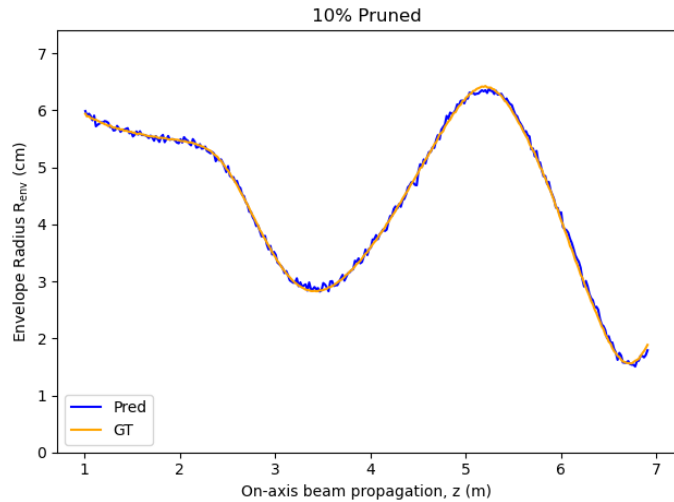
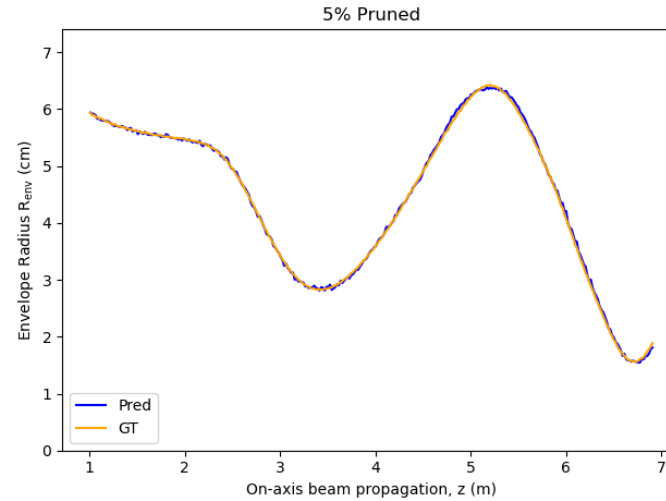
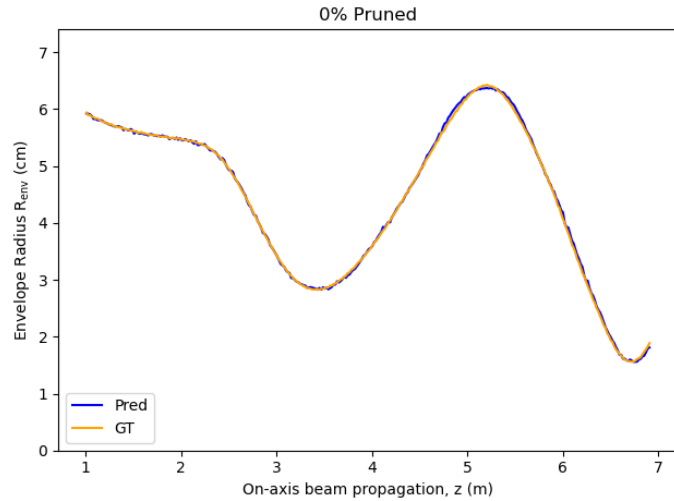
# Pruning

- ▶ Over 3.4 billion parameters
- ▶ Drop unimportant parameters to see how it affects model's performance
- ▶ L1 Regularization
- ▶ Pruning: 5%, 10%, 15%

	0% Pruned	5% Pruned	10% Pruned	15% Pruned
MSE	0.00263	0.00313	0.00729	0.0172
Weights dropped	0	174M	348M	523M

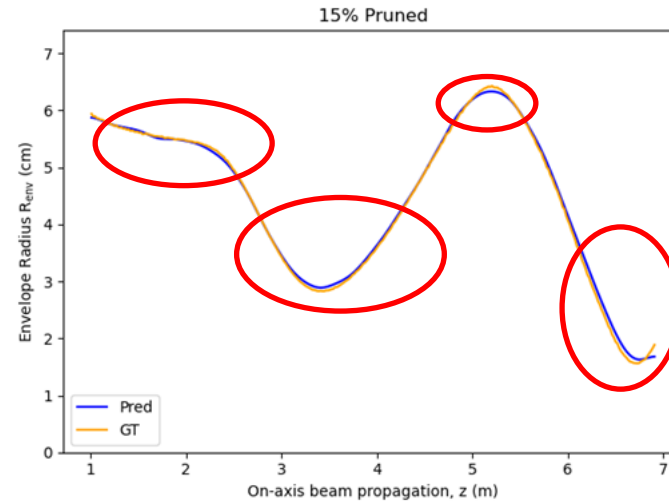
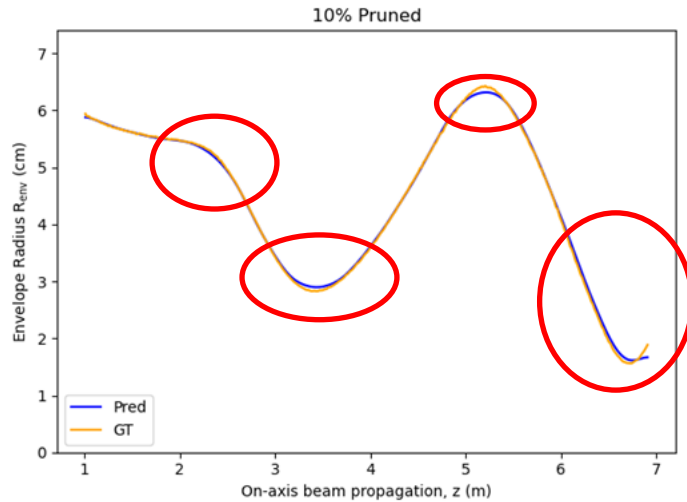
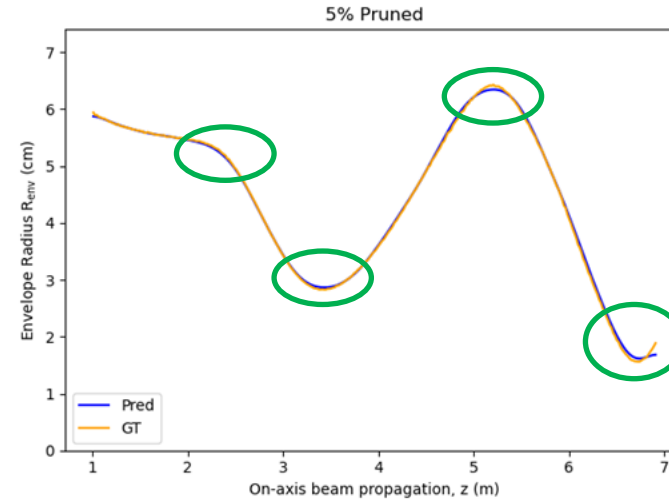
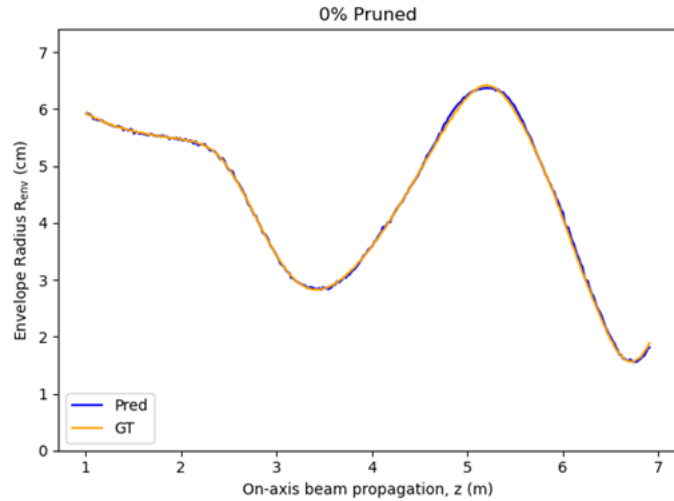
# Pruning

- Model output becomes noisier with as more weights are pruned

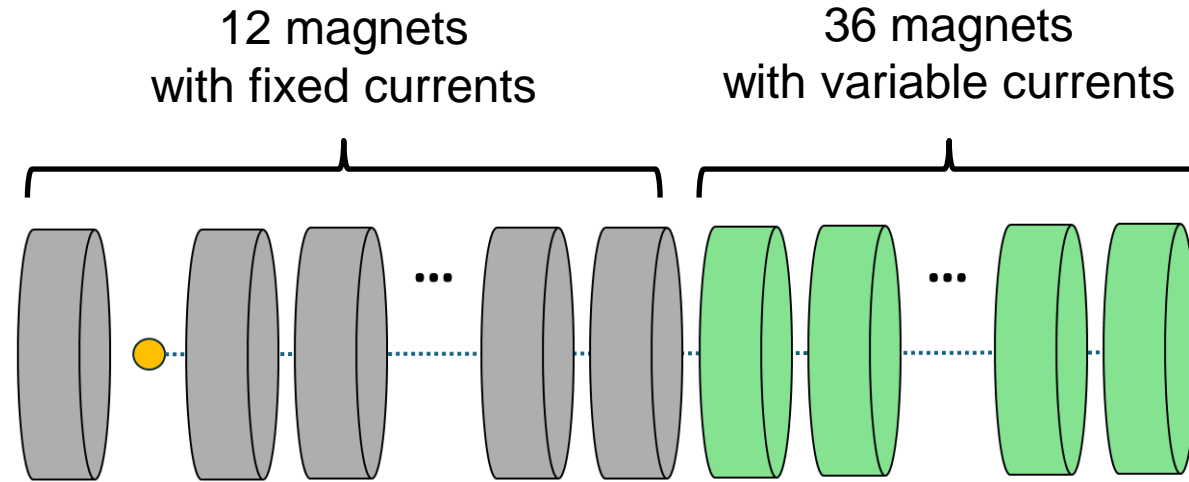


# Pruning

## ► Gaussian filter, $\sigma = 5$



- ▶ Extend for variable parameters for 12 fixed magnet currents



- ▶ Experimental data from LIAs
- ▶ Replace high-fidelity LSP simulation for magnetic lattice optimization
- ▶ First step for other optimization problems

# References

1. Ekdahl Jr., C. A., “The beam envelope equation and practical applications”, Los Alamos National Laboratory, 2019-08-2021
2. He, K. et al., “Deep Residual Learning for Image Recognition”, Computer Vision Pattern Recognition, 2016.

**Thank you!**

**Any questions?**

