Enhancing Neutron Scattering Experimentation

A Data Science and Machine Learning Approach to Predict Background Scattering

Petroula Karacosta

X-ray and Neutron Scattering Group, NBI

HAMLET
How to Apply Machine Learning to Experimental & Theoretical
PHYSICS
August 19 - 21, 2024
Copenhagen, Denmark

EUROPEAN SPALLATION SOURCE

McStas

UNIVERSITY OF COPENHAGEN

Master's Thesis:
Using McStas Union components to simulate a magnet sample environment & predicting background with machine learning

Kim Lefmann
Professor
NBI

Mads Bertelsen
Computational Neutron Scattering Scientist
ESS, DMSC

Alex Holmes
Scientific Engineer
ESS

What’s Neutron Scattering?

The ID of a Neutron

Spin	Charge	Magnetic Moment	Particle	Wave
1/2	0	-1.9130 $e\hbar/2m_p$	Yes	Yes

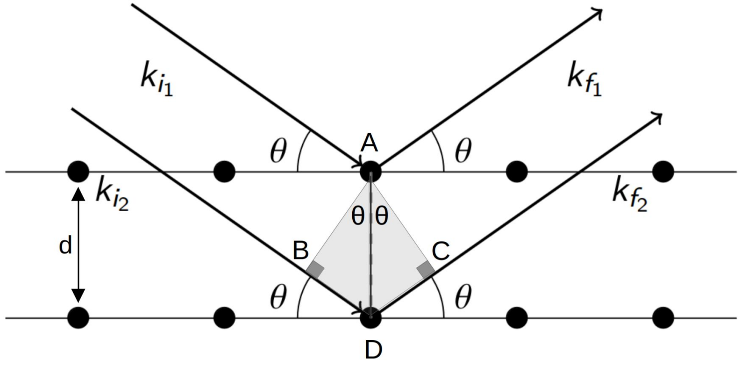
Interaction with matter

Absorption

Scattering

- Coherent & Incoherent
- Elastic & Inelastic

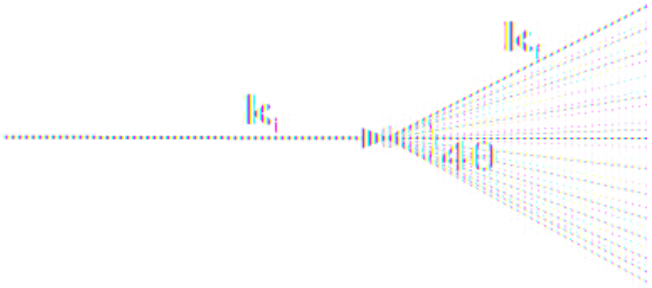
In Crystals



Bragg's Law

$$n\lambda = 2d\sin(\theta)$$

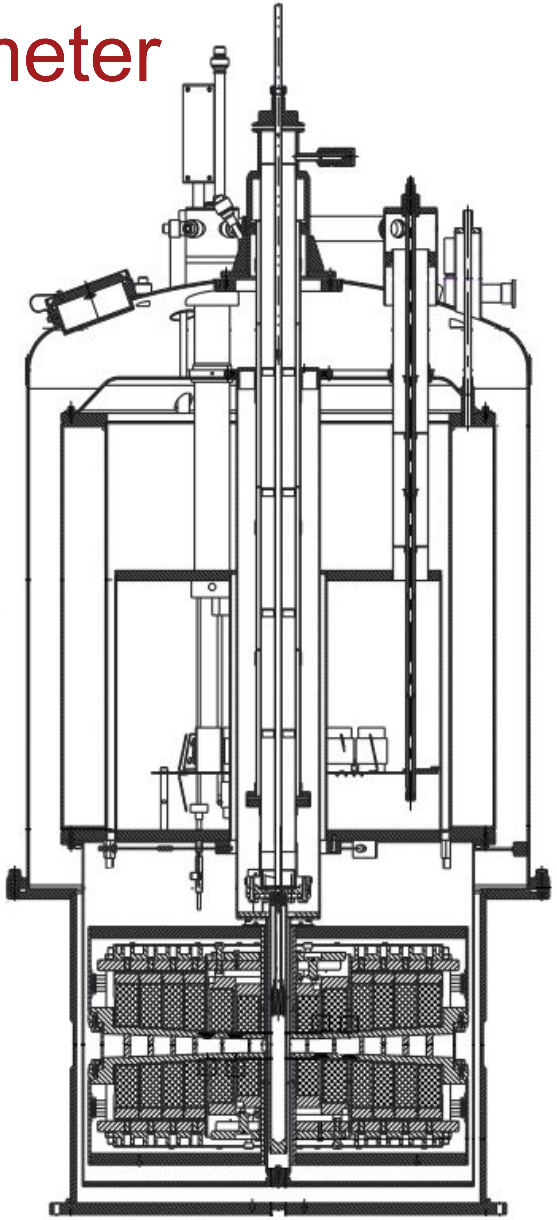
In Crystal Powders

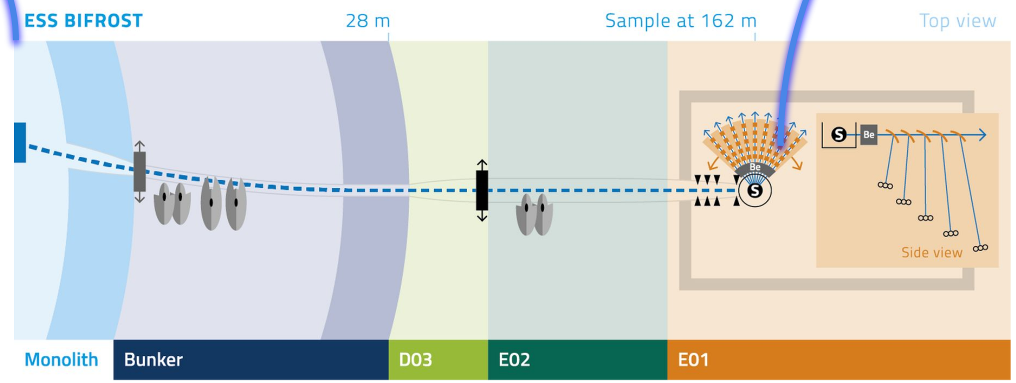


Debye-Scherre Cones

European Spallation Source - BIFROST Spectrometer

Sample Environment

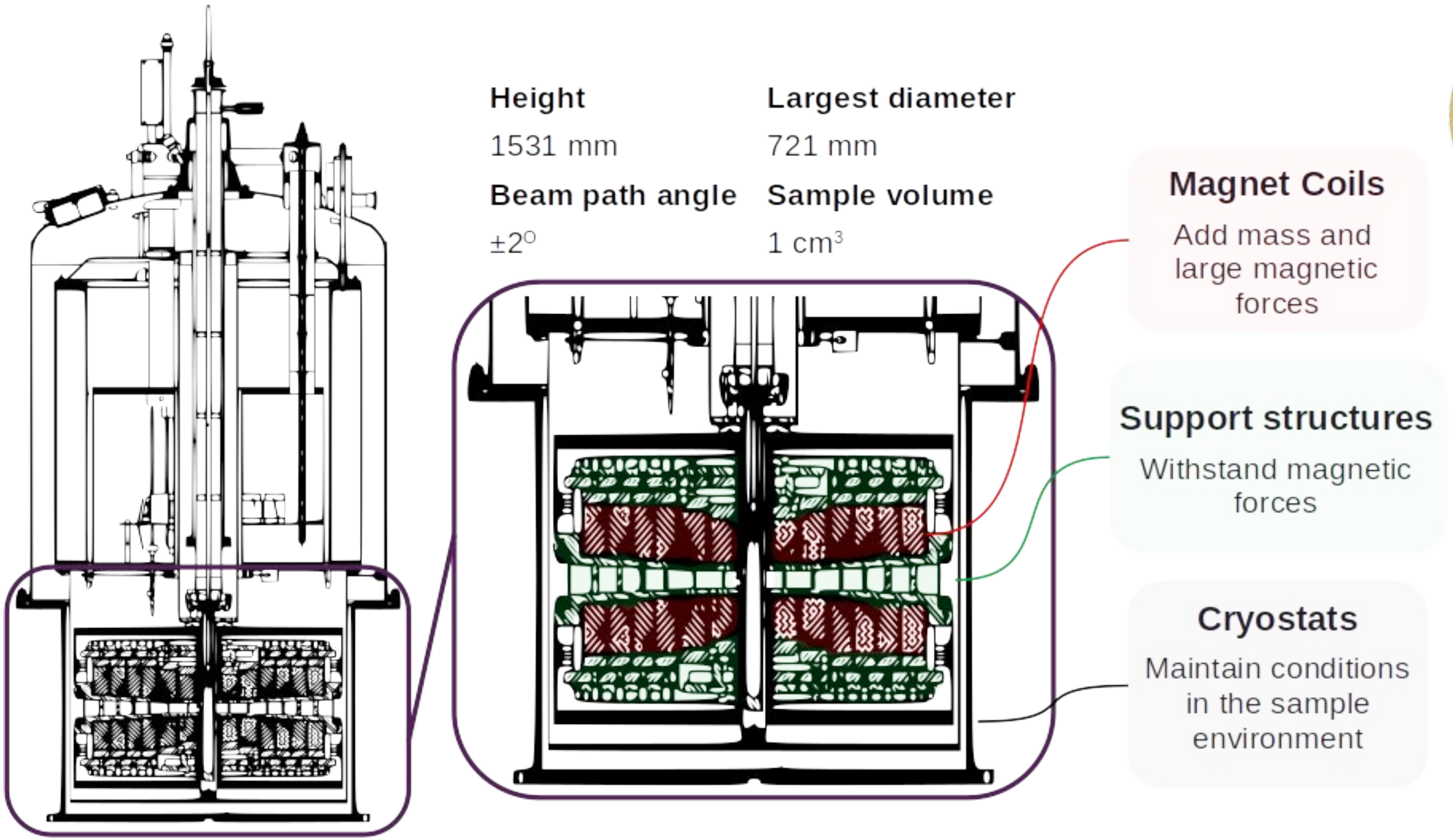




Creating synthetic data

Simulating the Sample Environment

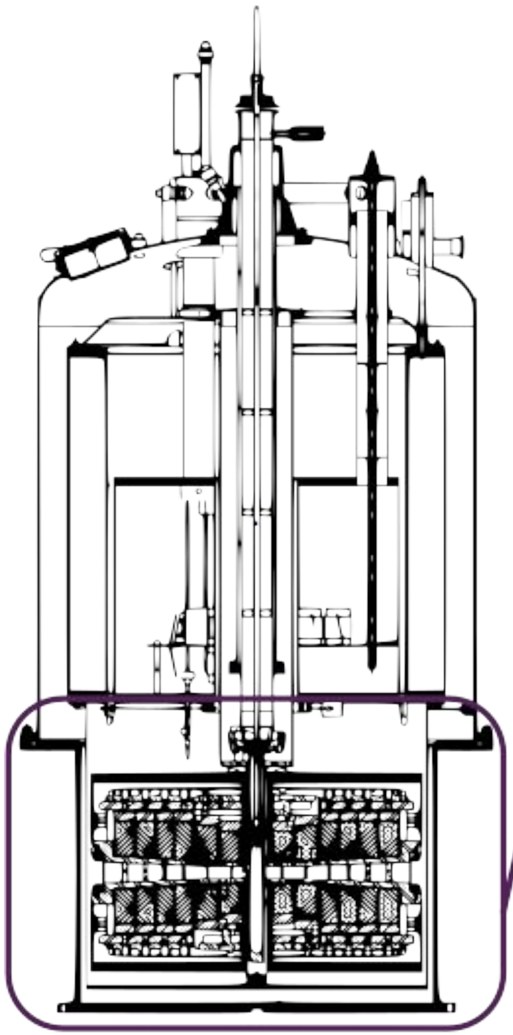
Monte Carlo Neutron
Ray-Tracing Simulation Package



McStas Union

Assign physical properties to desired geometries

Simulating the Sample Environment

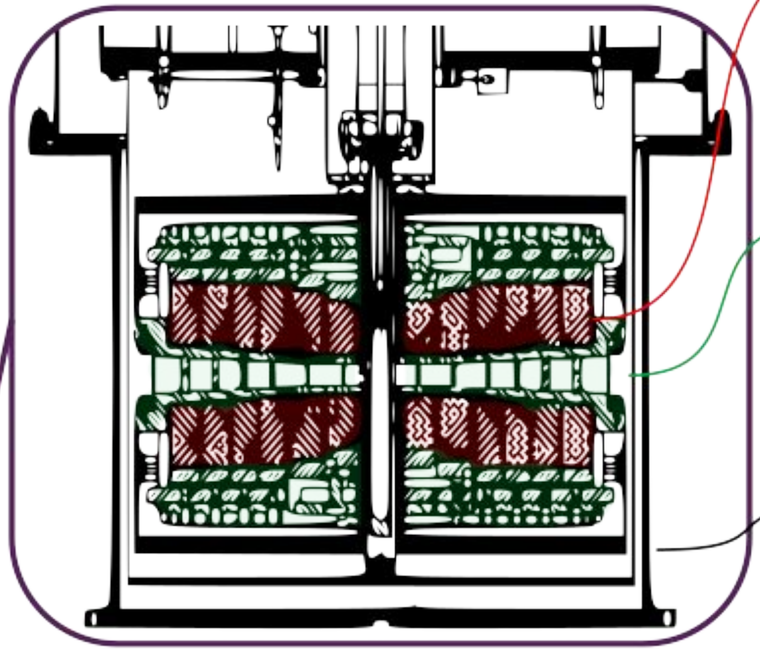


Height
1531 mm

Largest diameter
721 mm

Beam path angle
$\pm 2^{\circ}$

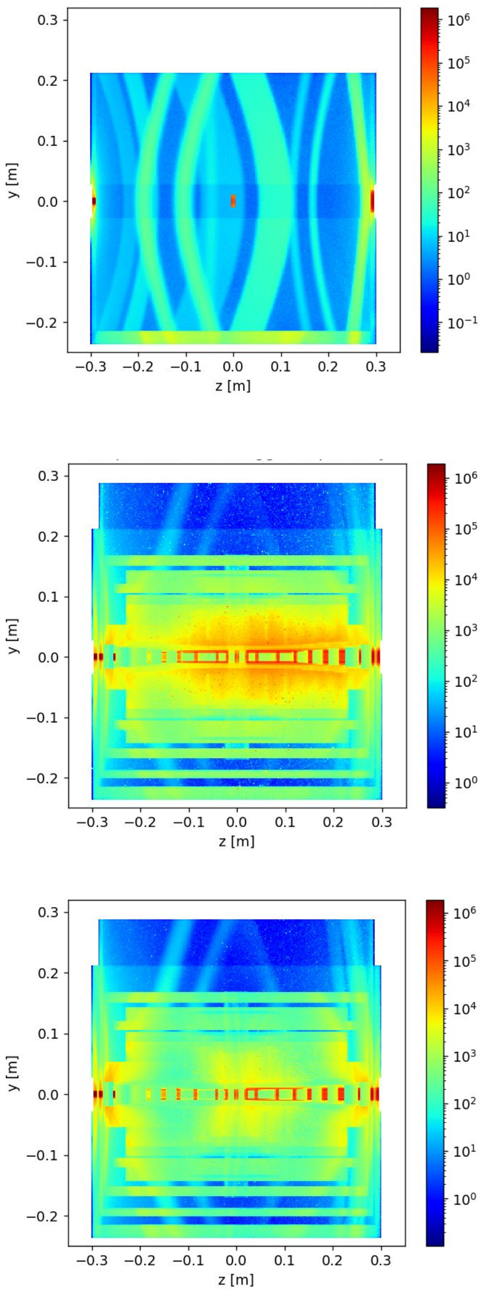
Sample volume
1 cm^3



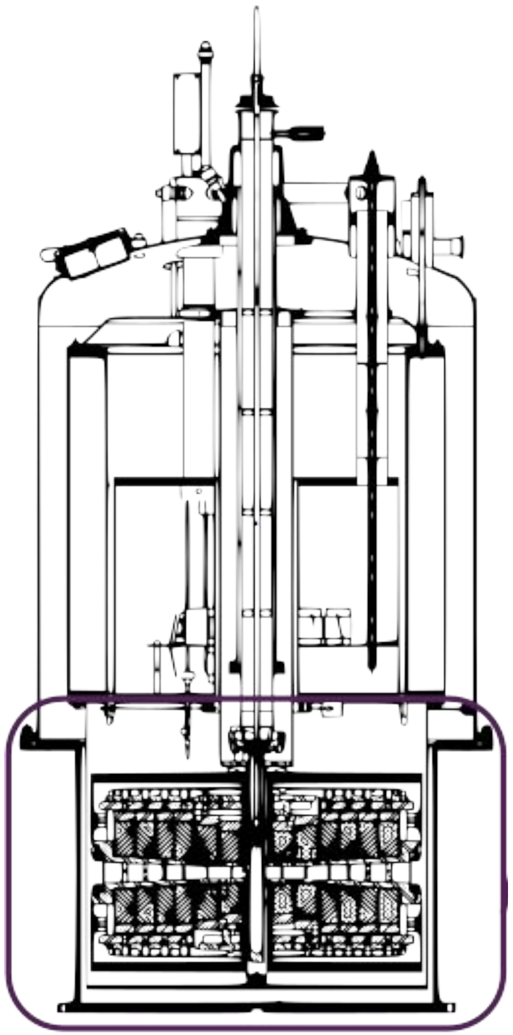
Magnet Coils
Add mass and large magnetic forces

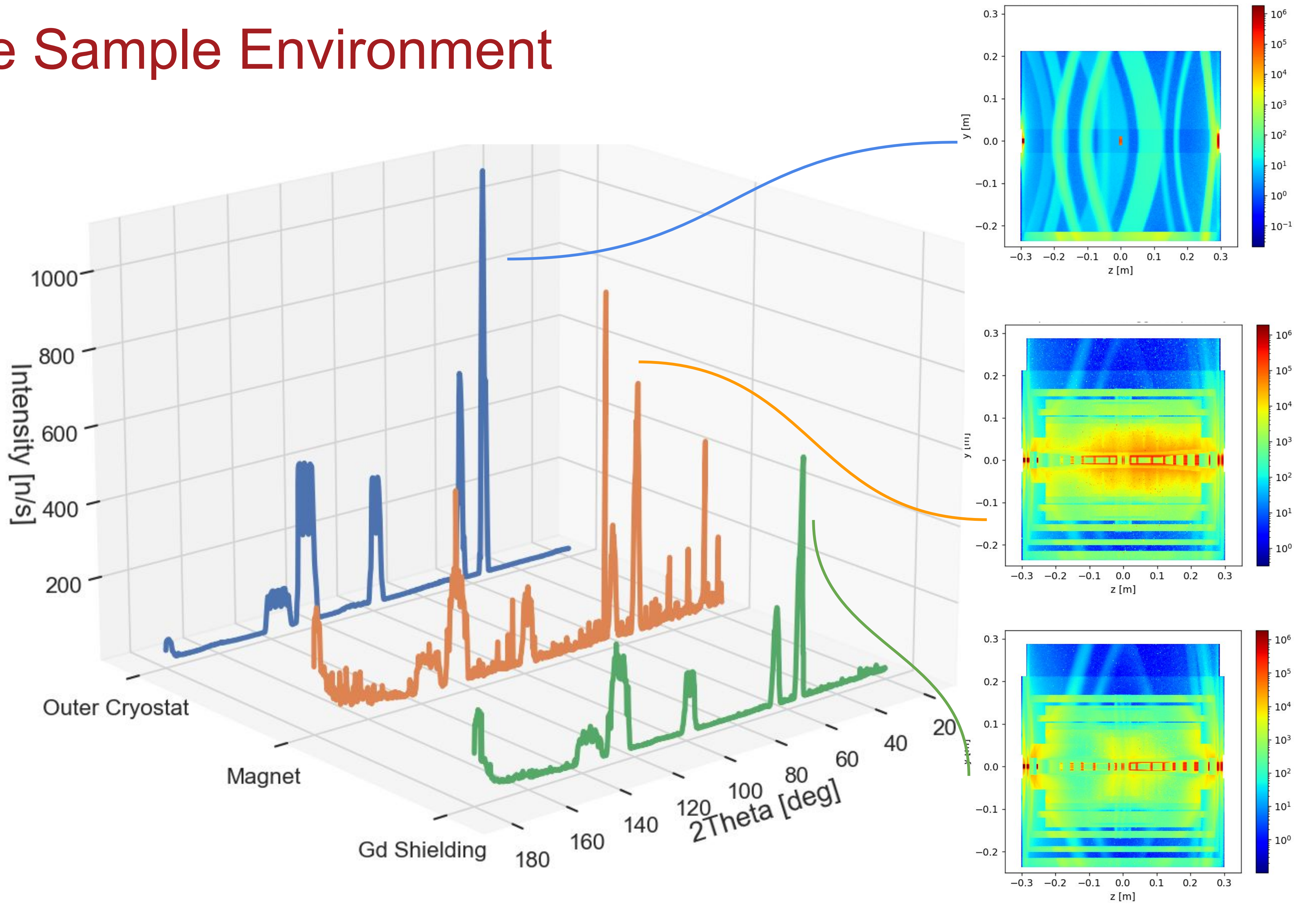
Support structures
Withstand magnetic forces

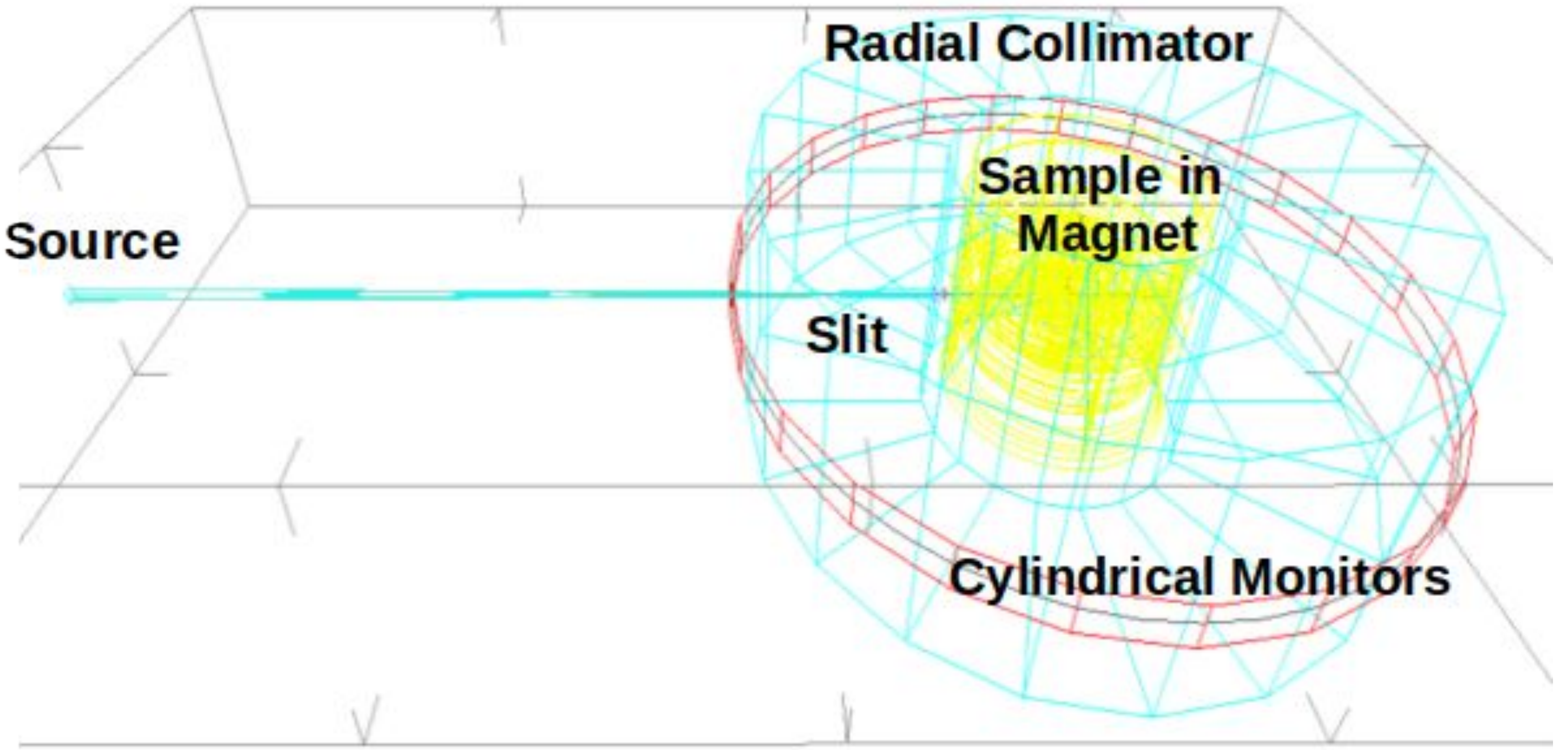
Cryostats
Maintain conditions in the sample environment

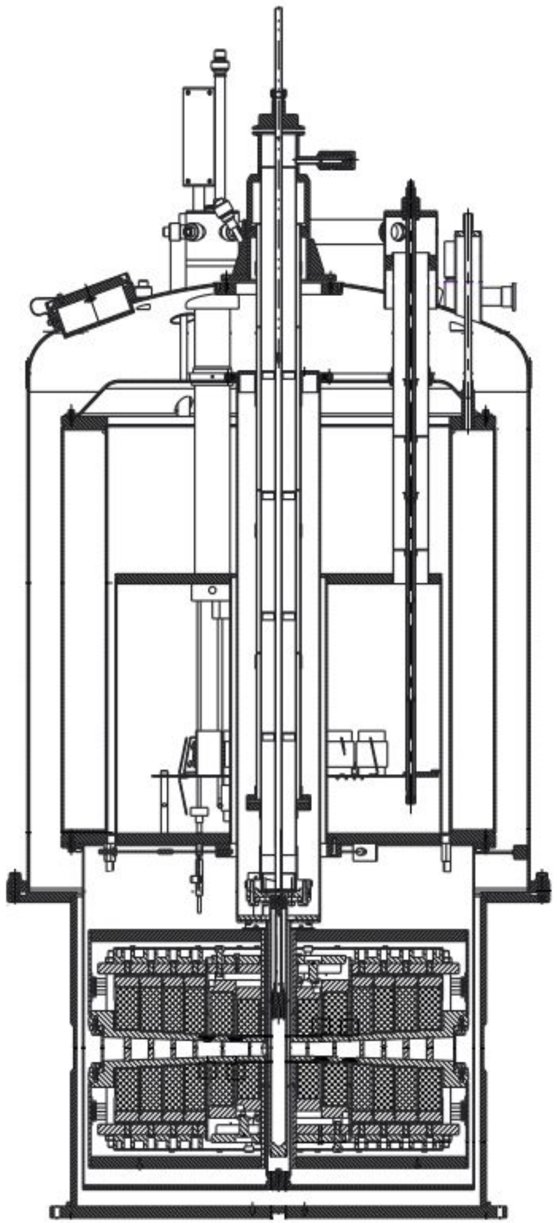


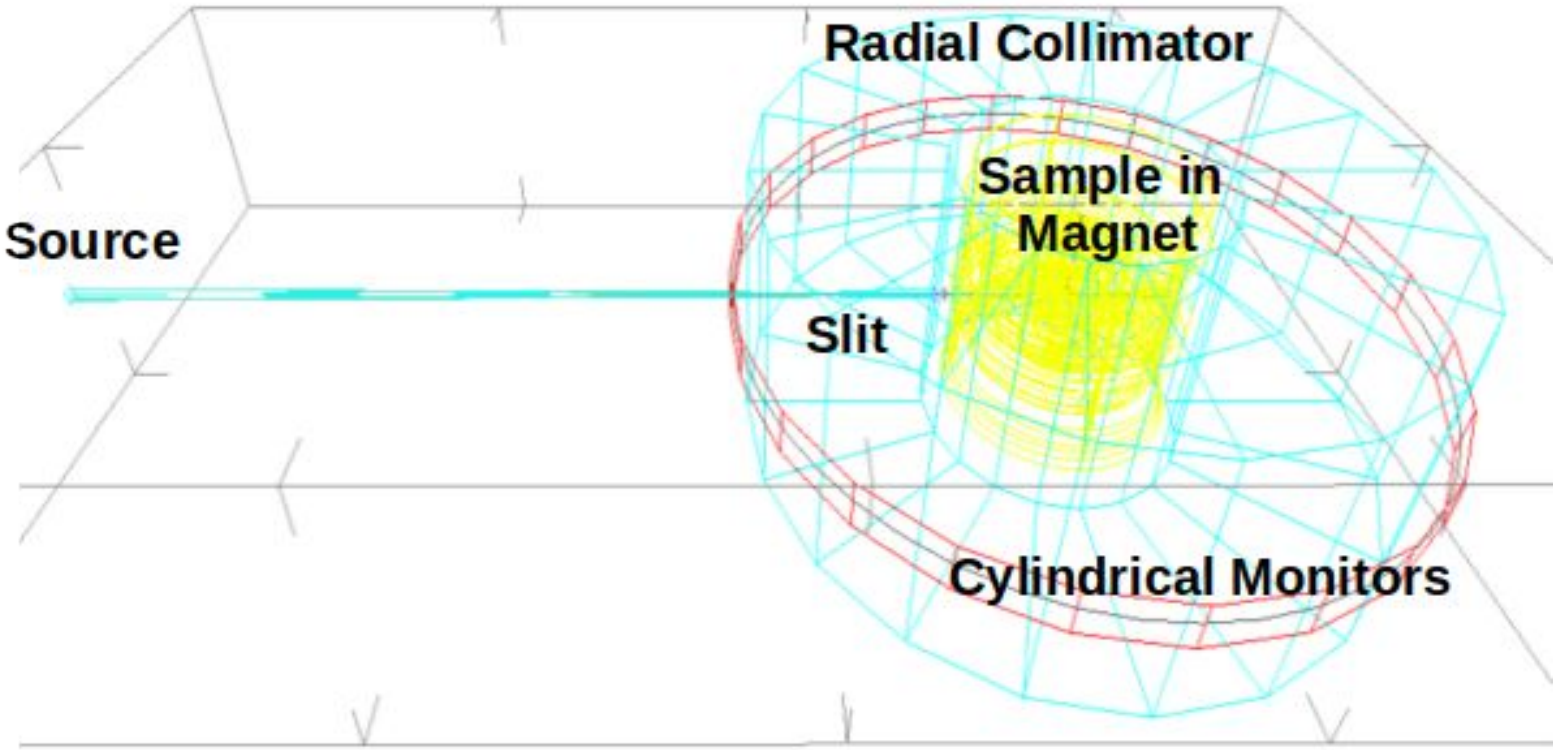
Simulating the Sample Environment

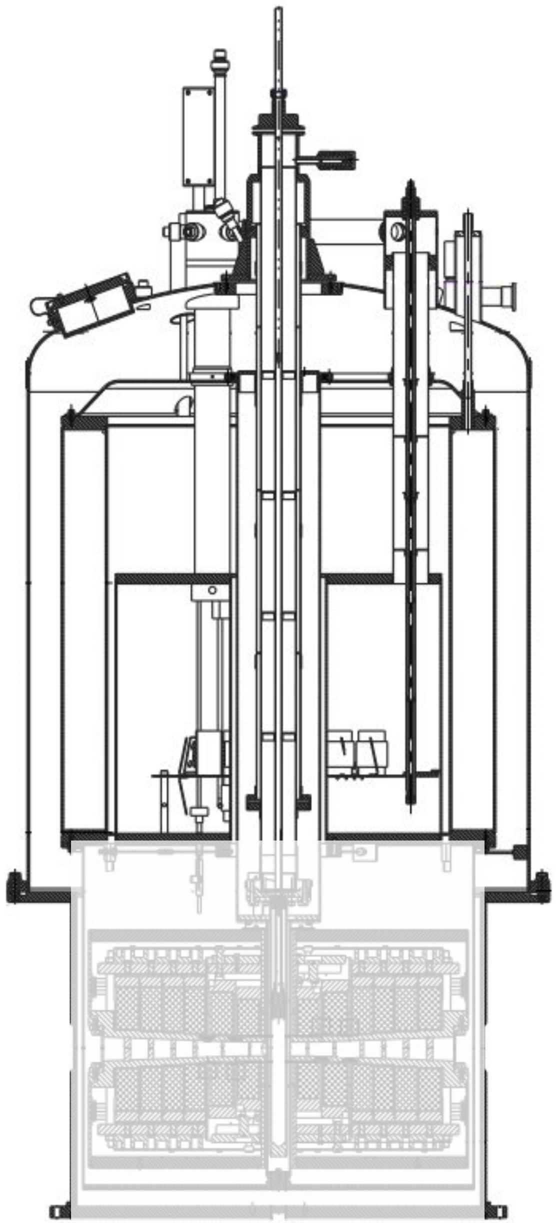





Radial Collimator
Sample in Magnet
Source
Slit
Cylindrical Monitors

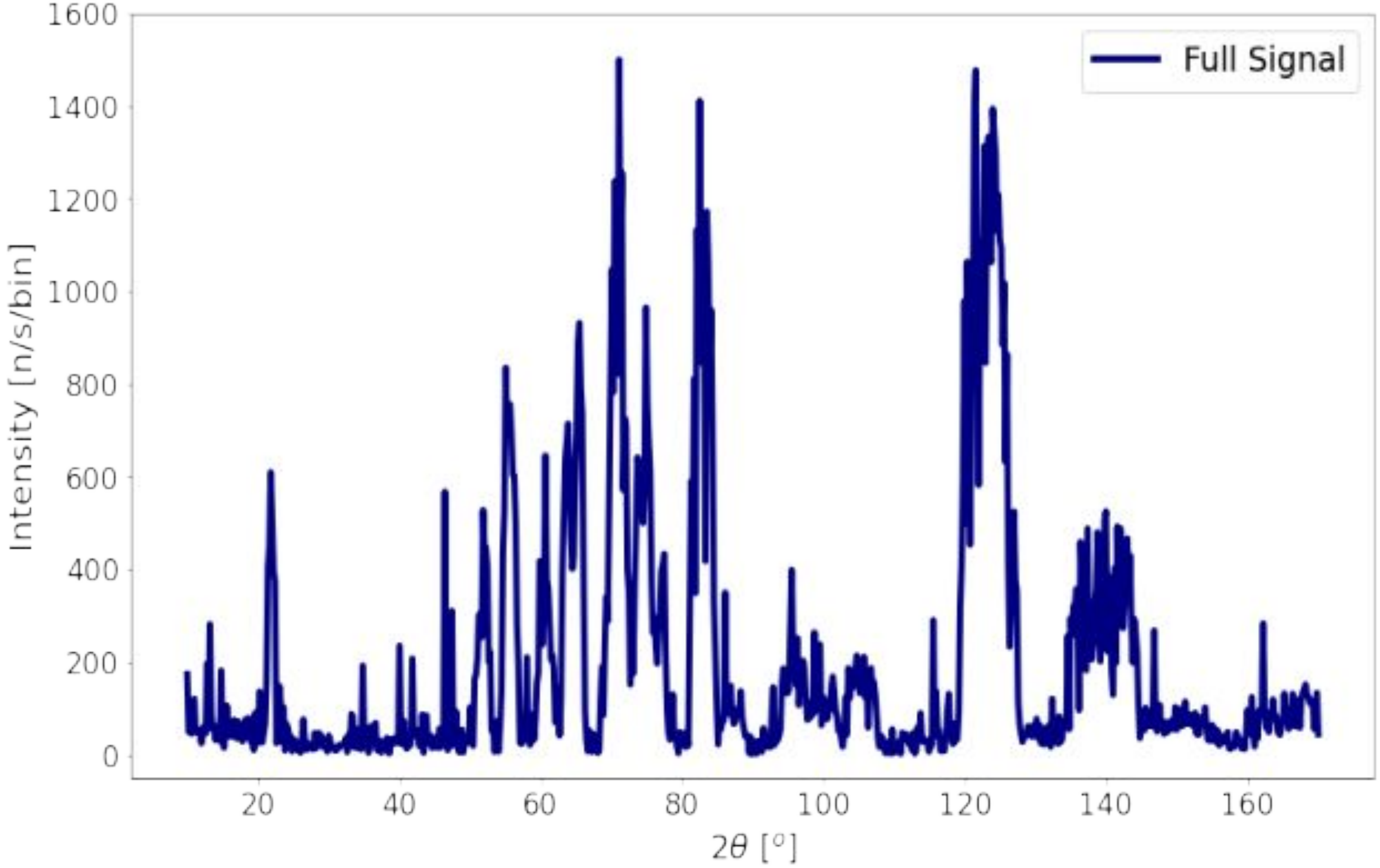



Radial Collimator
Sample in Magnet
Source
Slit
Cylindrical Monitors

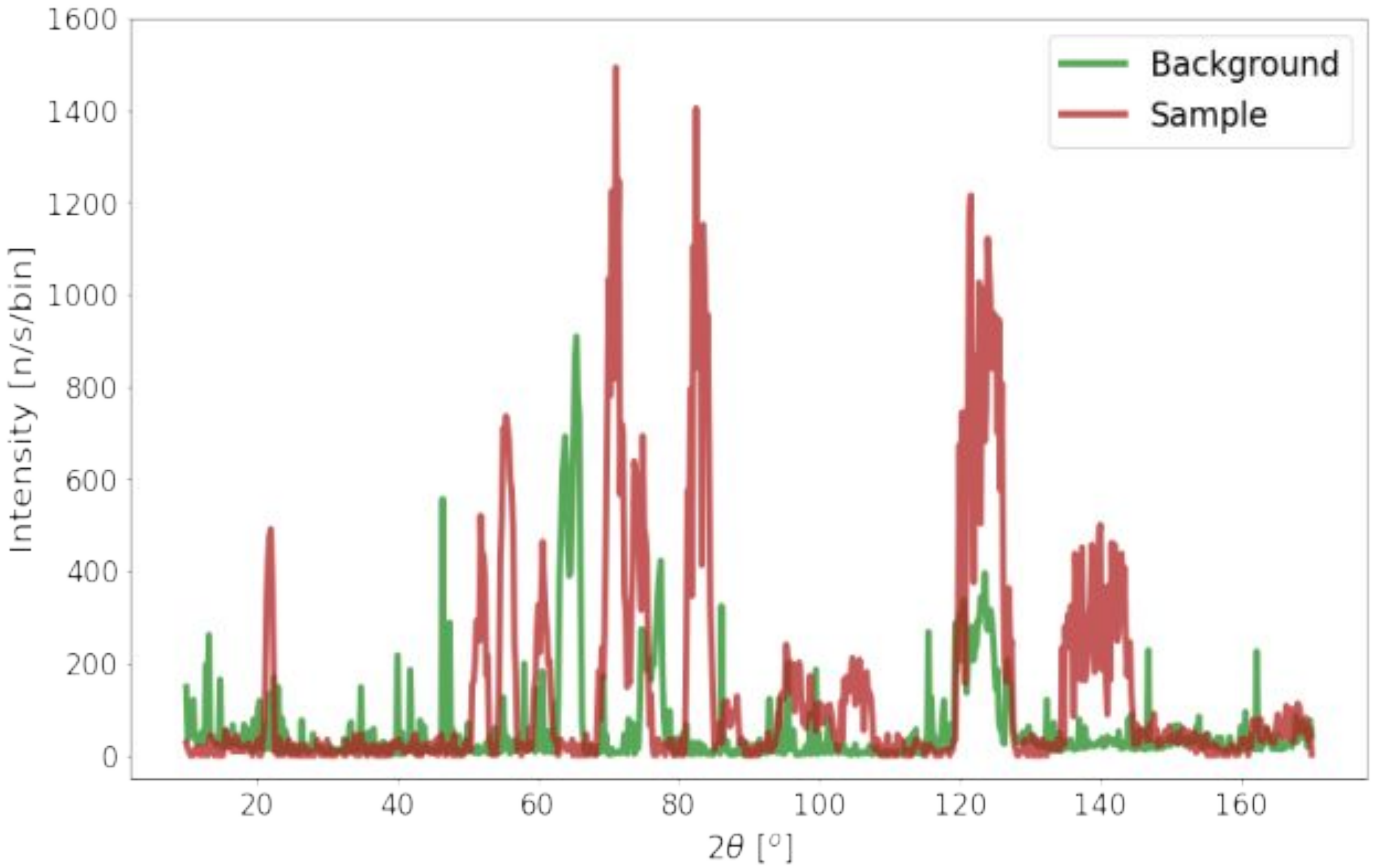


What problem do we need to solve?

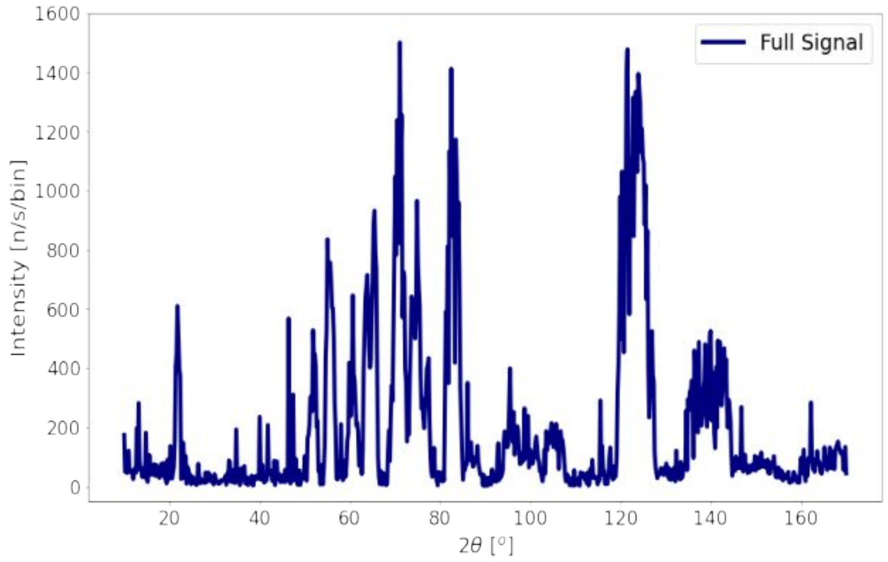
Background Scattering

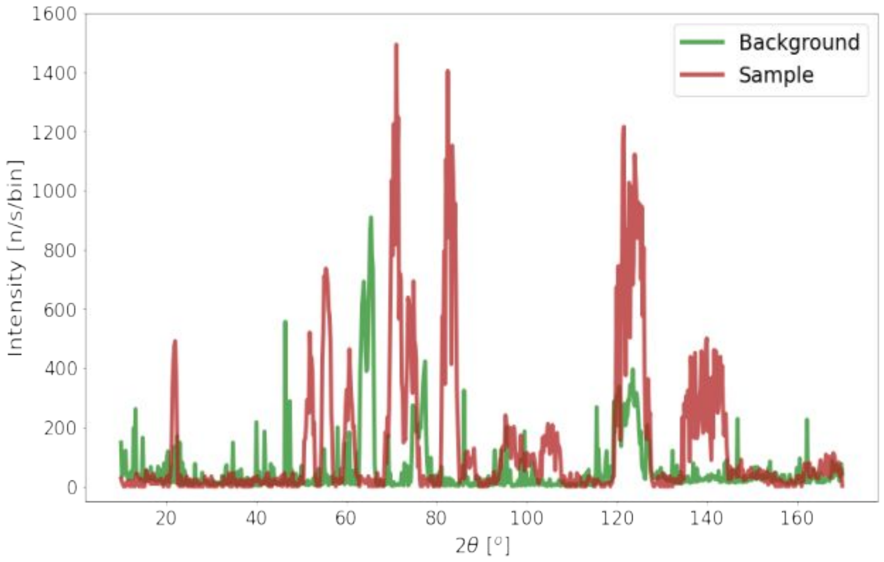


Background Scattering



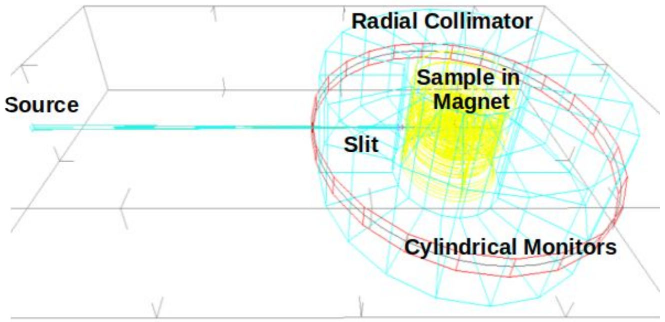
Example: $La_{2-x}Sr_xCuO_4$, λ=1.47 Å



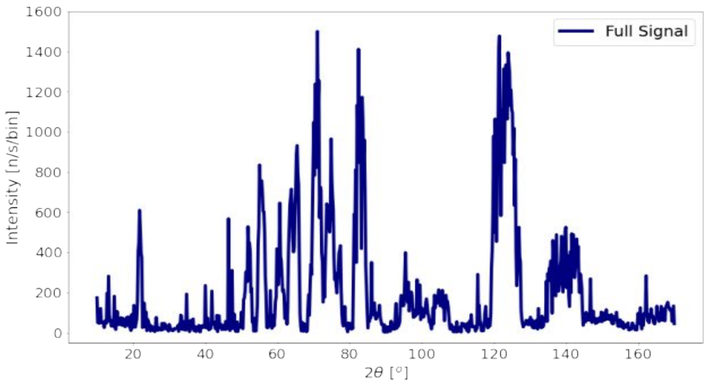


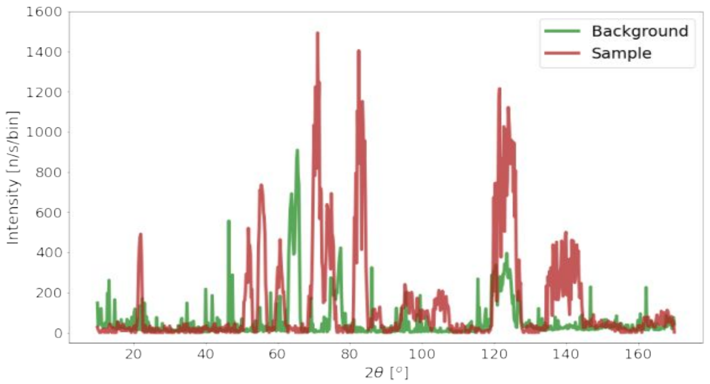
Produced over 24000 sets of synthetic data based on 7 parameters:

- Wavelength: λ, λd
- Beam divergence
- Sample dimension
- Sample - Detector distance
- Sample material



- Wavelength: λ, λd
- Beam divergence
- Sample height/radius
- Sample - Detector distance
- Sample material

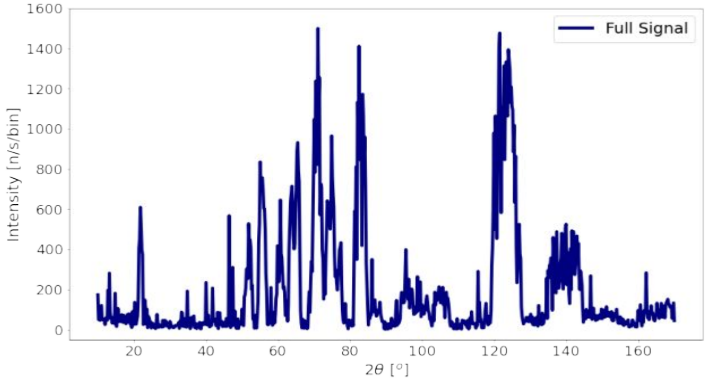


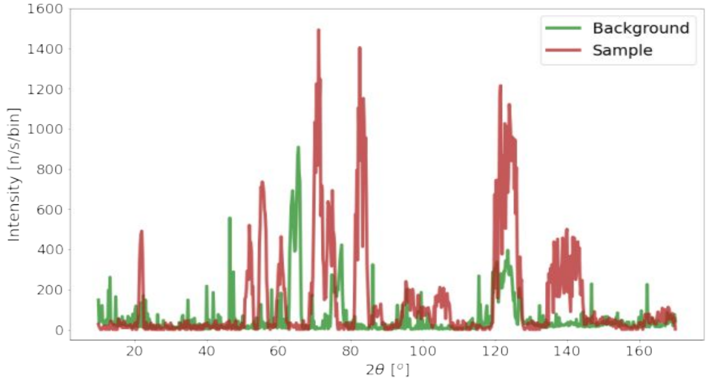


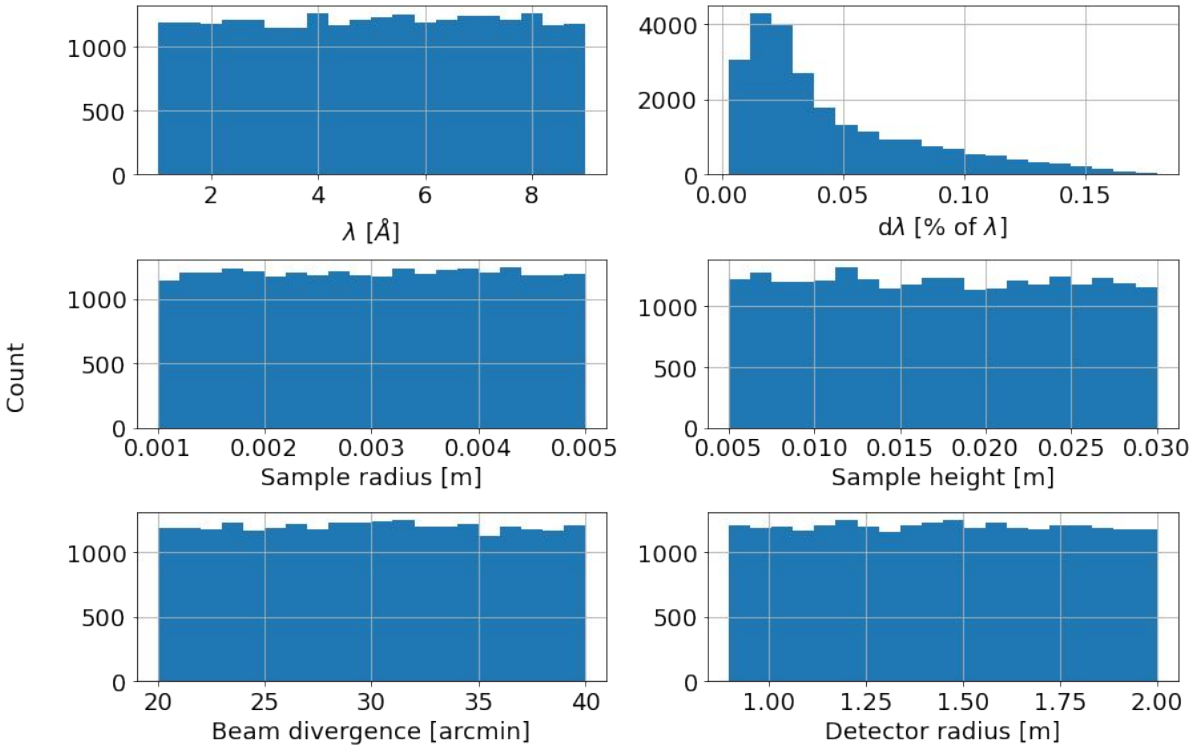
Features

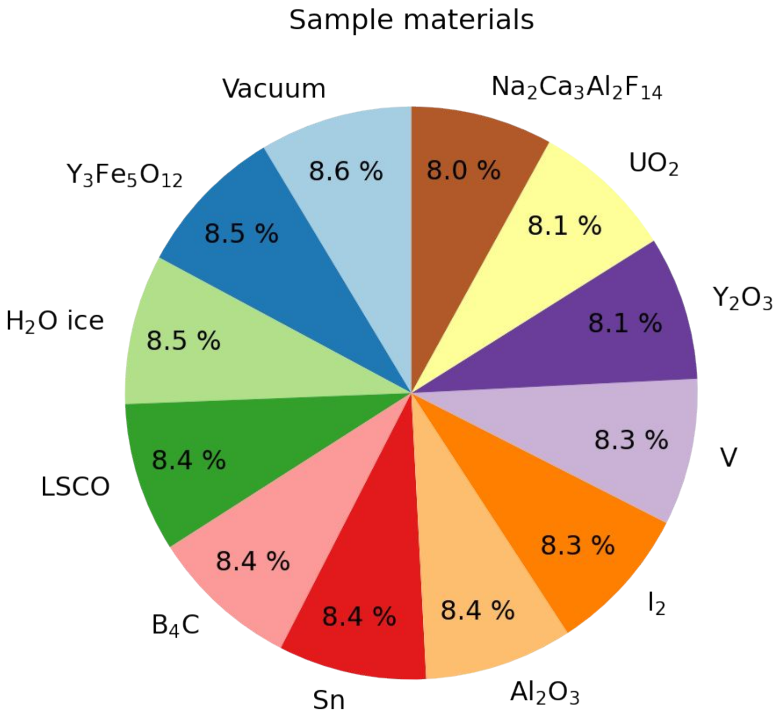
Target Values

- Wavelength: λ, λd
- Beam divergence
- Sample height/radius
- Sample - Detector distance
- Sample material







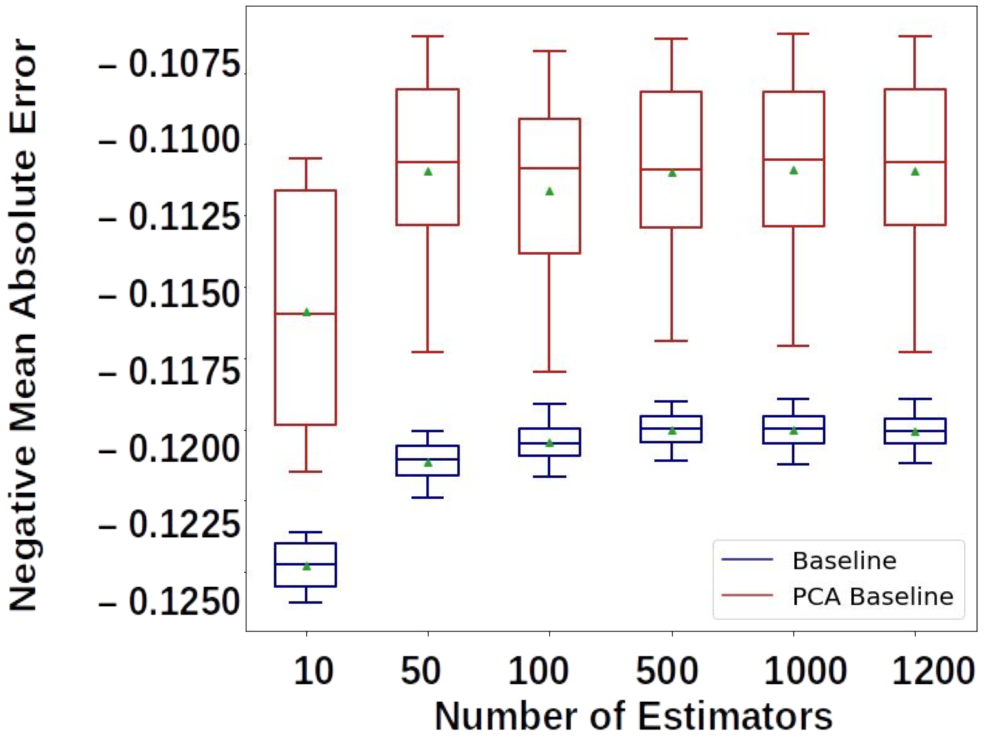


What’s the best way to predict background?

Starting with a Random Forest:

1. Features: High dimensional or reduced?
2. Information - Complexity trade-off
3. Background information - bias exploration

1. Dimensionality Reduction



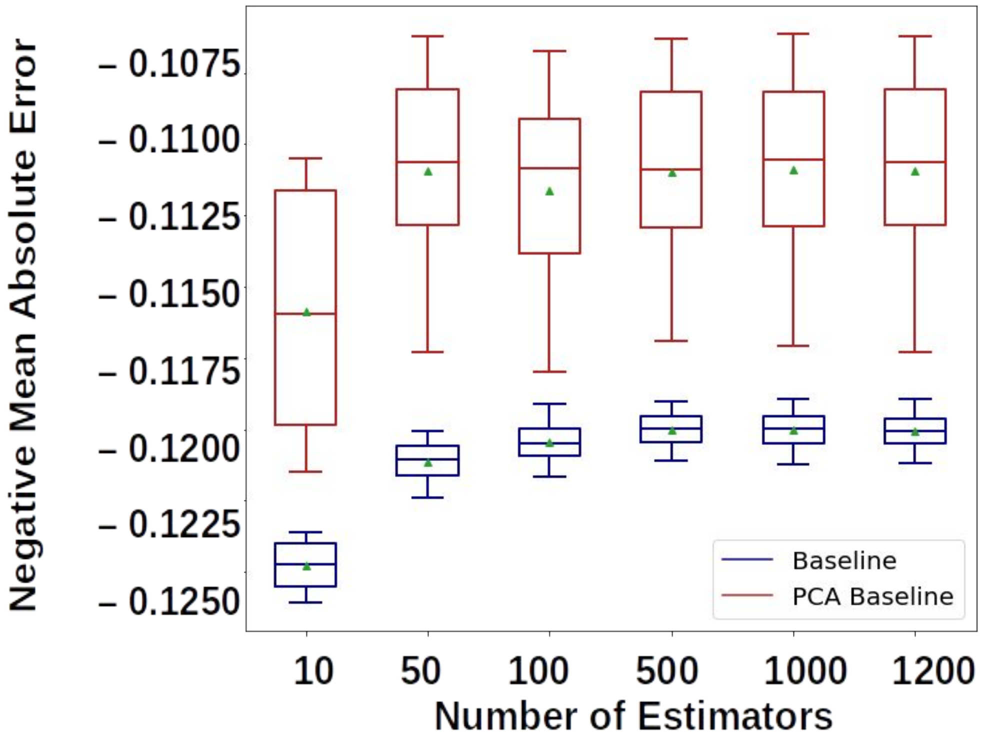
High dimensionality: 807 features

- 7 Instrument parameters
- 800 intensity/angle values

PCA: 81 features

- 89.96% of feature reduction
- 94.45% of information preserved

1. Dimensionality Reduction



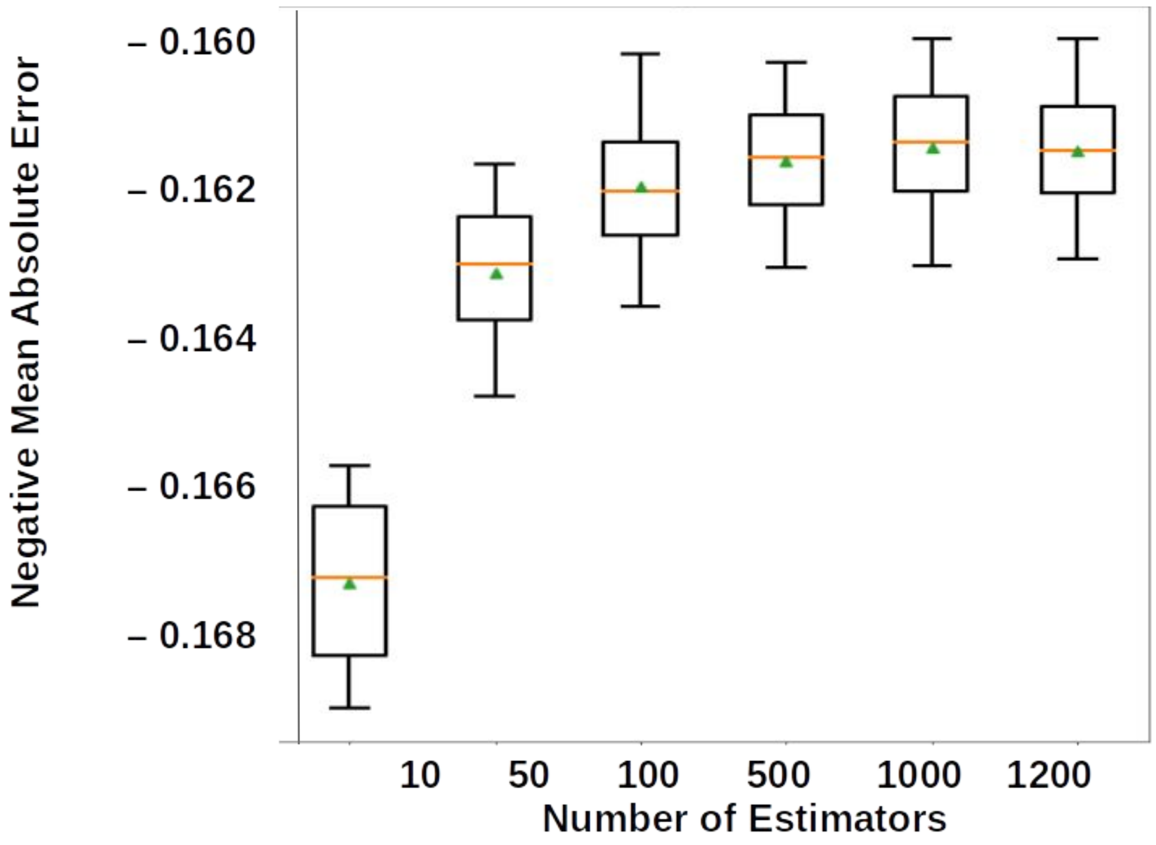
High dimensionality: 807 features

- 7 Instrument parameters
- 800 intensity/angle values

PCA: 81 features

- **89.96% of feature reduction**
- **94.45% of information preserved**

2. Information vs Complexity in target values



Normalised Intensity within 10° - 170°

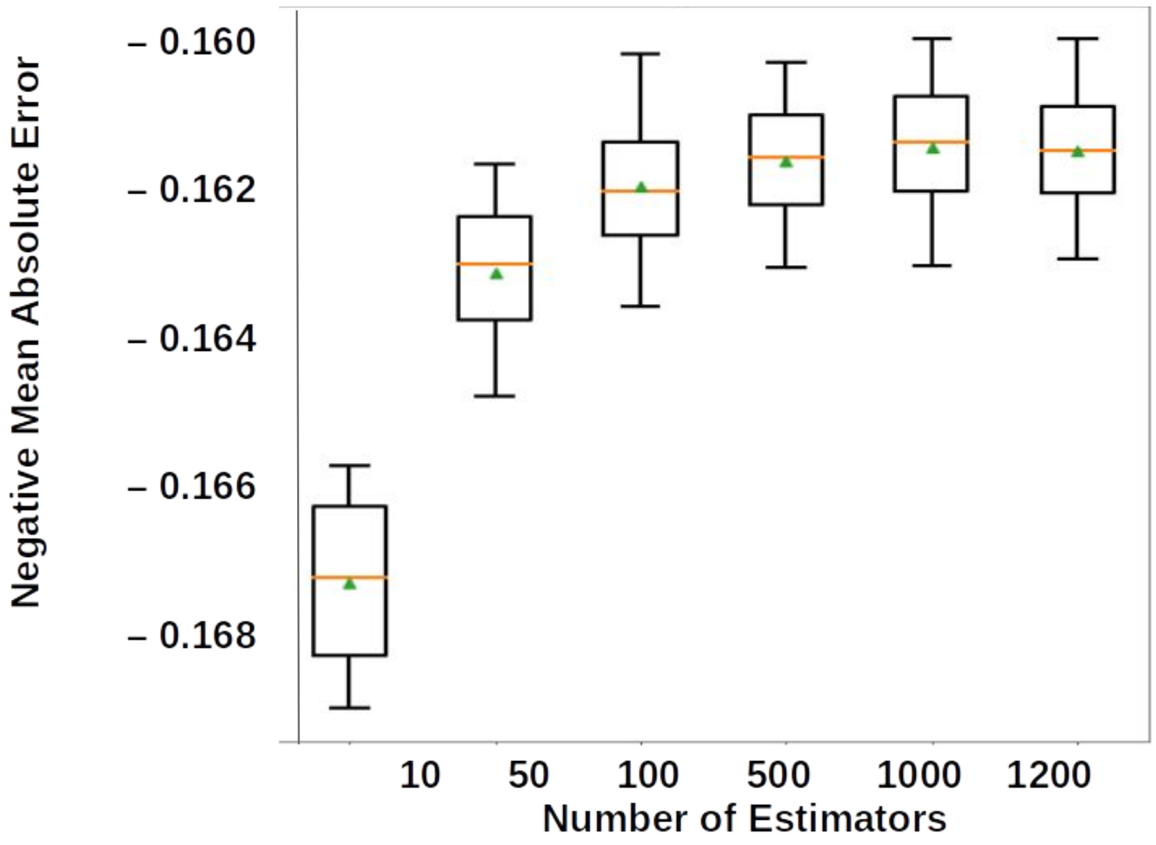
5 bins per degree: 800 targets

MAE: 0.161

1 bin per degree: 160 targets

MAE: 0.111

2. Information vs Complexity in target values

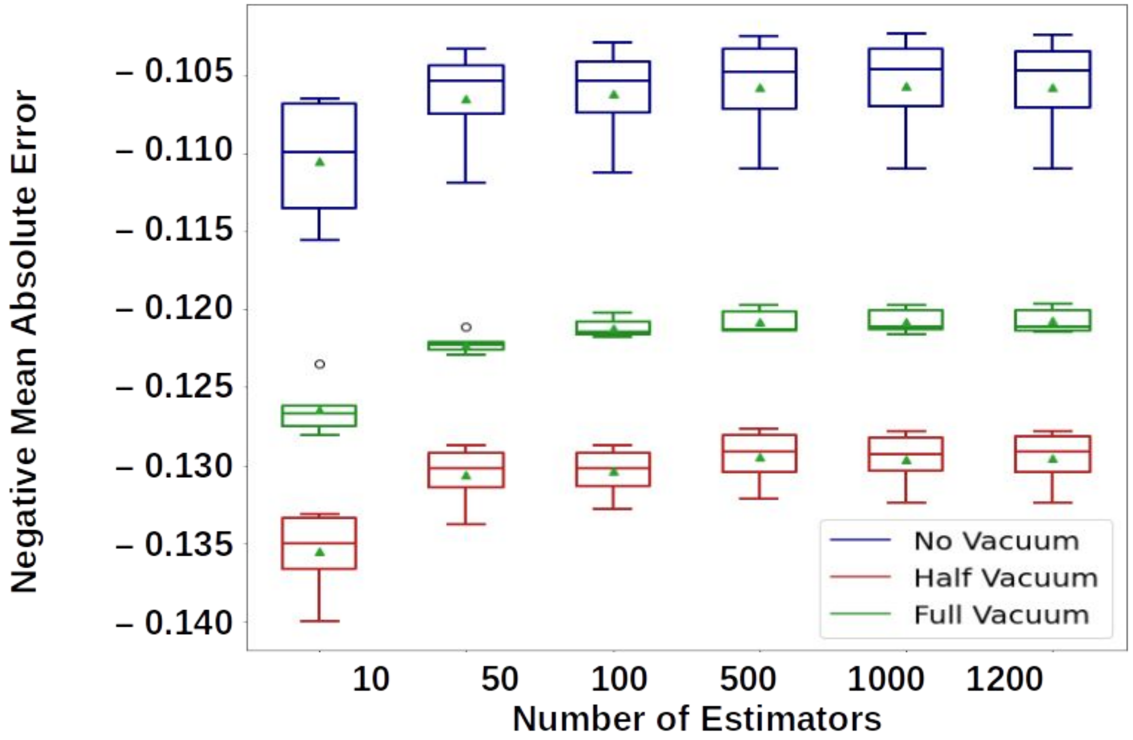


Normalised Intensity within 10° - 170°

5 bins per degree: 800 targets
MAE: 0.161

1 bin per degree: 160 targets
MAE: 0.111

3. Bias Exploration: Measurements of ***pure background***

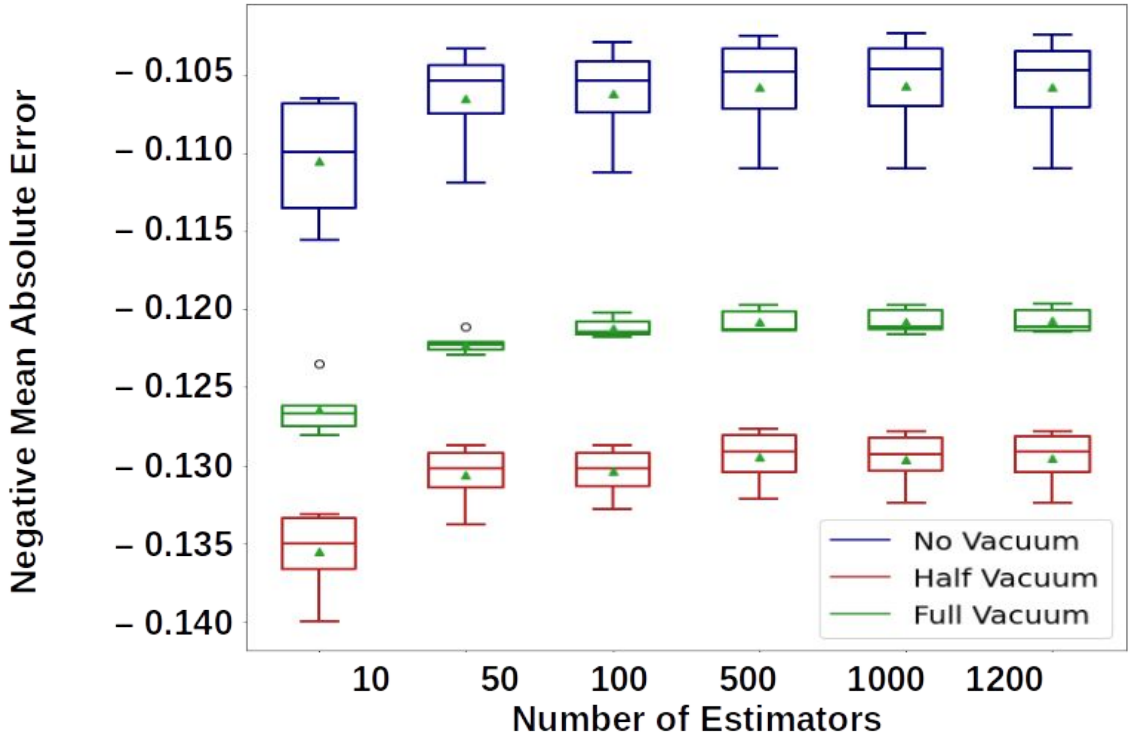


0% Background Measurements
MAE: 0.105

50% Background Measurements
MAE: 0.129

100% Background Measurements
MAE: 0.121

3. Bias Exploration: Measurements of ***pure background***

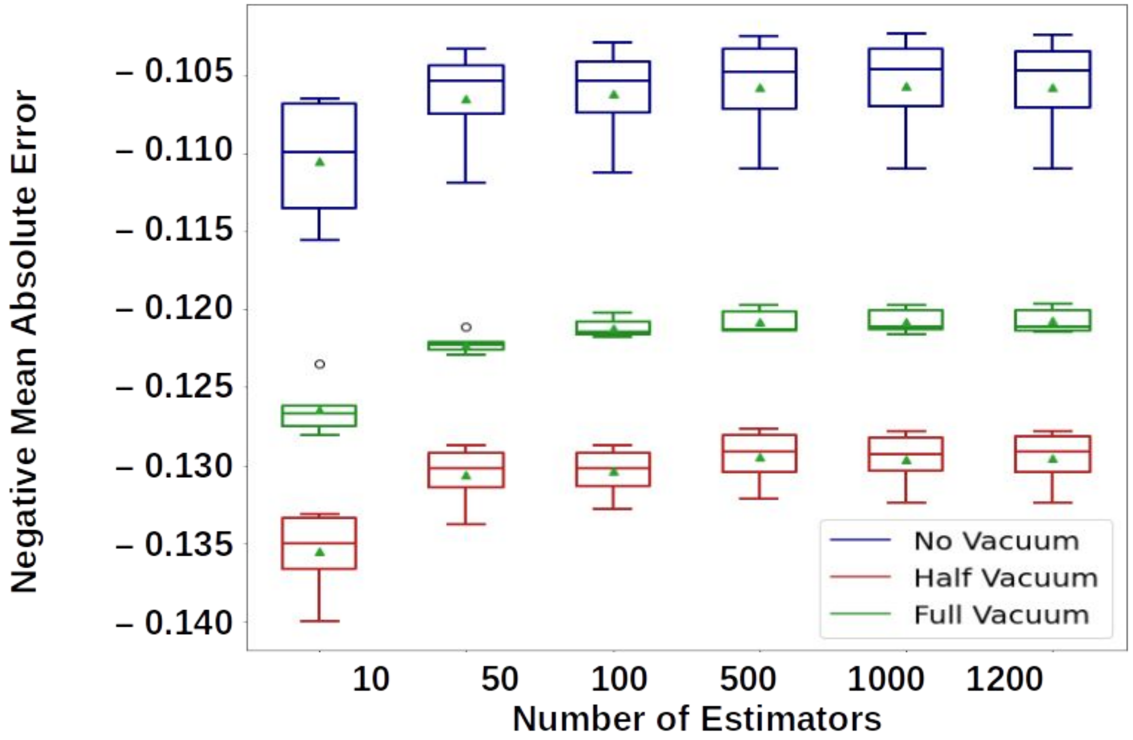


0% Background Measurements
MAE: 0.105

50% Background Measurements
MAE: 0.129

100% Background Measurements
MAE: 0.121

3. Bias Exploration: Measurements of ***pure background***



0% Background Measurements
MAE: 0.105

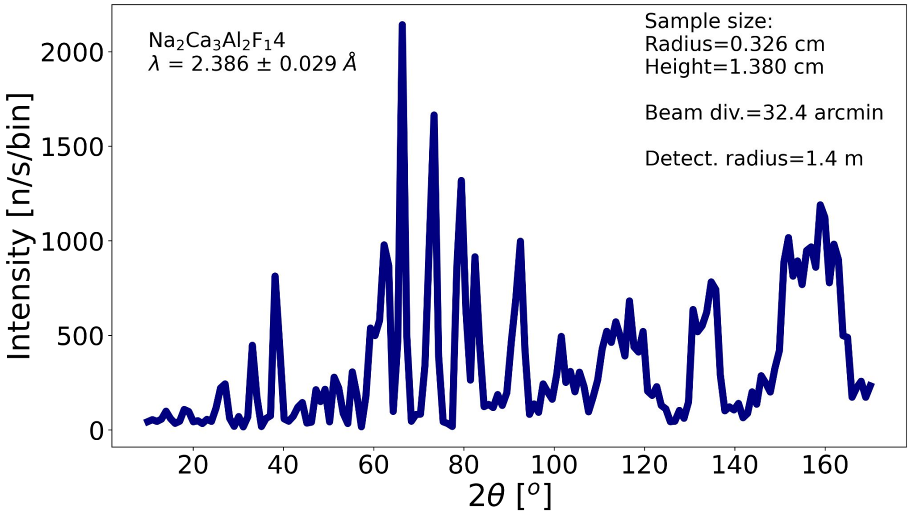
50% Background Measurements
MAE: 0.129

100% Background Measurements
MAE: 0.121

First results

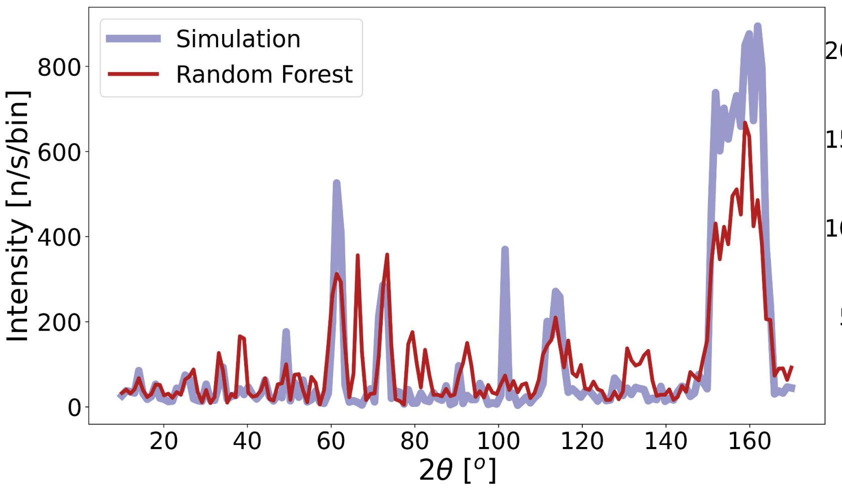
Random Forest vs Gradient Boost

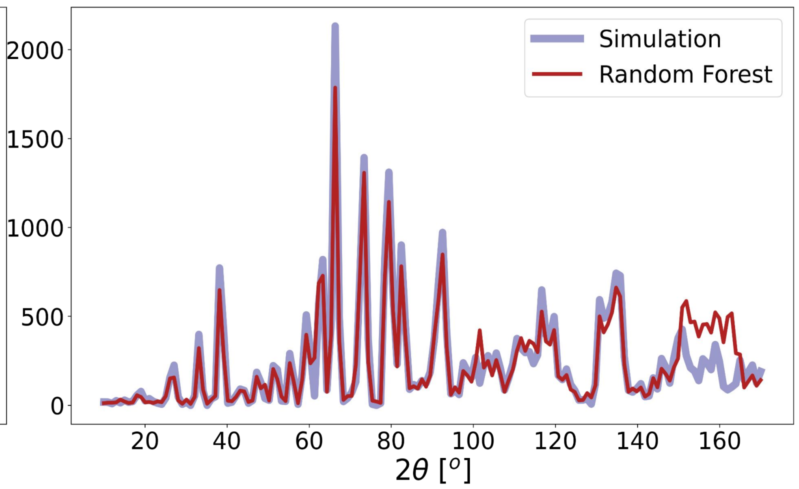
Simulation

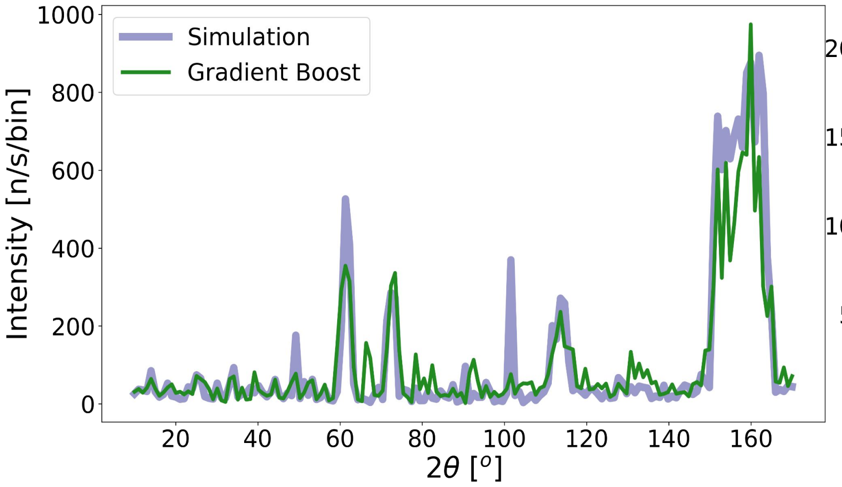


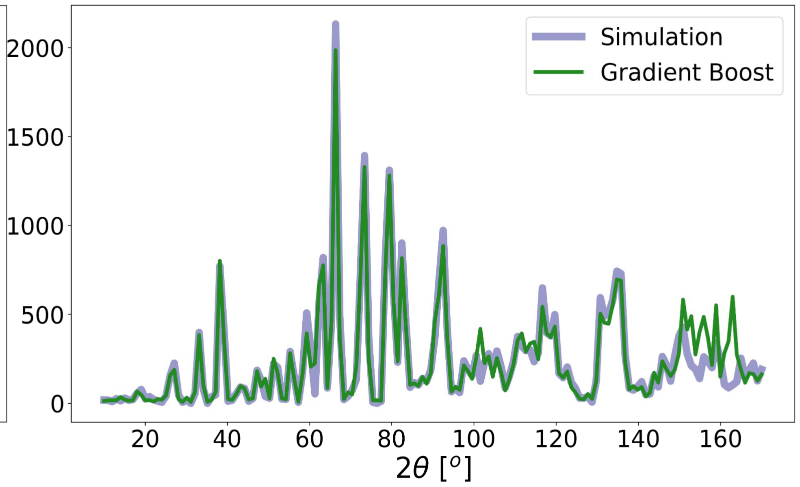
Predicted Background

Predicted Sample









Overall Performance

Model	Random Forest	Gradient Boost
MAE	0.1215	0.1192

Still a work in progress…

- Improve data: Simulation quality and better representation of materials
- Restructuring the database under the F.A.I.R. framework
- Try Neural Networks - open to suggestions
- Generalise: Material-agnostic model
- Synthetic + Real data hybrid

Nonetheless…

First step in background prediction in neutron scattering data

Thank you!

Questions?