FRIB

Towards AI Enabled High-Power Heavy Ion Facility with Confidence

Yue Hao and Steve Lidia

FRIB/NSCL

MICHIGAN STATE UNIVERSITY

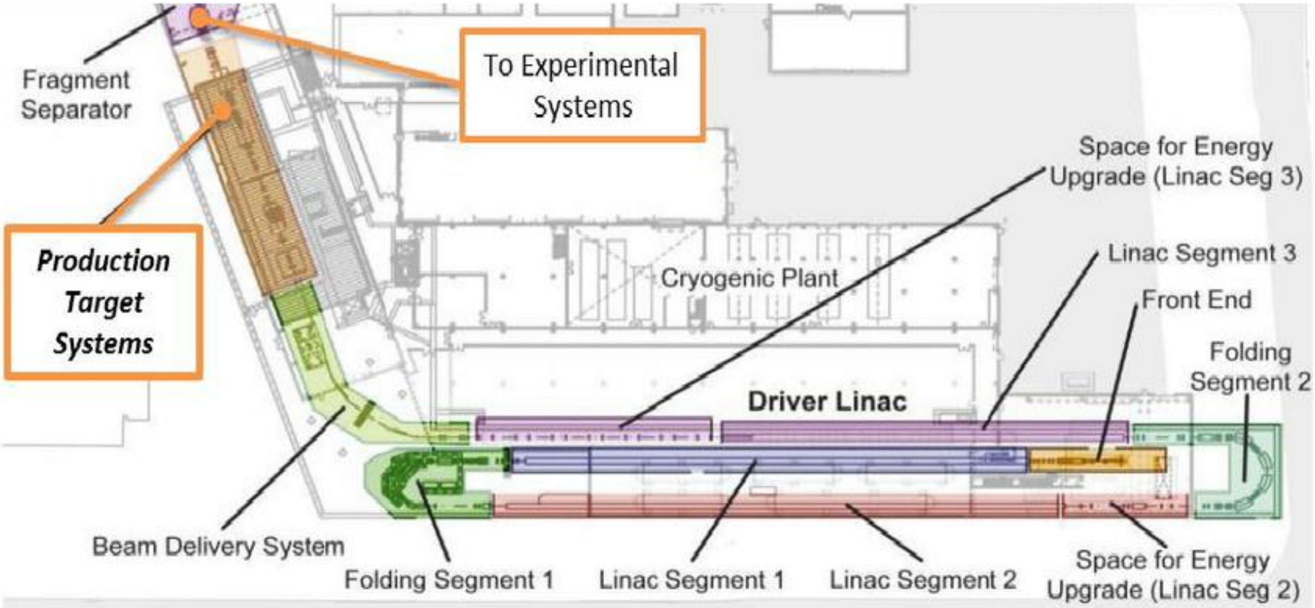
U.S. DEPARTMENT OF ENERGY | Office of Science

This material is based upon work supported by the U.S. Department of Energy Office of Science under Cooperative Agreement DE-SC0000661, the State of Michigan and Michigan State University. Michigan State University designs and establishes FRIB as a DOE Office of Science National User Facility in support of the mission of the Office of Nuclear Physics.

Overview

- Challenges at the heavy ion intensity frontier
- Current studies
- Future efforts
- Proposed areas for collaboration
- Summary

FRIB is an intensity-frontier facility



FRIB is expected to deliver the largest ion beam power.

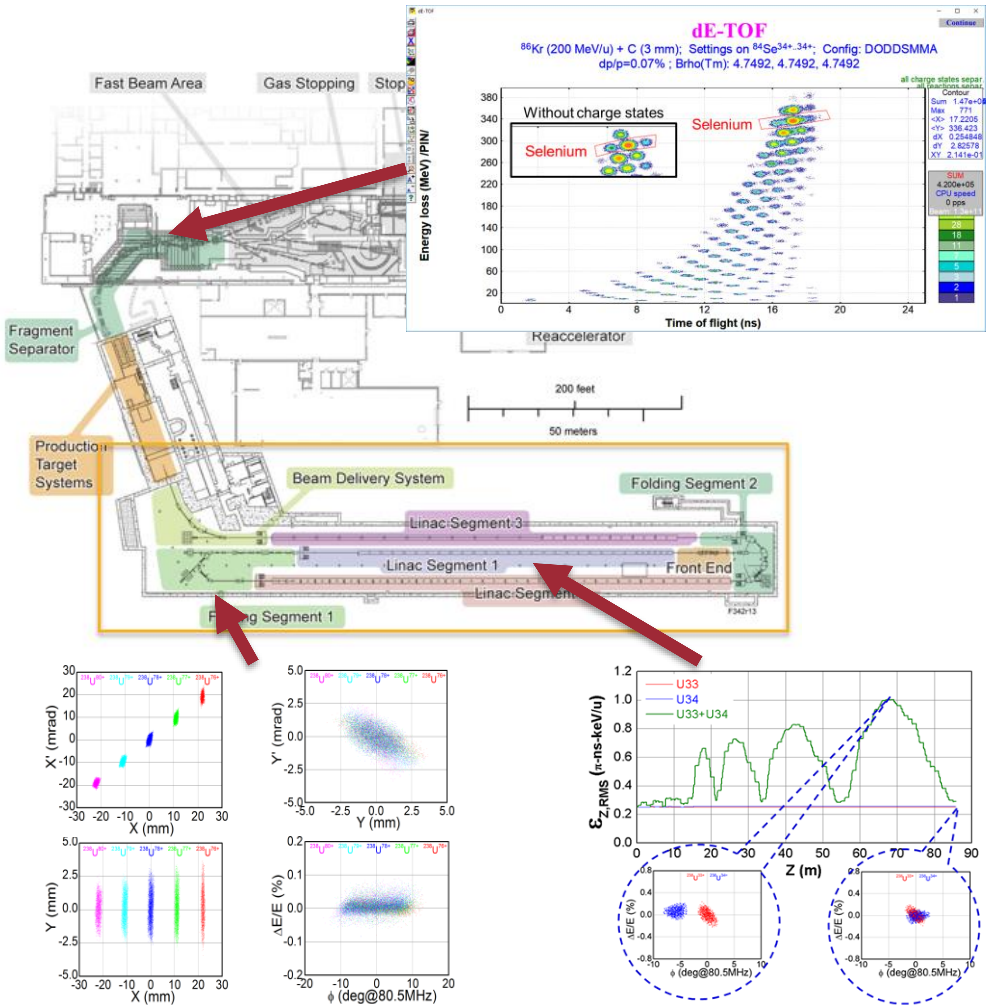
- 200 MeV/u Uranium
- 400 kW beam power
- High availability

Our efforts to achieve the goal:

- Build reliable hardware and precise installation
- Better model to describe and predict the machine
- Effective Machine and Personnel Protection

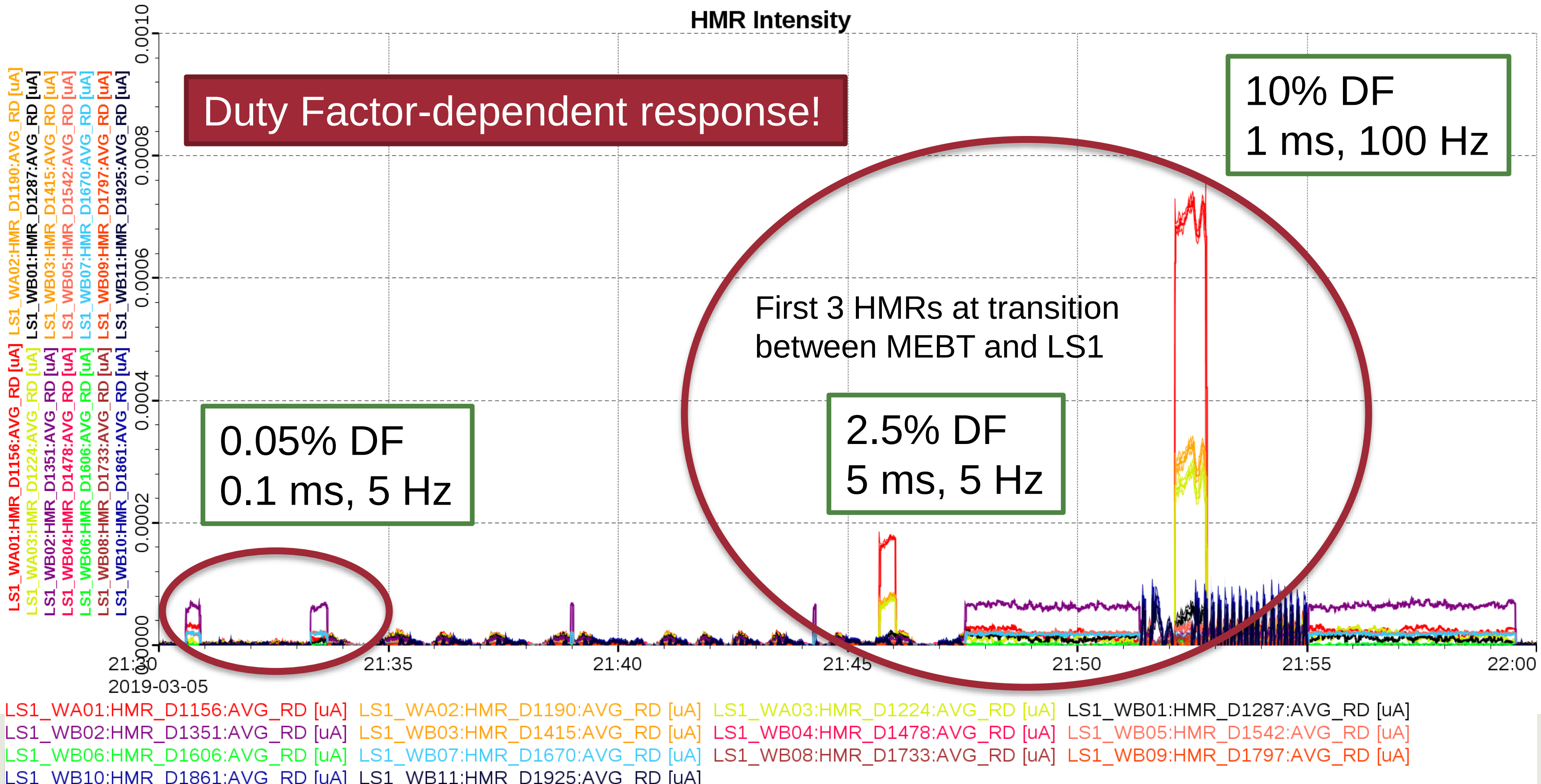
Technical Challenges to Realizing FRIB

- **Handling intense, low energy ion beams**
  - Multiple charge state beam dynamics with varying Q/A
  - Ensuring low beam losses with robust Machine Protection and diagnostics
  - Safe operation of liquid lithium charge stripper
  - 400 kW heavy ion beam target and pre-separator systems
- **Frequent retuning** to support many types of rare isotopes and experiments
  - Each run extends 1-2 weeks
  - Efficient tuning of ion source and injector
  - Efficient tuning of fragment separator



Beam Loss Changes with Duty Factor

- There are signals from the first 3 cryomodules Halo Monitor Rings.
- Higher sensitivity than Beam Current Monitors
- The highest shows beam fraction of $2 \cdot 10^{-4}$ was intercepted with HMRs



Application 1: How to control accelerator

Machine

Measurement, refinement →

← Optimization, tuning

Model

Measured data
- Beam position, intensity, energy and loss count
- Readback of settings
- Environment data (temp, pressure, vacuum, etc)

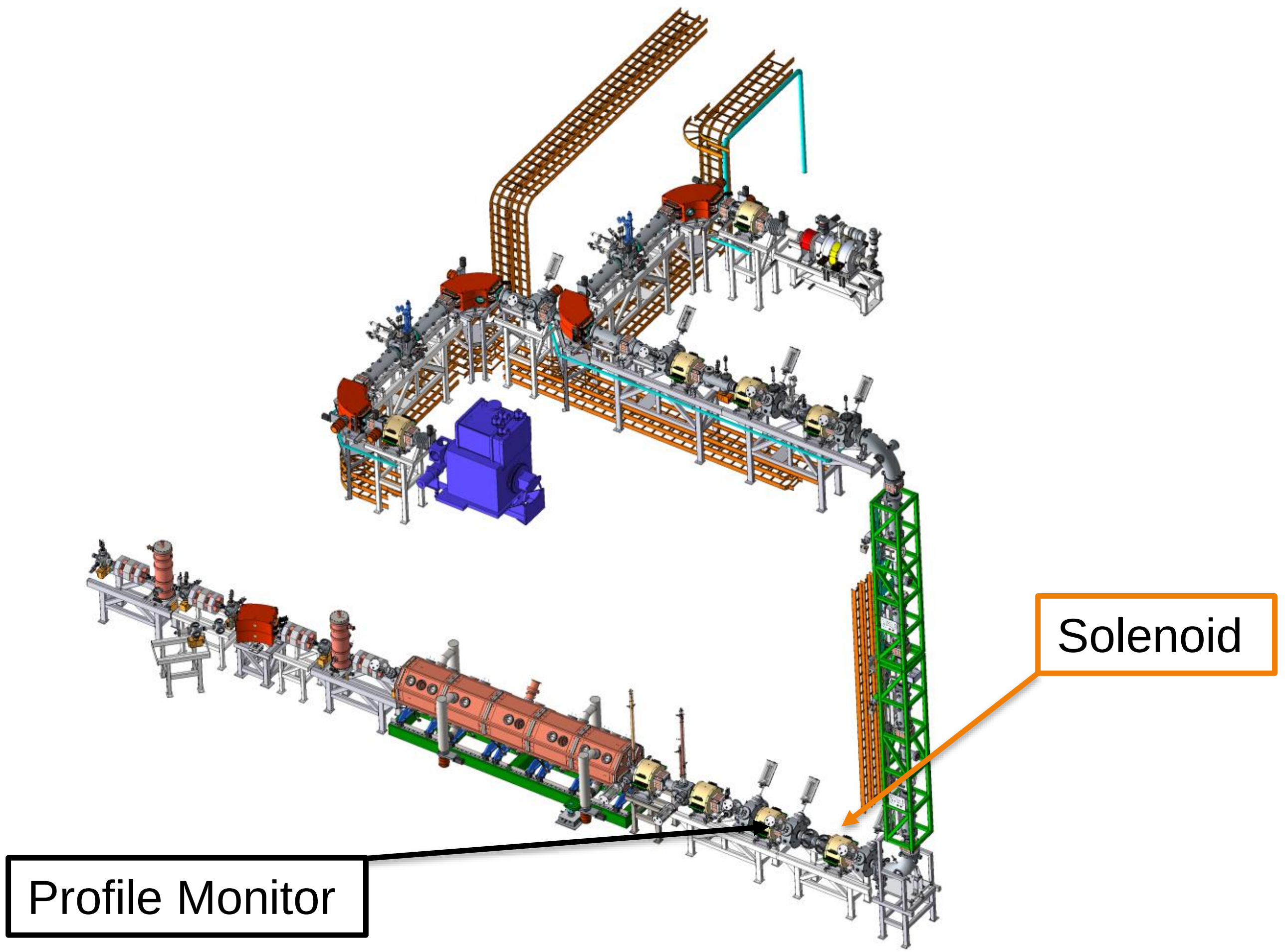
Parameters in Model:

- Beam parameters
- Element locations
- Machine settings
- E-/B- Field Profiles

Physics Model

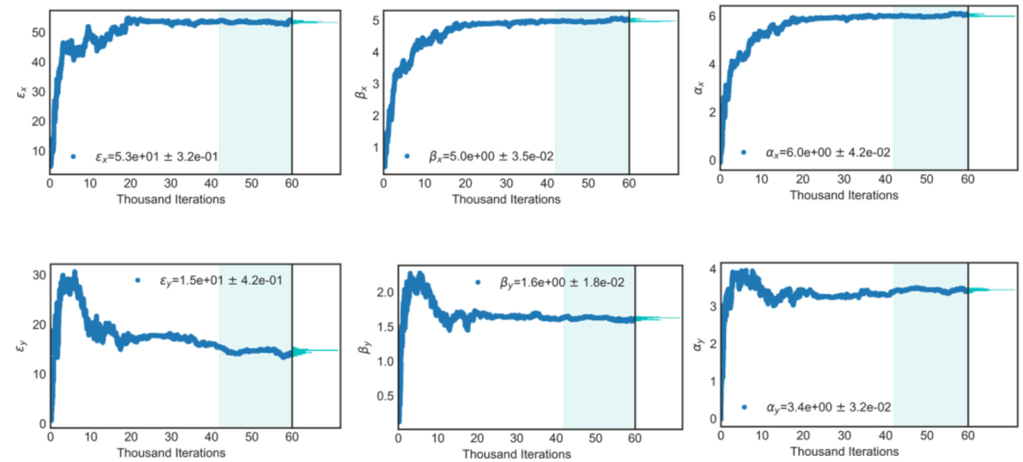
Data based Model

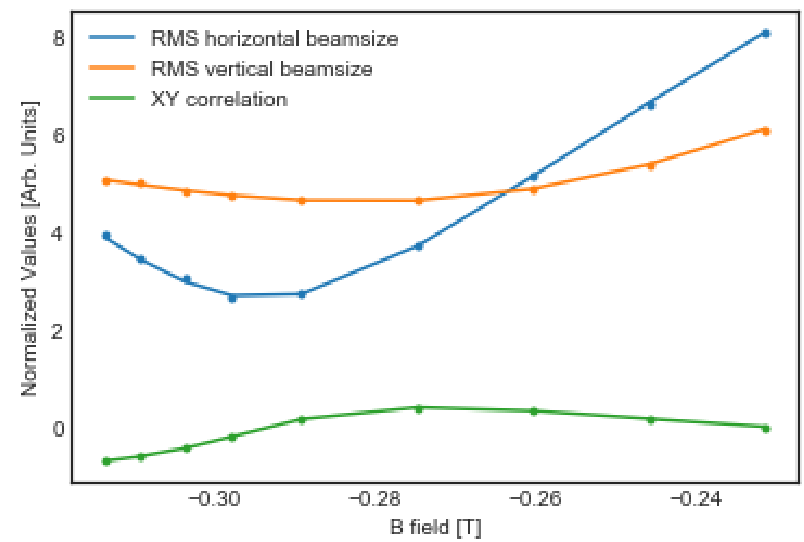
Solenoid Scan Example

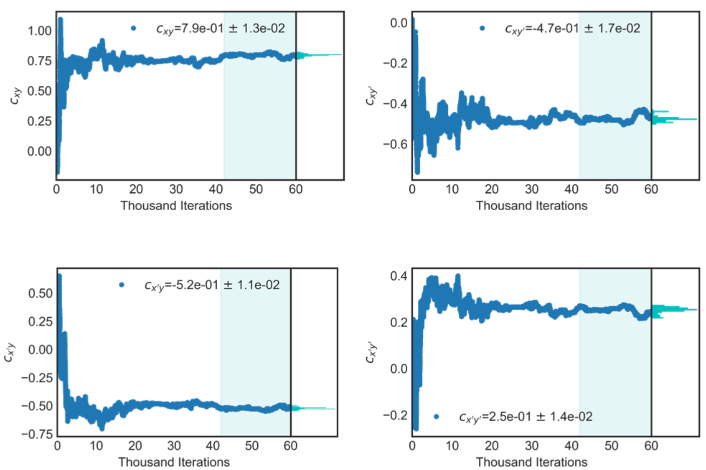


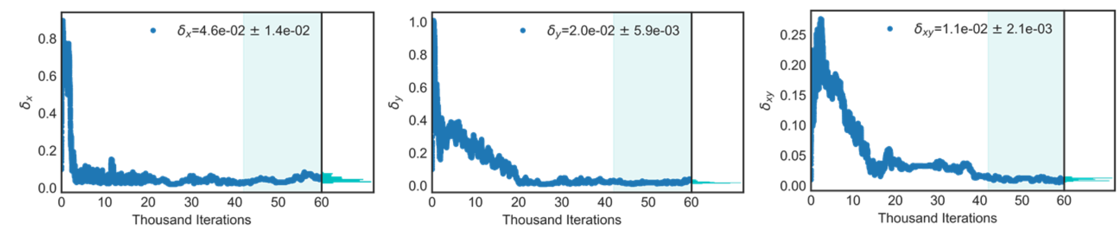
Solenoid Scan, Inference Results

Plot every 20 iterations, total 1.2M iterations







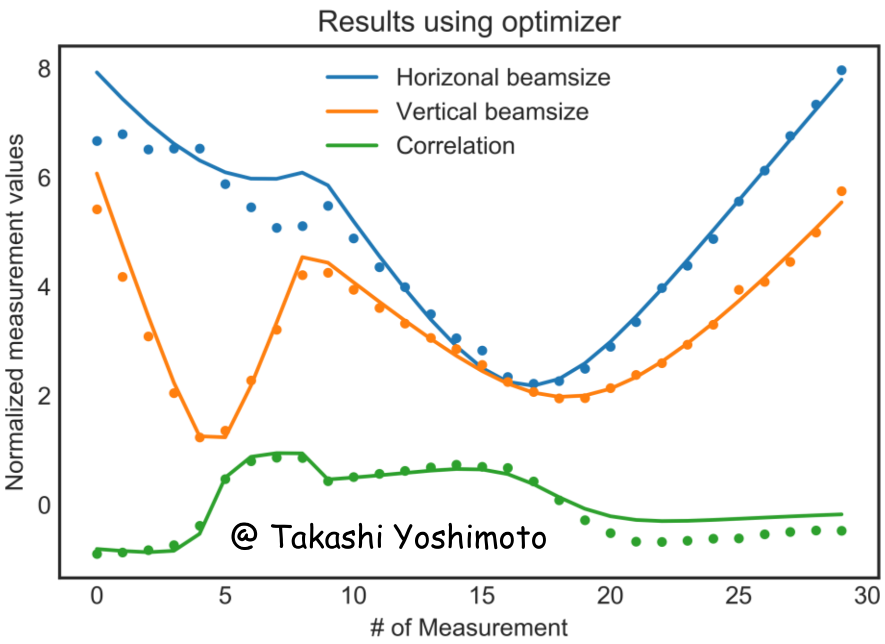


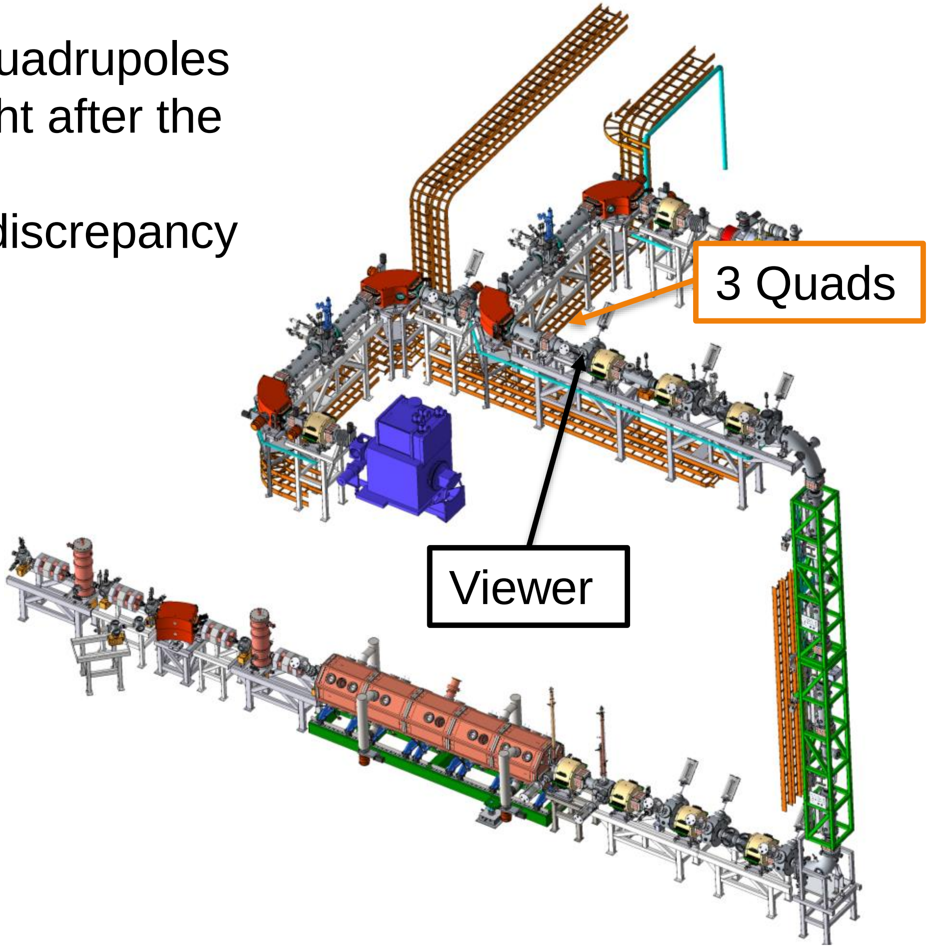
emit_x [mm-mrad]	β_x [m]	α_x	emit_y [mm-mrad]	β_y [m]	α_y	C_{xy}	$C_{xy'}$	$C_{x'y}$	$C_{x'y'}$
0.27	4.99	6.01	0.074	1.62	3.44	0.79	-0.47	-0.52	0.25

Quadrupole Scan Example

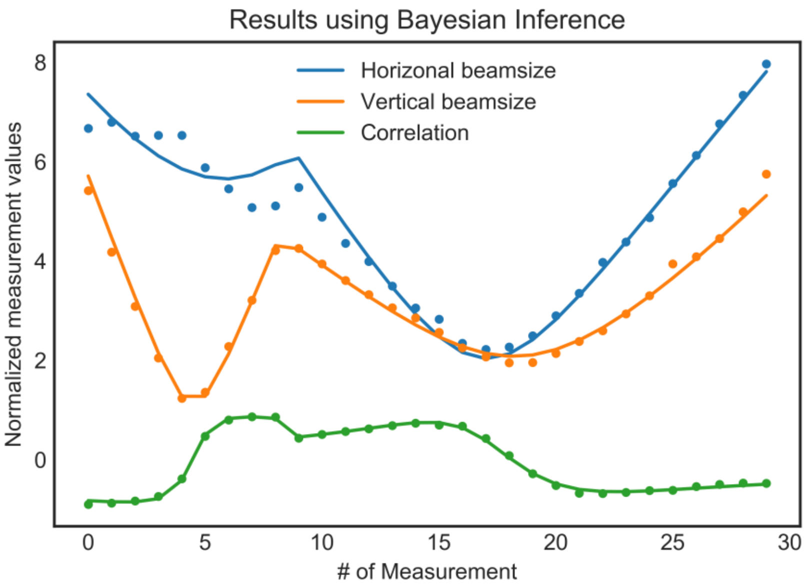
Measurement Scenario:

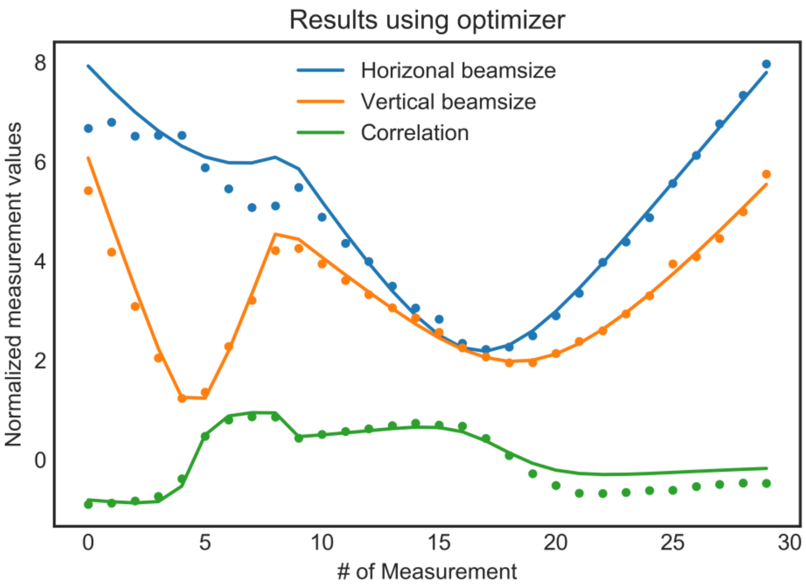
- Adjusting the voltage of the 3 upstream quadrupoles and measure the profile from a viewer right after the quadrupoles.
- Fitting routine found that there is always discrepancy between the model and measurement.





Quad Scan, Result Comparison





total 500K iterations

	emit_x [mm-mrad]	β_x [m]	α_x	emit_y [mm-mrad]	β_y [m]	α_y	C_{xy}	$C_{xy'}$	$C_{x'y}$	$C_{x'y'}$
Bayesian	0.104	2.76	-2.38	0.049	1.37	-0.31	0.26	0.63	-0.08	0.51
Optimizer	0.119	2.76	-2.26	0.05	1.51	-0.42	0.34	0.26	-0.06	0.01

Future Work based on Model Inference

- We have to go far beyond the simple examples to make it useful in FRIB.

- Involving larger amount of data and higher dimension.
  - If we take more measurement at different location, the result solenoid scan can be very different and does not have the suspicious behavior.
  - Can we get the indication of 'more measurement needed' from inference procedure?
    - Can we automate the measurement and design expert 'agents'
  - It seems that the convergence is harder to achieve in this example. An indicator?

- Deal with the difference between the model and machine
  - We infer the best parameter of the linear model
  - Model can not handle all situations by simply adjusting parameter.
  - The difference will modelled by Gaussian Process.
  - Comparison of 'Gaussian Process only' and 'Model + Gaussian Process'

Data Driven Models

- Beyond the particle motion, described by Maxwell Equation, we don't have very precise/quick models to predict the data.

- Modern accelerators with high speed interlink/network usually store gigantic amount of data of very high number of dimensions.

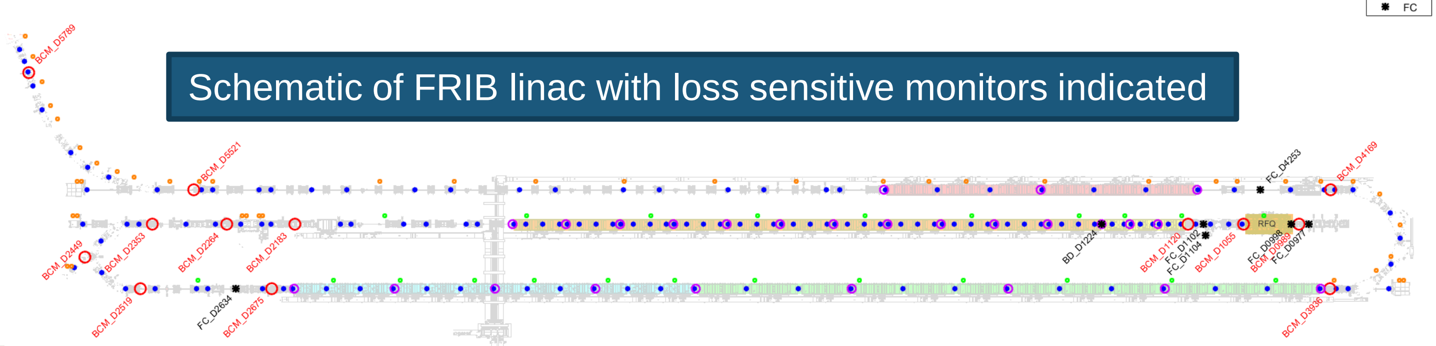
- The performance of the accelerator, on the contrary, are evaluated with small amount of parameters
  - Beam energy and its stability
  - Beam power and its stability
  - Radiation and beam loss control
  - Machine availability
- We plan to study the data driven model to predict the machine behavior and increase the performance.

Application 2: Identifying Sources of Beam Loss

- Beam Loss Monitors form a network to monitor beamlines for particle interception
- The BLM network has functions: *fault detection* and *fault diagnosis*
- *Fault detection* requires BLMs:
  - Fast response for acute losses
  - Location near sensitive components
  - Located at ‘critical positions’ to trigger Machine Protection response
- *Fault diagnosis* requires BLMs:
  - Sensitive to diagnose chronic losses and beam tuning errors
  - Ability to differentiate between controlled and uncontrolled losses
  - Located at ‘discrimination points’ to differentiate spatial loss patterns

We look to answer two questions:

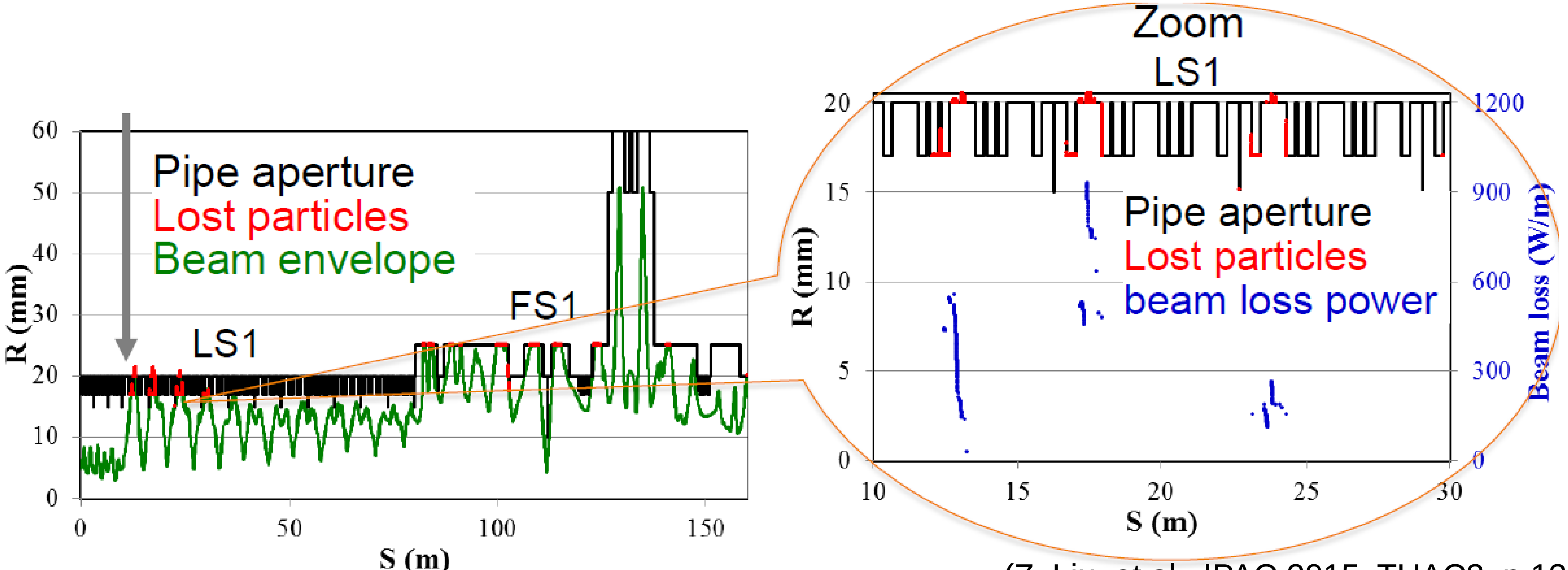
1. Do we have sufficient detectors in the correct places to catch all events?
2. Can we use pattern recognition with loss distribution to identify specific failures?



Schematic of FRIB linac with loss sensitive monitors indicated

Determining Critical Positions (CPs)

- Study correlations
  - Define critical positions by quantifying BLM positions against specific fault modes resulting in losses
- IMPACT simulations were generated in each case
  - 332 cavity faults generate 241 loss distributions
  - 69 solenoids each generate losses
  - Each accelerator element is a loss monitor (as far as IMPACT is concerned) – 572 elements



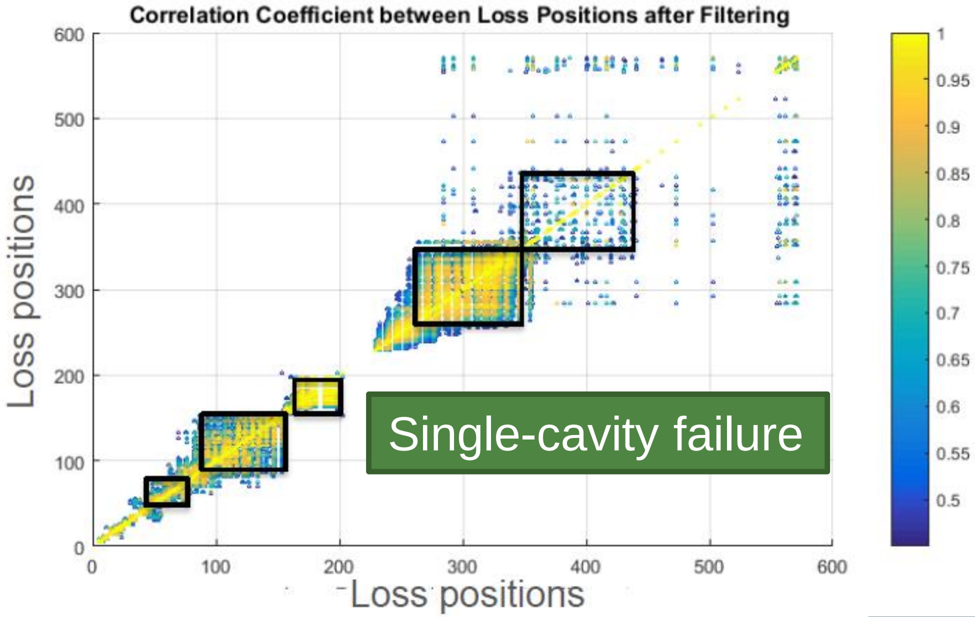
(Z. Liu, et al., IPAC 2015, TUAC3, p.1356)

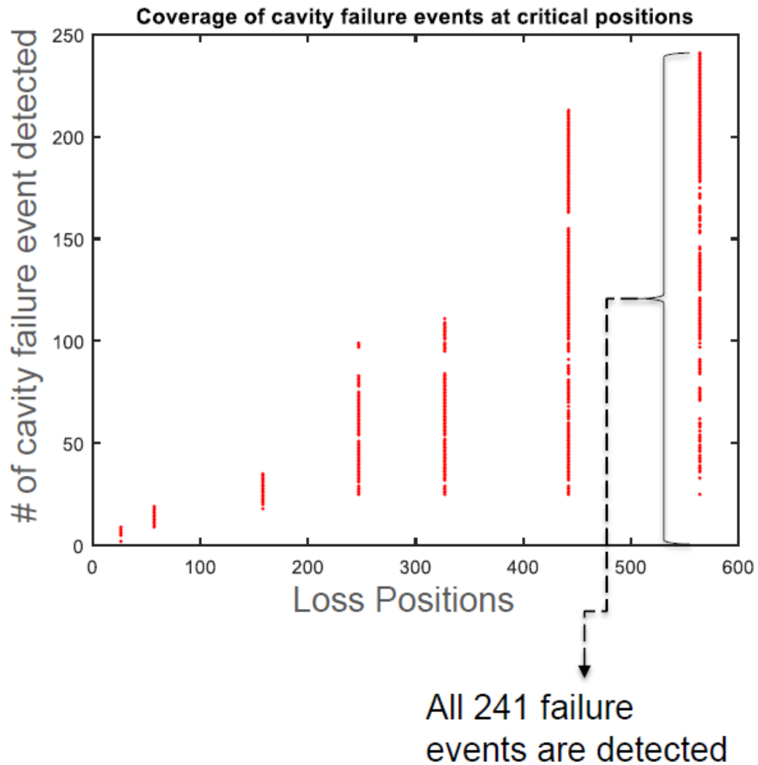
Use correlations to identify points of high sensitivity

Reduces required number of BLMs

- Calculate correlation. Filter for abs(R) > 0.45

$$R(i,j) = \frac{Cov(X_i,X_j)}{\sigma(X_i)\cdot\sigma(X_j)}$$





We can identify localized loss areas (outlined regions) that reduce required number of monitors (for this failure mode).

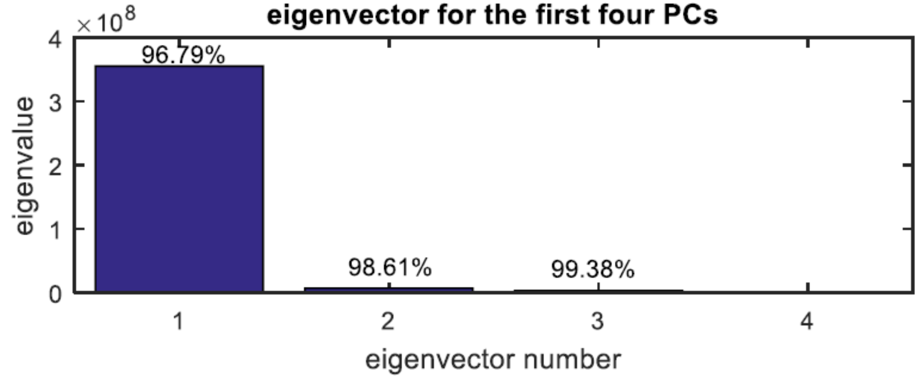
Analyze events to determine if Critical Position network observes all events.

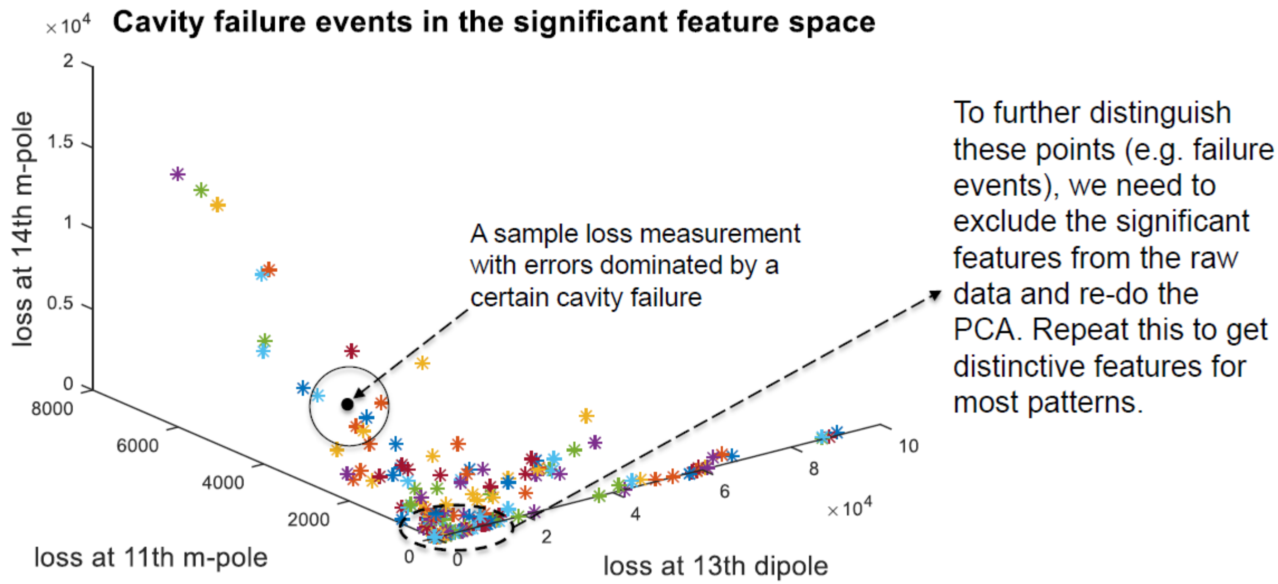
⟶ Network completeness with ~7 monitors!

Use PCA to find Discrimination Points

Maximizes variation between patterns

- PCA finds linear combinations of elements to diagonalize the space, remove degeneracies
- Use PCA on single-cavity failure loss matrix, $X_{572x241}$
  - Dramatic reduction in dimensionality 572 -> 1,2 !
- Pattern recognition proble
  - Finds ‘significant features’ for each fault mode
  - Differentiates ‘distinctive features’ between fault modes





Future efforts

- Accelerator health monitoring
  - LLRF, HPRF, cavity tuners
  - Diagnostic and instrumentation
  - Li stripper
  - High power target
- Beamline tuning
  - Ion sources
  - Fragment separator optics
  - Multi-user configurations
- Use data driven model to incorporate best understanding with measurement data

Areas for Collaboration

- JLAB/ODU effort on predicting SRF cavity trips or failures is directly applicable to FRIB reliability and uptime

- SLAC has developed tools for longitudinal phase space prediction
  - Work with authors to import and develop toolkits for NP linacs, EIC needs

Summary

- Work at FRIB focuses on several areas and techniques
  - Bayesian optimization is used to match model with measurements
  - PCA has been used for pattern recognition in loss measurements

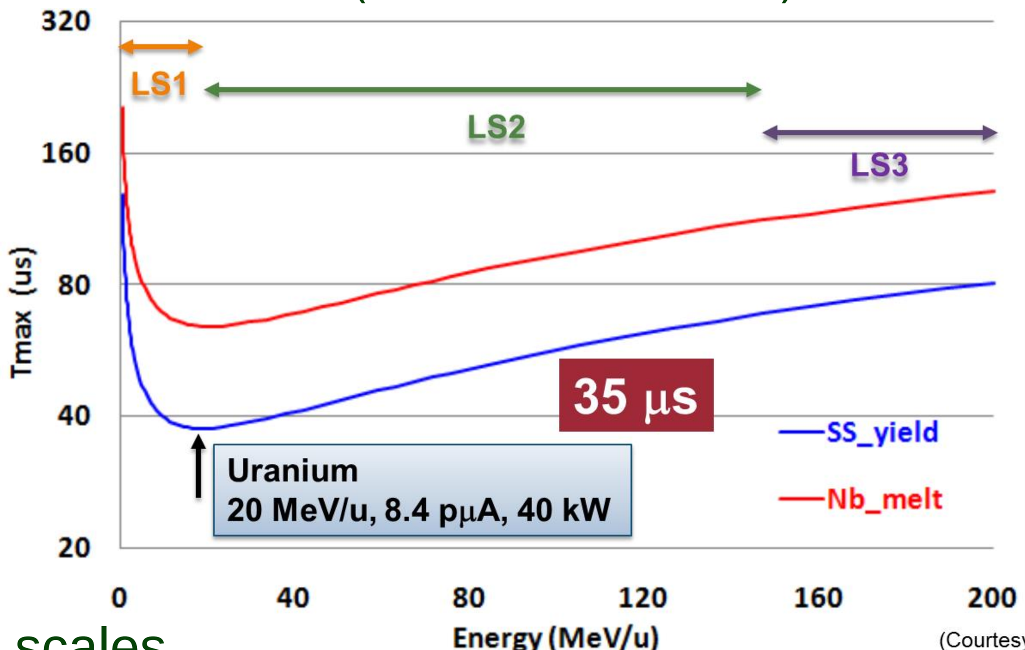
- We look to expand with ANN, Random Forest, and other methodologies
  - Assess for specific problems
  - Help develop toolkits for community use

- We are learning from the community and invite collaborative efforts

Thank you to Manouchehr and today's participants

Need for Robust Machine Protection Systems

- Protect against catastrophic fast beam losses
  - Melting threshold ~40 usec
  - Fast mitigation requirement 35 usec or faster
- Detect and mitigate chronic, low level losses ("1 W/m" standard)
  - Prolong SRF cavity lifespan
  - Reduce activation of beamline components
- Many sources of faults
  - LLRF
  - Timing, Chopper
  - Beam steering and focusing
  - Charge stripping and selection
  - Target failure
- MPS detection schemes multilayered in mode and time scales

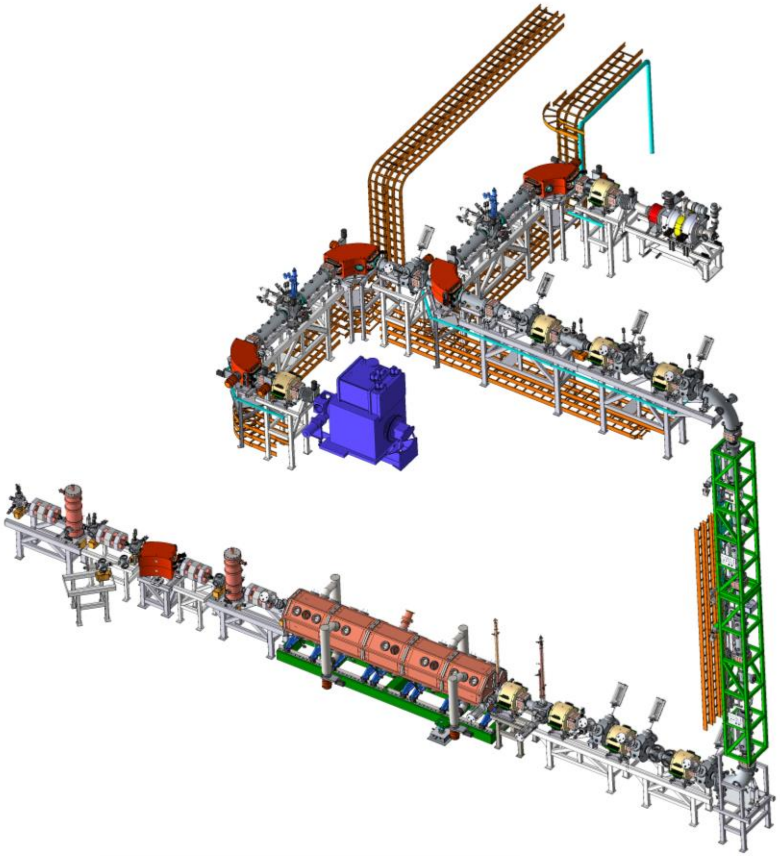


(Courtesy, Y. Zhang)

The worst case (uranium beam ~20MeV/u): may damage a SS bellow in less than 40 μs

Test of Applying Bayesian Inference

- Motivation:
  - Some parameters in accelerator can not be measured directly
  - These parameter can be used in accelerator model to predict the machine
  - This can be done with fitting the model to the measurement data using optimizers.
    - » Only results are given, how reliable?
    - » Prone to be have local minimum problem
    - » Hard to scale to high dimensional problem
    - » Depend on the definition of the penalty function



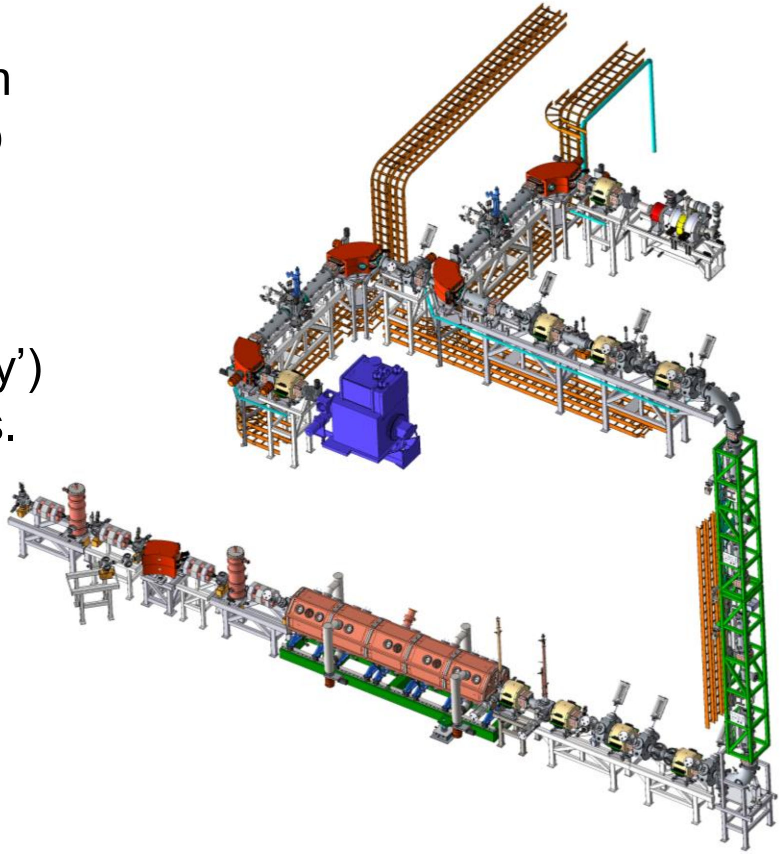
FRIB Front-End

- Expectations from Bayesian method
  - Provide statistics information on reliability
  - Better scaling to high dimensional problem
  - Less local minimum problem
  - Suggest the future experiment

The Problem to Solve

The accelerator model starts at the exit of the ion source. The description of initial beam is hard to measure directly.

The linear accelerator model need the following matrix to describe the beam property in (x, x', y, y') phase space, which contains 10 free parameters.



$$\begin{pmatrix} \langle x^2 \rangle & \langle xx' \rangle & \langle xy \rangle & \langle xy' \rangle \\ \langle x'x \rangle & \langle x'^2 \rangle & \langle x'y \rangle & \langle x'y' \rangle \\ \langle yx \rangle & \langle yx' \rangle & \langle y^2 \rangle & \langle yy' \rangle \\ \langle y'x \rangle & \langle y'x' \rangle & \langle y'y \rangle & \langle y'^2 \rangle \end{pmatrix}$$

In common accelerator language, we rewrite as:

$$\theta = (\epsilon_x, \beta_x, \alpha_x, \epsilon_y, \beta_y, \alpha_y, c_{xy}, c_{xy'}, c_{x'y}, c_{x'y'})$$

The Problem to Solve, cont'd

- We have to determine these 10 parameters using sets of measurements.
- The machine settings $V=(v_1, v_2, \ldots, v_n)$ are varied in the measurement and the beam profile (second order moments) are recorded by the profile monitor. In each machine setting, three observables are recorded:

$$\sigma = (\sigma_x, \sigma_y, \sigma_{xy})$$

Model

$$\sigma_{\mathrm{model}} = f(V, \theta)$$

Machine

$$\sigma_{\mathrm{measure}} = \sigma_{\mathrm{model}} + \delta\xi$$

$$\xi = \mathrm{randn}(); \delta = (\delta_x, \delta_y, \delta_{xy})$$

Bayesian Theorem

$$P\left(\theta, \delta \mid (\sigma_1, V_1), \cdots, (\sigma_i, V_i), \cdots\right) = \frac{P\left((\sigma_1, V_1), \cdots, (\sigma_i, V_i), \cdots \mid \theta, \delta\right) P(\theta, \delta)}{P((\sigma_1, V_1), \cdots, (\sigma_i, V_i), \cdots)}$$

The Problem to Solve, cont'd

The likelihood

$$P\left((\sigma_1, V_1), \cdots, (\sigma_i, V_i), \cdots \mid \theta, \delta\right) = \prod_i P\left((\sigma_i, V_i) \mid \theta, \delta\right)$$

$$\sim \prod_i \frac{1}{\delta_x \delta_y \delta_{xy}} \exp - \frac{(\sigma_{\text{measure}} - \sigma_{\text{model}})^2}{2\delta^2}$$

The prior

Uniform / Gaussian / beta prime distributions

We implement the Metropolis-Hasting method in Python to sample the posterior distribution.

The in-house linear model, FLAME, as the accelerator model. The linear model is fast enough to get converge result in ~10^3 seconds on laptop.