Modeling and Self-evaluation for Accelerator Control and Performance

Kevin Brown
Collider-Accelerator Department, BNL
brownk@bnl.gov

MODSIM 2022 August 12, 2022

Brookhaven National Laboratory
U.S. DEPARTMENT OF ENERGY

@BrookhavenLab

Outline

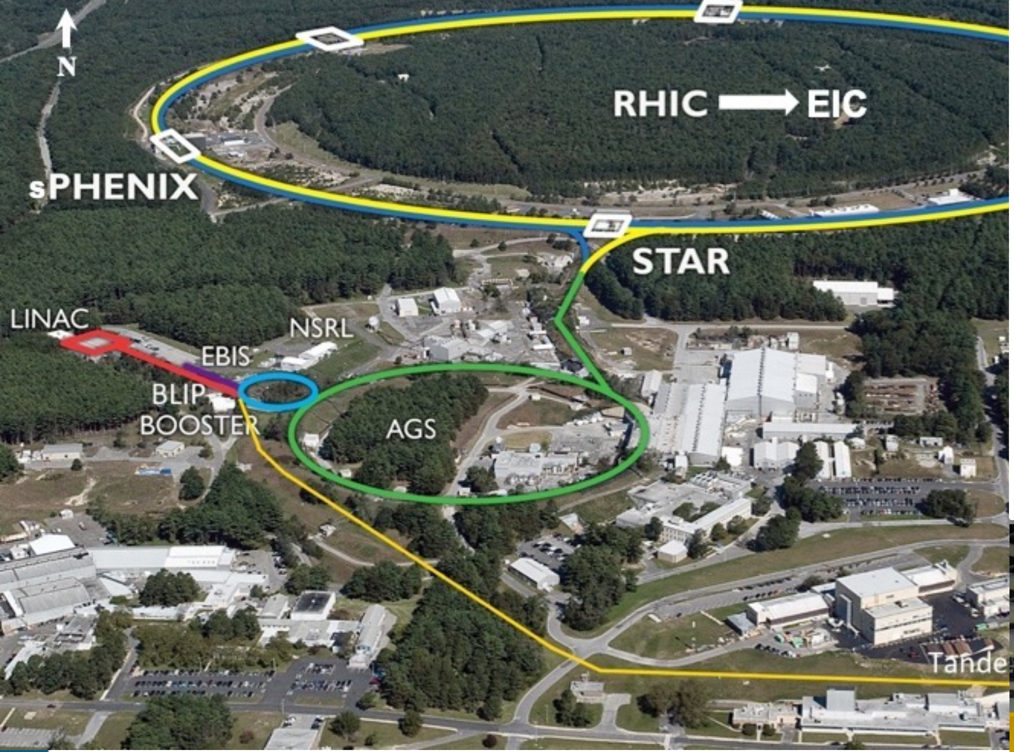
- Quick review of accelerators
- How accelerators are modelled (physics models)
- Methods used to validate models
- Self-evaluation in real-time
- How AI/ML will help
- Digital twins of accelerators
- Summary

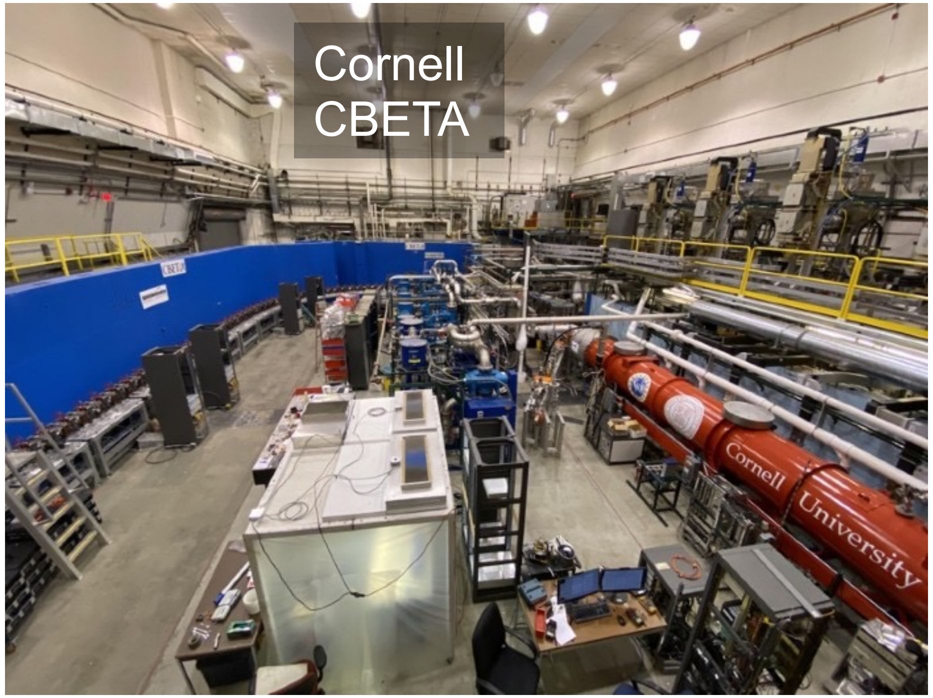
Brookhaven National Laboratory

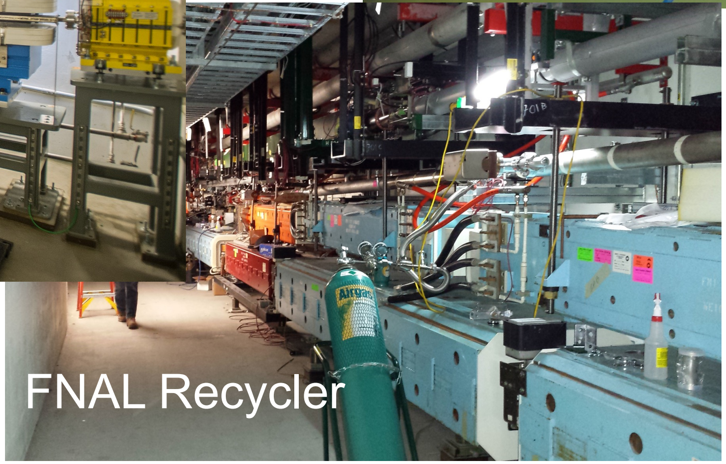
Quick review

Particle Accelerators

- >10,000 accelerators in use around the world; for industry, medicine, applied research, and basic research
- Basically, an accelerator is composed of the following:
  - A source or particles (electrons, protons, ions, etc.)
  - Stages of acceleration
    - Examples are Tandem van de Graaff's, Cockcroft-Walton generator, standing wave radio-frequency cavities (i.e., Wideröe's RF linear accelerator), traveling wave cavities, etc.
  - Elements to confine and control the trajectory of the particles
- Beyond that, there are many variations and specialized elements;
  - Higher order corrections (multipole magnets), helical dipoles (i.e., Siberian Snakes), superconducting magnets and RF cavities, insertion devices (e.g., wigglers and undulators), fast kickers, stripline kickers, etc.







Brookhaven National Laboratory

Some basic principles: cyclotron motion

To accelerate and control the trajectory of particles, they need either a charge or a magnetic moment (i.e., for molecules).

The Lorentz force is $\vec{F} = q\vec{E} + q\vec{v}\times\vec{B}$.

A cyclotron keeps a fixed B-field, and every turn adds energy to the charged particles by applying an electric field. The particles spiral out to larger radii as the energy increases. The basic relation for such an accelerator is $r = \frac{mv}{qB}$ and the stability condition is set by the equation of motion,

$$\ddot{x} + \omega_r^2 x = 0,$$

where $w_r = \sqrt{1-n}\, v_\theta/\rho$,

where there is only stable motion when $n < 1$.

Brookhaven National Laboratory

Some basic principles: AGS

Alternating Gradient Synchrotrons = strong focusing

Equation of motion follows Hill's differential equation,

$$\frac{d^2x}{ds^2} + K_x(s)x = 0, \qquad K_x \equiv \frac{B'}{B\rho} + \rho^{-2}, \qquad B' \equiv \partial B_y/\partial x$$

The field index is given by $n = -\left(\frac{r}{B}\right)\frac{\partial B}{\partial r}$, where $0 < n < 1$

For a synchrotron with circumference, C, $K_x(s + C) = K_x(s)$

Given N identical sections (unit cells), then we can say,

$$K(s + L) = K(s); \quad L = C/N$$

The solution of any linear second order differential equation such as Hill's equation, whether K is or isn't periodic, is uniquely determined by the initial values of x and its derivative, x'.

Some basic principles: AGS Cont.

Therefore, we can say

$$X(s) = \begin{bmatrix} X(s) \\ X'(s) \end{bmatrix} = M(s|s_0)X(s_0)$$

The determinant of M is unity, as Hill's equation does not contain first derivative terms. In the case when K is constant,

$$M(s_0|s) = \begin{bmatrix} \cos\varphi & K^{-1/2}\sin\varphi \\ (-K)^{1/2}\sin\varphi & \cos\varphi \end{bmatrix},$$

Where $\varphi = K^{1/2}(s - s_0)$. It is then an eigenvalue problem.

Here I skip a lot of interesting stuff …

The condition that the *Det M = 1* leads to $\beta\gamma - \alpha^2$=1, where the eigenvalue solutions have been reparametrized to define a simple symplectic condition. We can then write $M = I\ \cos\mu + J\ \sin\mu$, where I is the unit matrix and

$$J = \begin{bmatrix} \alpha & \beta \\ -\gamma & -\alpha \end{bmatrix} \ where\ J^2 = -I$$

Where is this leading us?

The particle motion from a point s_1 to a point s_2 in the lattice can be described by

$$\begin{pmatrix} x \\ x' \end{pmatrix}_{s_2} = M(s_1 \rightarrow s_2) \begin{pmatrix} x \\ x' \end{pmatrix}_{s_1}.$$

In terms of our lattice functions, (α, β, γ), for a phase advance between s_1 and s_2 of μ,

$$M(s_1 \rightarrow s_2) = \begin{pmatrix} \left(\frac{\beta_2}{\beta_1}\right)^{1/2} (\cos\mu + \alpha_1 \sin\mu) & (\beta_1\beta_2)^{1/2} \sin\mu \\ -\frac{1 + \alpha_1\alpha_2}{(\beta_1\beta_2)^{1/2}} \sin\mu & \left(\frac{\beta_2}{\beta_1}\right)^{1/2} (\cos\mu - \alpha_2 \sin\mu) \end{pmatrix}$$

Accelerator Simulations

Accelerator models

There are many physics codes for simulating accelerator beam dynamics. With >10,000 accelerators and well over 70 years of history, there are many codes. The ones most interesting to large accelerators (such as RHIC, EIC, and NSLS II) are;

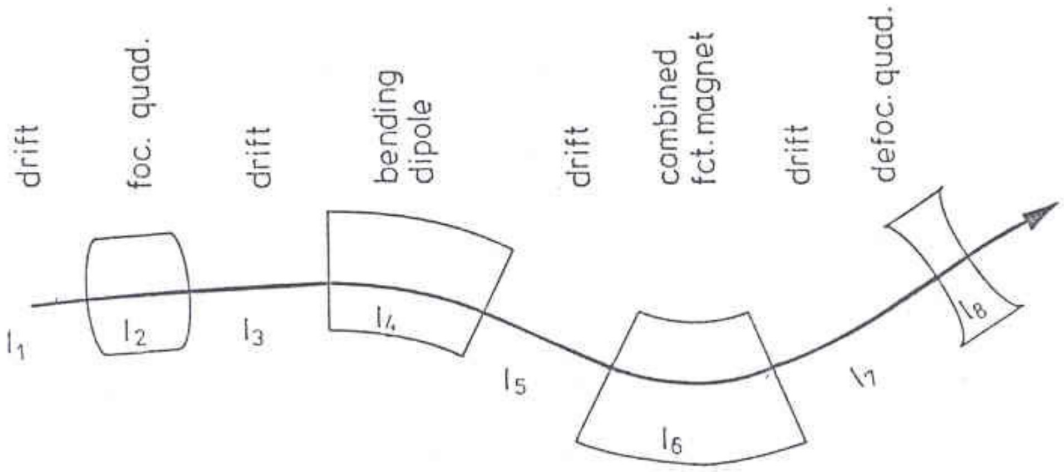
Name of code	What it does
Madx (CERN)	Single particle dynamics, expansions use Taylor maps, linear to 2^{nd} order, does handle higher order, can be made symplectic
Bmad (Cornell)	Single particle dynamics, spin tracking, expansions use Taylor maps, any order, very comprehensive (modern code), symplectic, synchrotron radiation
SixTrack (CERN)	Integrated 6D tracking
Zgoubi (BNL)	Single particle dynamics, 6D, spin tracking, numerically integrates the Lorentz equation based on Taylor maps, naturally symplectic, any order, synchrotron radiation
elegant	Often used for electrons and light sources, similar to above codes plus dynamic aperture maps, scanning lattice parameters, multistage/time dependent

Brookhaven National Laboratory

In general, eigenvalue solvers

The codes use various methods, including Lie algebra, differential algebra, polymorphic tracking, symplectic tracking, and more. Often codes are developed to run on HPC systems.

Electric, magnetic, and other elements (drift spaces, complex E-M fields, etc.) are simulated with matrix/tensor representations, typically in curvilinear coordinate systems (and others). Often elements are sliced to provide higher axial resolution.



$$M_T = M_8 M_7 M_6 M_5 M_4 M_3 M_2 M_1$$

Validation

Comparing Simulations to Reality

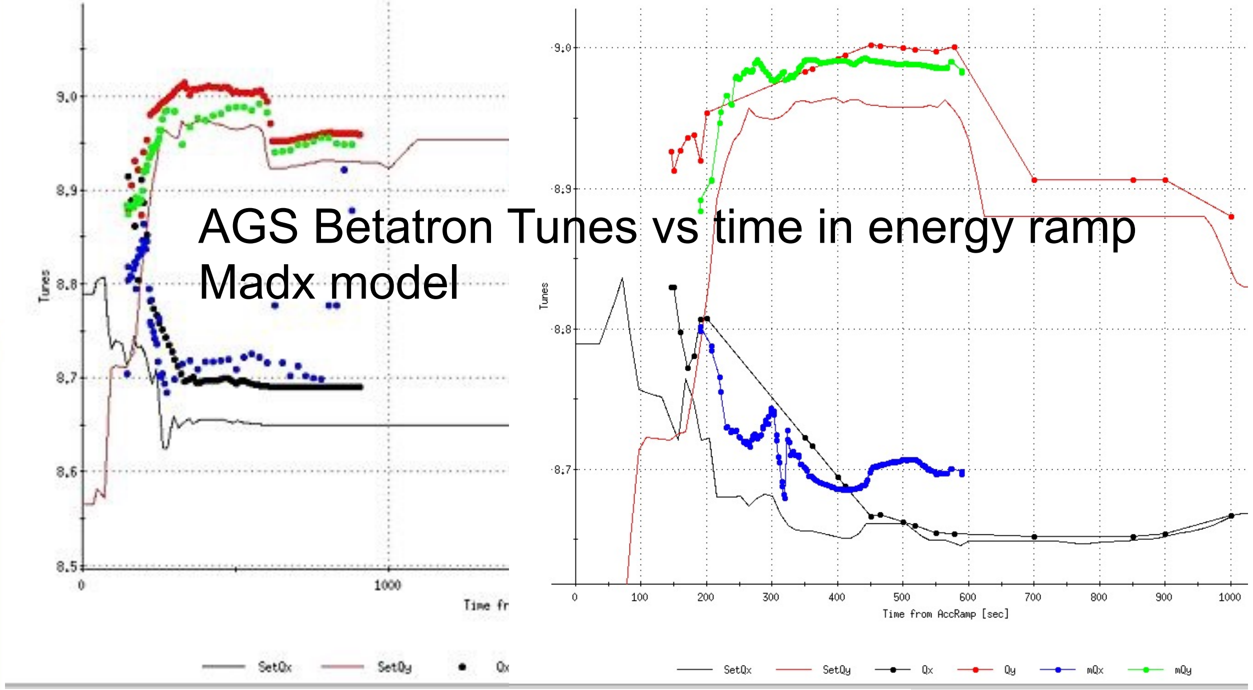
Measurable element parameters:

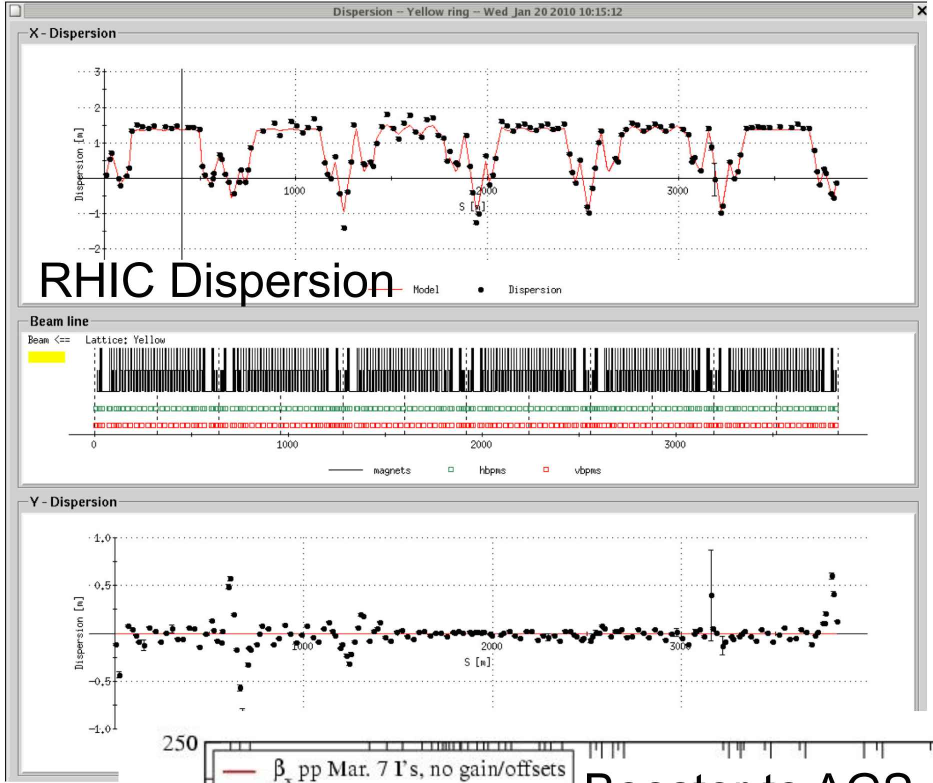
currents in magnets, timing for ramped elements, in some cases direct or indirect field measurements, cavity voltages, vacuum levels, etc.

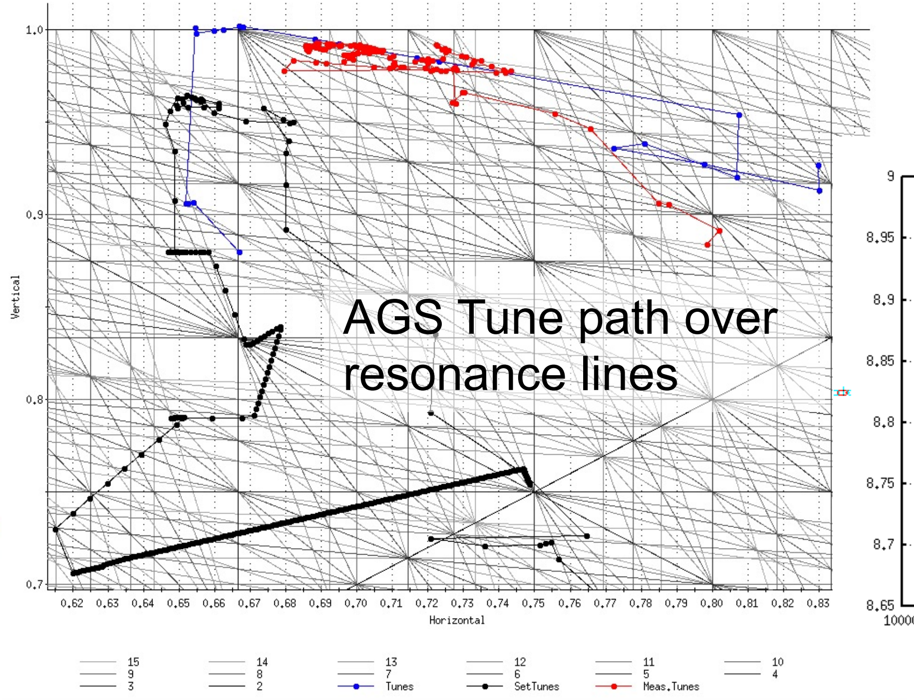
Measurable beam parameters:

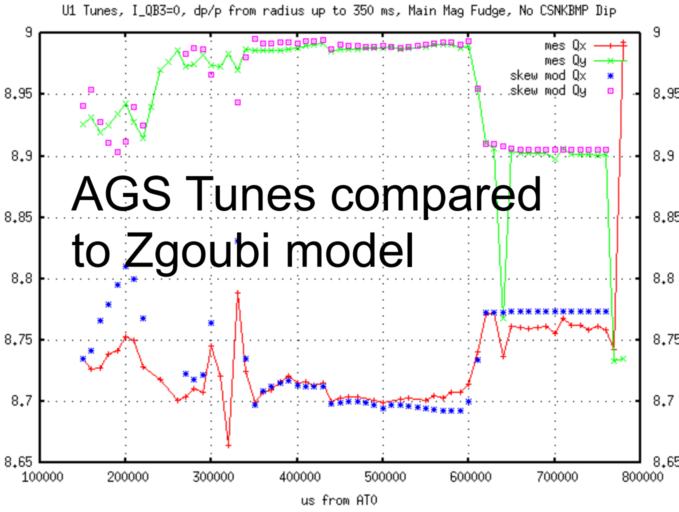
Directly: time averaged beam positions (orbit), time averaged beam size, betatron tunes, beam frequency, beam current, longitudinal profile

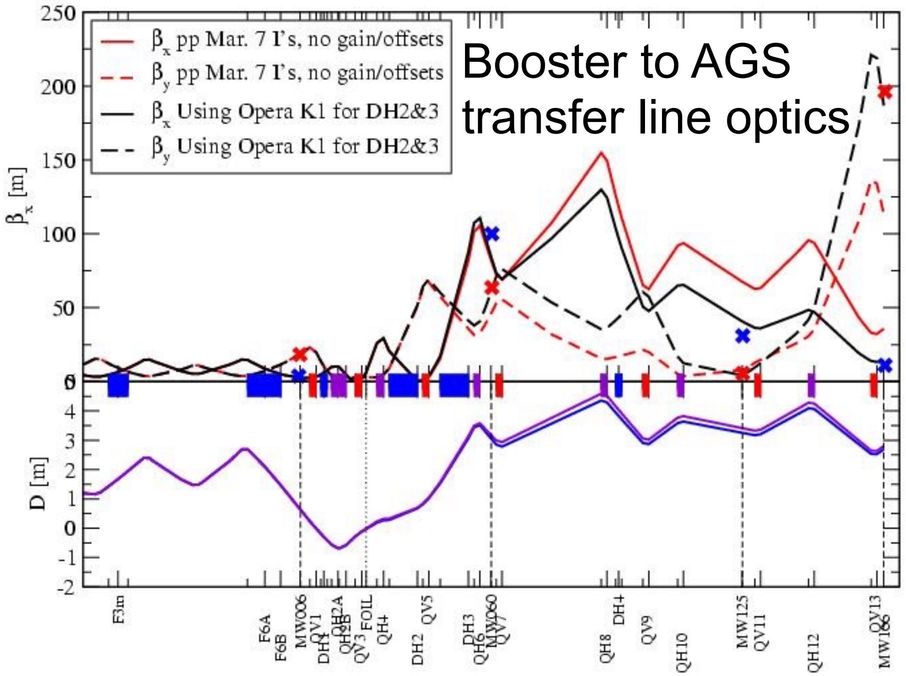
Indirectly: chromaticity (tune dependence on momentum), dispersion (orbit dependence on momentum), sometimes have turn-by-turn BPMs, Schottky analysis (tunes, chroms, coupling), coupling, resonance strengths, etc.


AGS Betatron Tunes vs time in energy ramp
Madx model
Tunes
Time fr
Time from AccRamp [sec]
SetQx
SetQy
Qx
Qy
mQx
mQy

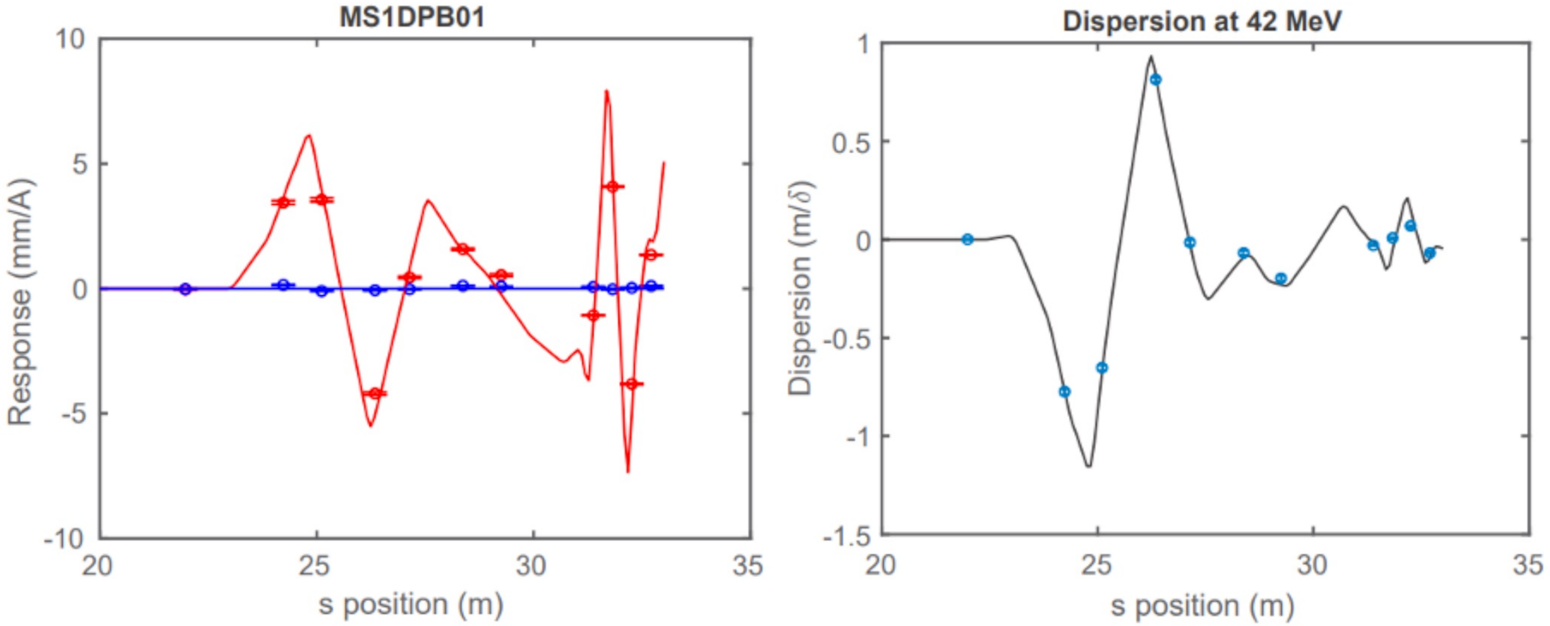

Dispersion -- Yellow ring -- Wed Jan 20 2010 10:15:12
X - Dispersion
RHIC Dispersion
Model
Dispersion
Beam line
Lattice: Yellow
magnets
hbpms
vbpms
Y - Dispersion
S [m]


AGS Tune path over
resonance lines
Vertical
Horizontal
Tunes
SetTunes
Meas.Tunes

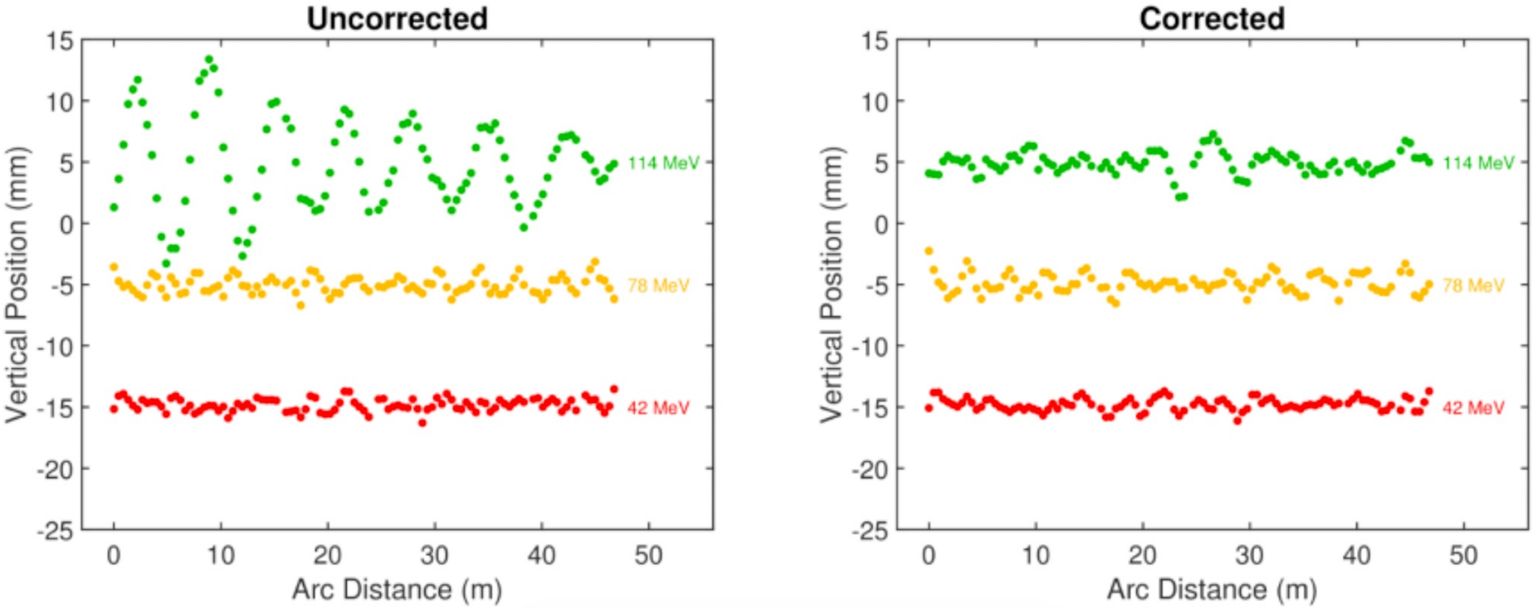

U1 Tunes, I_QB3=0, dp/p from radius up to 350 ms, Main Mag Fudge, No CSNKBMP Dip
AGS Tunes compared
to Zgoubi model
mes Qx
mes Qy
skew mod Qx
skew mod Qy
us from AT0


Booster to AGS
transfer line optics
β_x pp Mar. 7 I's, no gain/offsets
β_y pp Mar. 7 I's, no gain/offsets
β_x Using Opera K1 for DH2&3
β_y Using Opera K1 for DH2&3
β_x [m]
D [m]

Cornell CBETA



Online display of measured data and VM prediction of (left) an orbit bump produced from a single dipole magnet, and (right) the dispersion as measured by a small change in beam energy.



Before (left) and after (right) applying a simultaneous orbit correction of the vertical orbit through the CBETA return loop, where there are up to 7 overlapping beams of 4 different energies.

Self-evaluation

Orbit Response

Imagine we add/sub bend angle, at some location k, to the one turn trajectory of the beam. Averaged over time, a particle will oscillate around some equilibrium trajectory to form a closed orbit. If this added angle is $\Delta\theta_k$, then for a linear system the position of the beam at a location m is

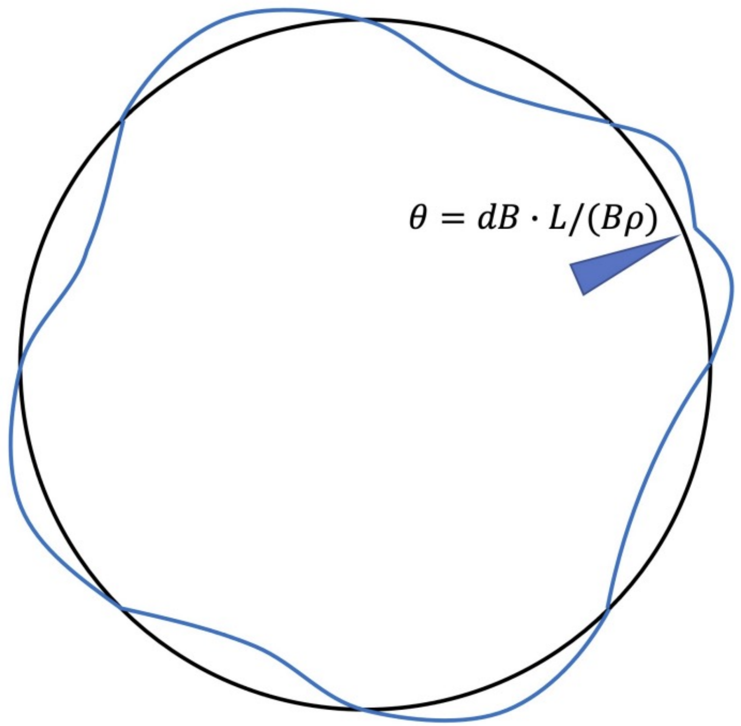
$$\Delta x_m = R_{km}\Delta\theta_k$$

R_{km} is the linear response due to this increased or decreased angle.

$$R_{km} = \frac{\sqrt{\beta_k\beta_m}}{2\sin\pi\nu}\cos(|\varphi_k - \varphi_m| - \pi\nu),$$

where we have introduced the betatron tune, which is the phase advance for one turn,

$$\nu = \frac{1}{2\pi}\int_0^C \frac{ds}{\beta}$$



Orbit Response Matrix (ORM)

- Mapping $\vec{R}$ between closed orbit measurements and corrector settings
- AGS Orbit measured at 72 pick-up electrodes (PUE), 6 in each super-period
- AGS has 48 horizontal and vertical corrector pairs, 4 in each super-period
- Linear orbit response to corrector change: calculate R matrix by changing each corrector pair separately
- Corrector current $I \rightarrow$ angle θ by conversion factor

$$\begin{pmatrix} \Delta\vec{x} \\ \Delta\vec{y} \end{pmatrix} = \vec{R} \begin{pmatrix} \Delta\vec{\theta}_x \\ \Delta\vec{\theta}_y \end{pmatrix}$$

$$\frac{\Delta x_i}{\Delta\theta_j} = R_{ij}$$

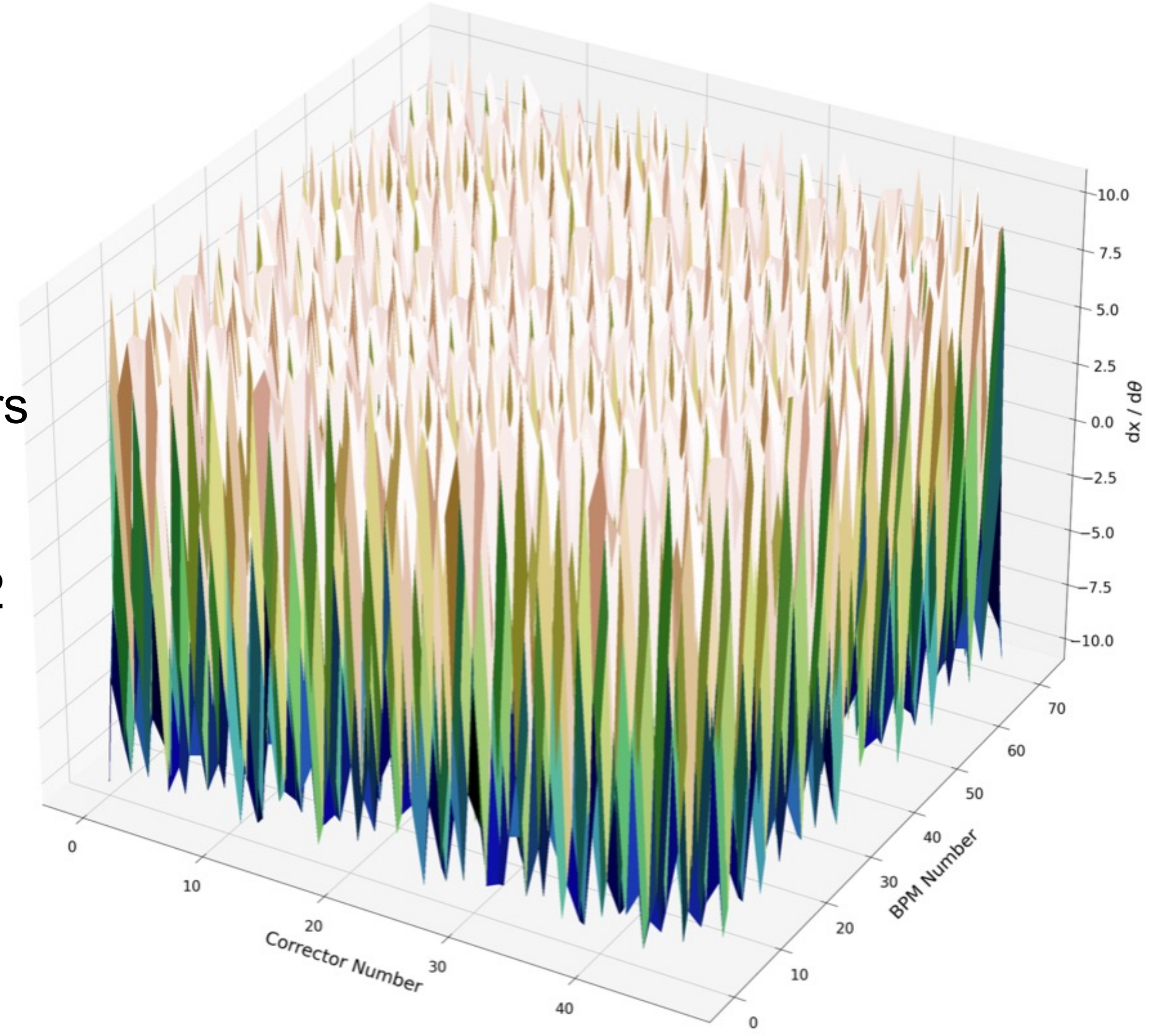
Reference $\overrightarrow{R_x}$ matrix

- Reference = bare machine, no error

Test Exercise
Treat a simple model of the AGS (bare) as the reference data.

Create a second model with 12 known errors by turning on 12 quadrupoles.

Test the process for reconstructing those 12 errors using an ORM.



Use ORM to identify machine errors

- Actual machine with errors (e.g. quadrupole gradient errors, corrector calibration errors, etc.) produce different $\vec{R}_{measured}$ from model/reference machine $\vec{R}_{model}$

$$\Delta R_{ij} = R_{ij}^{model} - R_{ij}^{measured}$$

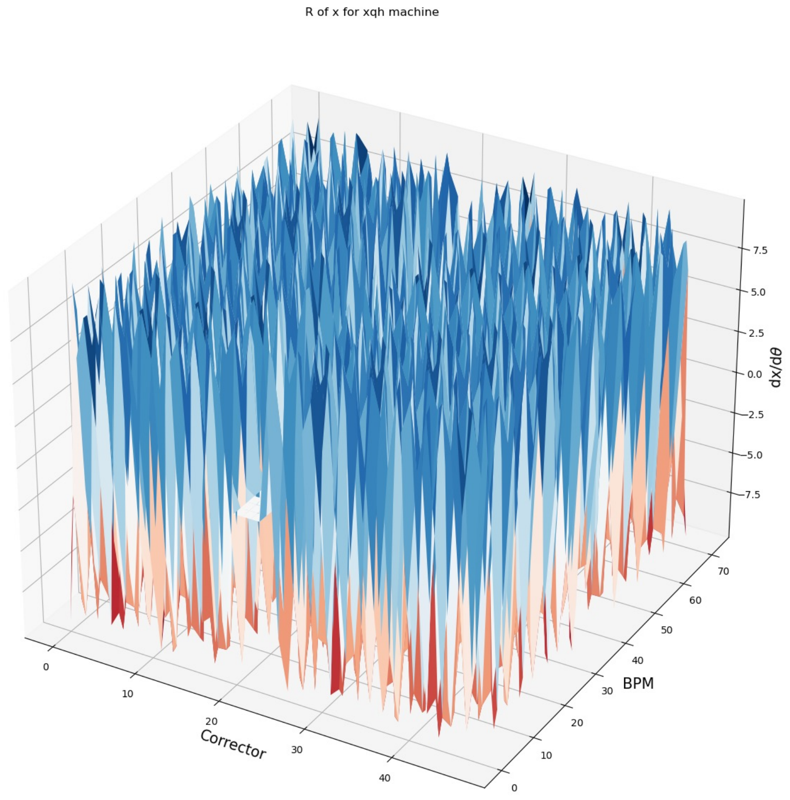
- Considering all possible sources of errors as a vector $\vec{\nu}$, build response error model $\vec{J}_{model}$

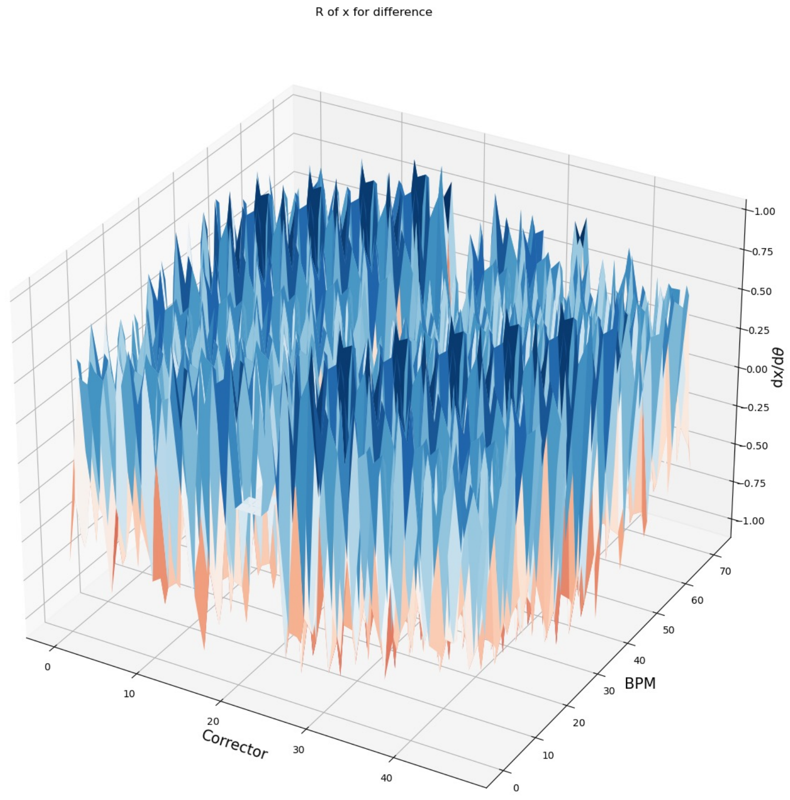
$$\begin{pmatrix} \Delta R_{11} \\ \Delta R_{12} \\ \cdots \\ \Delta R_{n(m-1)} \\ \Delta R_{nm} \end{pmatrix} = J_{model} \begin{pmatrix} \Delta \nu_1 \\ \Delta \nu_2 \\ \cdots \\ \Delta \nu_{N-1} \\ \Delta \nu_N \end{pmatrix}$$

- Reconstruct any $\vec{\nu}$ given known $\Delta\vec{R}$ and $\vec{J}_{model}$

R matrix with 12 quadrupole ‘errors’

$R_{diff} = R_{ref} - R_{err}$





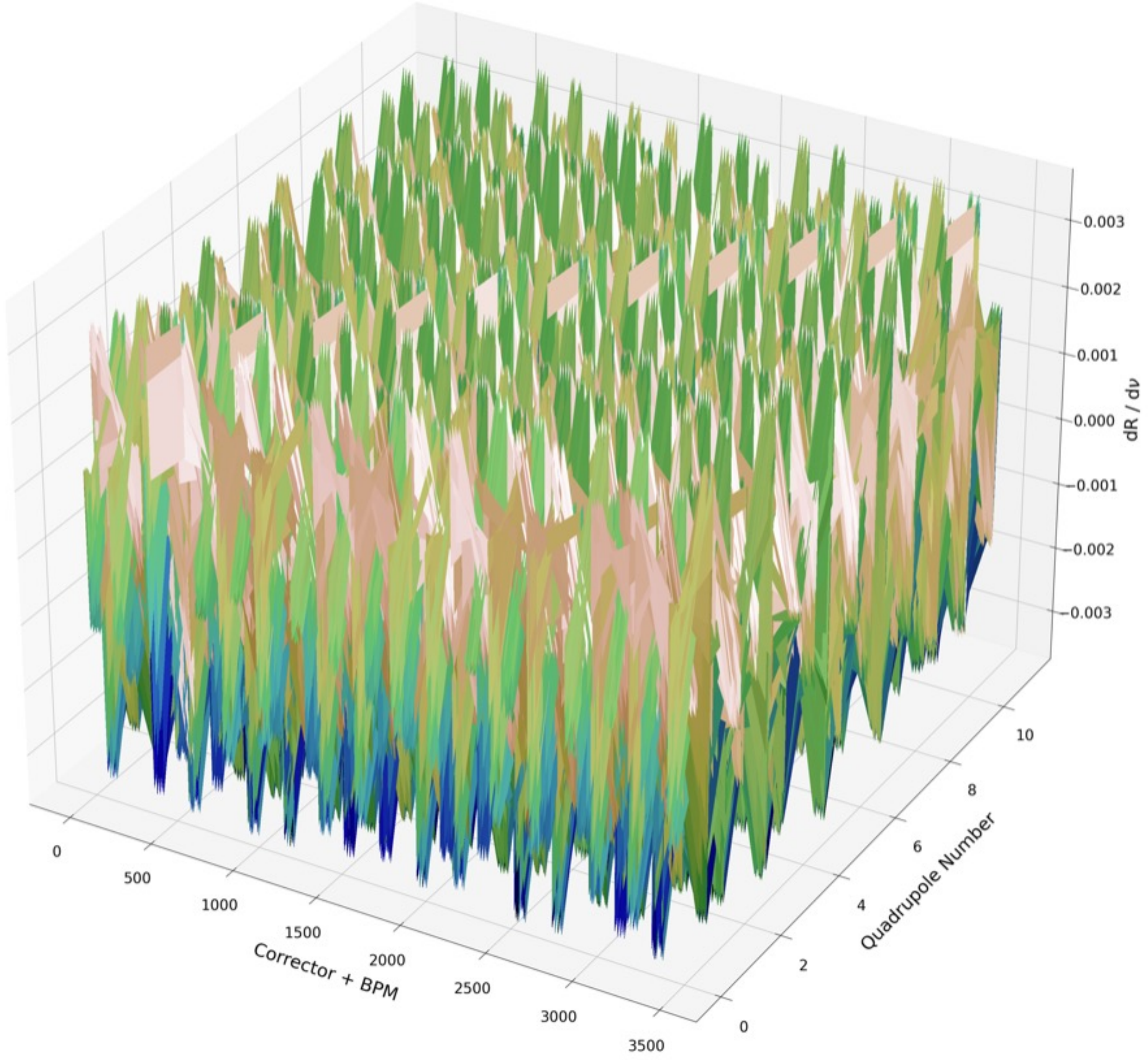
We will show a real example later, but the idea is the R_{err} is a real machine being measured and R_{ref} would be our best model of that machine.

Brookhaven National Laboratory

Test case $\vec{J}_{model}$ matrix (horizontal)

- Calculated using $\Delta v = 40$ for each quadrupole

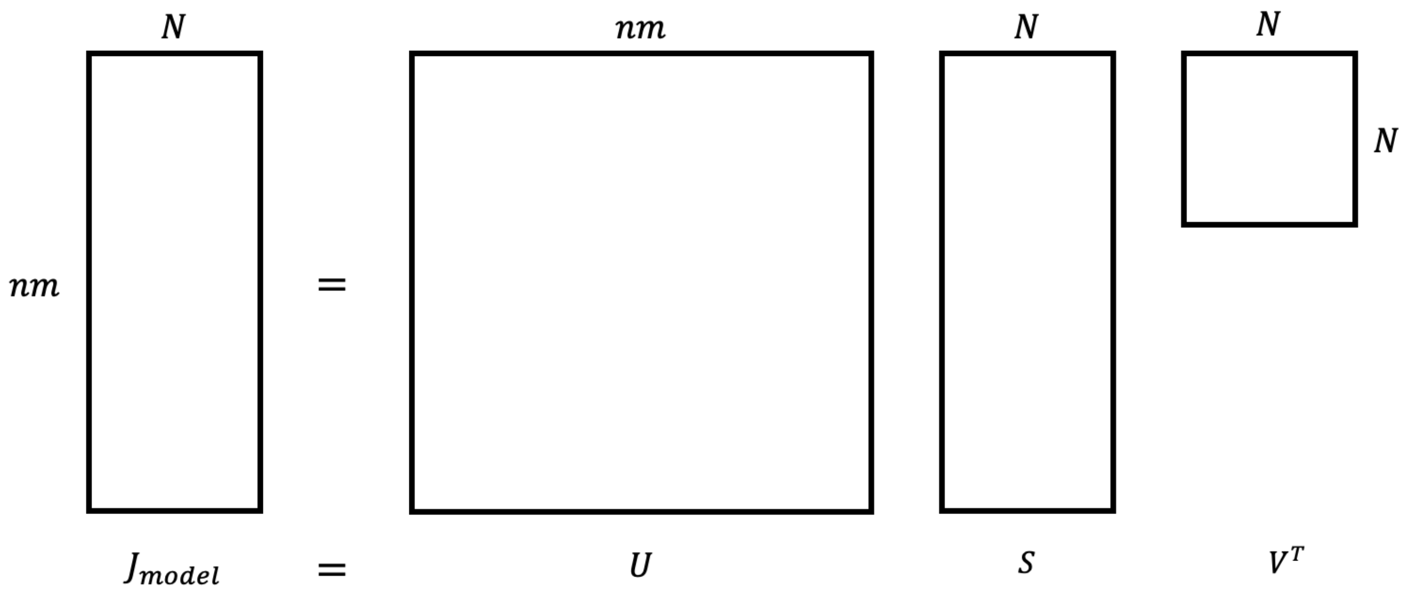
- Agreement with MAD-X model (redefined every quad individually) was obtained



Reconstruct errors using SVD

- Solve for $\Delta\vec{v}$ using $\Delta\vec{R} = \vec{J}_{model}\,\Delta\vec{v}$, where $\vec{J}_{model}$ is not a square matrix
- Perform singular value decomposition (SVD) on $\vec{J}_{model}$

$$J_{model} = USV^T$$



Reconstruct errors using SVD

- $\vec{U}$ and $\vec{V}$ are square orthogonal matrices: $UU^T = VV^T = I$
- $\vec{S}$ is an $nm \times N$ matrix whose first N diagonal elements are singular values σ of $\vec{J}_{model}$

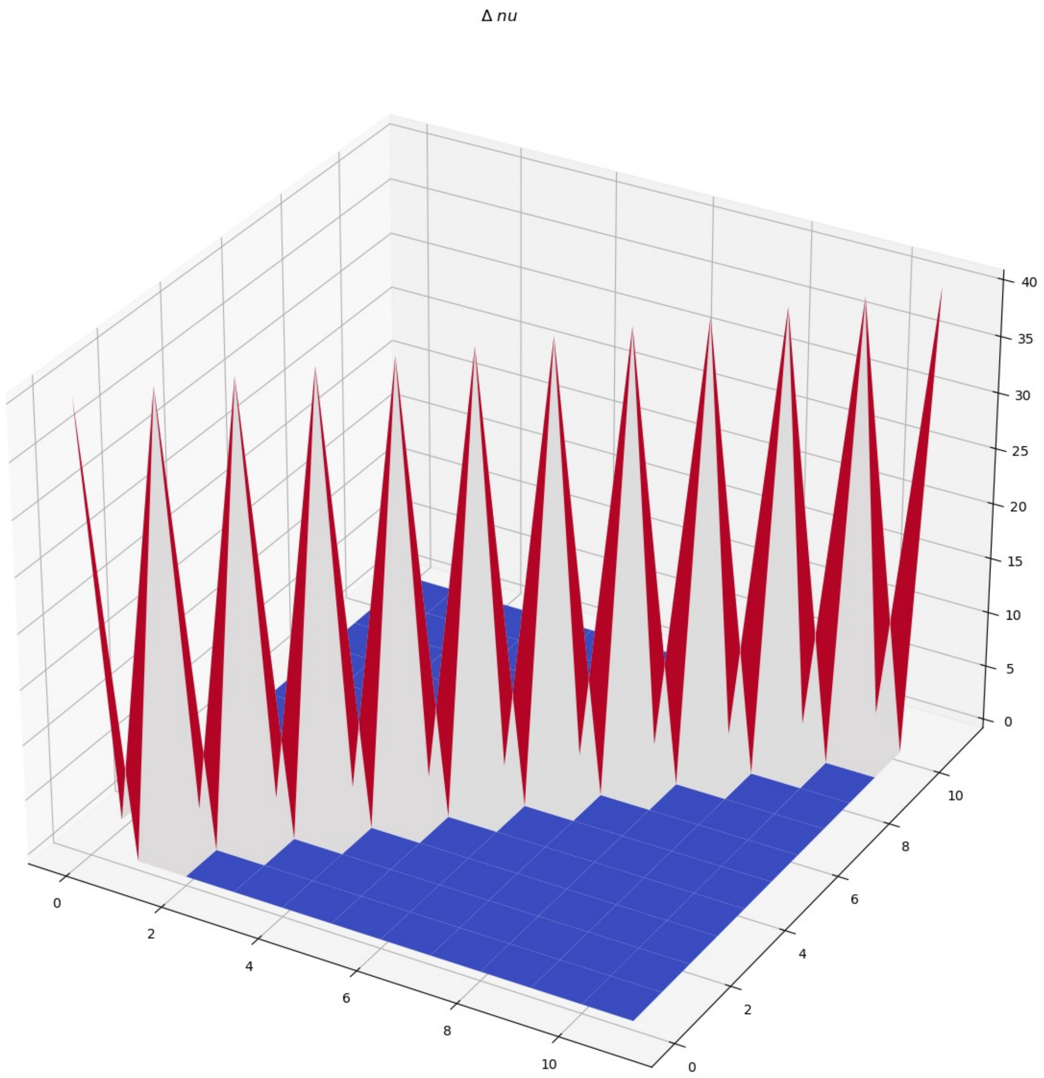
$$S = \begin{pmatrix} S_N \\ 0 \end{pmatrix} \in \mathbb{R}^{nm \times N}, \;\; S_N := diag(\sigma_1, \ldots, \sigma_N, 0, \ldots, 0) \in \mathbb{R}^{N \times N}$$

- $\vec{S}^+$ is pseudoinverse of $\vec{S}$ whose first N diagonal elements are $\frac{1}{\sigma}$

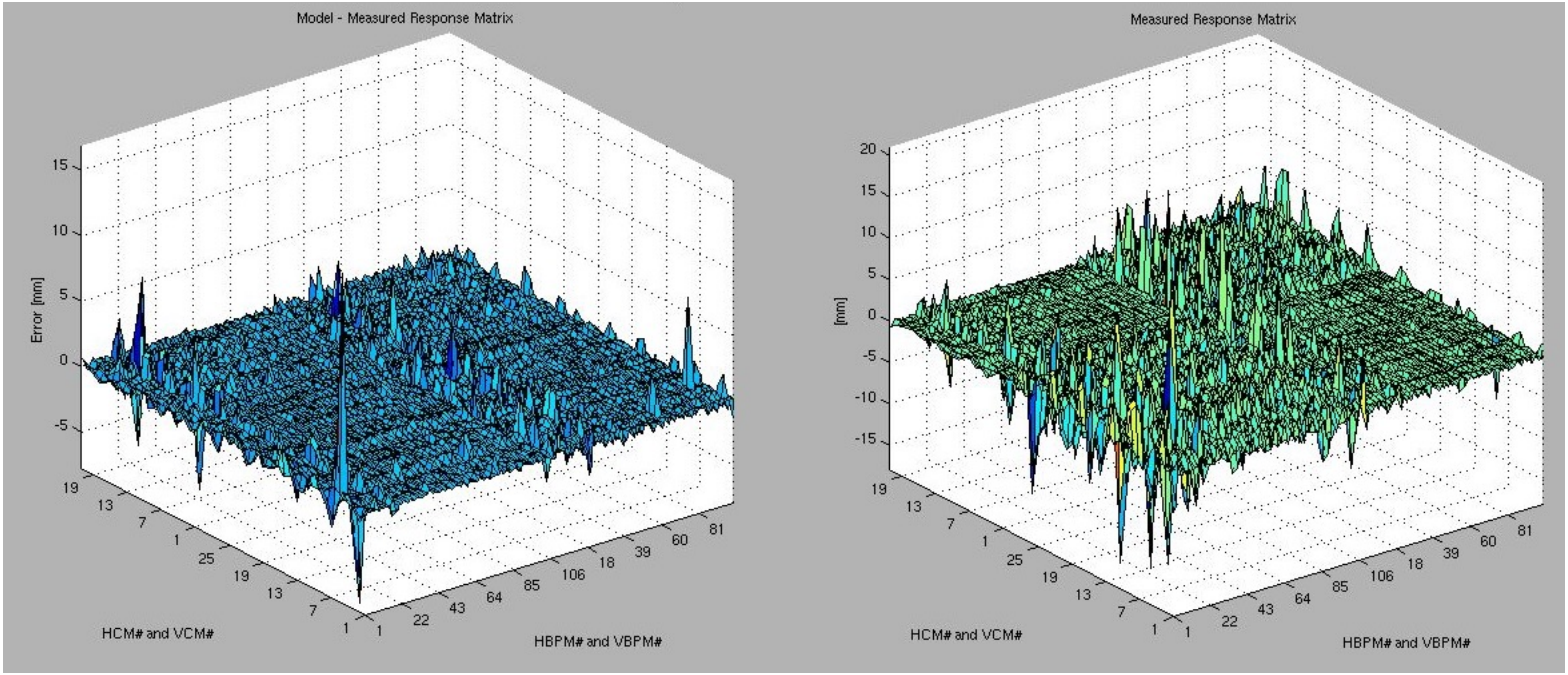
$$S^+ = \begin{pmatrix} S_N^+ \\ 0 \end{pmatrix} \in \mathbb{R}^{N \times nm}, \;\; S_N^+ := diag(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_N}, 0, \ldots, 0) \in \mathbb{R}^{N \times N}$$

$$\begin{pmatrix} \Delta\nu_1 \\ \Delta\nu_2 \\ \cdots \\ \Delta\nu_{N-1} \\ \Delta\nu_N \end{pmatrix} = J^+_{model} \begin{pmatrix} \Delta R_{11} \\ \Delta R_{12} \\ \cdots \\ \Delta R_{n(m-1)} \\ \Delta R_{nm} \end{pmatrix} = VS^+U^T \begin{pmatrix} \Delta R_{11} \\ \Delta R_{12} \\ \cdots \\ \Delta R_{n(m-1)} \\ \Delta R_{nm} \end{pmatrix}$$

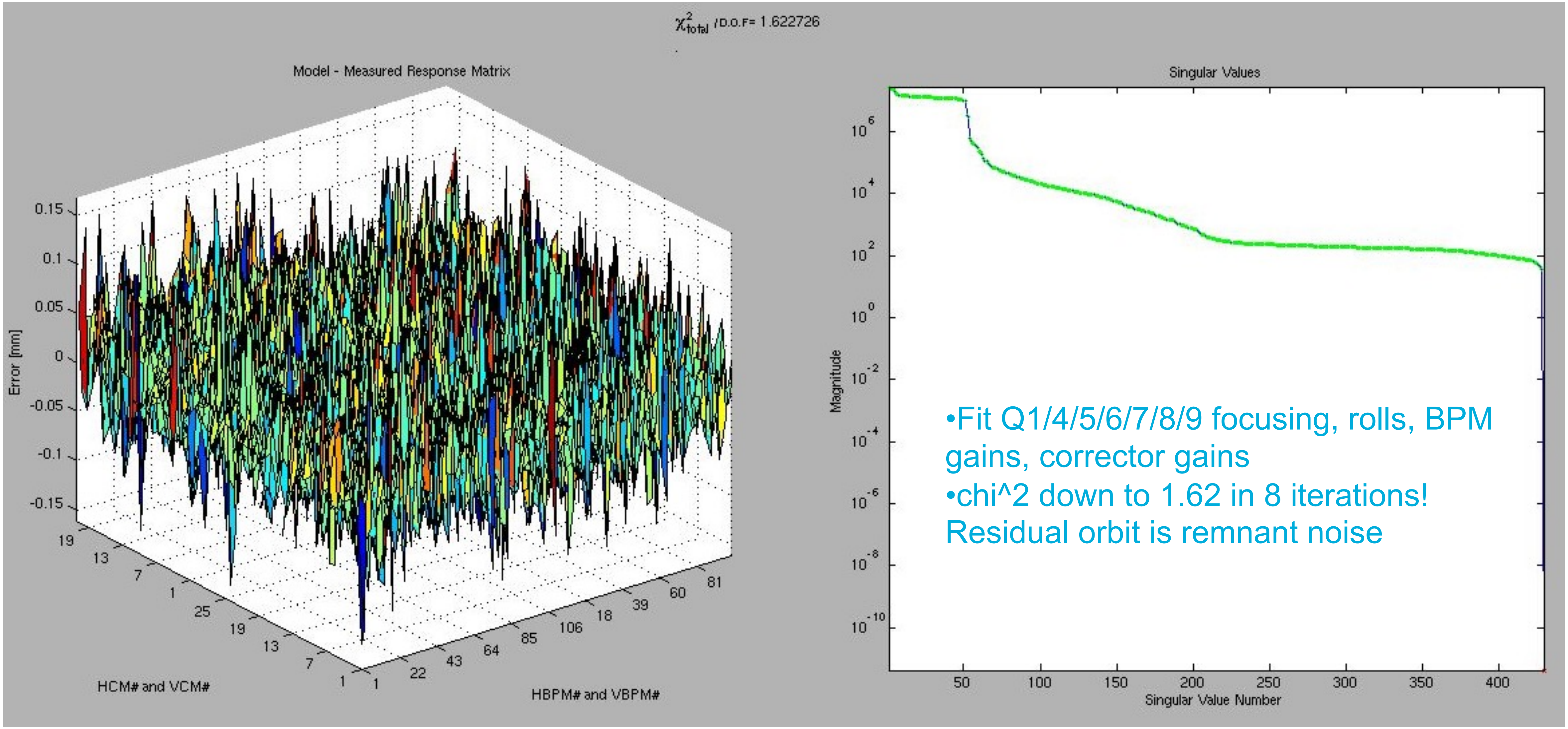
Reconstructed Errors



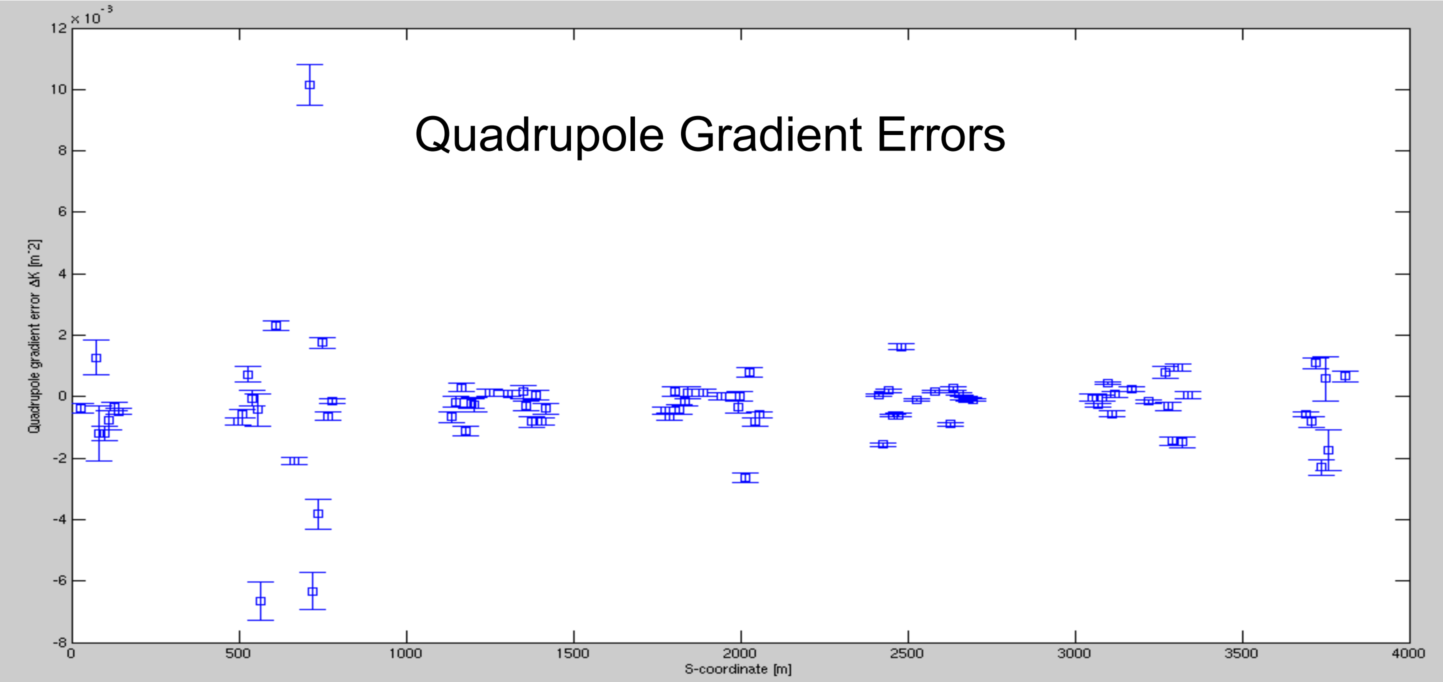
Beta-beat measurements in RHIC

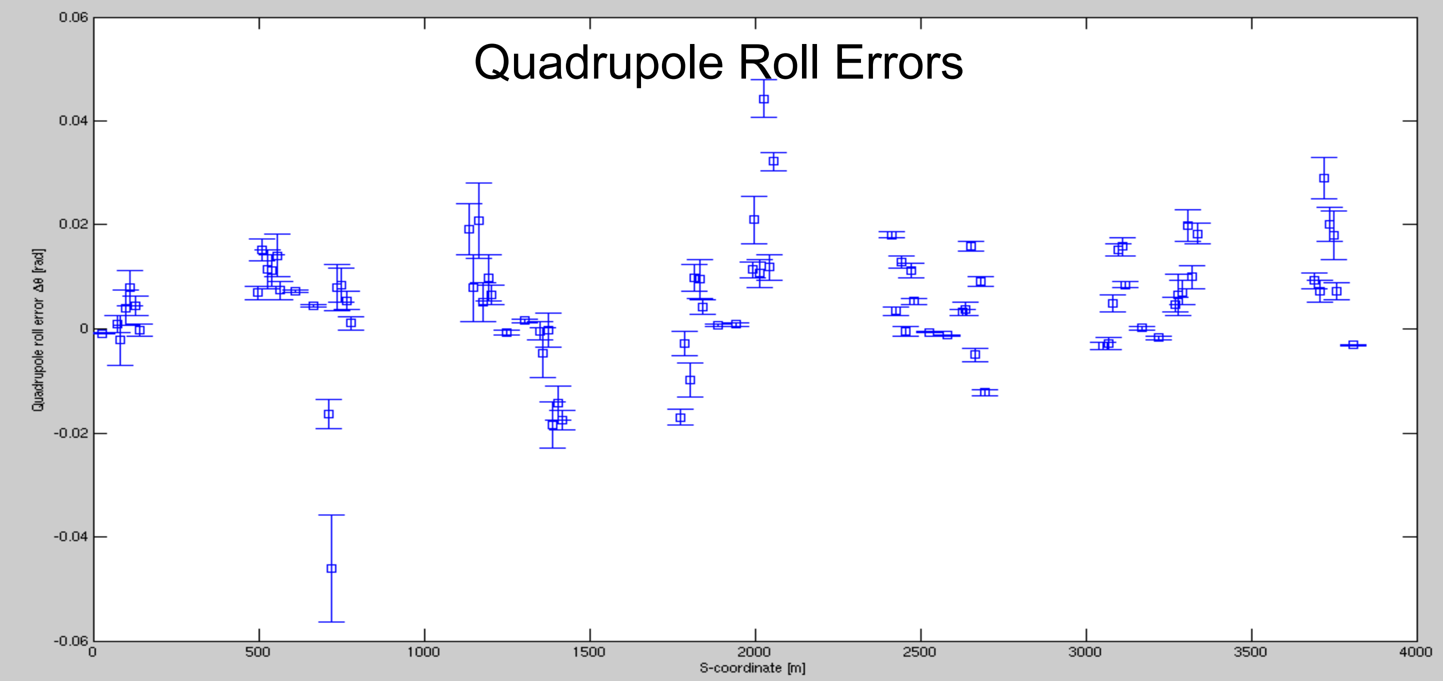


Brookhaven National Laboratory



- Fit Q1/4/5/6/7/8/9 focusing, rolls, BPM gains, corrector gains
- chi^2 down to 1.62 in 8 iterations! Residual orbit is remnant noise





Brookhaven National Laboratory

AI/ML

The process works, but …

It takes dedicated beam experiments to collect the data (hours).

It takes days to analyze the data (filter, reduce, etc.)

We only see a single temporal snapshot of the accelerator.

The bottlenecks

Before solving to learn the errors,

1. Must analyze the BPMs and determine if there are any gain/offset errors, or other problems
2. Must also analyze the corrector magnets and build trust in their behavior
3. Must test different $J_{model's}$ to learn where possible errors may be

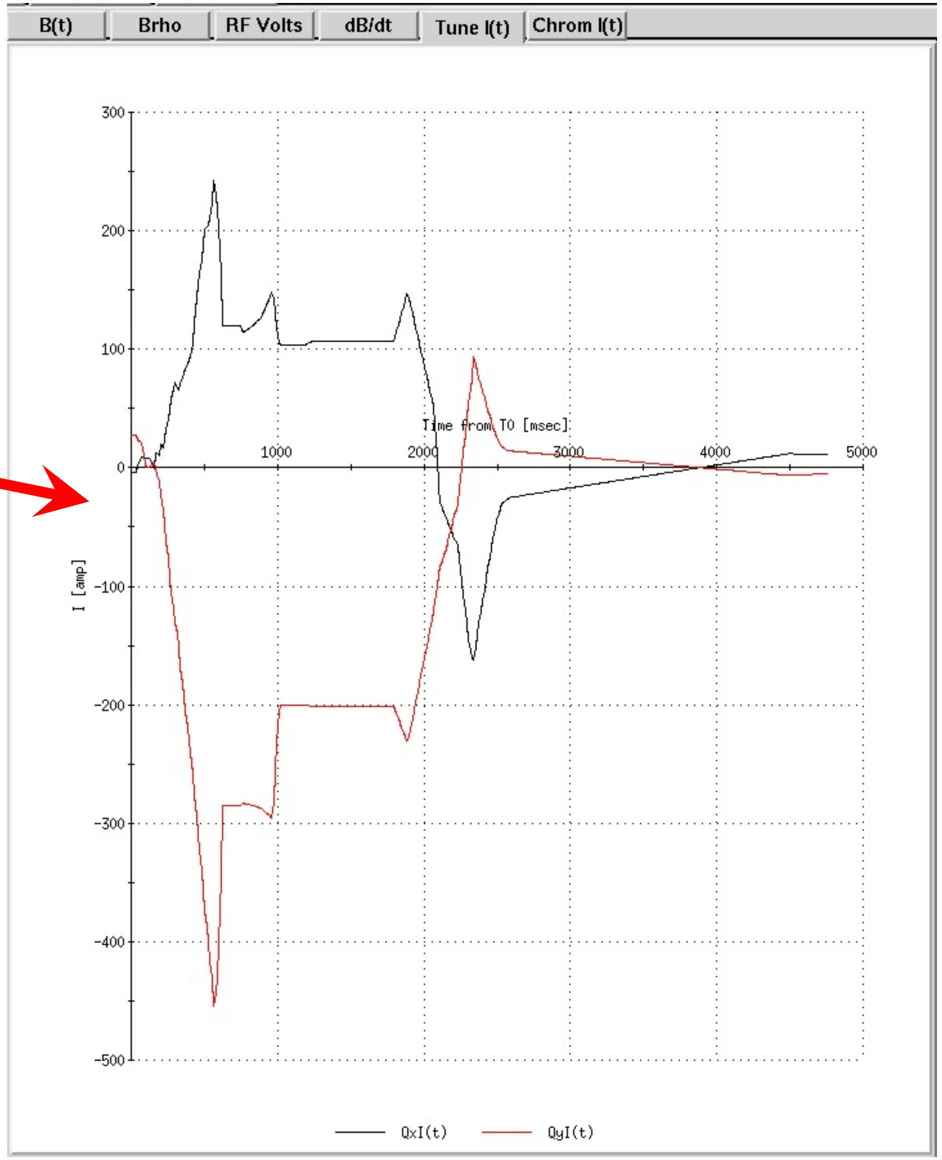
Where will ML help?

Learn bpm behavior and response = develop trust

Learn corrector behavior and response = develop trust

In general, learning how well devices follow instructions.

Magnets often need to follow complex waveforms.



From the control I/O points we can build a twin of the live machine that can reflect

- Stability
- Standard deviations
- Anomalous deviations
- Points of repeated anomalous behavior

Brookhaven National Laboratory

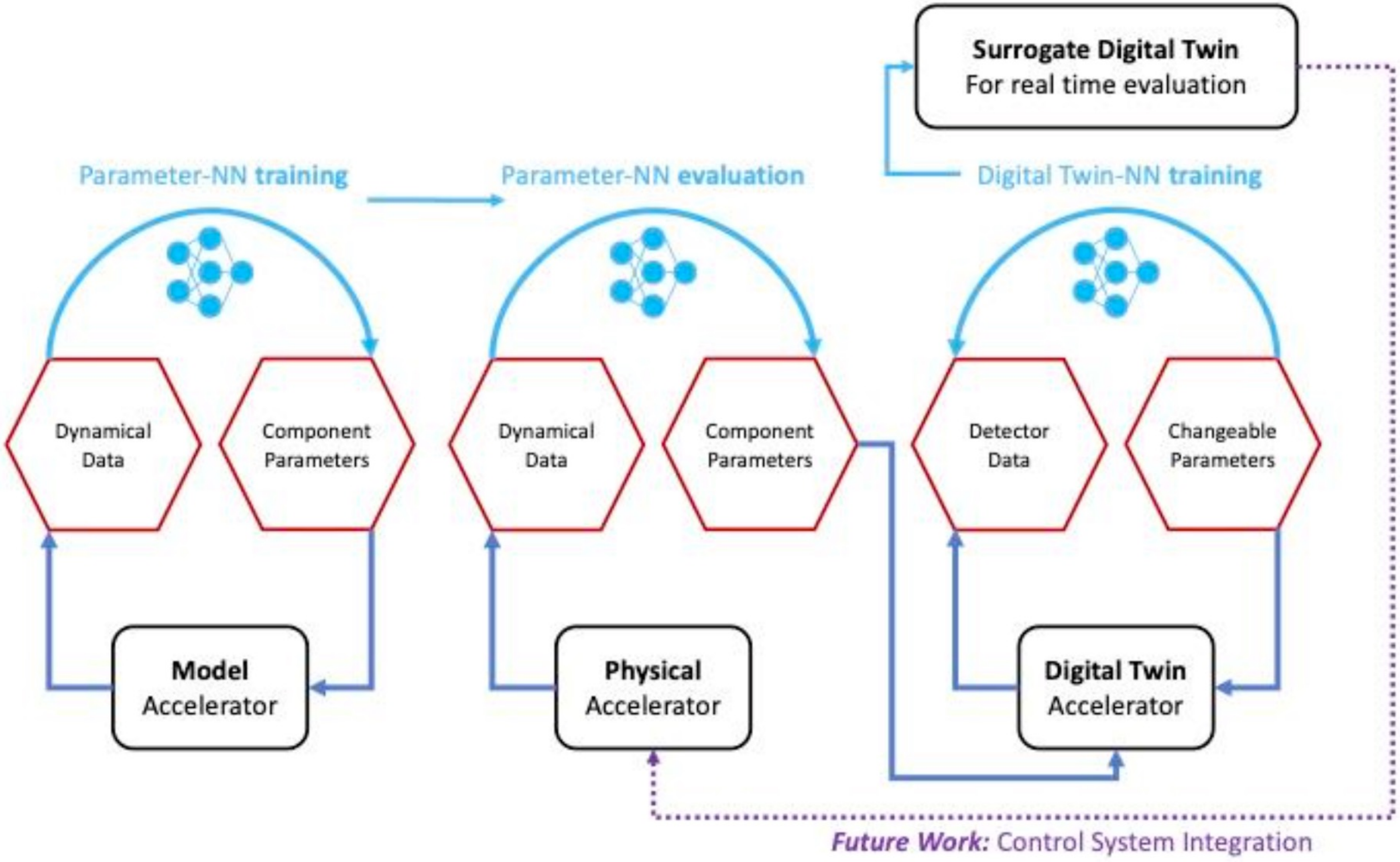
Where does Online Model fit in?

- physics is in the models – although imperfect
  - The model tells us how sensitive physics parameters are to device behavior
  - The model predicts how changes will affect the accelerator performance
- What is called the model is a suite of models that serve different purposes but are kept consistent (linear optics model, dynamic aperture model, longitudinal model, corrected orbit model, etc.)
- Is the 'right' data going into the model?
  - If we use Power Supply measured current – how good is the calibration? Does it change over time?
  - If we use Power Supply reference current – won't know of any 'true' drifts or changes
  - If a power supply gets replaced, how do we ensure the calibrations, etc. are closely enough corrected (e.g., a change made at 2am may be done quickly to get ops back).
- Combining uncertainty analysis/quantification from the twin will enhance the insights by reflecting those confidence levels into physics predictions

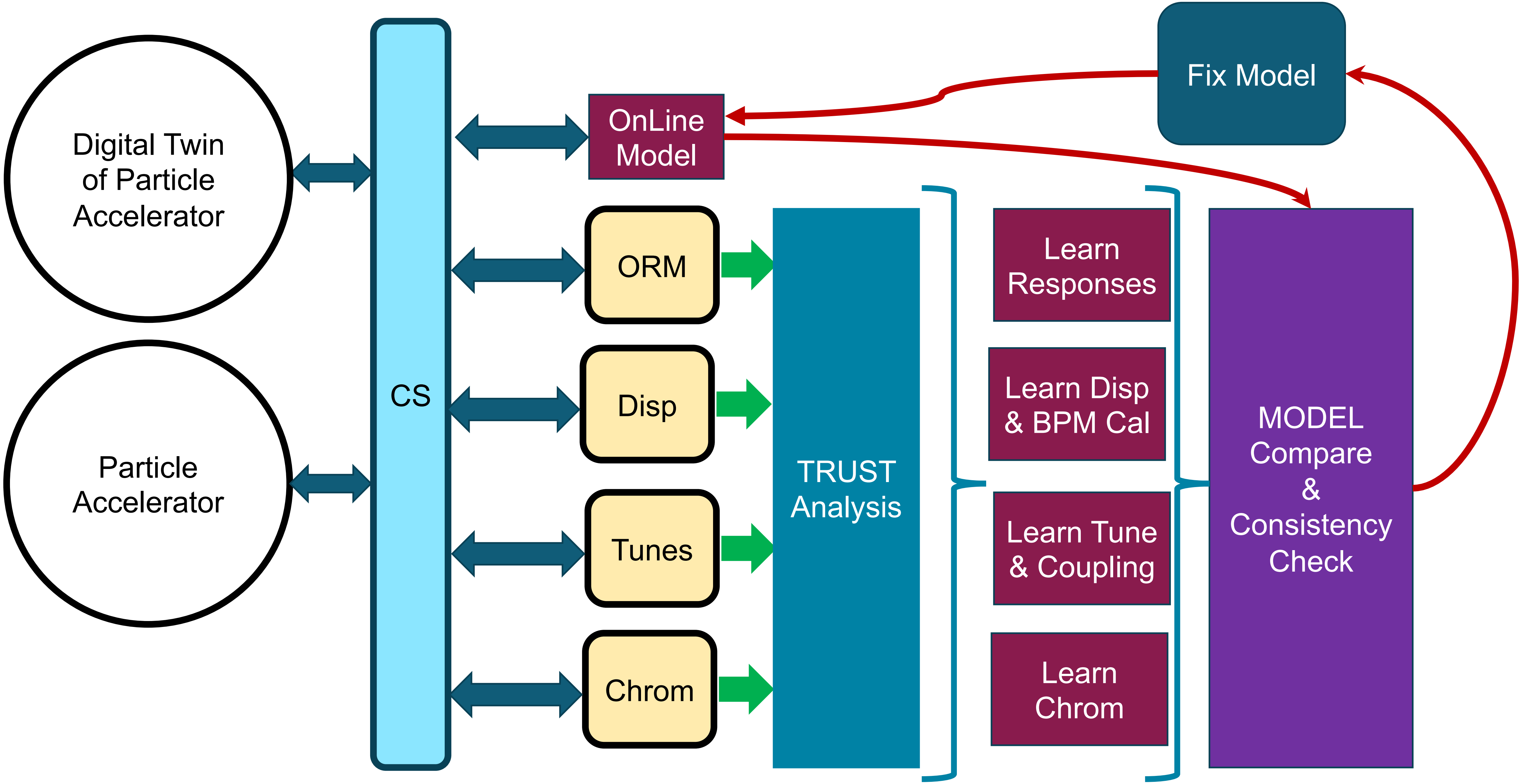
Where do we put AI/ML?

- ORM will give us
  - BPM and Corrector Anomalies (Trust Analysis)
  - Gradient errors for given conditions
  - Beta-deviations from model
- Dispersion measurements give us
  - BPM Consistency check for given dp/p (BPM Anomalies)
  - Coupling through longitudinal motion (very slow, typically)
- Tune measurements
  - Betatron tune and coupling = destructive measurement in Booster/AGS
  - Tune, Chrom, coupling, emittance, dp/p from RHIC Schottky (parasitic)
- Chromaticity measurements – need to change energy and measure tune
- Orbit Measurements – parasitic

Digital twins of accelerators



Schematic of how the accelerator model, physical accelerator, and DT of the accelerator are related. The Parameter-NN is trained on the model accelerator dynamical data. This NN is then used to map the dynamic data of the physical accelerator to component parameters of the DT. A separate NN is trained on the output data of the DT, acting as a quick-to-evaluate surrogate of the DT. This Digital-Twin NN maps simulated component parameters to physical accelerator parameters.


Digital Twin of Particle Accelerator
Particle Accelerator
CS
OnLine Model
Fix Model
ORM
Disp
Tunes
Chrom
TRUST Analysis
Learn Responses
Learn Disp & BPM Cal
Learn Tune & Coupling
Learn Chrom
MODEL Compare & Consistency Check

Computing?

Physics models have different computing requirements

- Offline Lattice analysis = single cpu (fast with deep memory)
- Online Lattice analysis = complete computation in ~10msec for real-time feedback, multiple fast independent cpu's
- Dynamic aperture = gpu's, HPC level
- Spin tracking = HPC, can still take days

AI/ML models can also vary

- Physics model informed Bayesian Optimization is very fast – single fast cpu works most of the time
- Deep NN requires HPC level resources

Accelerator control systems do not use HPC resources. Our paradigm needs to shift to combine Online (fast but simple) models with Offline (slow, includes more physics) models.

Summary

- Accelerators are highly complex systems that run at and beyond design specs. To operate they require precise physics models to guide the operators and push for higher performance.
- Using well trusted beam-based measurements these models can be improved. This will eventually lead to much higher degree of automation.
- Self-diagnosis utilizing AI/ML methods along with improved physics models will inform operators, physicists, and accelerator designers in new ways leading to new innovations and higher performance.

Thank you.

BNL

Kevin Brown, Sam Clark, Yuan Gao, Natalie Isenberg, Yongjun Li, Jennefer Maldonado, François Meot, Vincent Schoefer, Reid Smith, Nathan Urban

Cornell University

Georg Heinz Hoffstaetter, Lucy Lin, David Sagan

Stony Brook University

Weining Dai, Bohong Huang, Thomas Robertazzi