Fermilab U.S. DEPARTMENT OF ENERGY Office of Science

# Recent applications of machine learning for particle accelerator control

Daniel Bowring

FNAL Accelerator Physics & Technology Seminar

5 June 2018

# Lots of ground to cover in one talk. Some goals:

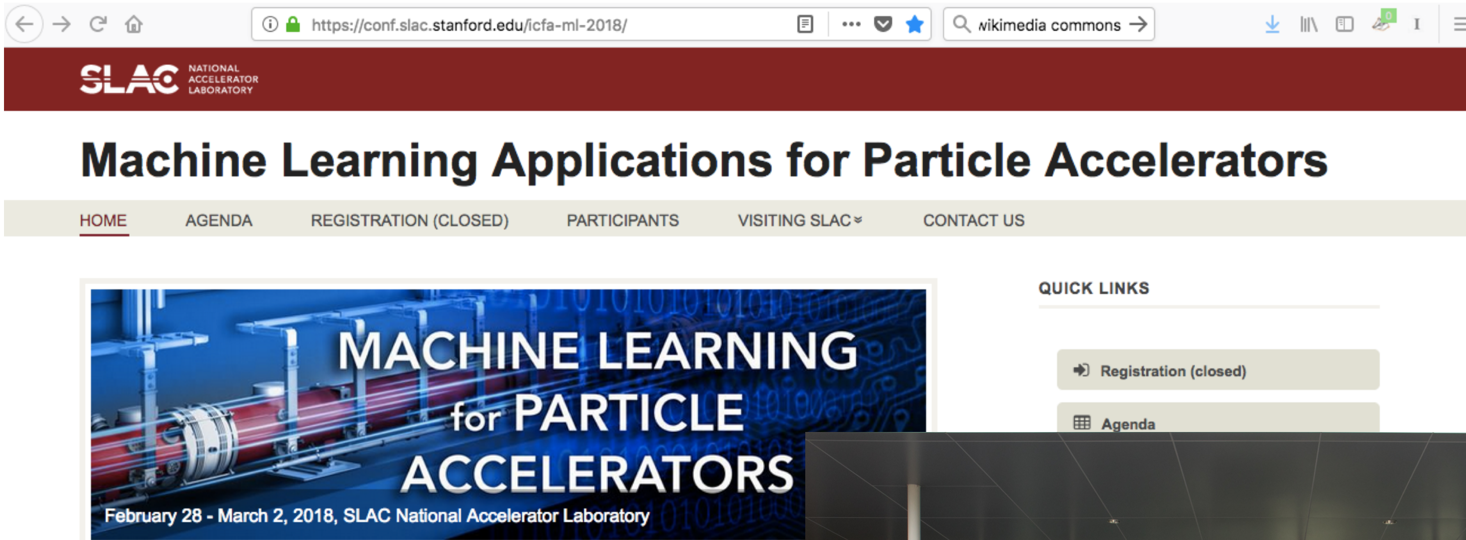
- What is machine learning and why should the accelerator community care?

- Review 2018 ICFA workshop content.
  - I will be discussing others' work. Normal caveats apply.
  - Much of this work is in progress. Watch out for an Arxiv white paper with references.

- Observation: computer scientists and accelerator physicists talk about these things very differently.

- I assume I'm addressing mainly AD staff who want to know whether machine learning might be useful for their work.

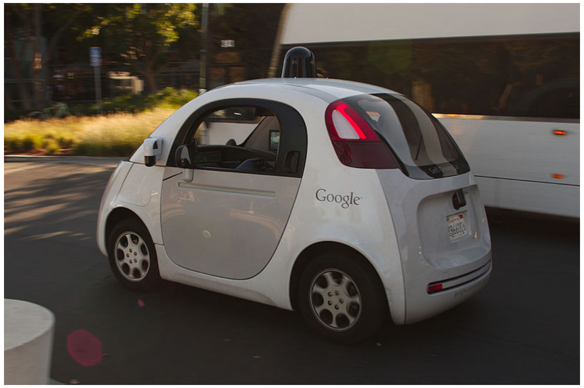
# This talk summarizes the 2018 ICFA workshop.




- 65 Participants
- 20+ Institutions
- Attendees' backgrounds: computer science, physics, controls, operations, industry

# Why is machine learning everywhere all of a sudden?


1. Technological/commercial advances
2. Availability of HPC resources
3. Progress in algorithm development
4. Availability of open-source, high-level frameworks (TensorFlow, scikit-learn, etc.)
5. Private-sector demand




# What machine learning (ML) is / is not:

- Machine learning is not:
  - a magic solution to all your current problems;
  - globally applicable;
  - easy/trivial to implement.
- Machine learning might be described by:
  - computational statistics;
  - big data sets;
  - non-intuitive and/or non-analytic modeling.

# What can machine learning do for us?

- More uptime
  - better/faster optimization & tuning
  - improved fault monitoring, prediction, and diagnosis
- Less waste
  - improved control of power-hungry components
  - see above

# Taxonomy: Systems / Methods

- Supervised learning: data somehow labeled
  - Classification
    - Spam ID
    - Credit fraud detection
  - Regression
- Unsupervised learning: unlabeled data
  - Clustering
    - Netflix recommendations
  - Dimensionality reduction
- Semisupervised learning
  - Face tagging in photo software

# Taxonomy: Systems / Methods

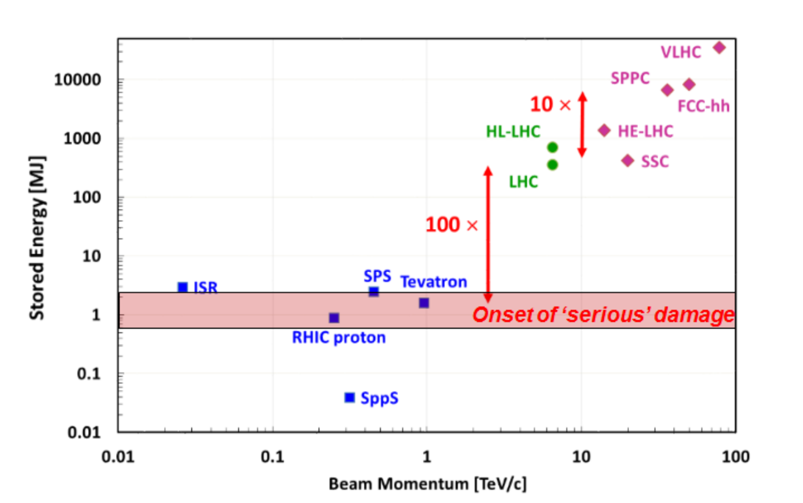
- Reinforcement learning
  - policy developed based on rewards/penalties from environment
    - robot training
- Batch vs online learning
  - batch: train model on ”all” data
  - online: update model on the fly
- Instance- vs model-based learning
  - instance: new data compared with known entities
  - model: new data compared with model

# Workshop Synopsis

# Session: Facility Needs

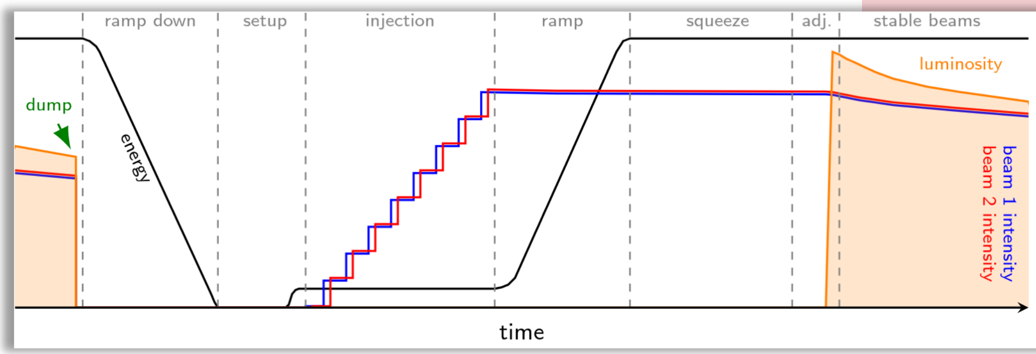
- What are the current problems faced by various accelerator labs?
- Can (should) these problems be addressed through machine learning?
- Common themes were:
  - Failure ID / prediction
  - Faster simulations and online models
  - How to best use our "big data"?
  - **Where does ML add value?**

# K. Fuchsberger: Operational challenges at LHC



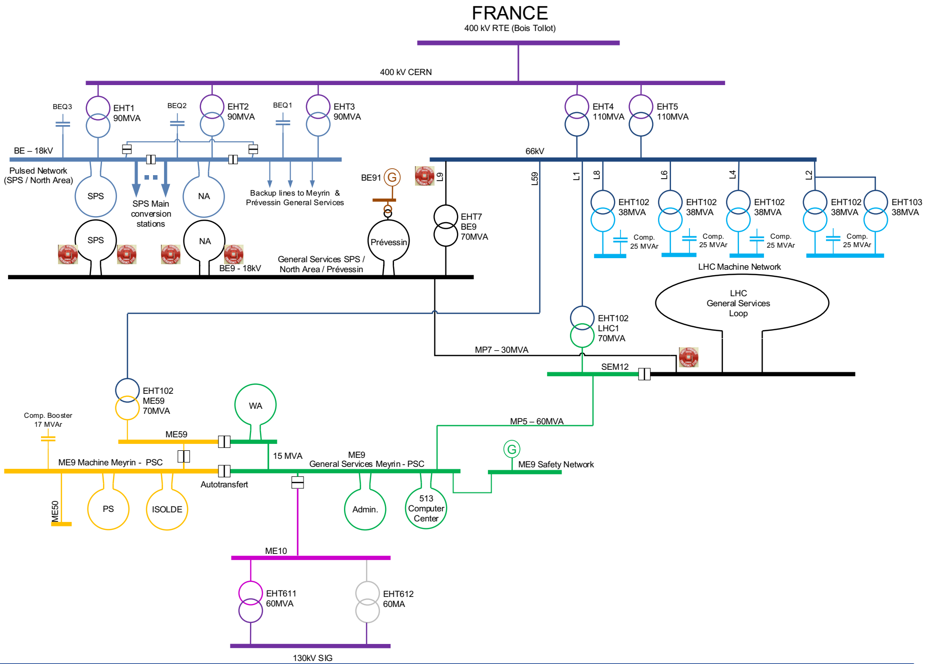

360 MJ → No "playing around" possible;
Highly reproducible operation desired…

- Also:
  - Old machines drift, are insufficiently instrumented, have narrow optima, ...
  - Stability requirements change with a new cycle every 1.2 seconds x hundreds of cycles per year in PSB.




Dumps are Costly!
→ Downtime Minimization

# J. Nielsen (LHC): Total Power Outage




- How to correlate problem (outage, system failure) with a specific cause?
- Fault anticipation & faster diagnostics yield **less downtime.**
- How are faults diagnosed now? Operator/electrician instincts play a role.

**What happened?**

- Faulty emergency stop button
- Complete Power Outage

**What could we do with machine learning?**

- Detect safety alarm (emergency stop)
- Correlate safety alarm with power cut

- Propose to create event in logbook
- Propose possible causes (simple here..)

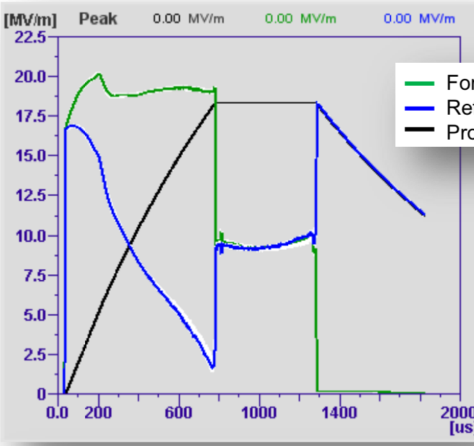
- Propose critical installations that would need special attention

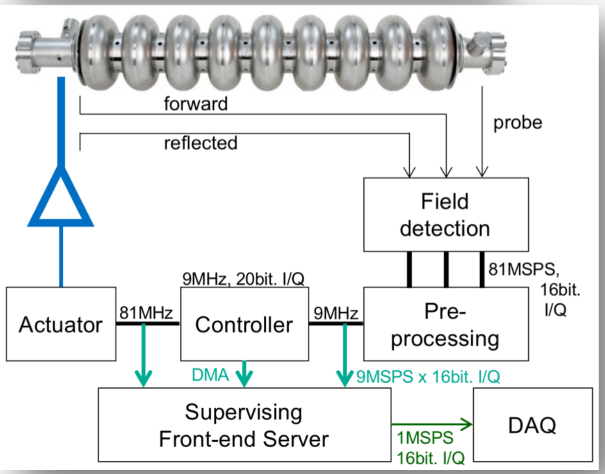
# R. Kammering (DESY): SRF cavity anomaly detection

- Cavity fault detection requires: $U_{for}$, $U_{ref}$, $U_{probe}$
- Data rates to DAQ per cavity per pulse:
  - 2048 x 2 x 3 x 16bit = 24.6kB
  - Pulses per Day = 864000
  - 700 cavities → **604 Mio** events/day
  - Total data/day = **14.8 TB**
- **Good statistics** (ensemble & events)

**Questions we like to address:**

- How many cav./pulses behave normally
- Cav/Pulses out of nominal operation range
- Reliably quench detection and reaction
- Anomalies: due to parameter changes
  due to digital / communication/ readout




Primary goal is cryomodule health, but a window into cavity physics issues would be helpful too.

# Session: Tuning

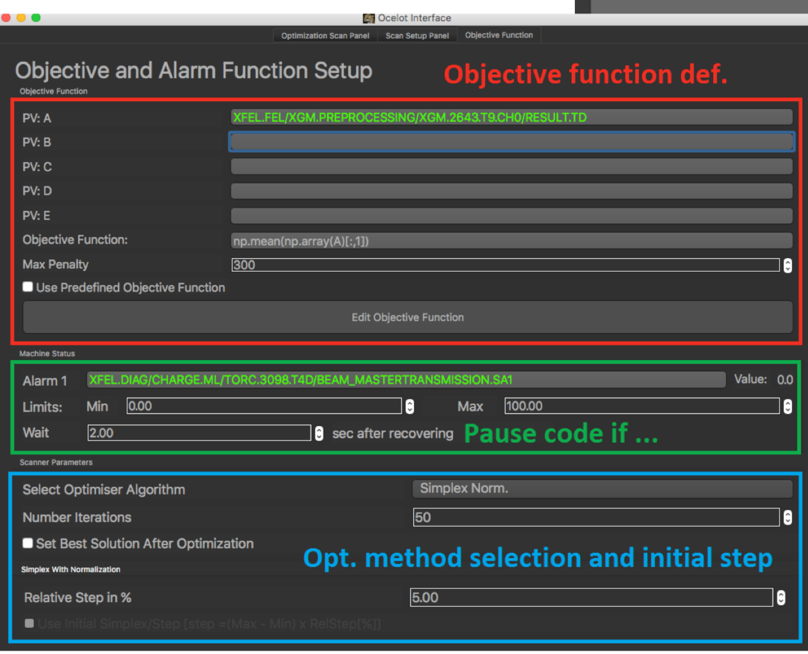
- Automated tuning would be a big boost to operations.
- Should this tuning be model-based, or model-independent?

# S. Tomin, I. Agapov (DESY/XFEL)

**Tuning platforms:**

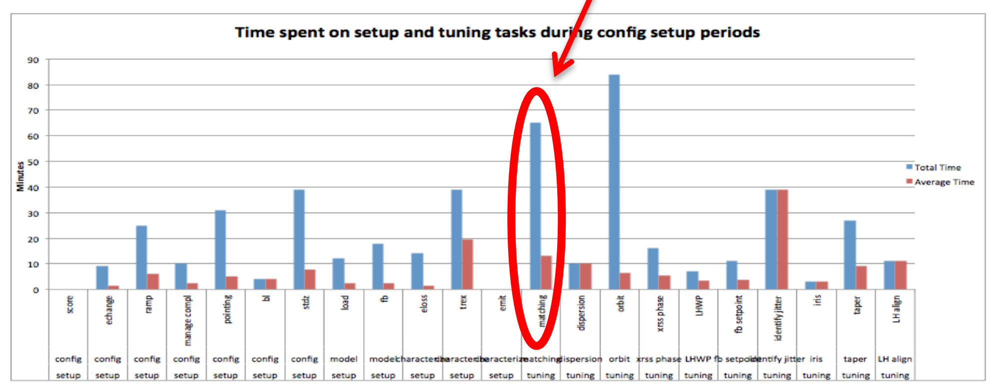
Ocelot (DESY) provides generic base for accelerator optimization/simulation

➔ Now multi-lab collaboration

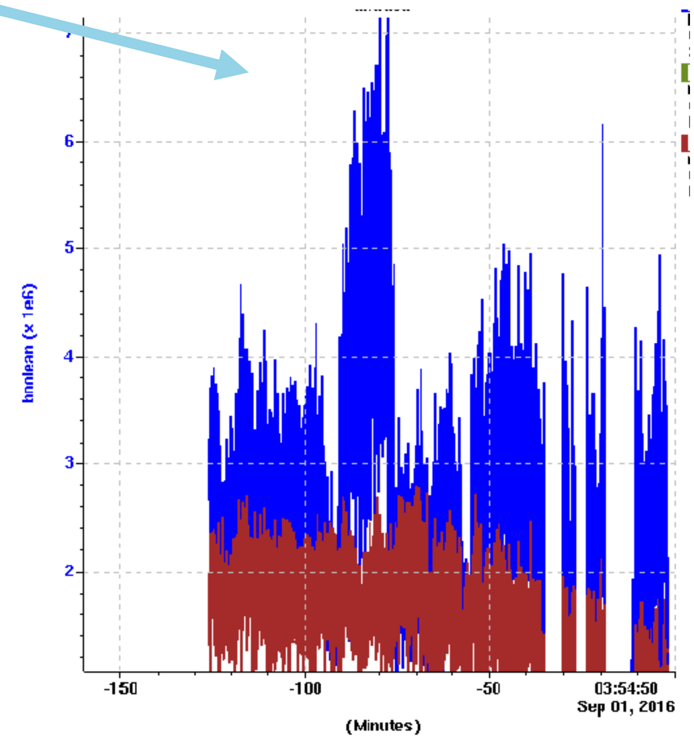

At LCLS: 450 hours/year in 2016

➔ Automated tuning cut avg time by half in 2017!



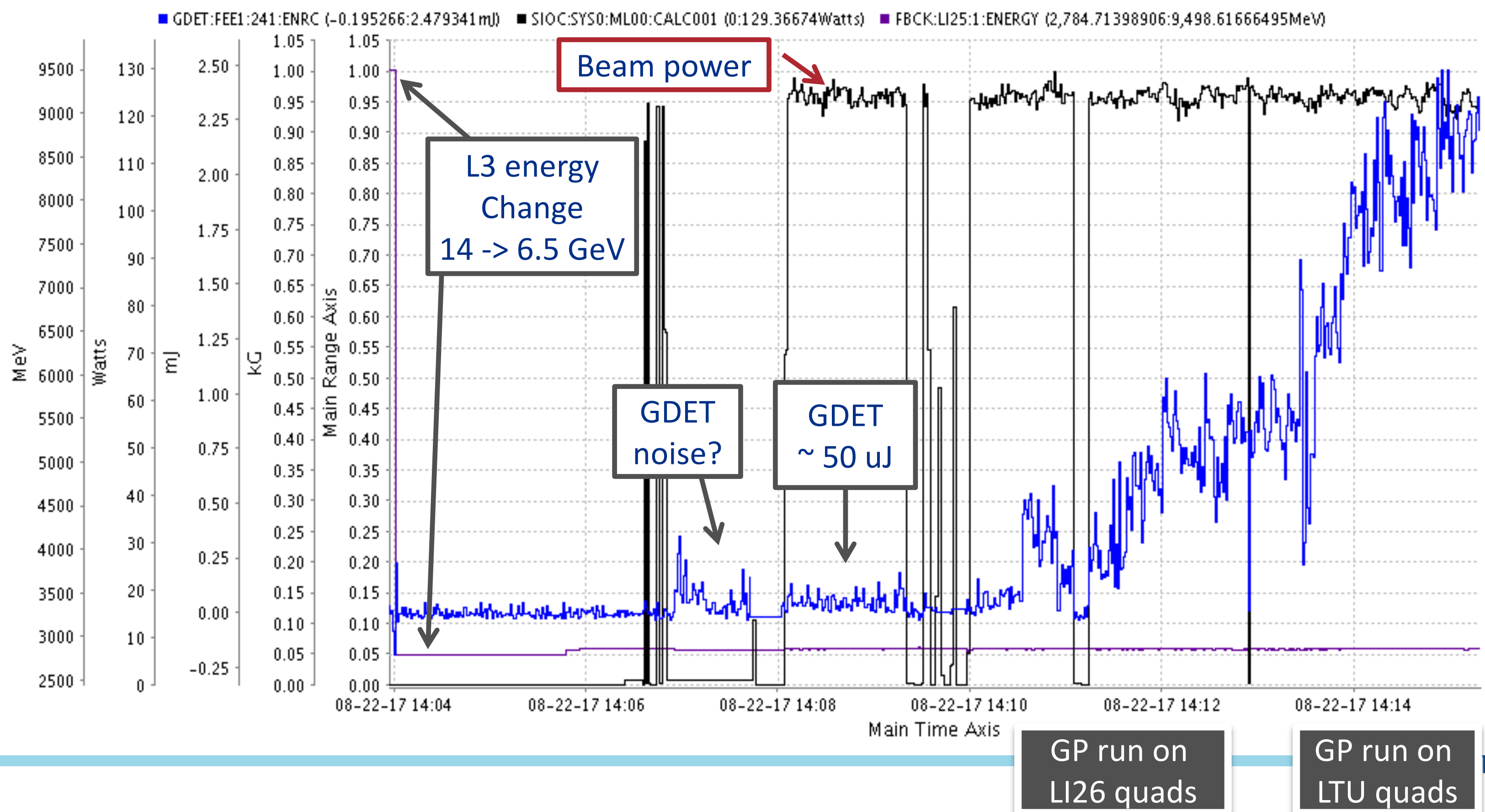
# From J. Wu (SLAC): LCLS taper optimization

- 5.5 keV self-seeding FEL
- Clustering / reinforcement learning to characterize system state(s)
- ML **online** optimization of taper profile doubled peak energy
- publication pending



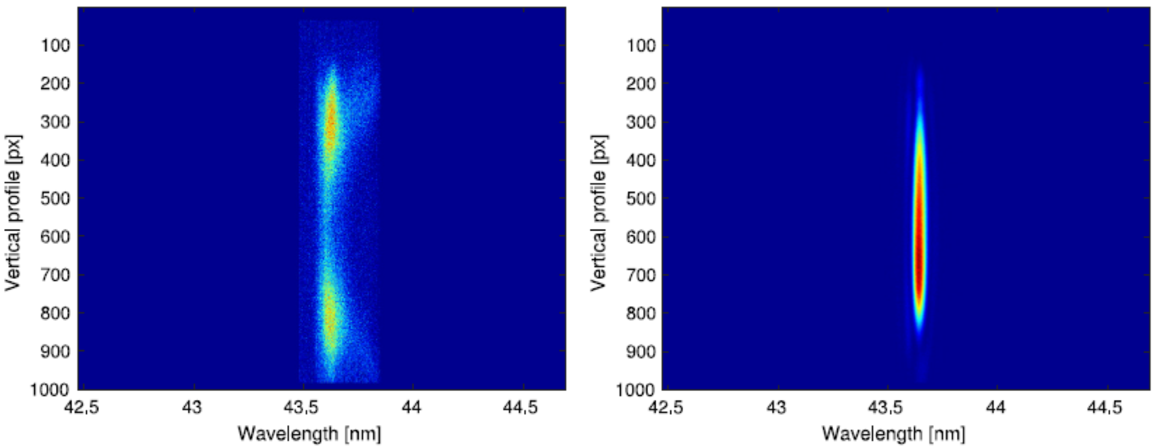
# J. Duris (SLAC): Bayesian optimization to tune from noise

We used a Bayes prior from summer 2017 data to tune up a brand new config from noise. Simplex could not do this as it needs signal to tune on.



# G. Gaio (FERMI@Elettra): Spectrum Optimization

- The **FEL Quality Factor** (*FELQFactor*) is an index which summarizes in a number the **most important features of the photon energy spectrum**: intensity, spectral purity and number of modes.
- The *FELQFactor* algorithm is capable of **evaluating the spectrum image as an expert** does
- It has been used in some preliminary machine optimization tests; two optimization algorithms have been used: **Ascent Gradient** and **Extremum Seeking.**
- The actuators were the seed laser delay line and the dispersive section.




FEL spectrum before (left) and after (right) optimization based on the *FELQFactor* index

**High resolution photon spectrometer**

- the horizontal axis is the photon energy
- the vertical axis represents the vertical photon beam distribution

"Free-electron Laser Spectrum Evaluation and Automatic Optimization", Nuclear Inst. and Methods in Physics Research, A 871 (2017) 20 29

# Session: Modeling and simulation

- Using neural networks, GANs to model complex systems
  - including systems here at Fermilab!
- ”Surrogate modeling”: using a slow, offline model to train a fast-executing neural network

# Temperature Control for the RF Photoinjector at FAST

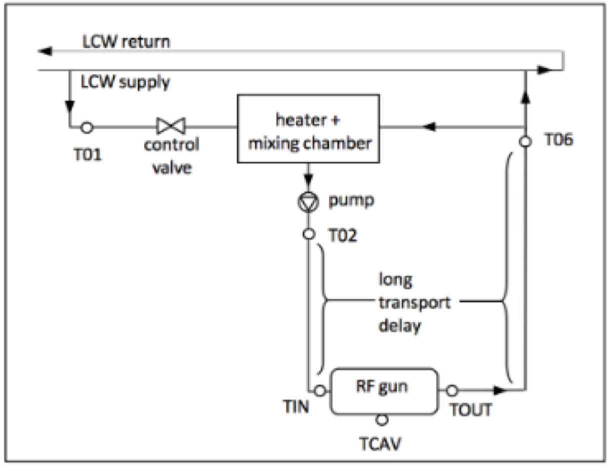
Resonant frequency controlled via temperature

PID control is undesirable in this case:

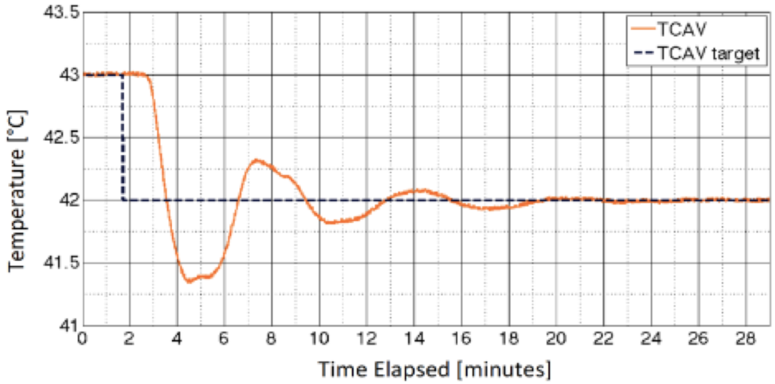
- Long transport delays and thermal responses
- Recirculation leads to secondary impact of disturbances
- Two controllable variables: heater power + valve aperture

Applied model predictive control (MPC) with a neural network model trained on measured data: ~ 5x faster settling time + no large overshoot

Gun Water System Layout

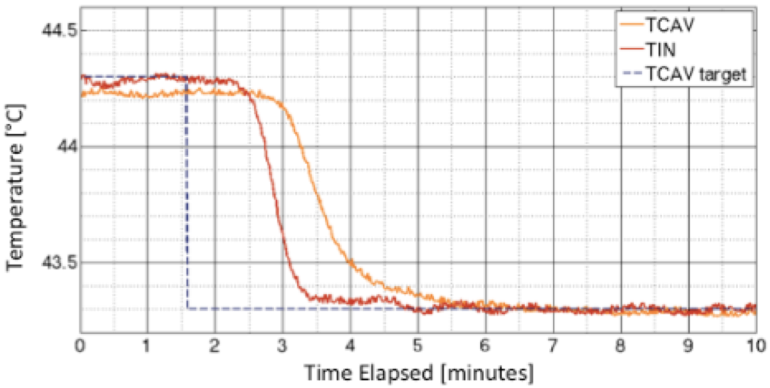



Existing Feedforward/PID Controller




*Note that the oscillations are largely due to the transport delays and water recirculation, rather than PID gains*

Model Predictive Controller



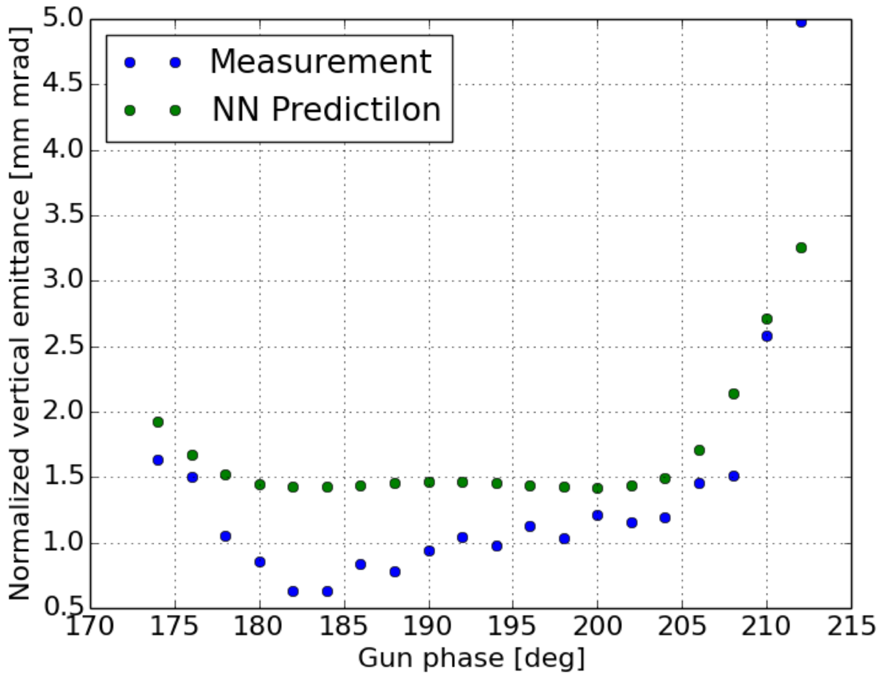

*A. L. Edelen et al., TNS, vol. 63, no. 2, 2016* *A.L. Edelen et al., IPAC '15*

# J. Edelen (RadiaSoft): NN modeling and virtual diagnostics at FAST

- Using Neural Networks
  - *Online modeling*
    - High fidelity physics simulations run off line on HPC resources
    - Use dataset to train a neural network surrogate model
    - Fast executing, can be used to model upstream components when downstream machine configurations are changing
  - *Virtual diagnostics*
    - In many ways this is similar to online modeling
    - Intercepting diagnostics are very useful for characterizing the beam
    - Pre-train a neural network model using simulation data and update with measurements

# J. Edelen (RadiaSoft): NN modeling and virtual diagnostics at FAST

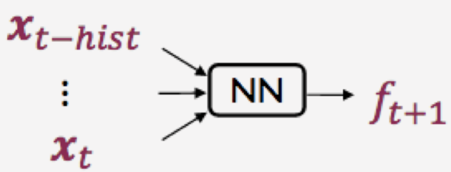
- Using Neural Networks
  - *Online modeling*
    - High fid resources
    - Use dat odel
    - Fast exe omponents when downstr g
  - *Virtual dia*
    - In many
    - Intercep terizing the beam
    - Pre-trair data and update with me



# Initial NN Modeling for RFQ: Same as for FAST

*wanted to make sure we **could** model the response before moving forward*

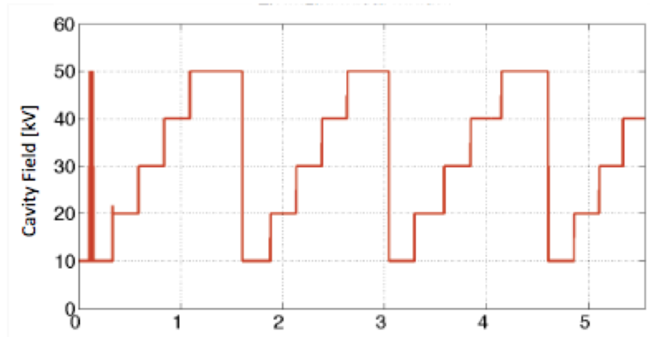
## Feed-forward, Fully-connected Network

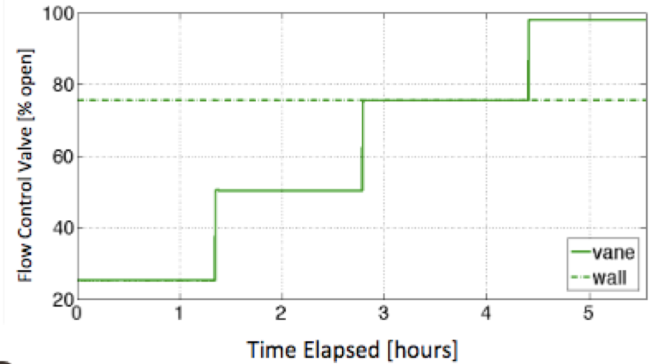


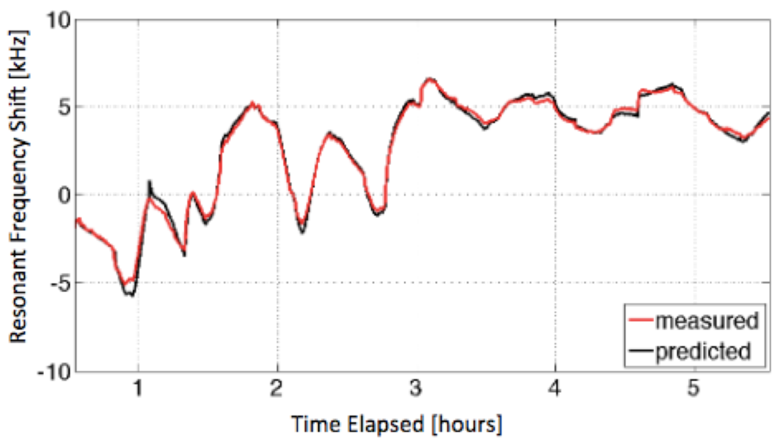

Two hidden layers: 25 and 7 nodes

$hist$: 30 minutes at 1 Hz

$x$ includes:
- Valve settings
- Average RF power
- Water temperatures
- Cave temperature
- Cave humidity




## Mean Absolute Error

346 Hz – test set
98 Hz – validation set
115 Hz – across all sets

## Training Data

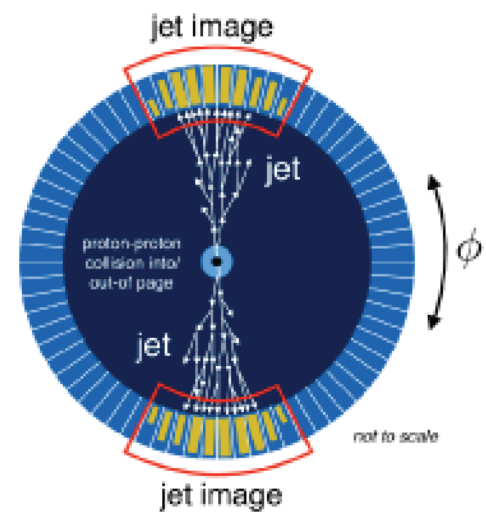
~ 64 hours of measurements
Scanned average RF power, valves
Includes RF trips, startup/shutdown

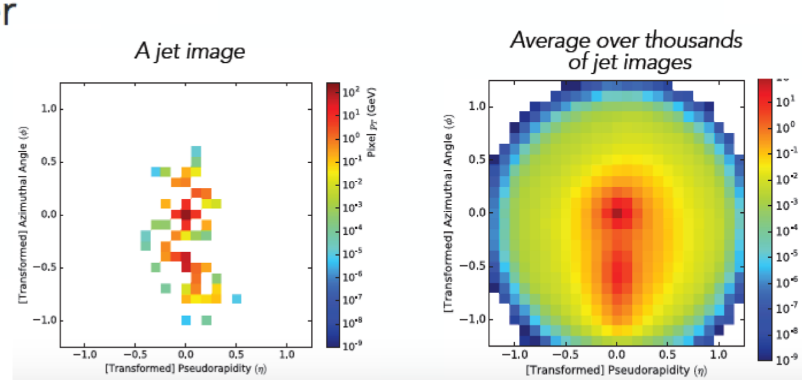
*A.L. Edelen et al., IPAC '16*

# L. De Oliveira (LBNL): GANs for HEP simulation

## GANs in HEP

- Ultimate goal: quickly and accurately simulate particles interacting with individual detector components - speed up slow simulation
- Intermediary goal: can we speed up calorimeter simulation, which is the current bottleneck?
- First step: can we learn to generate jet images using a Generative Adversarial Network?



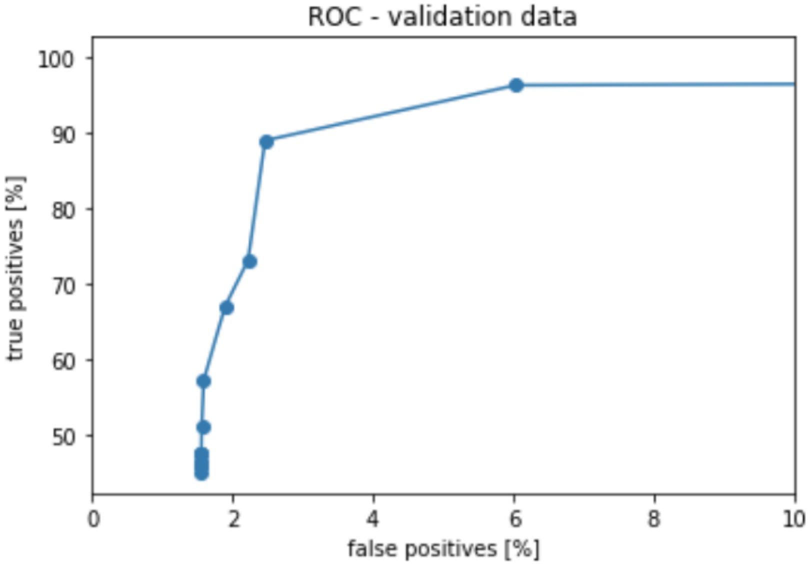



# Session: Prognostics

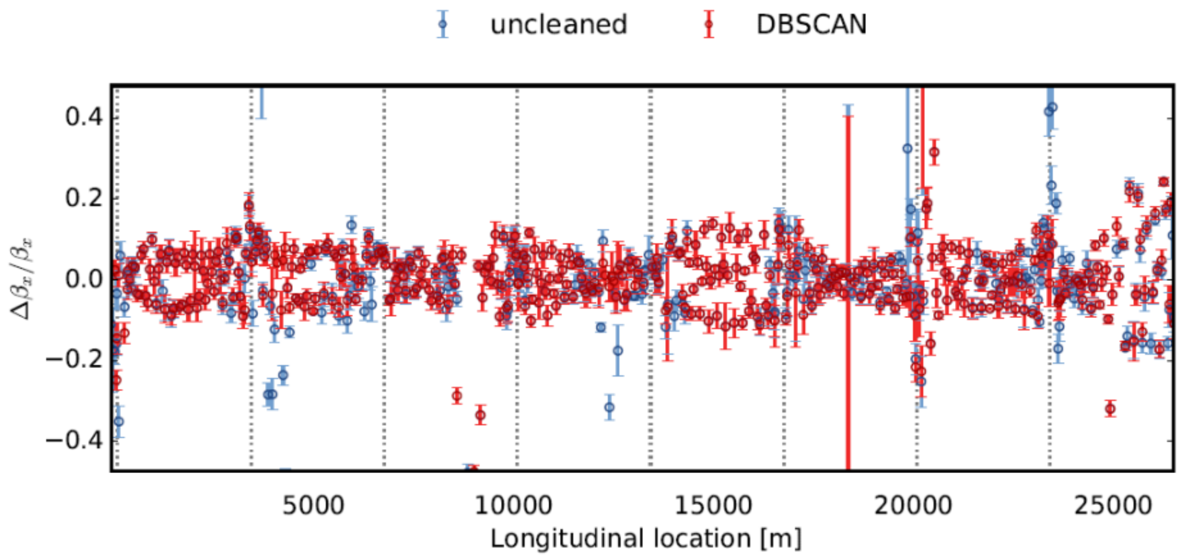
- Use available data to identify anomalous behavior.
  - Identify faulty devices, predict imminent failure.
  - Recall discussion of SRF cavity quench prediction.
  - See also: Wielgosz et al., ”Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets”, https://www.sciencedirect.com/science/article/pii/S016890021730668X

# A. Adelmann and J. Snuvernik (PSI): Forecasting interlocks

- At PSI, interlocks cause O(2%) of beam time loss.
- Try to predict interlocks through classification on 57 parameters (magnet currents, non-intercepting diagnostic data, etc.)
- Very preliminary results:



# E. Fol (CERN): Detection of faulty beam position monitors

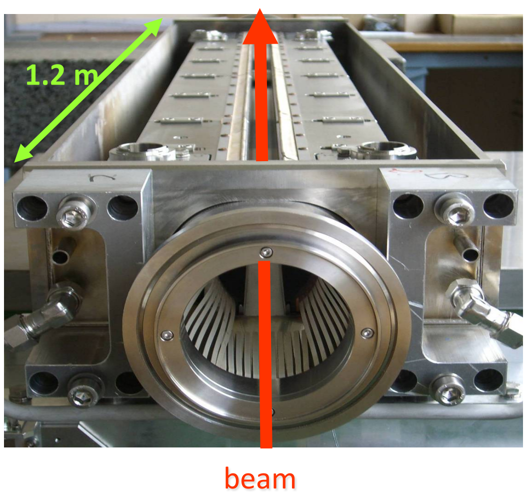


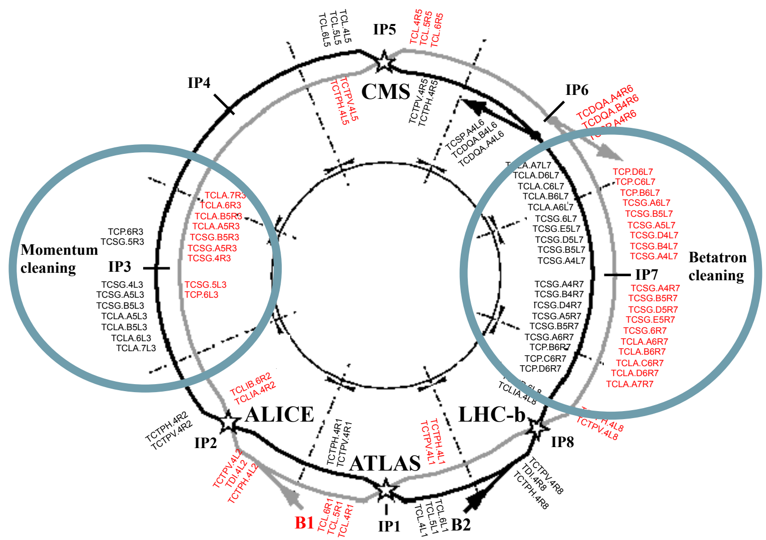

- 1024 BPMs per beam. Used for optics reconstruction & to guide corrections.
- ~10% of BPMs are faulty. These should be scrubbed from the correction process.
- DBSCAN clustering analysis to find outliers.

# G. Valentino (CERN/U. Malta): Beam loss plane detection

- The LHC is equipped with a multi-stage collimation system to protect it from normal and abnormal beam losses.
  - **Normal losses:** ensure that proton leakage to superconducting magnets is minimal, preventing quenches
  - **Abnormal losses:** protection against fast failure scenarios such as asynchronous beam dump

- The collimation system cleans particles with large betatron  and off-momentum offsets



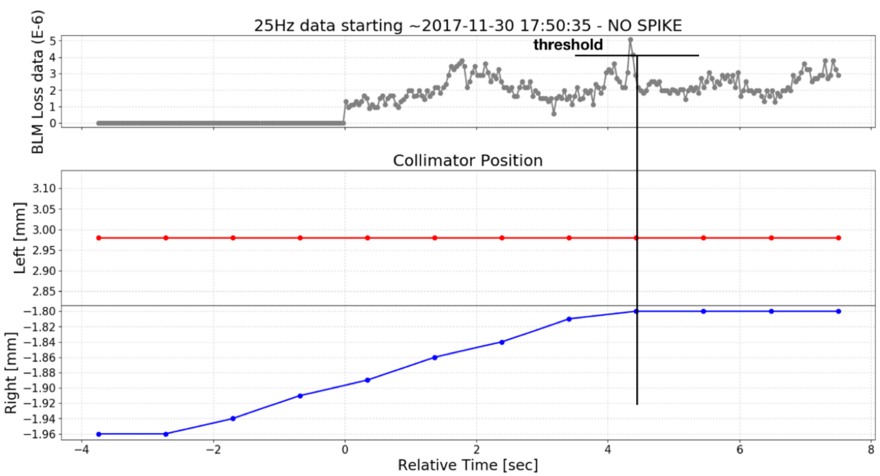



# G. Valentino: BPM Spike Classification

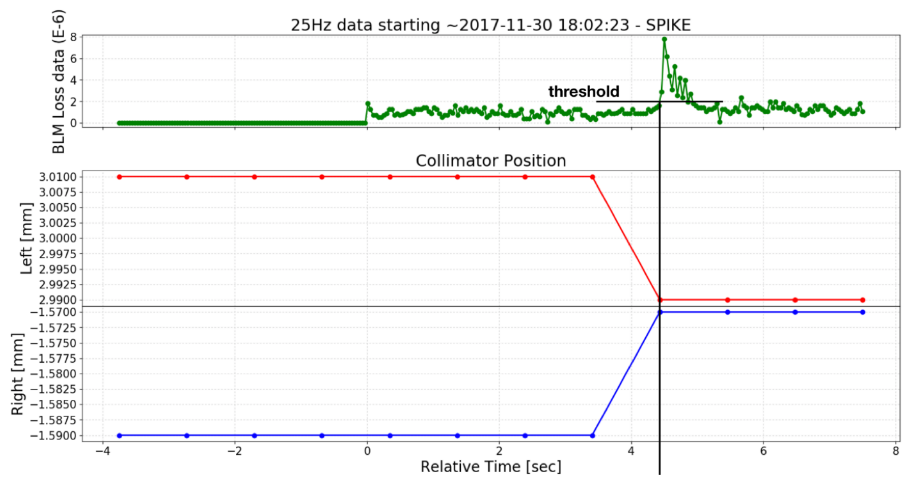
The ML model was tested during an MD and achieved 50/52 correct classifications

Examples:

<u>*Correctly classified No Spike*</u>




<u>*Correctly Classified Spike*</u>




G. Azzopardi, “Spike Pattern Recognition for Automatic Collimation Alignment” No. CERN-ACC-NOTE-2018-0010.

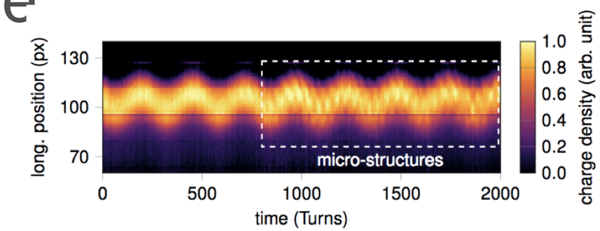
Courtesy: G. Azzopardi

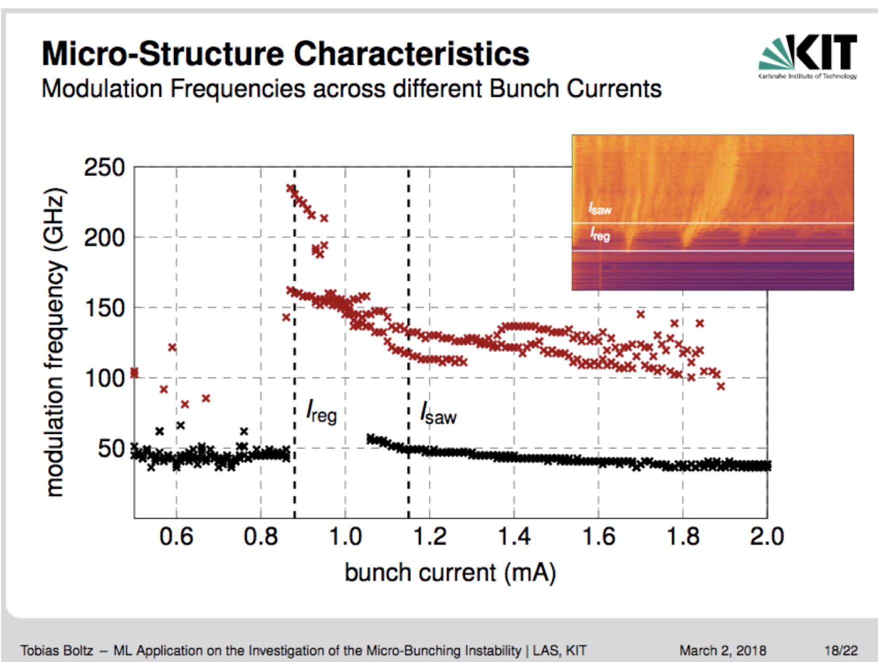
# Session: Data analysis

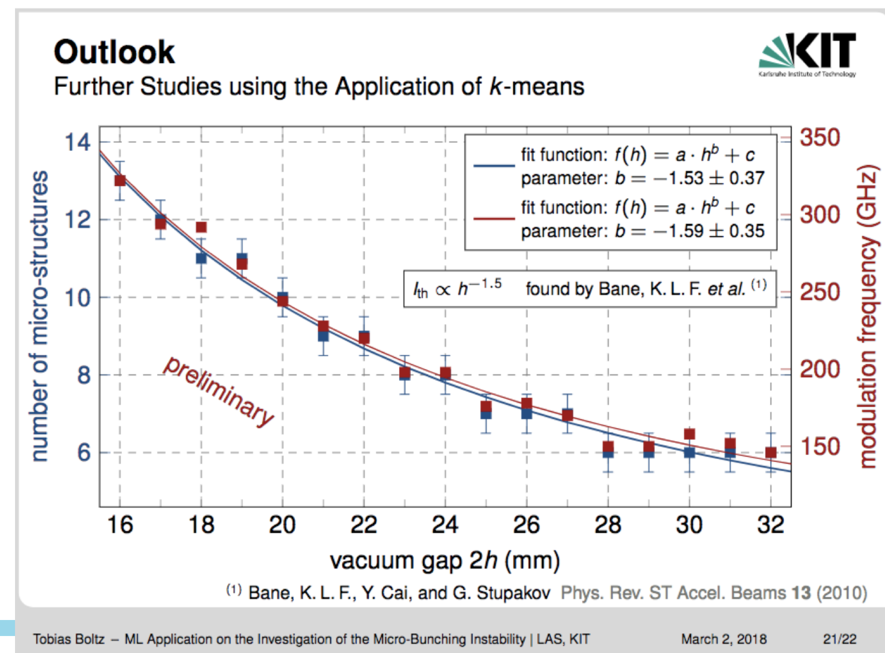
- What can we do with all the data we collect?

# T. Boltz (KIT): MBI structure in storage rings

- Micro-bunching instability (MBI) in storage rings with short e- bunches
- → Fluctuations in emitted CSR power
- Apply clustering algorithms to 1.5M bunch profiles:
  - Distinct micro-structures identified.



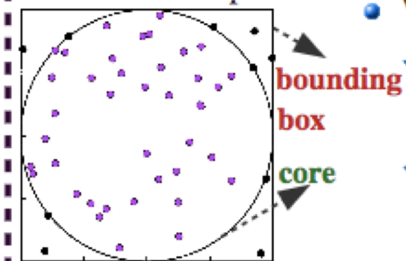



# T. Mohayai (IIT): Emittance measurement via KDE

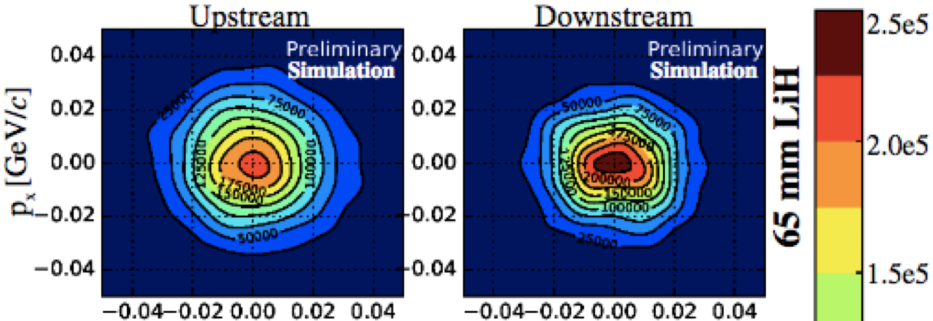
## KDE Density and Volume – MICE Baseline

- **Density**:
  - ★ Predict density using kernels centered at each muon in 4D phase space
  - ★ Extract (cluster/classify) the core contour (9th percentile)

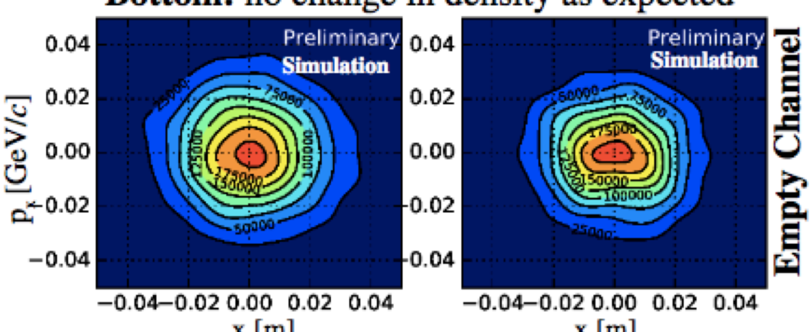
2D volume example

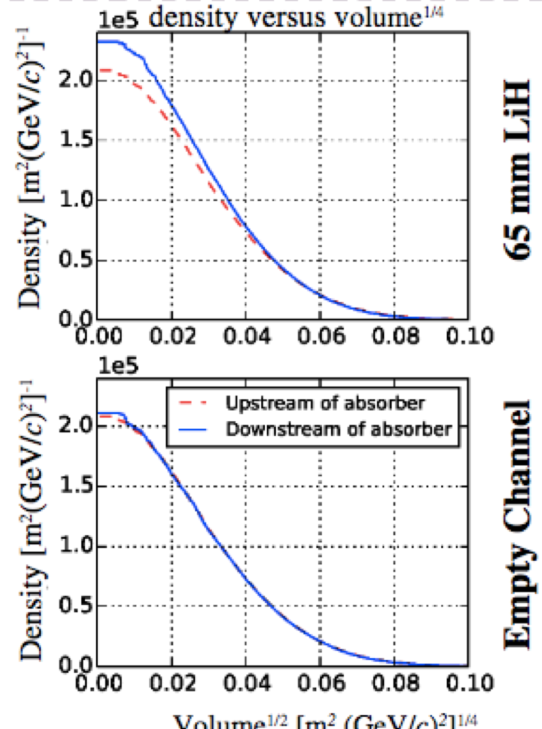



- **Volume**:
  - ★ Generate MC (Monte Carlo) points inside the box bounding the core
  - ★ Compute core volume as a fraction of MC points inside the box

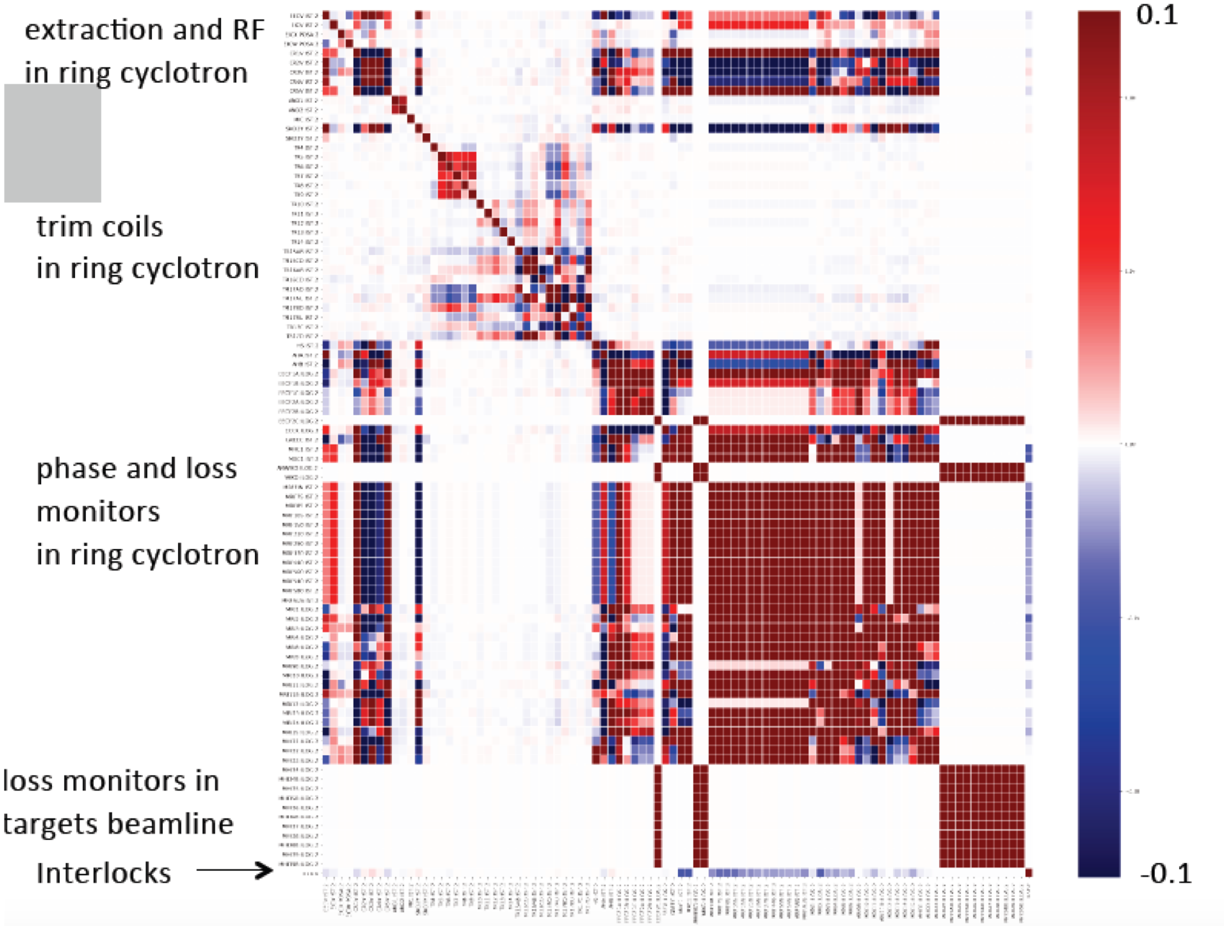



**Top:** density increase in beam core
**Bottom:** no change in density as expected



# J. Snuvernik (PSI): Data visualization




- Correlation matrix for beamline elements
- Ongoing work

# The next workshop: March 2019 at PSI

**A synopsis of the 2018 workshop will be published on the Arxiv soon.**

# Reviewing ML-assisted performance boosts:

- Automated tuning reduces tuning time by 50% at LCLS
- Online taper optimization doubled FEL peak power (J. Wu)
- Tune FEL up from noise w/ Bayesian priors (J. Duris)
- Optimize beam spectrum (G. Gaio)
- Improve control / reduce settling times for RFQ, gun (A. Edelen)
- Identify interlocks / component failures (A. Adelmann, E. Fol)
- Automate collimator alignment (G. Valentino)
- Identify & characterize micro-bunching structures (T. Boltz)
- Precisely measure beam emittances (T. Mohayai)
- Visualize complex datasets (J. Snuvernik)
- **This is not an exhaustive list!**

# ML experts are looking for collaboration

- If you have problems amenable to ML solutions, let's discuss.
- Modest resources available at Fermilab.

**Thanks for your attention!**