ICFA Mini-Workshop on Machine Learning for Accelerators

March 5, 2018

D. Ratner, et al.
SLAC National Accelerator Laboratory

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

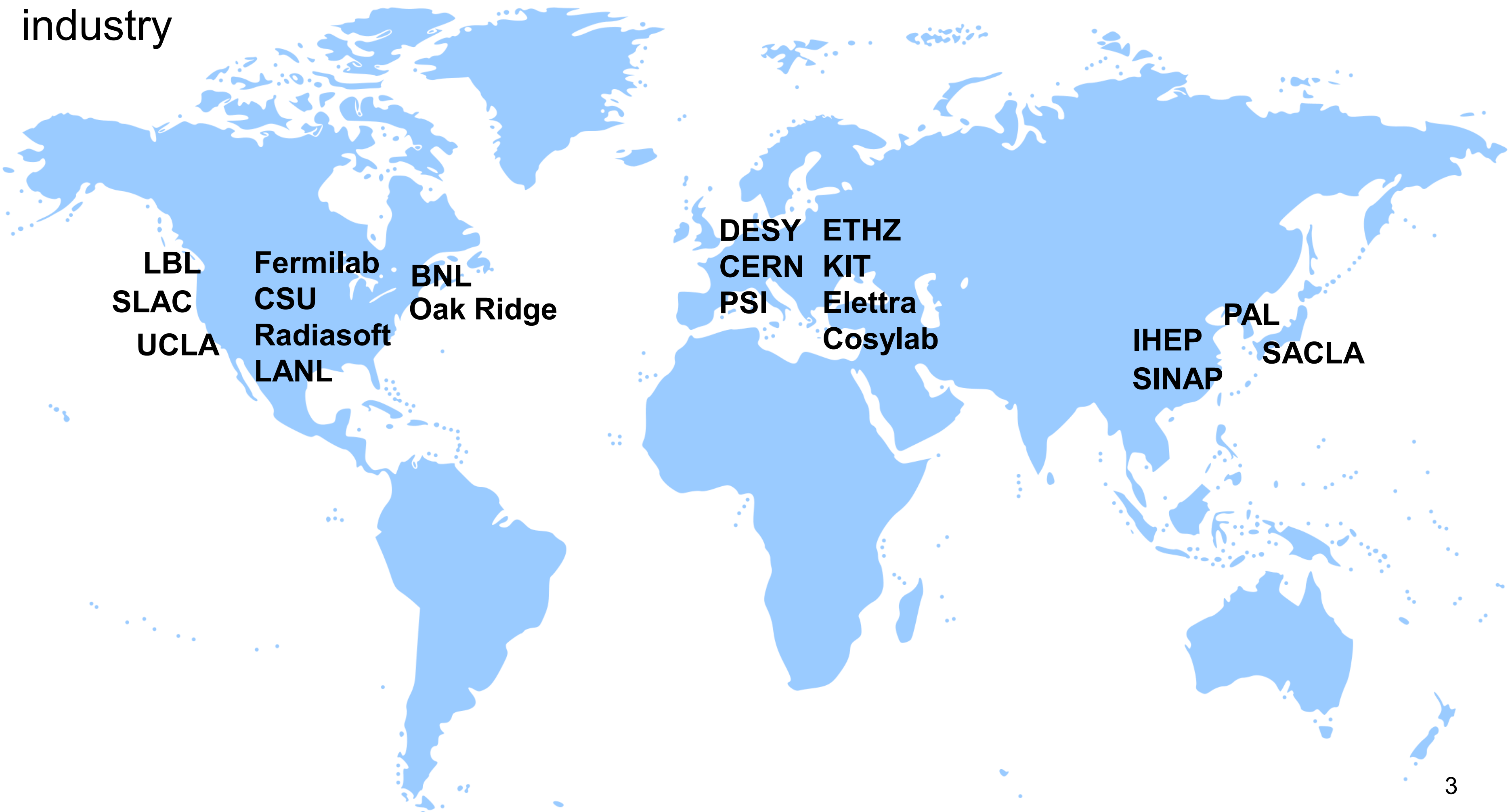
Topics

1. Facility needs
2. Optimization/tuning
3. Simulation/Modeling
4. Prognostics
5. Data analysis

2/27 TUESDAY	2/28 WEDNESDAY	3/1 THURSDAY	3/2 FRIDAY
Beginner/Intermediate Tutorial			
	Workshop intro (10 min), Session intro (15 min) (Kevin	NN Modeling (A. Edelen)	ML for MBI (Boltz)
	Facility LHC (Kajetan/Jorg)	Collective effects (Adelmann)	Non-parametric density estimators (Mohayai?)
	Facility XFEL (Raimund)	Light source simulations (Tomin)	Data mining at HIPA (Snuverink)
Coffee	Facility Synch (Xiaobiao)	Coffee/Poster session	Coffee/Poster session
	Coffee	Coffee/Poster session	Coffee/Poster session
	Fault detection (Nielsen)	machine modeling at FAST (J. Edelen)	Session summary
	Facility Other Side (Candel)	GANs (Oliviera)	Session summary
	Facility Discussion	Simulations/Modeling discussion 2	White paper/ Collaboration planning
Lunch	Lunch	Lunch	Lunch
Lunch	Lunch	Lunch	Lunch
	Reinforcement learning (Wu)	Prognostics: ML for anomaly detection in distributed	White paper/ Collaboration planning
Ocelot satellite meeting/ Tutorial	Gaussian Process 1 (Kirschner)	Beam loss plan recognition (Valentino)	White paper/ Collaboration planning
	Gaussian Process 2 (Duris)	Detection of bad bpms (Fol)	White paper/ Collaboration planning
	ML Tuning discussion	DESY zoom talk?	White paper/ Collaboration planning
	Poster blitz	Prognostics discussion	
	Poster blitz	Coffee/Poster session	
	Coffee/Poster session	Coffee/Poster session	
	Coffee/Poster session	Tour	
	XFEL tuning (Agapov)	Tour	
	Online opt (Scheinker)	Tour	
	opt with GA (Bazarov)	Tour	
	Model Free Discussion	Tour	
	Dutch Goose	Reception at the Meyer-Buck Estate	

Attendees

65 Participants from 20+ Institutions:
Specialties spanning computer science, physics, controls, operations, industry



Tutorial

Full day tutorial for machine learning novices

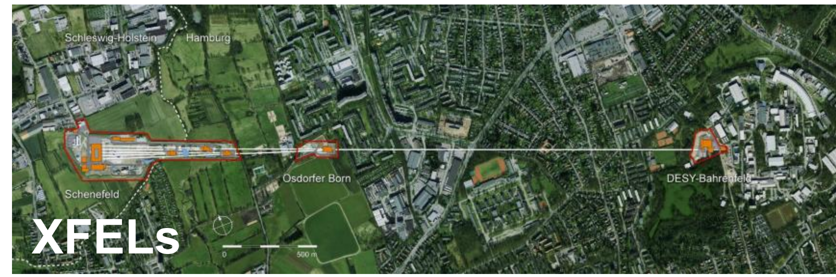
~60 participants learning basic ML and Ocelot platform

D. Bowring, A. Edelen, C. Mayes, I. Agapov, S. Tomin

Highlights: Facility needs

Some recurrent needs:

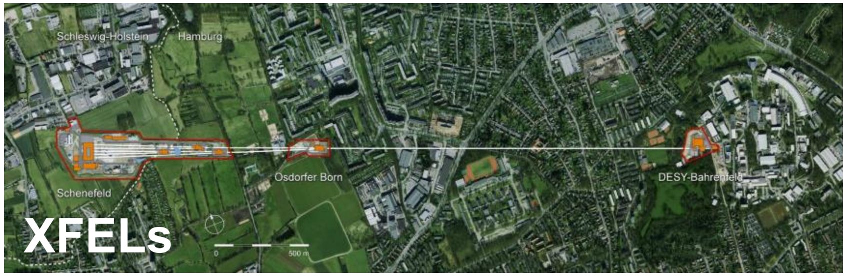
1. Identify broken parts, predict failures
2. Faster simulations, online models
3. Digging through large data sets for correlations, new physics



Highlights: Facility needs

Some recurrent needs:

1. Identify broken parts, predict failures
2. Faster simulations, online models
3. Digging through large data sets for correlations, new physics



General themes:

1. How do we identify where ML adds value?
2. Look for opportunities to collaborate

Highlights: Optimization

Wide agreement on need for automated tuning
Question: model-based or model-independent?

Highlights: Optimization

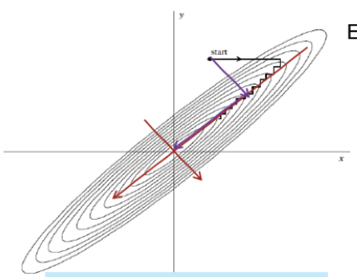
SLAC

Wide agreement on need for automated tuning
Question: model-based or model-independent?

Search over conjugate directions

RCDS: X. Huang

SLAC



Inefficient search directions

Efficient search directions: conjugate directions

A search over conjugate direction does not invalidate previous searches.

Directions $\mathbf{u}$ and $\mathbf{v}$ are conjugate if

$$\mathbf{u}^{\mathrm{T}} \cdot \mathbf{H} \cdot \mathbf{v} = 0$$

with $\mathbf{H}$ being the Hessian matrix of function $f(\mathbf{x})$, $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Around the minimum

$$f(\mathbf{x}_m + \Delta\mathbf{x}) = f(\mathbf{x}_m) + \frac{1}{2}\Delta\mathbf{x}^{\mathrm{T}} \cdot \mathbf{H} \cdot \Delta\mathbf{x}.$$

It takes many tiny steps to get to the minimum when searching along x and y directions.

Powell's method can update the directions using past search results to develop a conjugate set.

*W.H. Press, et al, Numerical Recipes

*M.J.D. Powell, Computer Journal 7 (2) 1965 155

X. Huang (SLAC), Online optimization, ML Workshop 2018 4

SPEAR3: coupling correction, kicker bump matching, dynamic aperture, transport line steering and optics, etc
X. Huang, et al, NIMA 726 (2013) 77; X. Huang et al, PRSTAB 18, 084001 (2015)

LCLS: taper optimization
J. Wu, et al, FEL 2017

ESRF: storage ring coupling correction, beam lifetime, injection efficiency
S. M. Liuzzo, et al, IPAC'2016, THPMR015

IHEP: BEPC-II luminosity optimization, CSNS/RCS collimation system
H.-F. Ji, et al, Chinese Physics C 39, 12, 127006 (2015); H.-F. Ji, et al, HB2016

DIAMOND: nonlinear beam dynamics optimization
I. Martin, et al, IPAC'2016, THPMR001

NSLS-II: Injection kicker bump matching
G.M. Wang, et al, IPAC'2017, THPVA095

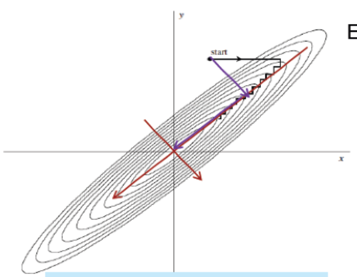
SPS: Injection optimization, coupling optimization
T. Pulampong, et al, IPAC'2017, THPAB151

Highlights: Optimization

SLAC

Wide agreement on need for automated tuning
Question: model-based or model-independent?

Search over conjugate directions — RCDS: X. Huang



Inefficient search directions

It takes many tiny steps to get to the minimum when searching along x and y directions.

Efficient search directions: conjugate directions

A search over conjugate direction does not invalidate previous searches.

Directions $\mathbf{u}$ and $\mathbf{v}$ are conjugate if

$$\mathbf{u}^{\mathrm{T}} \cdot \mathbf{H} \cdot \mathbf{v} = 0$$

with $\mathbf{H}$ being the Hessian matrix of function $f(\mathbf{x})$, $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Around the minimum

$$f(\mathbf{x}_m + \Delta\mathbf{x}) = f(\mathbf{x}_m) + \tfrac{1}{2}\Delta\mathbf{x}^{\mathrm{T}} \cdot \mathbf{H} \cdot \Delta\mathbf{x}.$$

Powell's method can update the directions using past search results to develop a conjugate set.

*W.H. Press, et al, Numerical Recipes

*M.J.D. Powell, Computer Journal 7 (2) 1965 155

X. Huang (SLAC), Online optimization, ML Workshop 2018 4

SPEAR3: coupling correction, kicker bump matching, dynamic aperture, transport line steering and optics, etc
X. Huang, et al, NIMA 726 (2013) 77; X. Huang et al, PRSTAB 18, 084001 (2015)

LCLS: taper optimization
J. Wu, et al, FEL 2017

ESRF: storage ring coupling correction, beam lifetime, injection efficiency
S. M. Liuzzo, et al, IPAC'2016, THPMR015

IHEP: BEPC-II luminosity optimization, CSNS/RCS collimation system
H.-F. Ji, et al, Chinese Physics C 39, 12, 127006 (2015); H.-F. Ji, et al, HB2016

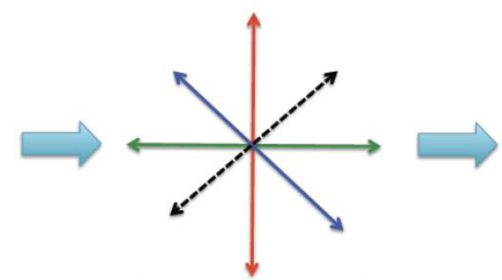
DIAMOND: nonlinear beam dynamics optimization
I. Martin, et al, IPAC'2016, THPMR001

NSLS-II: Injection kicker bump matching
G.M. Wang, et al, IPAC'2017, THPVA095

SPS: Injection optimization, coupling optimization
T. Pulampong, et al, IPAC'2017, THPAB151

Extremum seeking: A. Scheinker

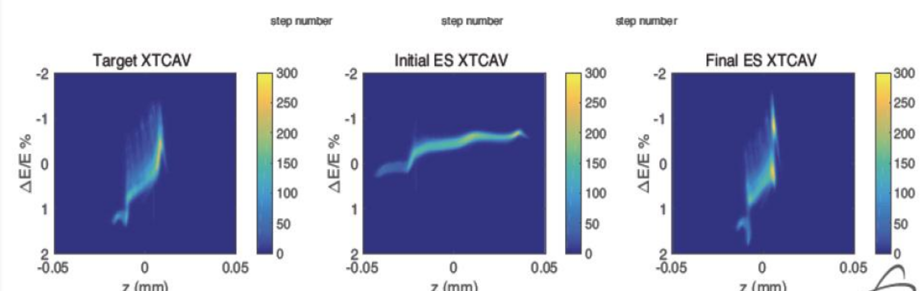
$$\frac{dp_1}{dt} = \sqrt{\alpha\omega_1}\cos(\omega_1 t + ky)$$
$$\frac{dp_2}{dt} = \sqrt{\alpha\omega_2}\cos(\omega_2 t + ky)$$
$$\frac{dp_3}{dt} = \sqrt{\alpha\omega_3}\cos(\omega_3 t + ky)$$
$$\vdots$$
$$\frac{dp_m}{dt} = \sqrt{\alpha\omega_m}\cos(\omega_m t + ky)$$



$$\frac{d\mathbf{p}}{dt} = -\frac{k\alpha}{2}\left(\nabla_{\mathbf{p}} V(\mathbf{x},t)\right)^T$$

On average, the system performs a gradient descent of the unknown, time-varying function V(x,t)

Allows simultaneous tuning of ALL parameters in parallel.



Los Alamos NATIONAL LABORATORY

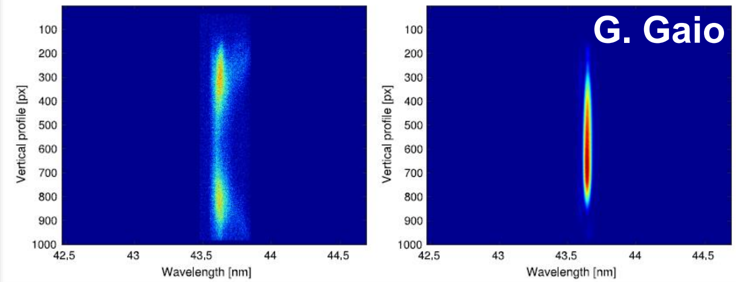
Alexander Scheinker (ascheink@lanl.gov) Model-independent Feedback Control & Optimization

Highlights: Optimization

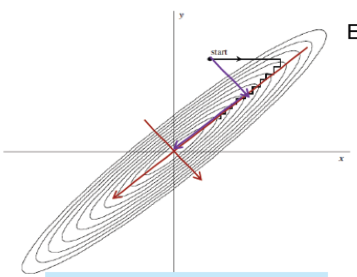
SLAC

Wide agreement on need for automated tuning
Question: model-based or model-independent?

Example: FEL "quality" at FERMI



Search over conjugate directions — RCDS: X. Huang



Inefficient search directions

It takes many tiny steps to get to the minimum when searching along x and y directions.

Efficient search directions: conjugate directions

A search over conjugate direction does not invalidate previous searches.

Directions $\mathbf{u}$ and $\mathbf{v}$ are conjugate if

$$\mathbf{u}^T \cdot \mathbf{H} \cdot \mathbf{v} = 0$$

with $\mathbf{H}$ being the Hessian matrix of function $f(\mathbf{x})$, $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$.

Around the minimum

$$f(\mathbf{x}_m + \Delta\mathbf{x}) = f(\mathbf{x}_m) + \frac{1}{2}\Delta\mathbf{x}^T \cdot \mathbf{H} \cdot \Delta\mathbf{x}.$$

Powell's method can update the directions using past search results to develop a conjugate set.

*W.H. Press, et al, Numerical Recipes

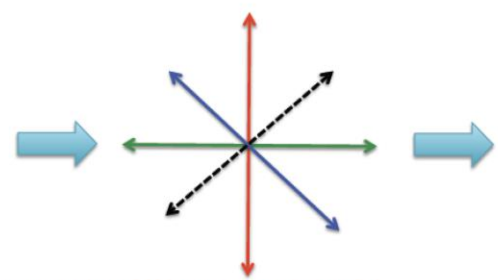
*M.J.D. Powell, Computer Journal 7 (2) 1965 155

X. Huang (SLAC), Online optimization, ML Workshop 2018 4

SPEAR3: coupling correction, kicker bump matching, dynamic aperture, transport line steering and optics, etc
X. Huang, et al, NIMA 726 (2013) 77; X. Huang et al, PRSTAB 18, 084001 (2015)
LCLS: taper optimization
J. Wu, et al, FEL 2017
ESRF: storage ring coupling correction, beam lifetime, injection efficiency
S. M. Liuzzo, et al, IPAC'2016, THPMR015
IHEP: BEPC-II luminosity optimization, CSNS/RCS collimation system
H.-F. Ji, et al, Chinese Physics C 39, 12, 127006 (2015); H.-F. Ji, et al, HB2016
DIAMOND: nonlinear beam dynamics optimization
I. Martin, et al, IPAC'2016, THPMR001
NSLS-II: Injection kicker bump matching
G.M. Wang, et al, IPAC'2017, THPVA095
SPS: Injection optimization, coupling optimization
T. Pulampong, et al, IPAC'2017, THPAB151

Extremum seeking: A. Scheinker

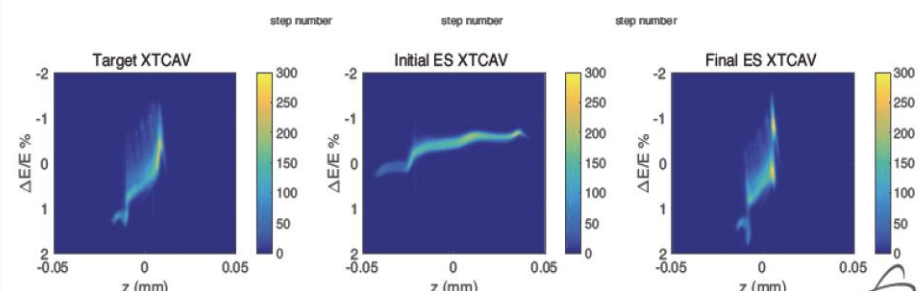
$$\frac{dp_1}{dt} = \sqrt{\alpha\omega_1}\cos(\omega_1 t + ky)$$
$$\frac{dp_2}{dt} = \sqrt{\alpha\omega_2}\cos(\omega_2 t + ky)$$
$$\frac{dp_3}{dt} = \sqrt{\alpha\omega_3}\cos(\omega_3 t + ky)$$
$$\vdots$$
$$\frac{dp_m}{dt} = \sqrt{\alpha\omega_m}\cos(\omega_m t + ky)$$



$$\frac{dp}{dt} = -\frac{k\alpha}{2}\left(\nabla_{\mathbf{p}} V(\mathbf{x},t)\right)^T$$

On average, the system performs a gradient descent of the unknown, time-varying function V(**x**,t)

Allows simultaneous tuning of ALL parameters in parallel.



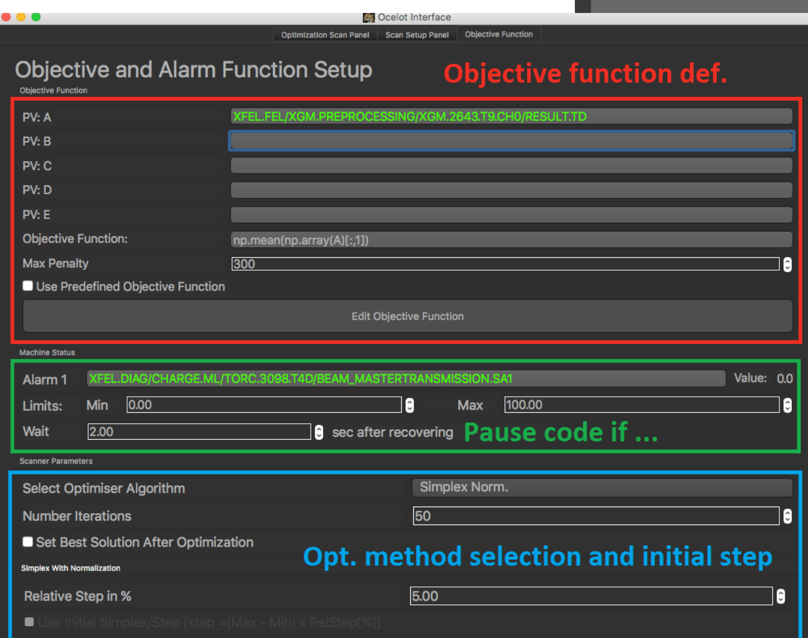
Alexander Scheinker (ascheink@lanl.gov) Model-independent Feedback Control & Optimization

Los Alamos NATIONAL LABORATORY

Highlights: Optimization

SLAC

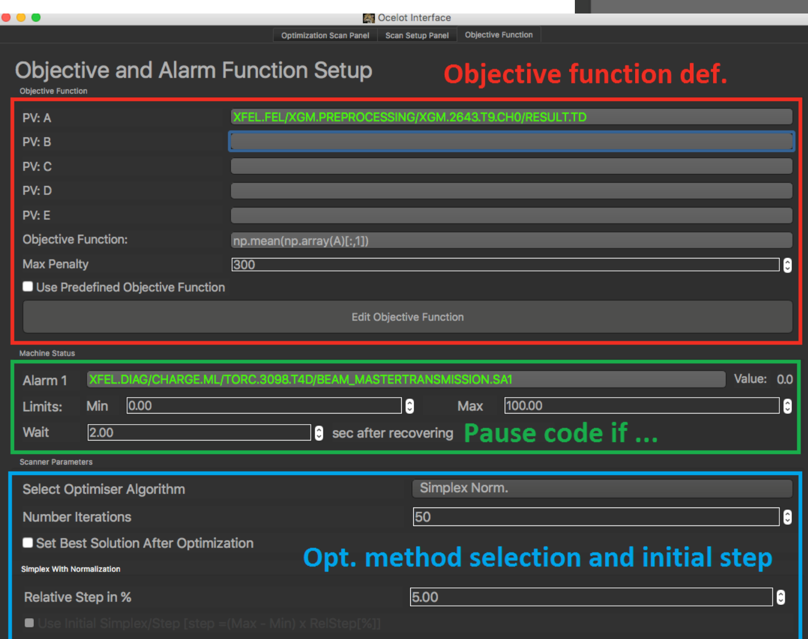
Tuning platforms:
Ocelot (DESY) provides generic base for accelerator optimization/simulation
➔ Now multi-lab collaboration



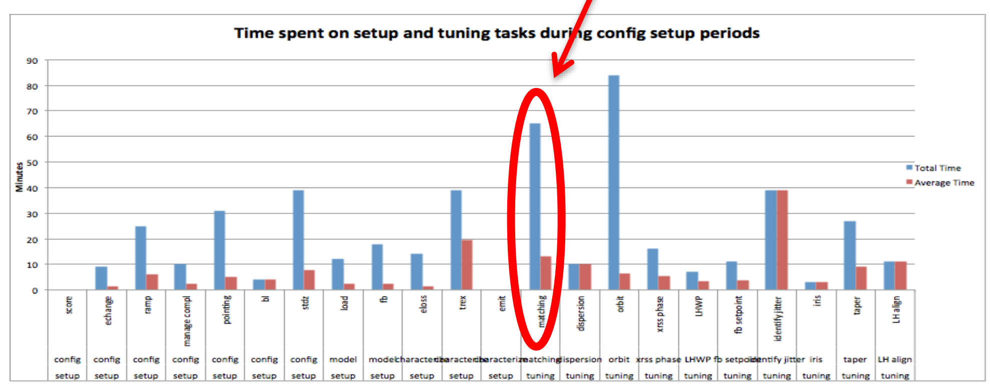
S. Tomin, I. Agapov, DESY/XFEL

Highlights: Optimization

Tuning platforms:
Ocelot (DESY) provides generic base for accelerator optimization/simulation
➔ Now multi-lab collaboration



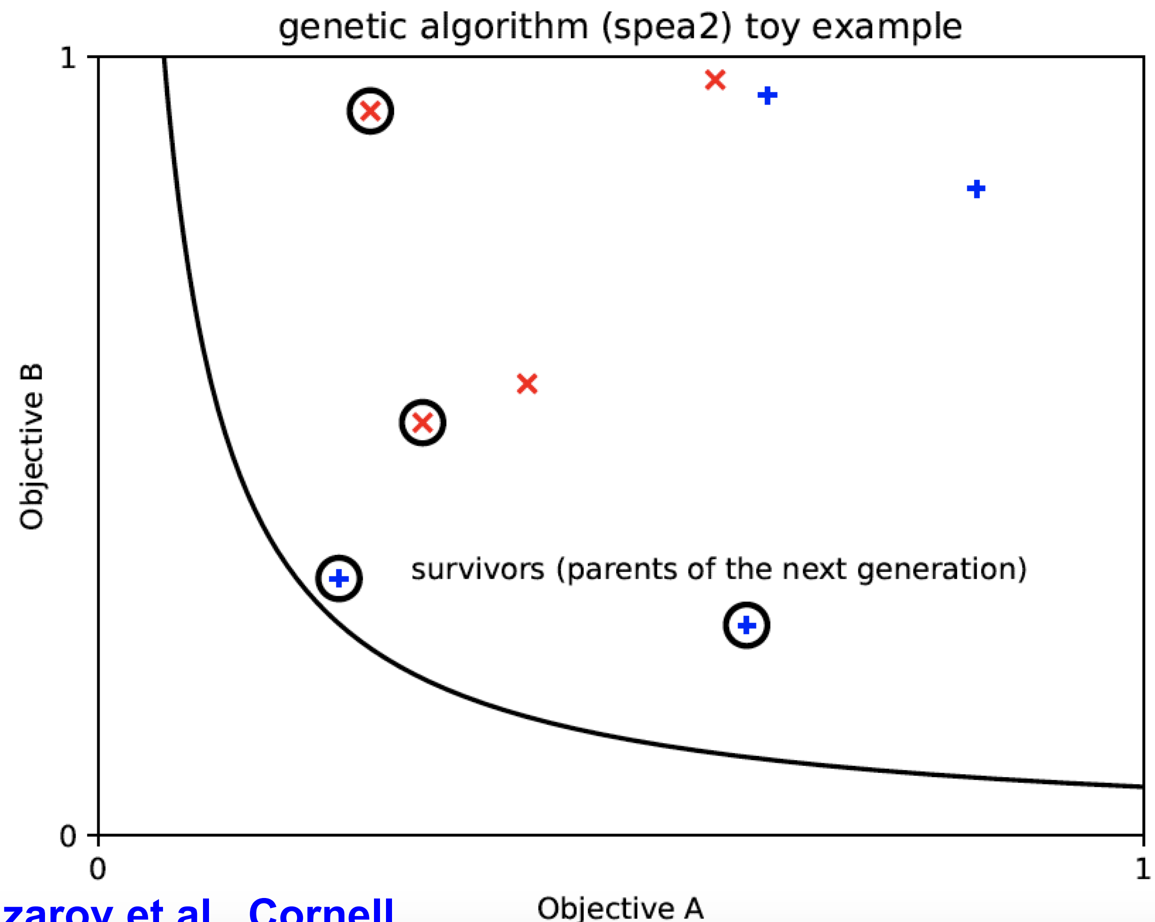
At LCLS: 450 hours/year in 2016
➔ Automated tuning cut avg time by half in 2017!



S. Tomin, I. Agapov, DESY/XFEL

Highlights: Optimization

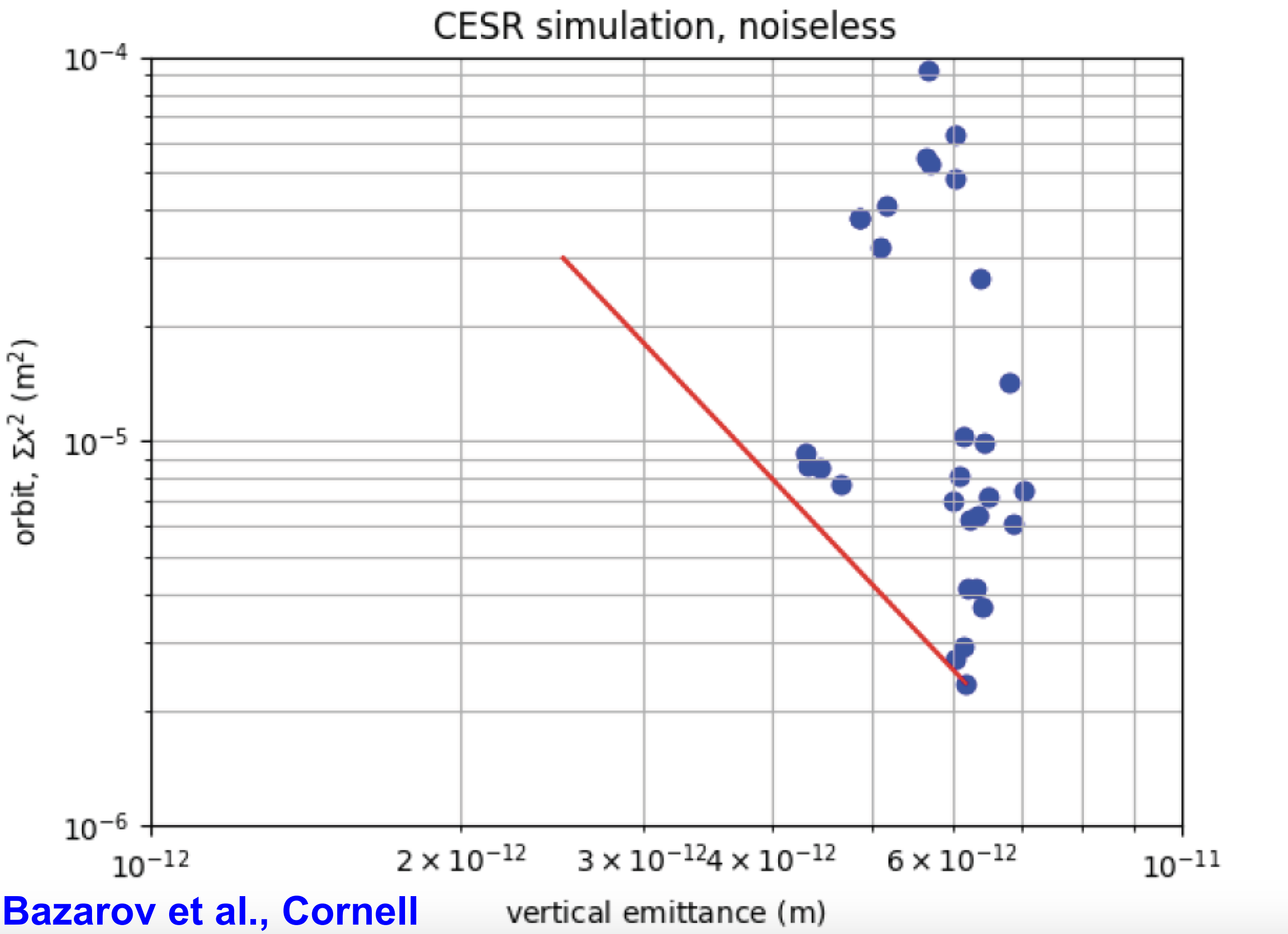
Genetic algorithms to find optimize and find pareto frontier



I. Bazarov et al., Cornell

Highlights: Optimization

Genetic algorithms to find optimize and find pareto frontier

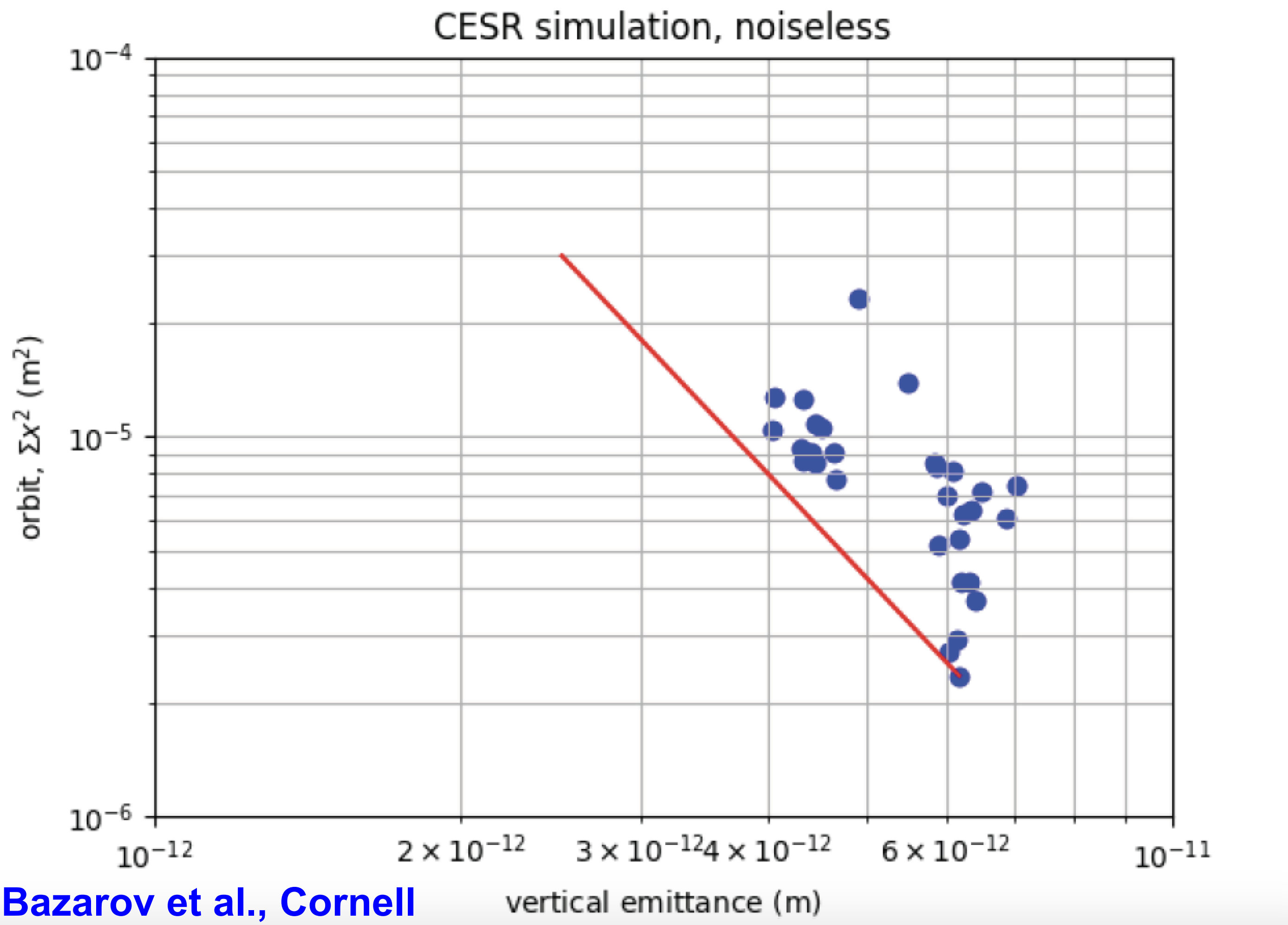


I. Bazarov et al., Cornell

Highlights: Optimization

SLAC

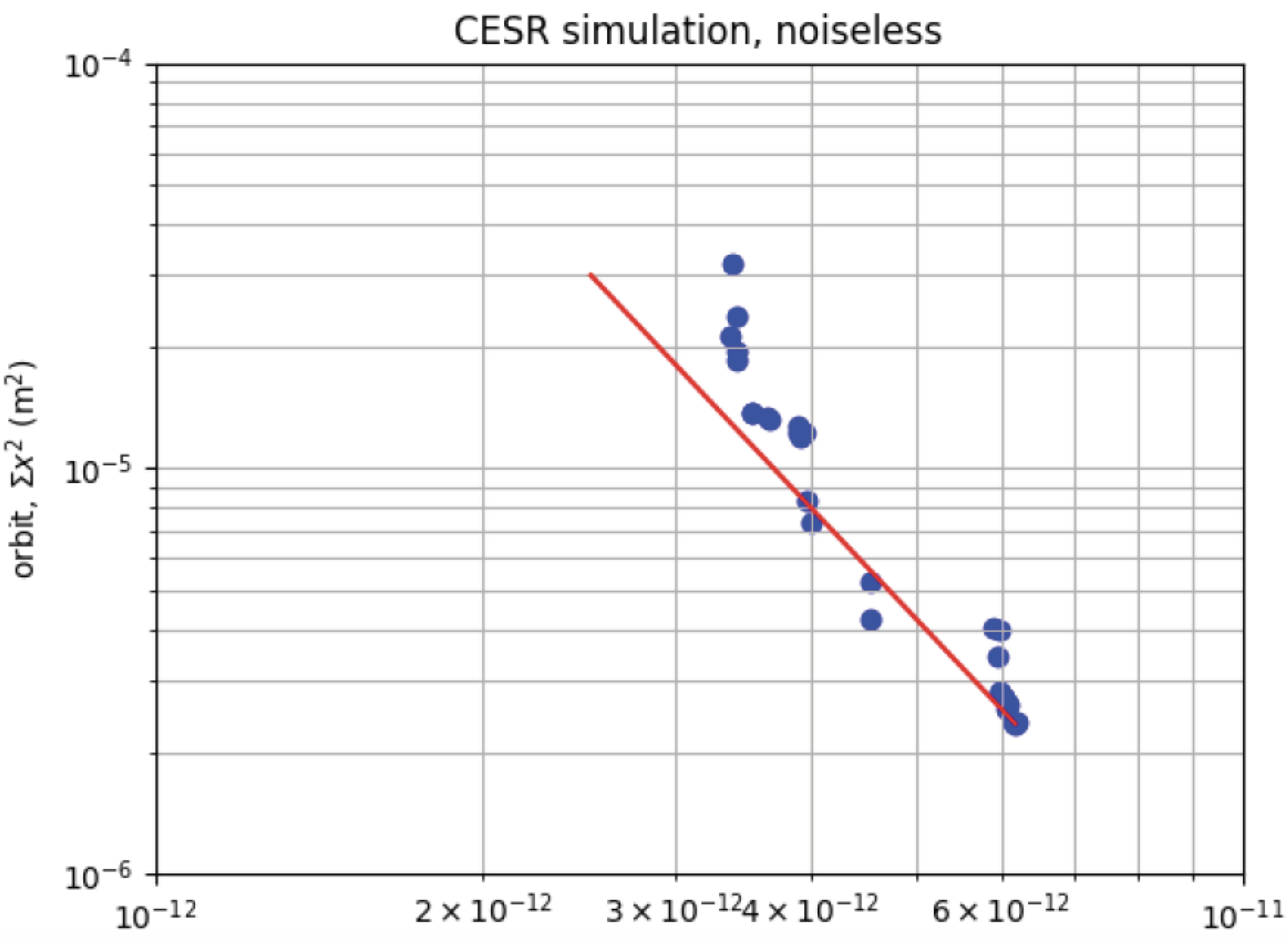
Genetic algorithms to find optimize and find pareto frontier



I. Bazarov et al., Cornell

Highlights: Optimization

Genetic algorithms to find optimize and find pareto frontier

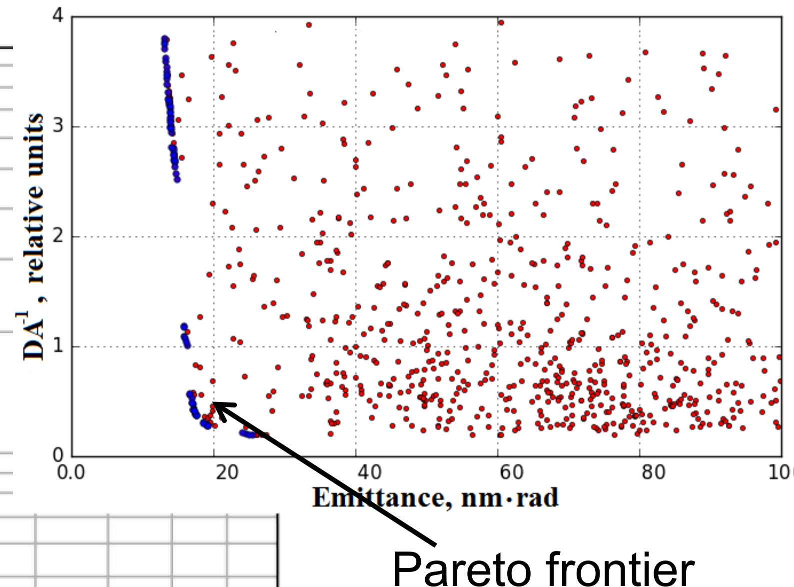


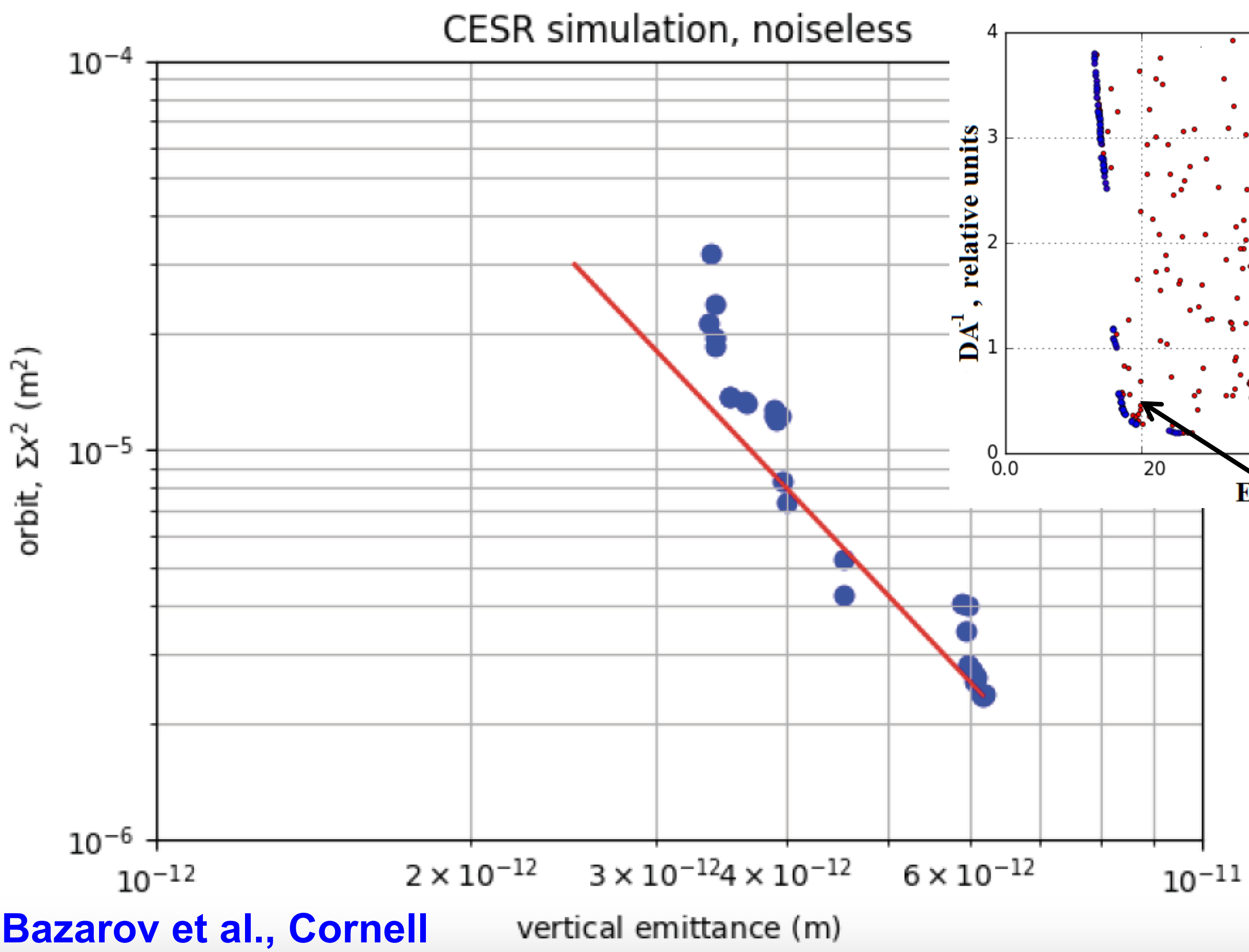
I. Bazarov et al., Cornell

Highlights: Optimization

Genetic algorithms to find optimize and find pareto frontier

I. Agapov, DESY

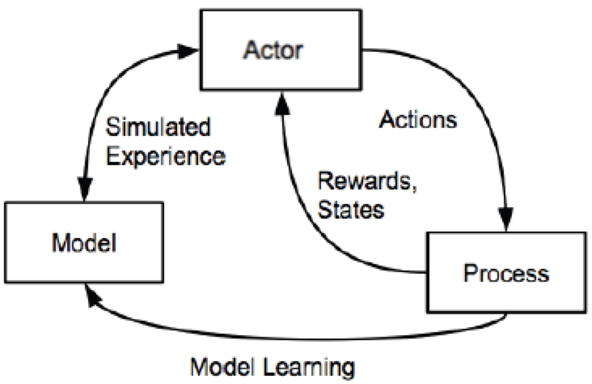




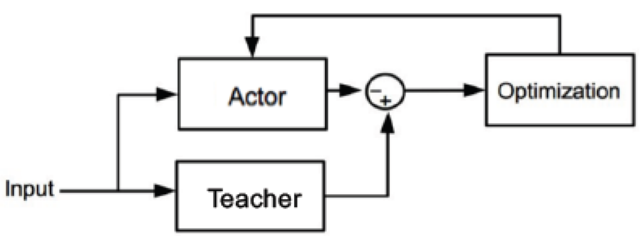
I. Bazarov et al., Cornell

Highlights: Optimization

Model-based: Reinforcement learning (used by alpha-GO)



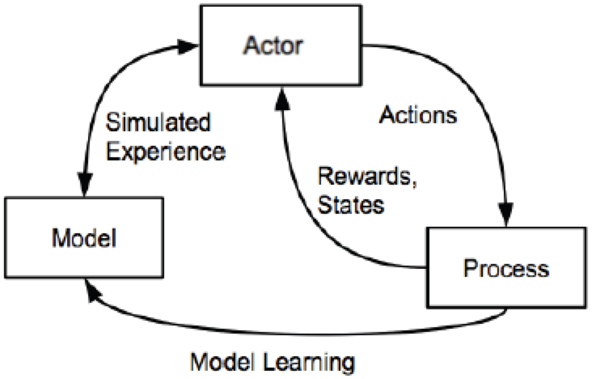
Can train on models first to get a good initial solution before deployment



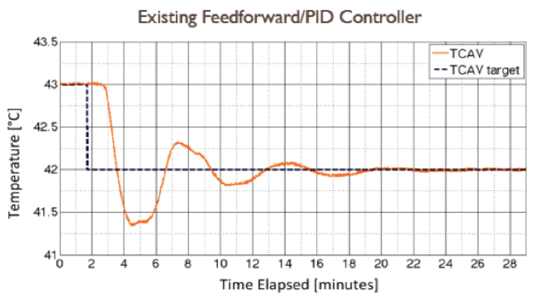
Can use supervised learning to first approximate the behavior of a different control policy

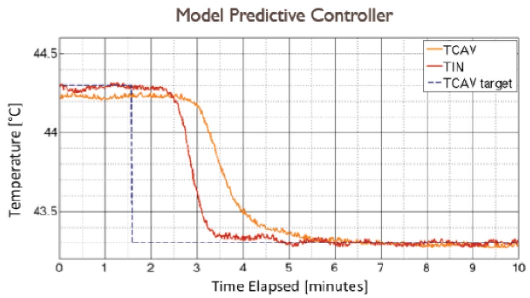
A. Edelen, CSU

Model-based: Reinforcement learning (used by alpha-GO)



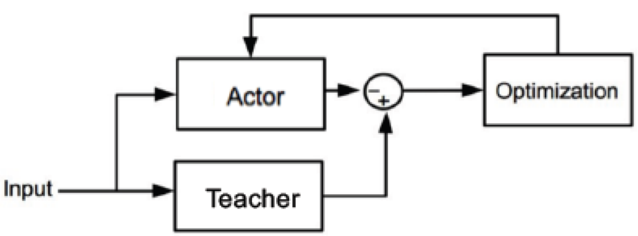
Can train on models first to get a good initial solution before deployment





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

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

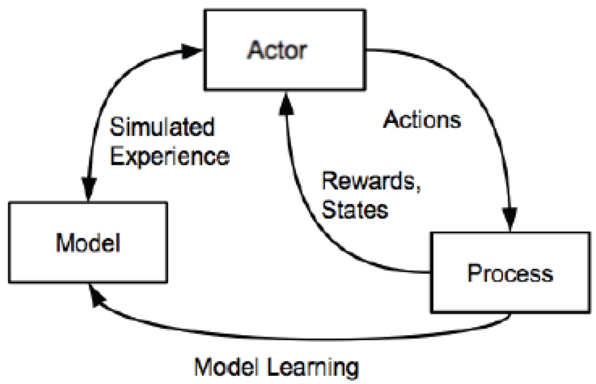


Can use supervised learning to first approximate the behavior of a different control policy

A. Edelen, CSU

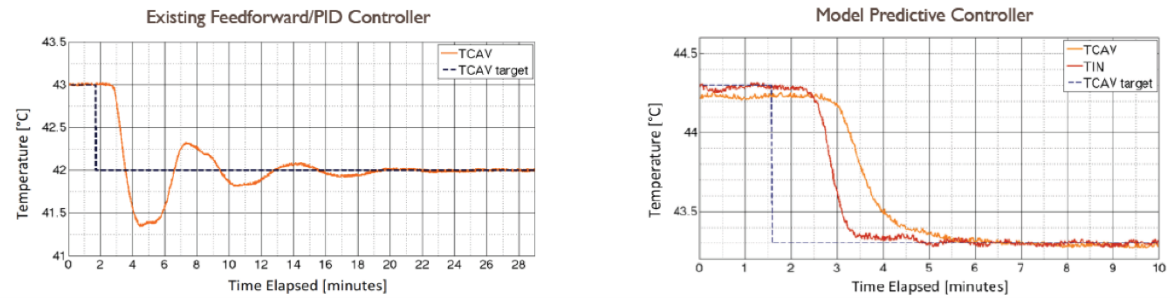
Highlights: Optimization

Model-based: Reinforcement learning (used by alpha-GO)



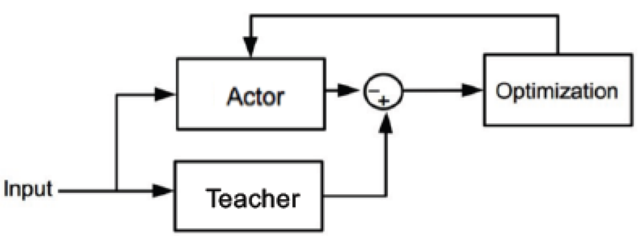
Can train on models first to get a good initial solution before deployment

Temperature control, RF photo-injector at FAST



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

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



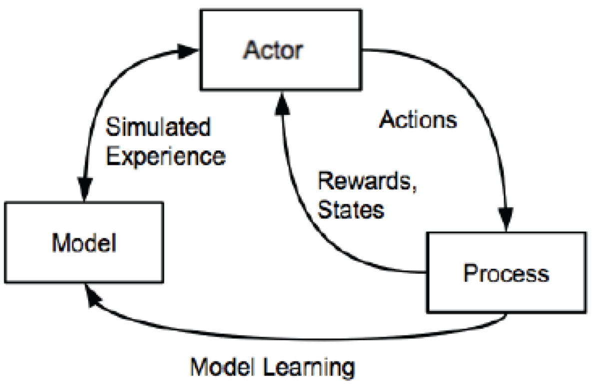
Can use supervised learning to first approximate the behavior of a different control policy

A. Edelen, CSU

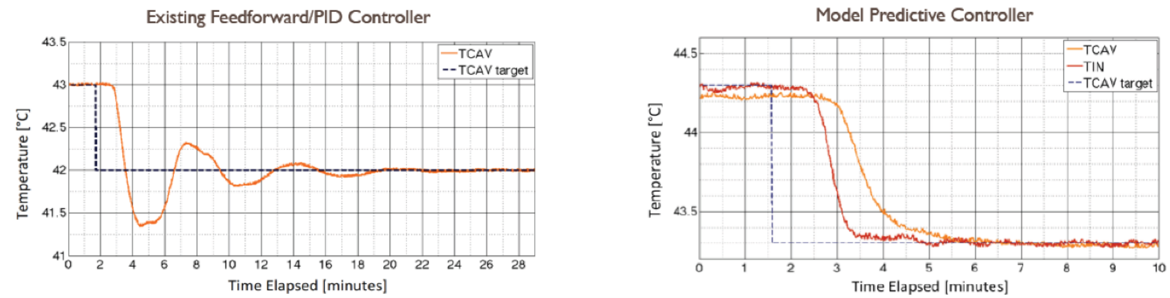
Highlights: Optimization

SLAC

Model-based: Reinforcement learning (used by alpha-GO)



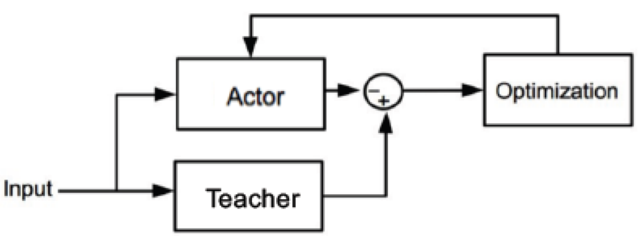
Temperature control, RF photo-injector at FAST



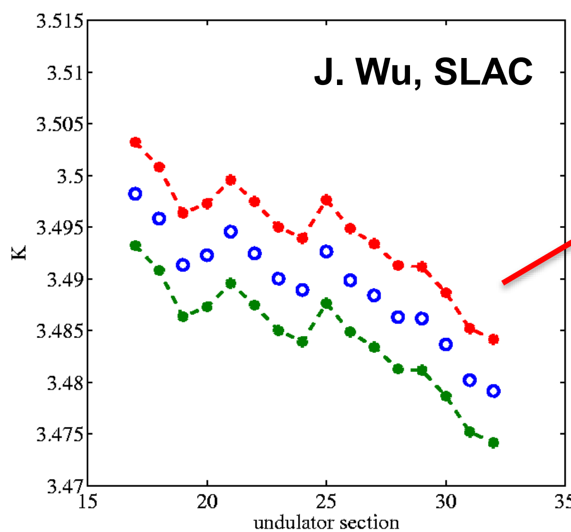
Note that the oscillations are largely due to the transport delays and water recirculation, rather than PID gains

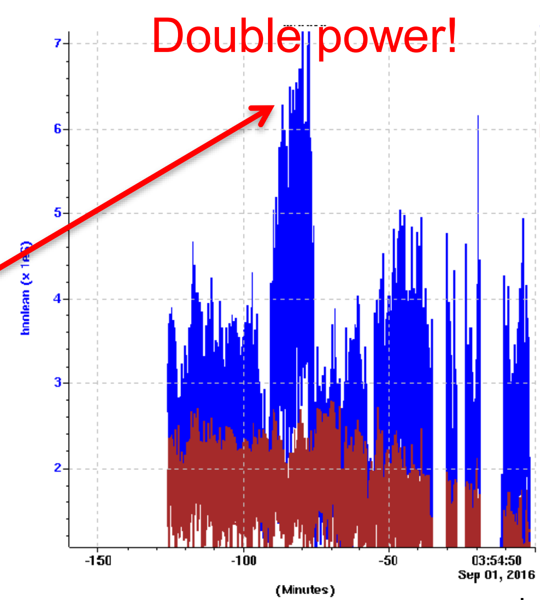
A. L. Edelen et al., TNS, vol. 63, no. 2, 2016　*A.L. Edelen et al., IPAC '15*

Can train on models first to get a good initial solution before deployment



Can use supervised learning to first approximate the behavior of a different control policy

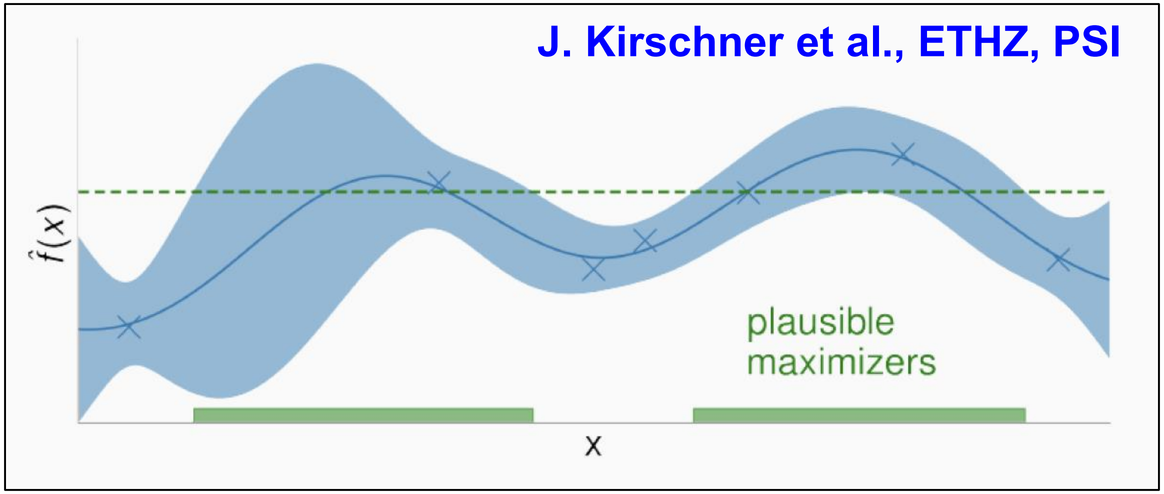


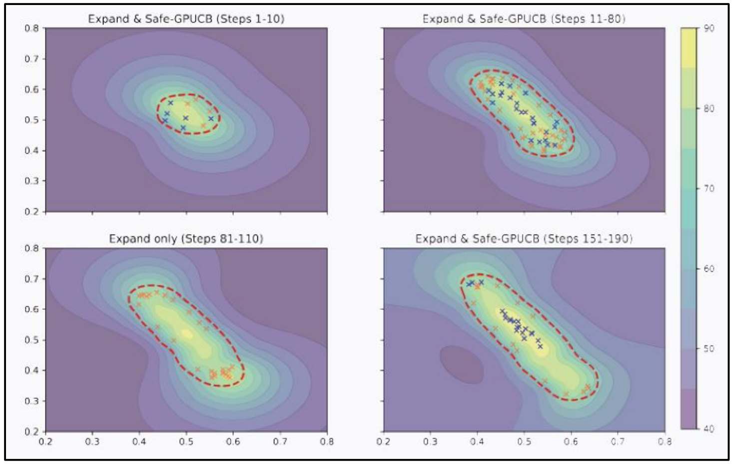


A. Edelen, CSU

Highlights: Optimization

Model-based: Bayesian optimizers (Gaussian Process Model)



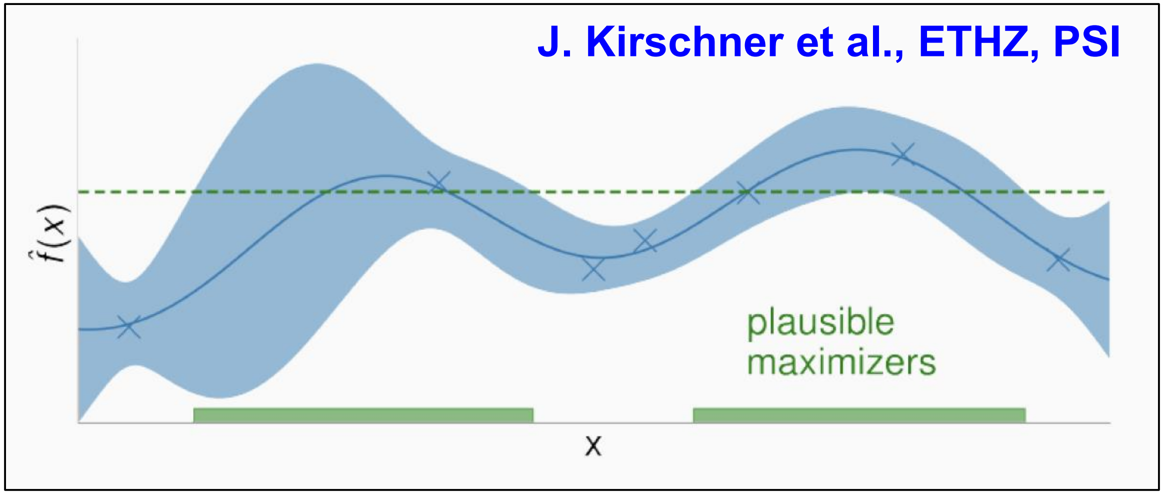


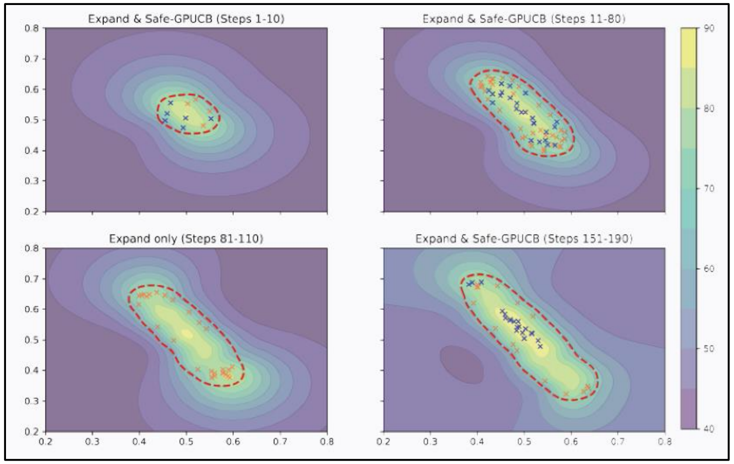
Safe optimization

Highlights: Optimization

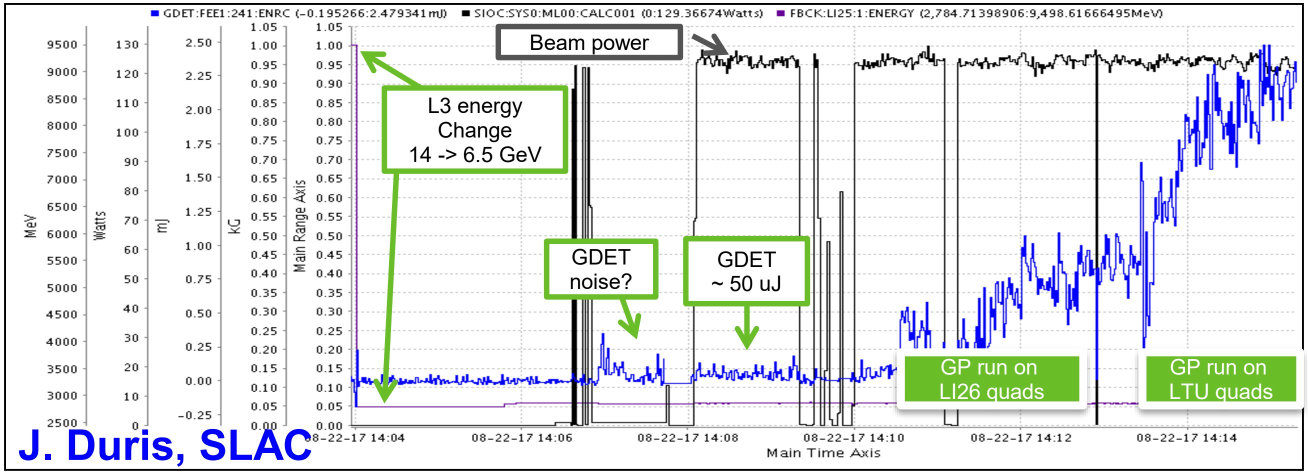
Model-based: Bayesian optimizers (Gaussian Process Model)

J. Kirschner et al., ETHZ, PSI





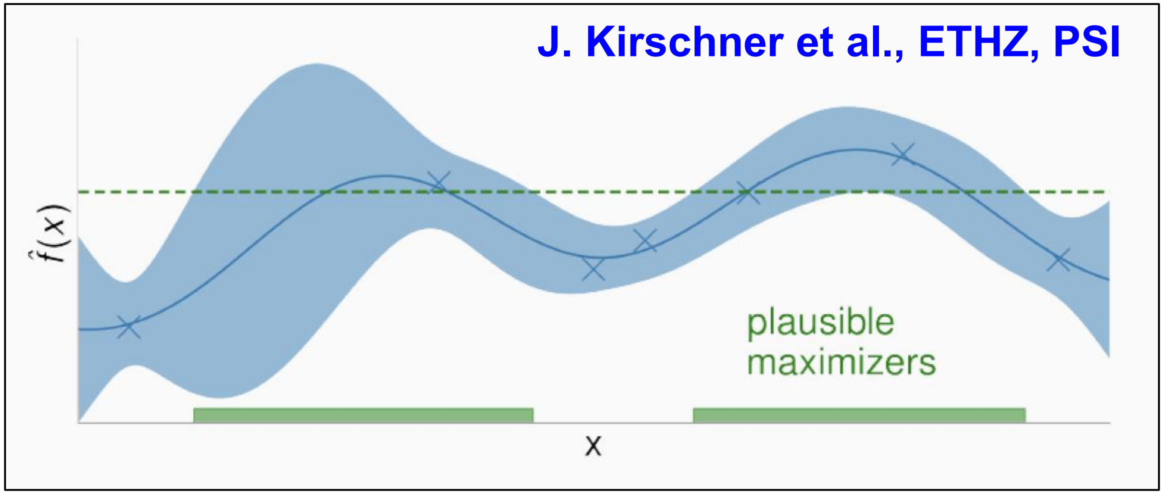
Safe optimization

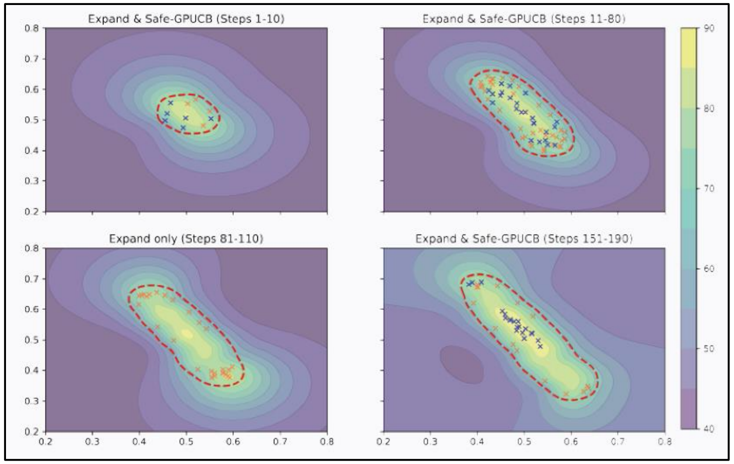


J. Duris, SLAC

Highlights: Optimization

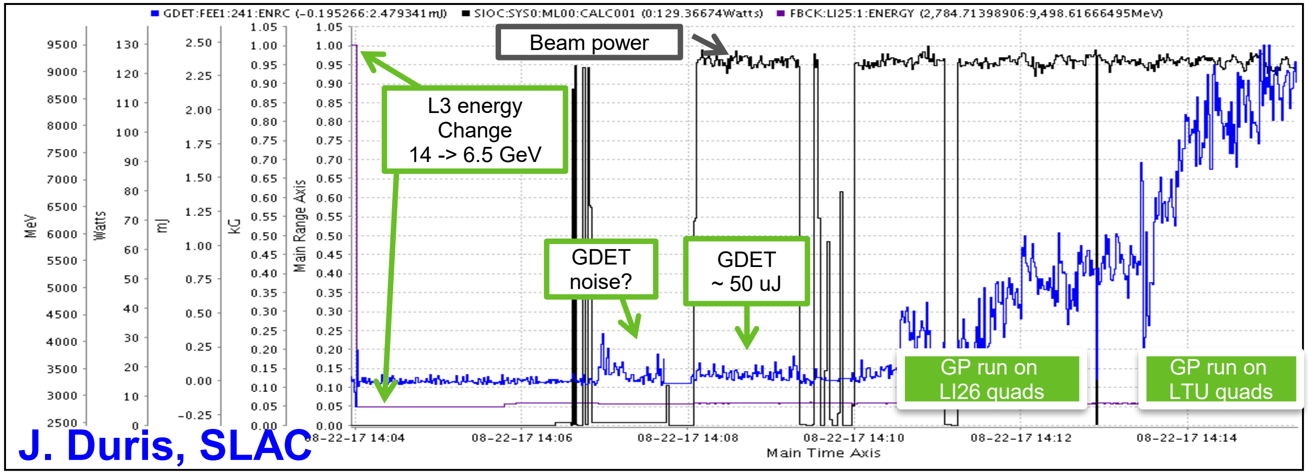
Model-based: Bayesian optimizers (Gaussian Process Model)





Safe optimization

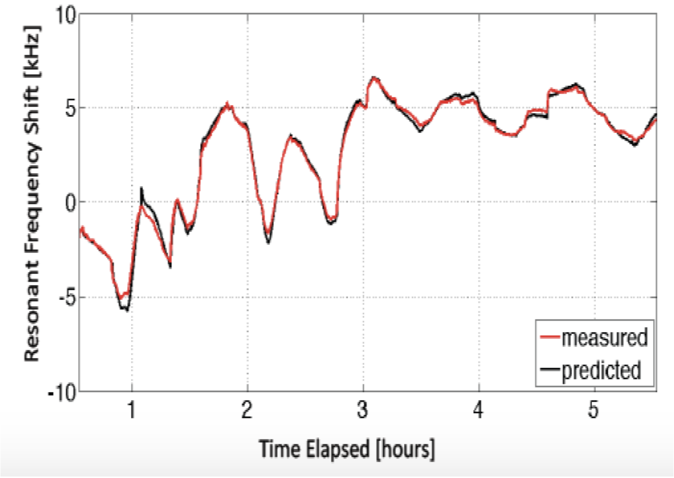
Train on archive
➔ Tune from noise!

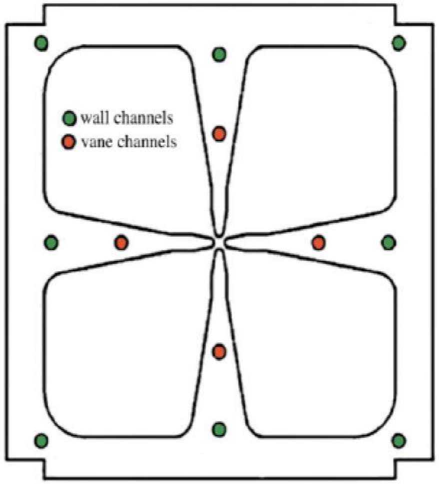


Highlights: Modeling/Simulations

Modeling systems with neural networks: Examples at FAST

Predict RFQ frequency



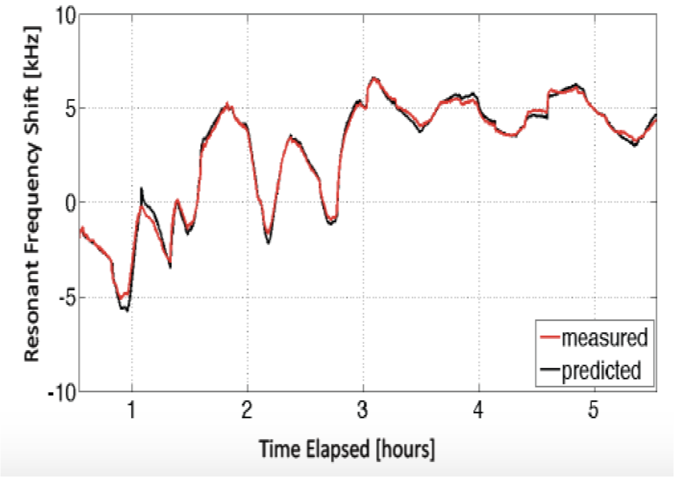


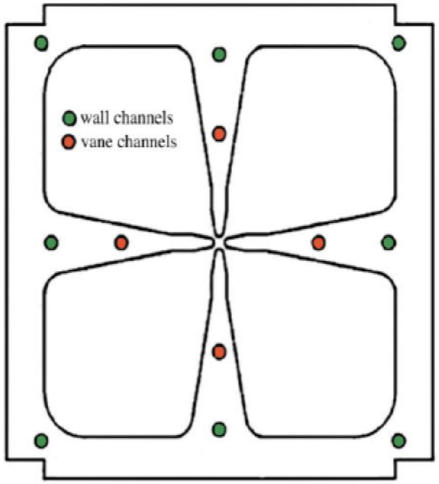
A. Edelen

Highlights: Modeling/Simulations

Modeling systems with neural networks: Examples at FAST

Predict RFQ frequency



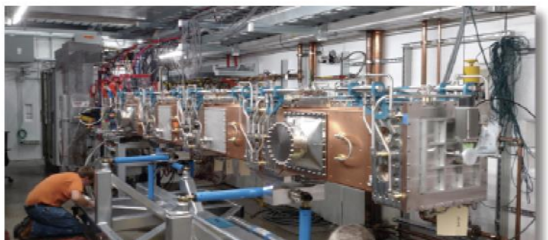


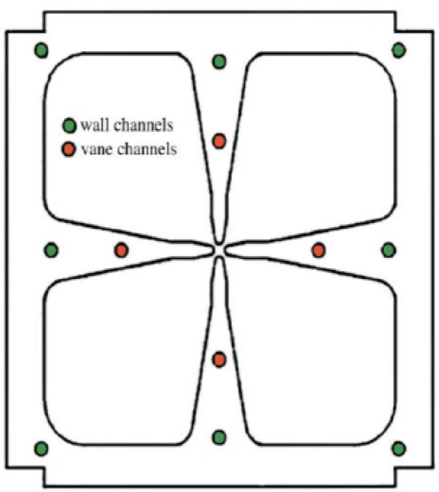
A. Edelen

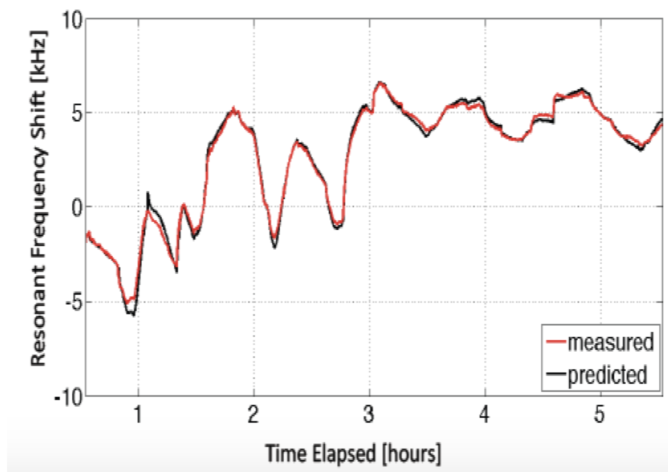
J. Edelen

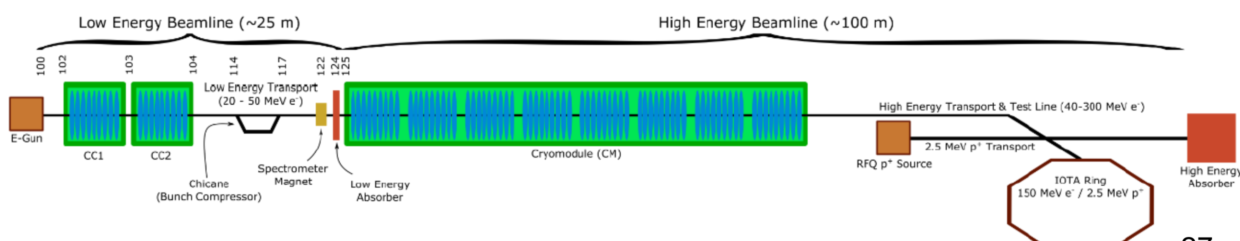
Modeling systems with neural networks: Examples at FAST

Predict RFQ frequency









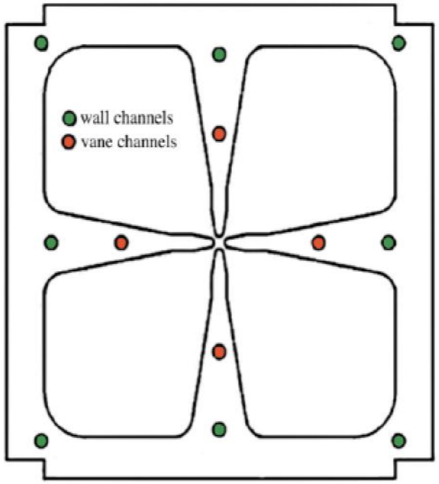
A. Edelen

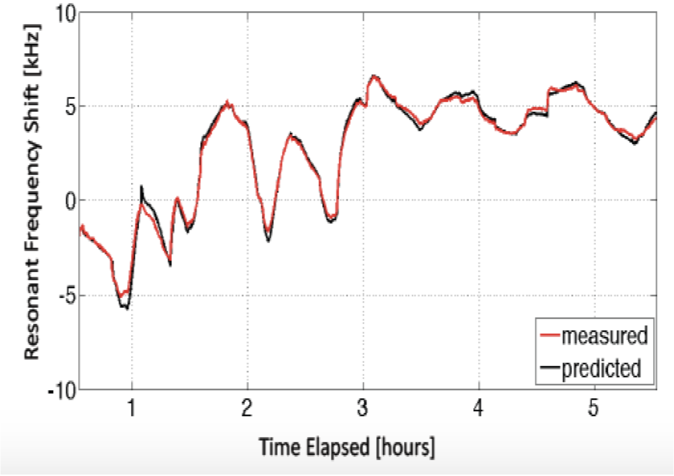
J. Edelen

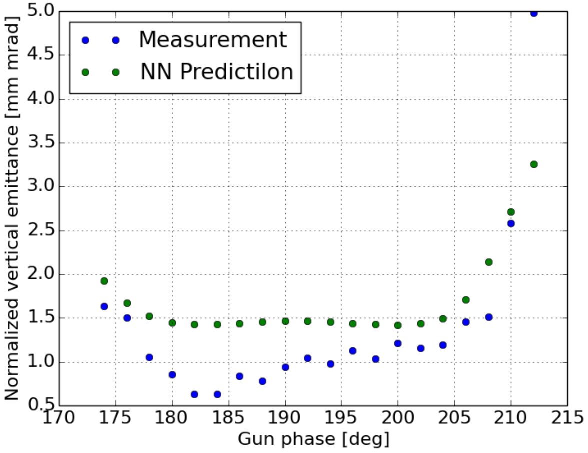
Highlights: Modeling/Simulations

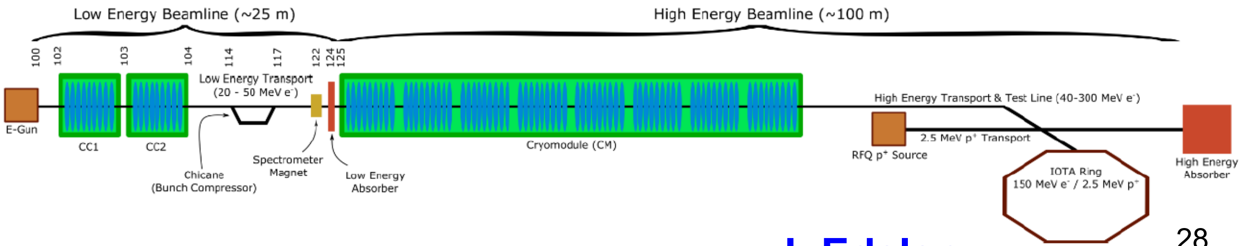
Modeling systems with neural networks: Examples at FAST

Predict RFQ frequency









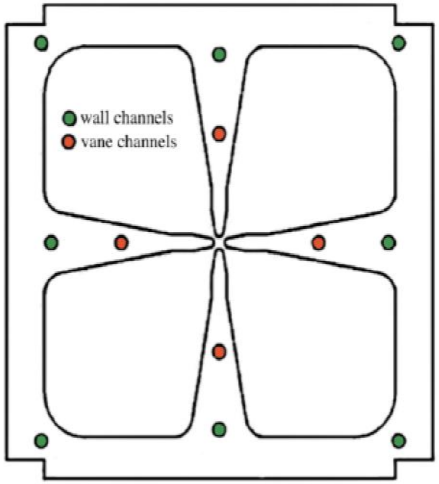
A. Edelen

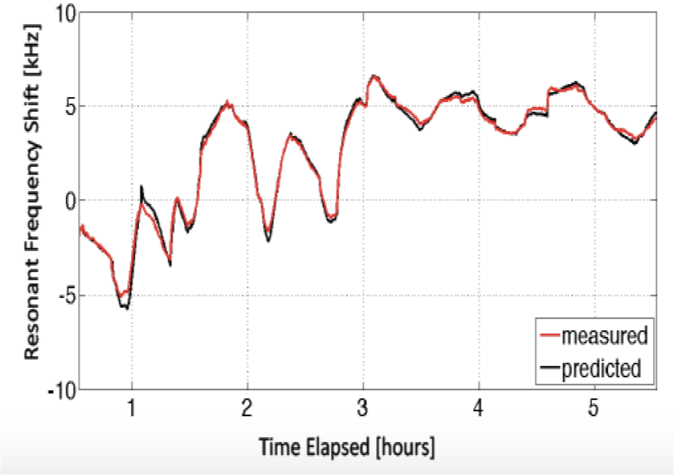
J. Edelen

Highlights: Modeling/Simulations

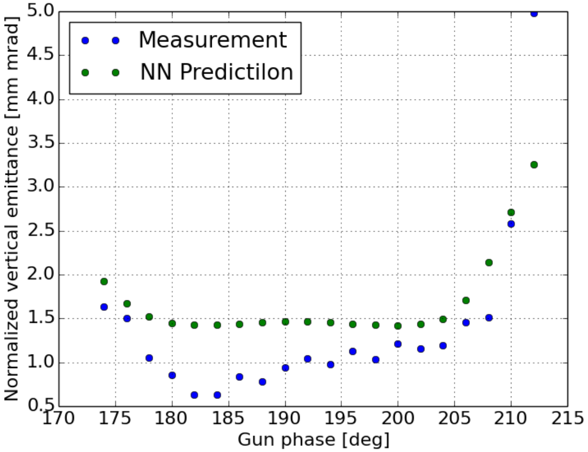
Modeling systems with neural networks: Examples at FAST

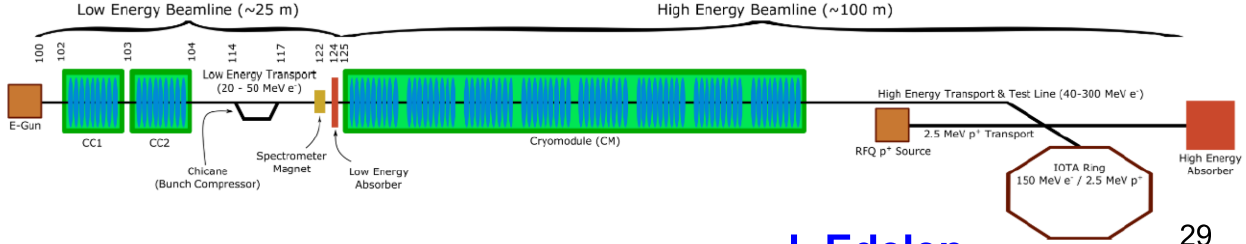
Predict RFQ frequency





Predict emittance



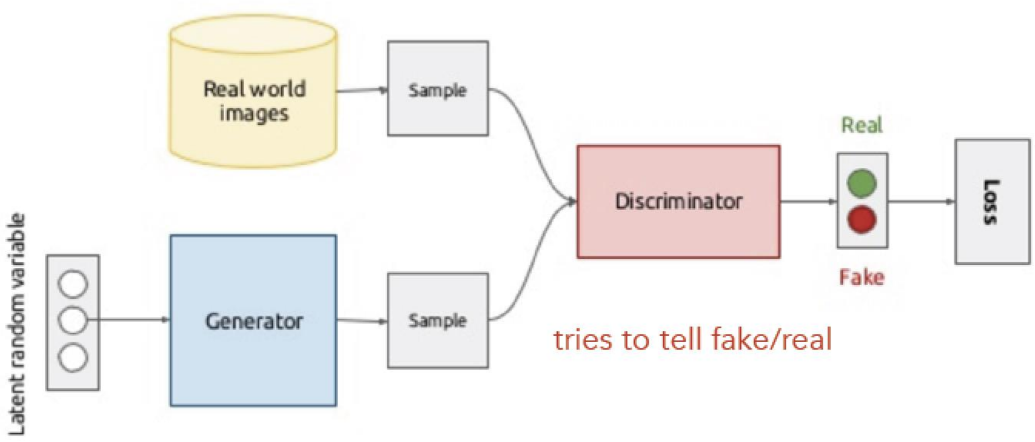


A. Edelen

J. Edelen

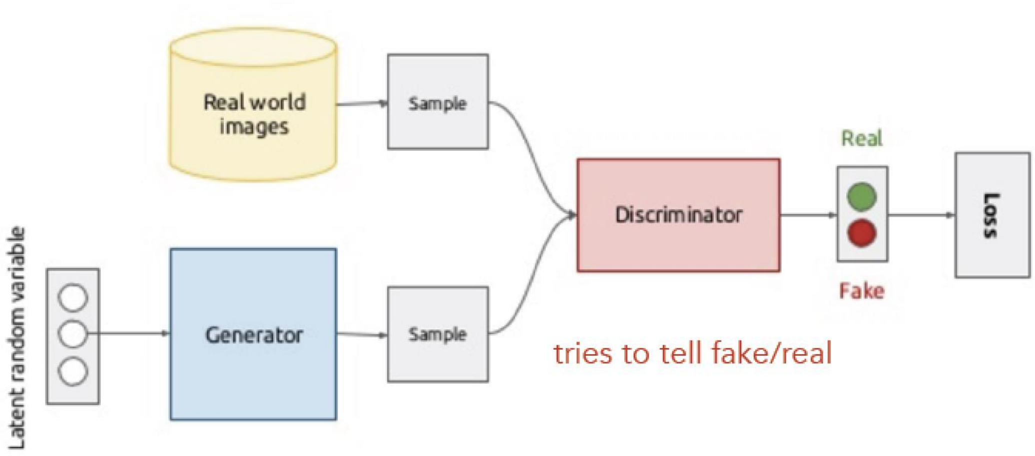
Highlights: Modeling/Simulations

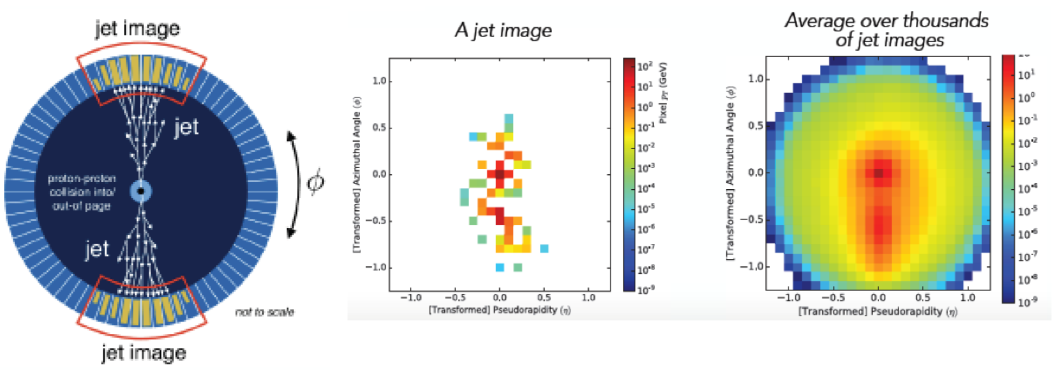
Generative adversarial networks (GANs): mimicking simulations



L. Oliviera, LBNL

Generative adversarial networks (GANs): mimicking simulations

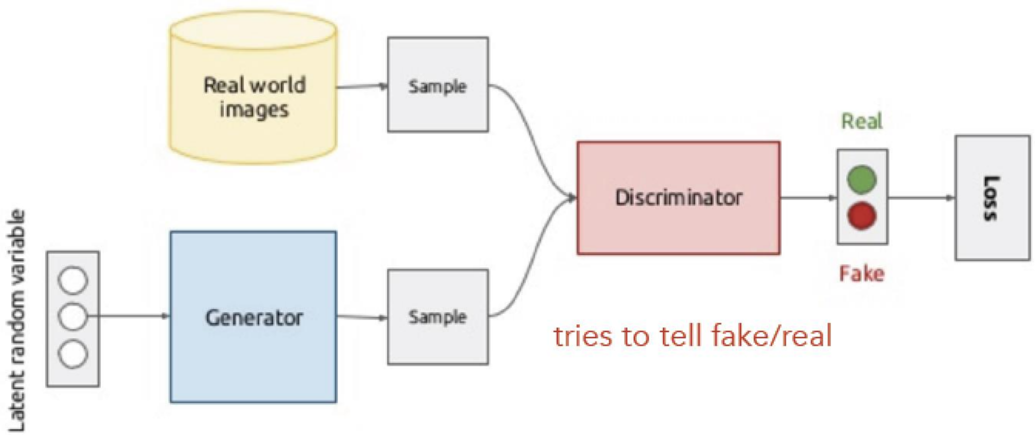




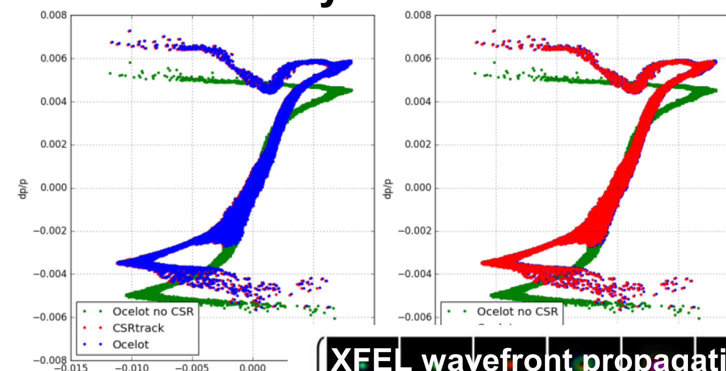
L. Oliviera, LBNL

Highlights: Modeling/Simulations

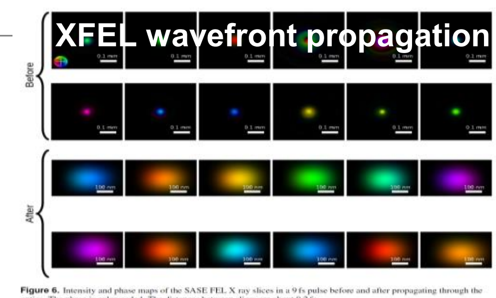
Generative adversarial networks (GANs): mimicking simulations

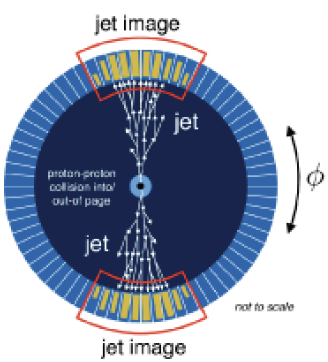


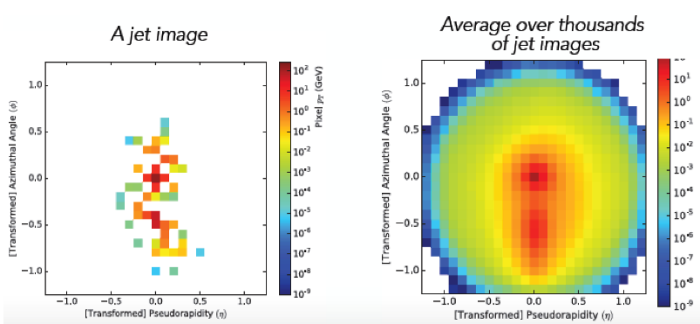
Coherent synchrotron radiation

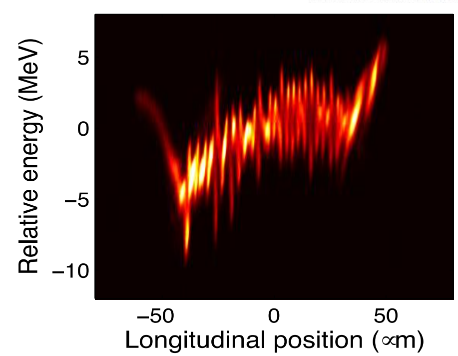


I. Agapov, DESY









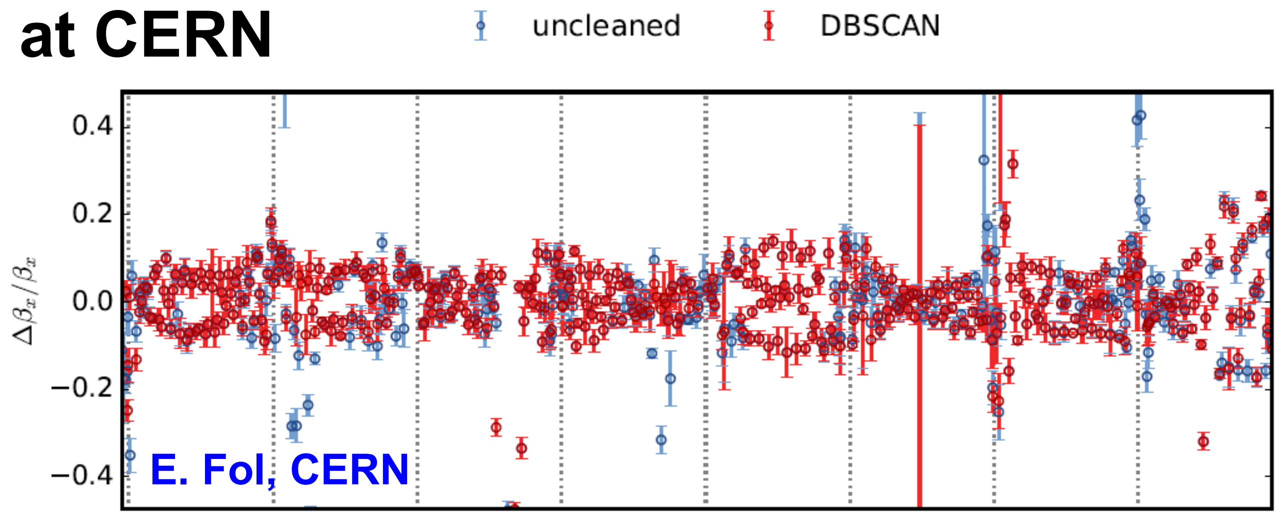
Microbunching instability, SLAC

L. Oliviera, LBNL

Highlights: Prognostics

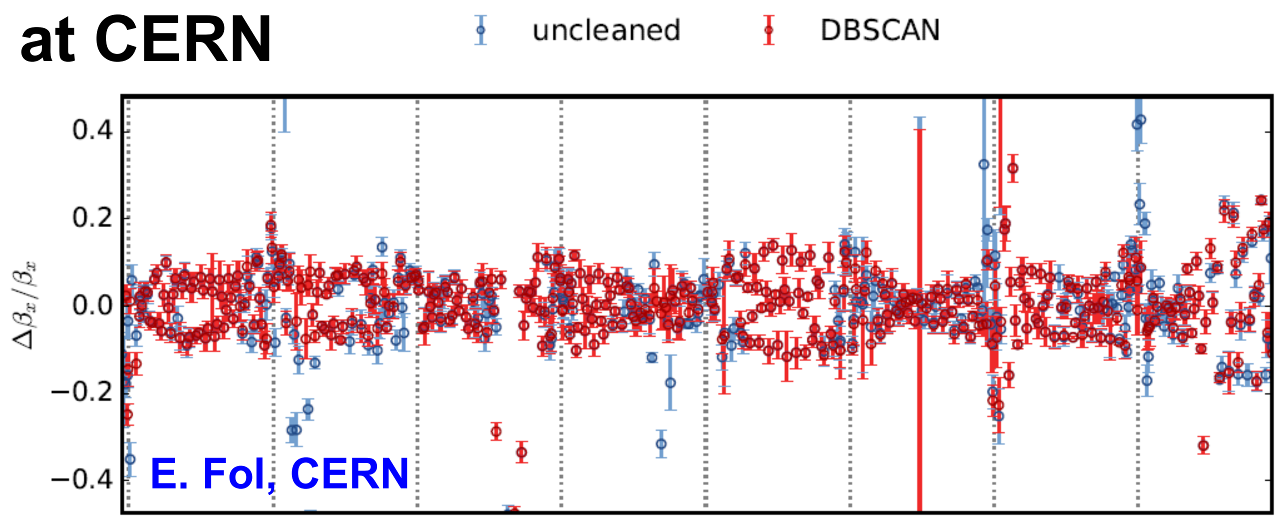
Outlier detection at CERN

find faulty BPMs:



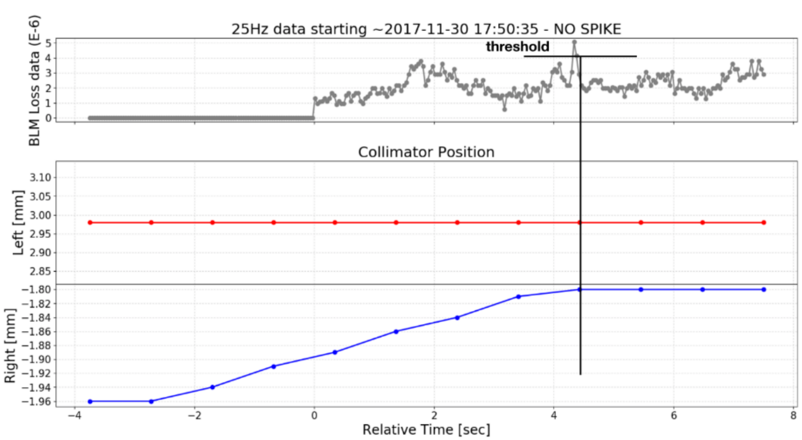
Outlier detection at CERN

find faulty BPMs:

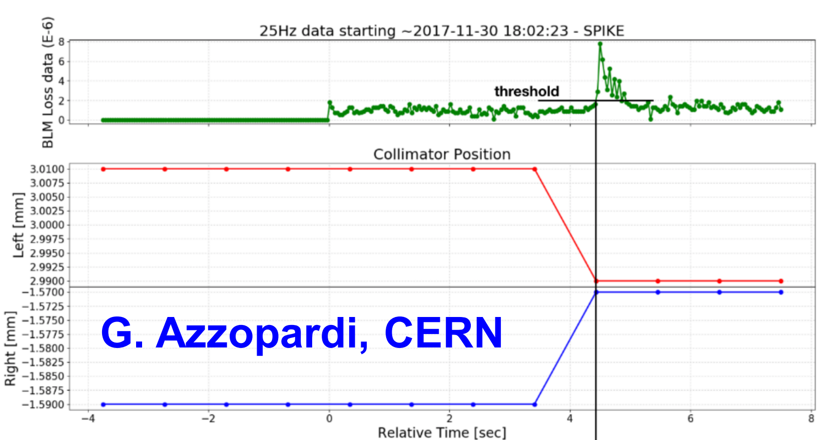


Distinguish true/false collimator events

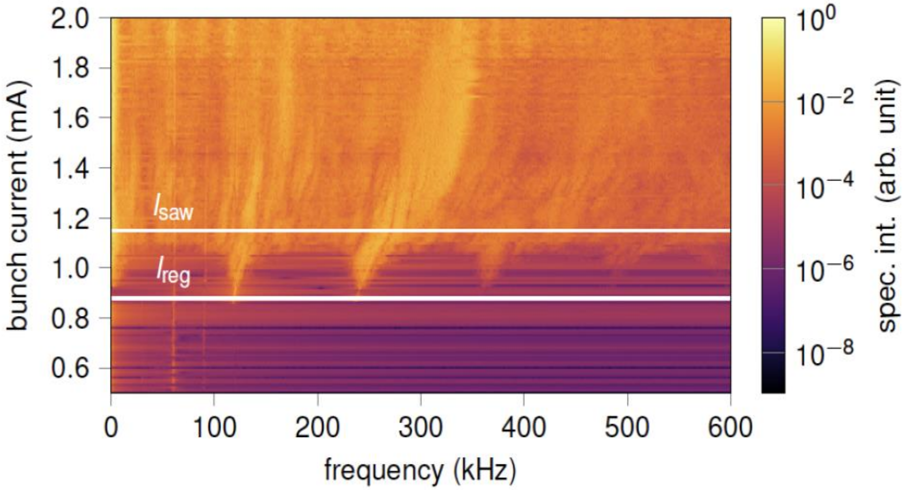
Correctly classified NO spike

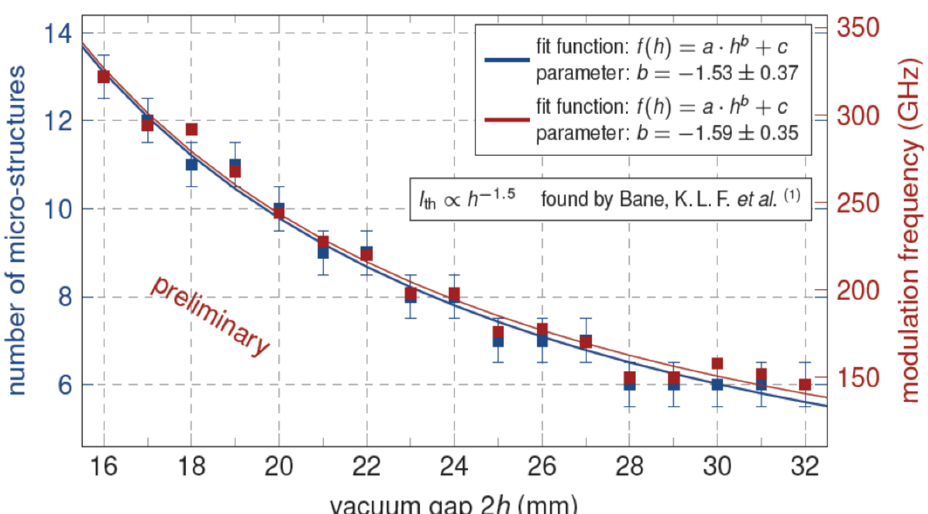


Correctly classified YES spike



K-means to understand MBI structures

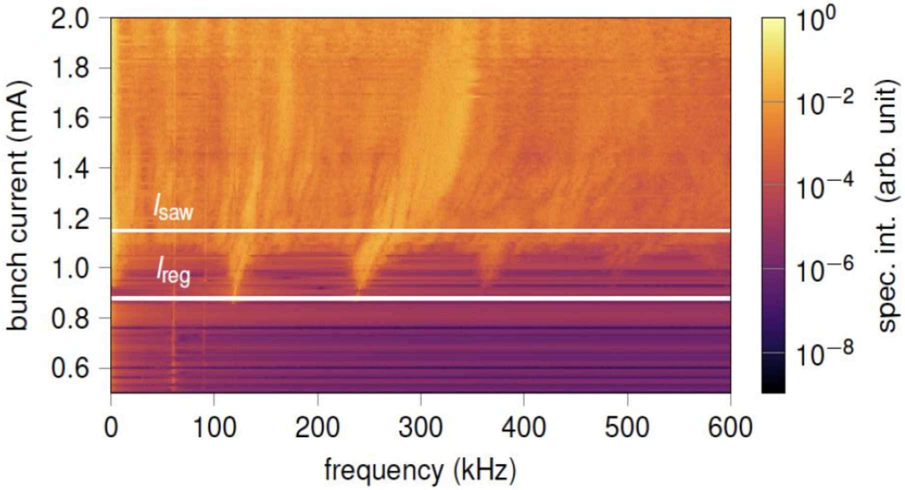


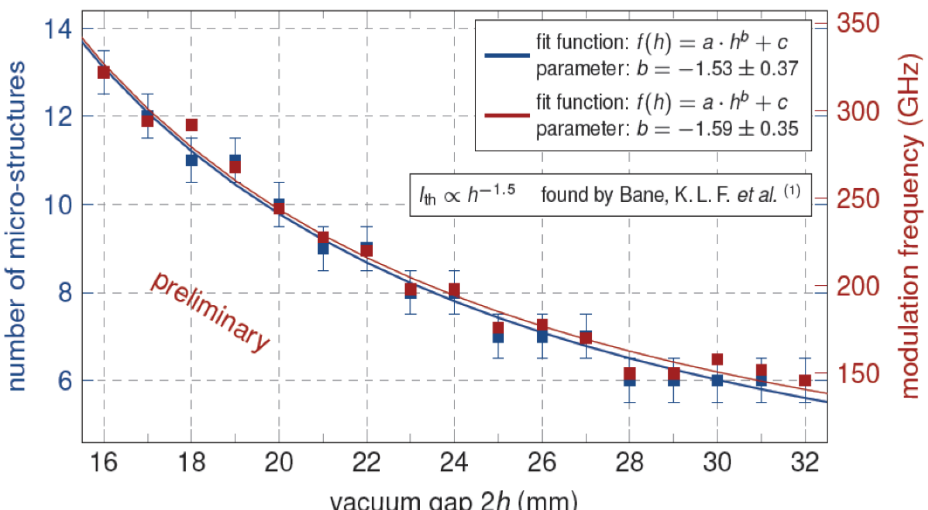


T. Boltz, KIT

Highlights: Data Analysis

K-means to understand MBI structures



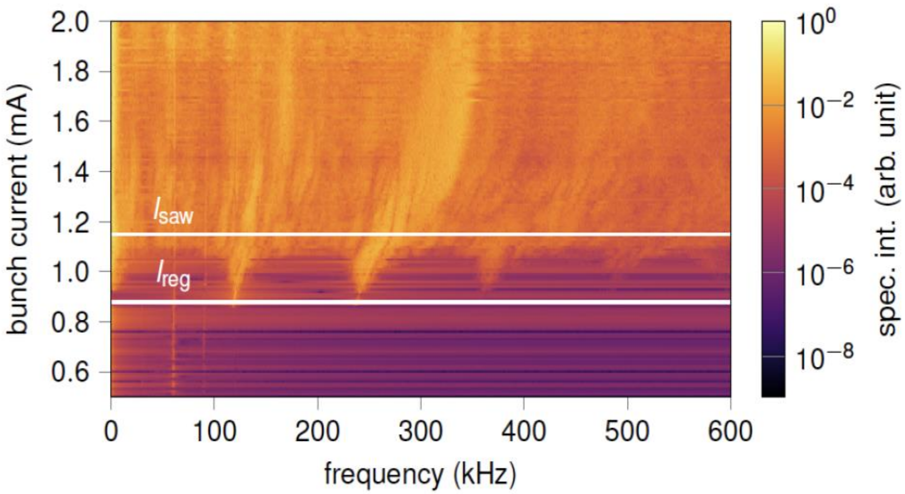


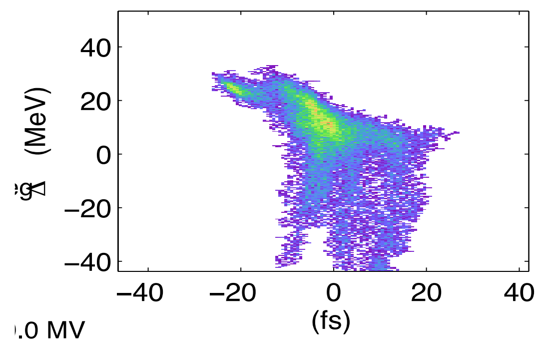
T. Boltz, KIT

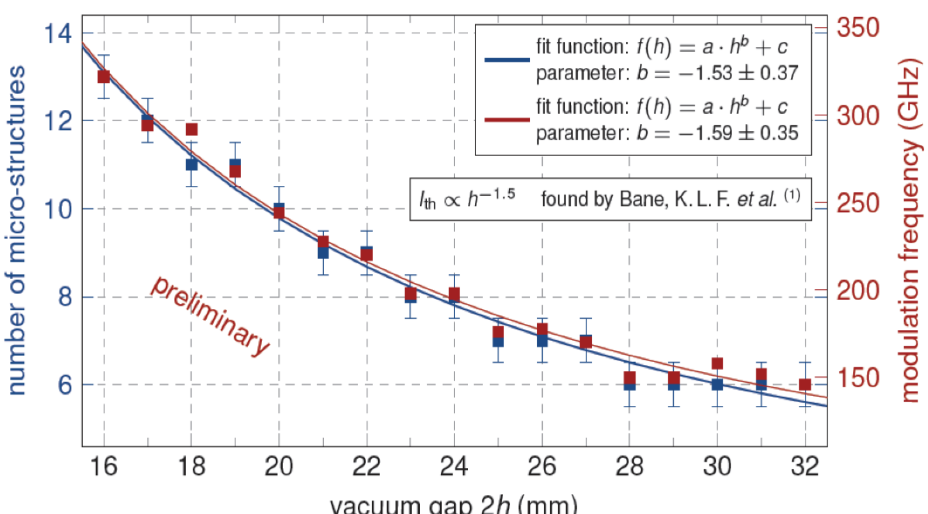
X. Zhang, SLAC

Highlights: Data Analysis

K-means to understand MBI structures





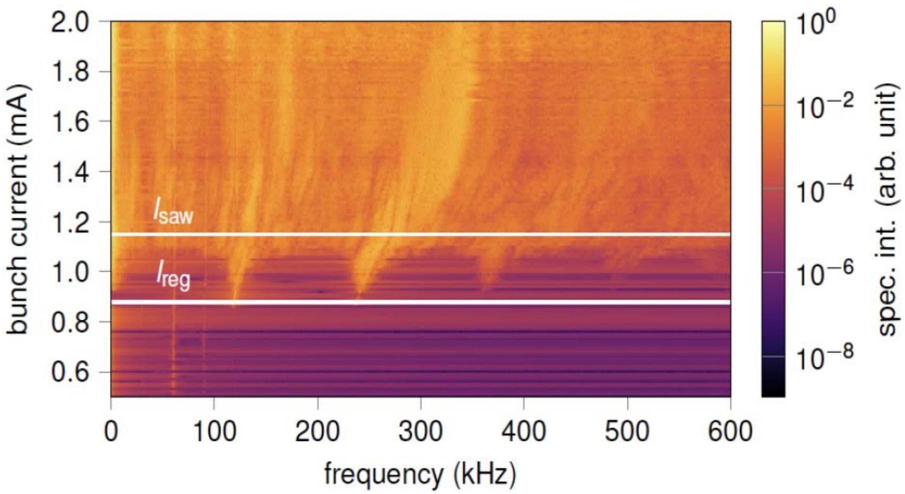


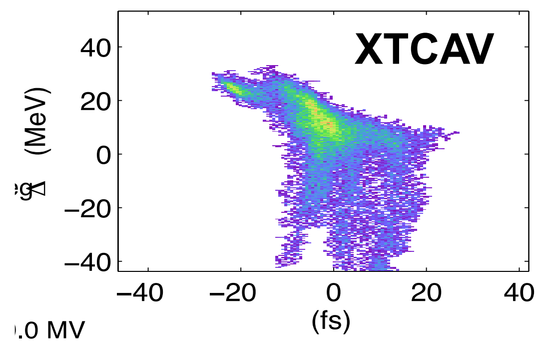
T. Boltz, KIT

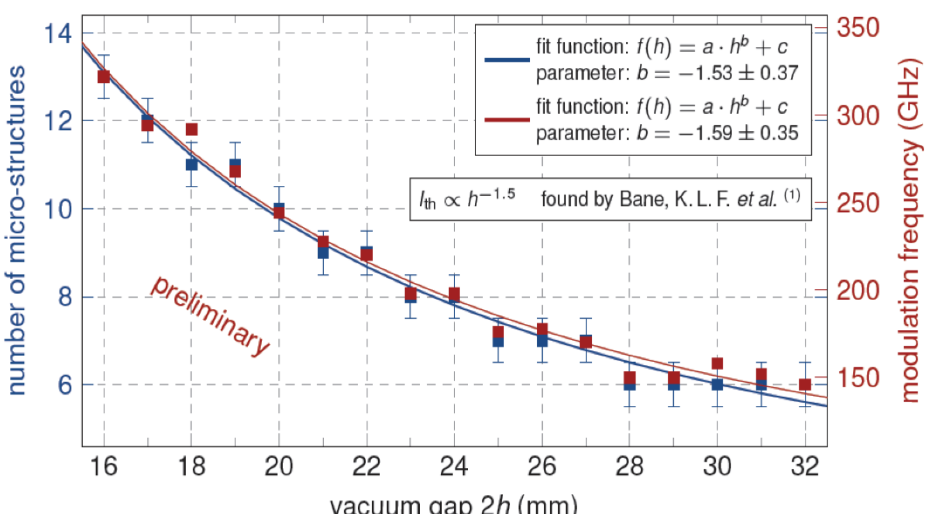
X. Zhang, SLAC

Highlights: Data Analysis

K-means to understand MBI structures







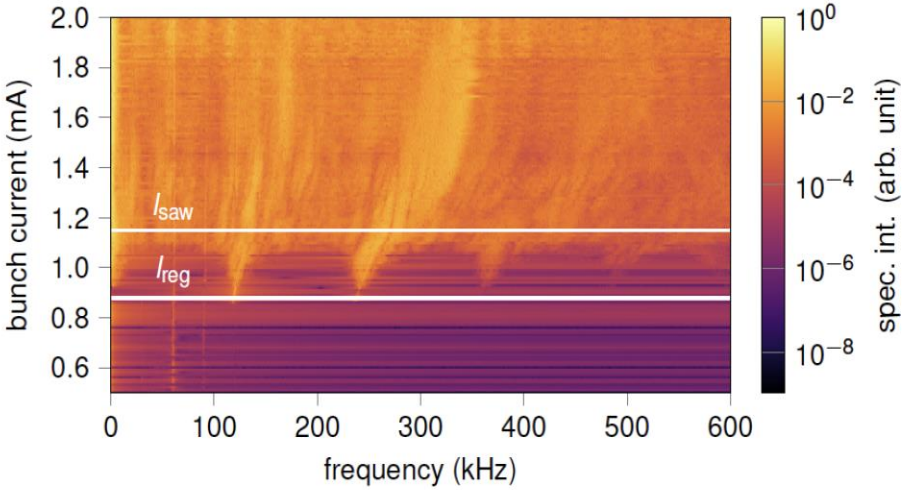
T. Boltz, KIT

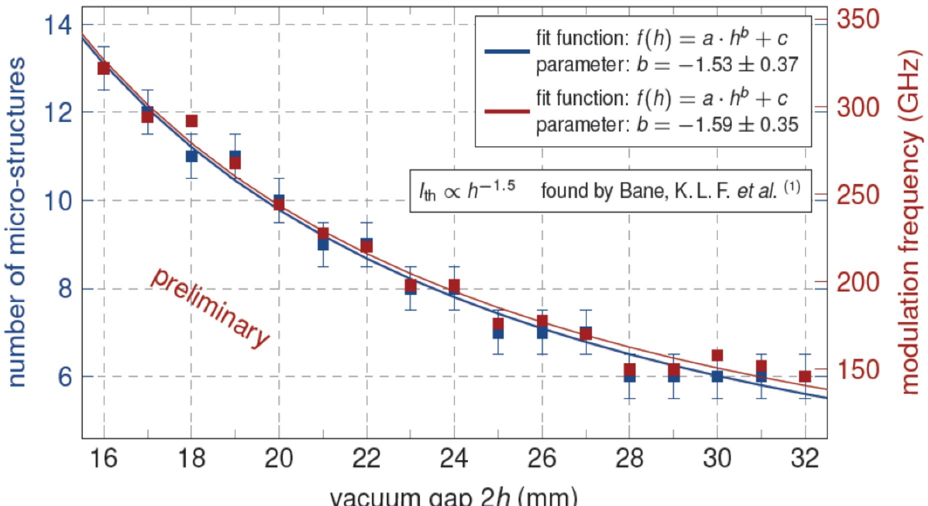
X. Zhang, SLAC

Highlights: Data Analysis

SLAC

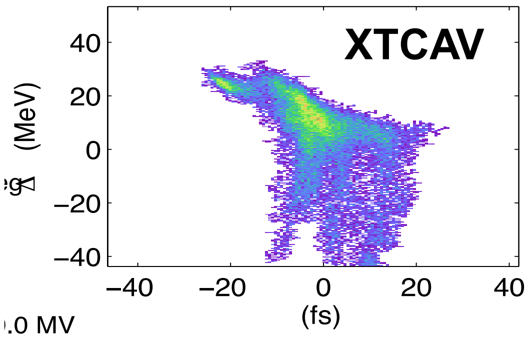
K-means to understand MBI structures





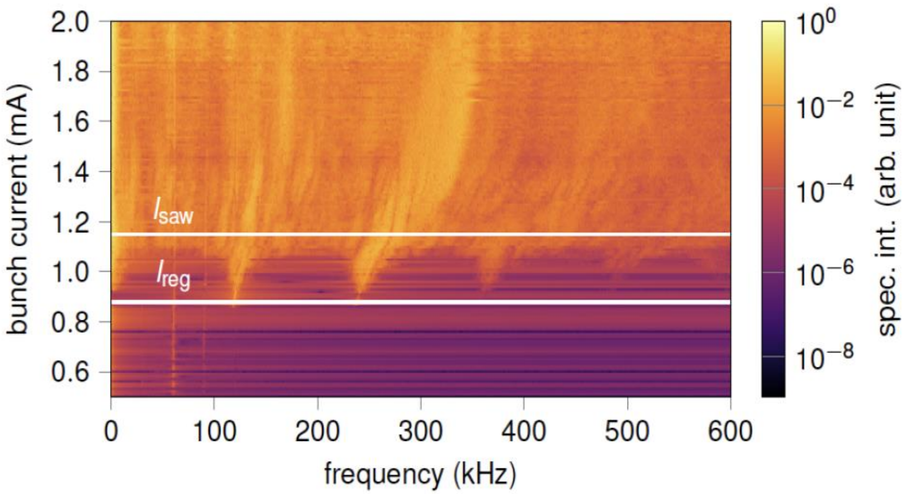
T. Boltz, KIT

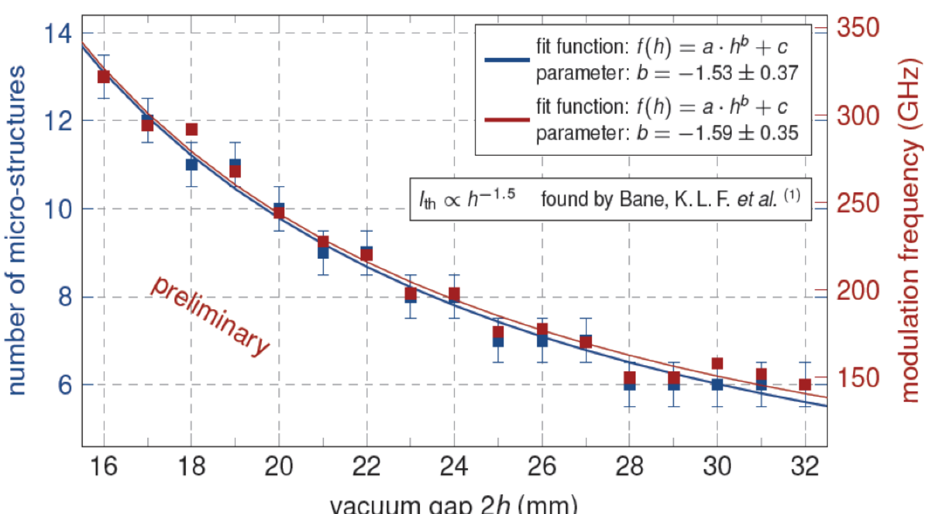
Reconstructing FEL Pulses



X. Zhang, SLAC

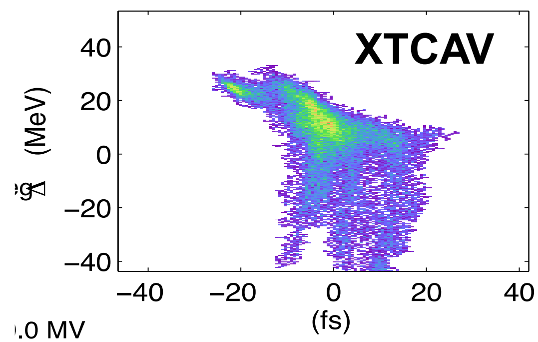
K-means to understand MBI structures

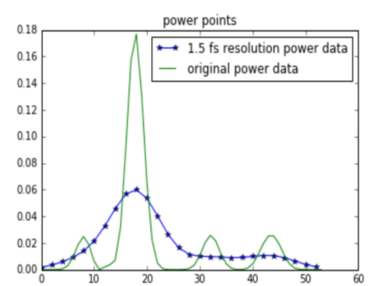




T. Boltz, KIT

Reconstructing FEL Pulses

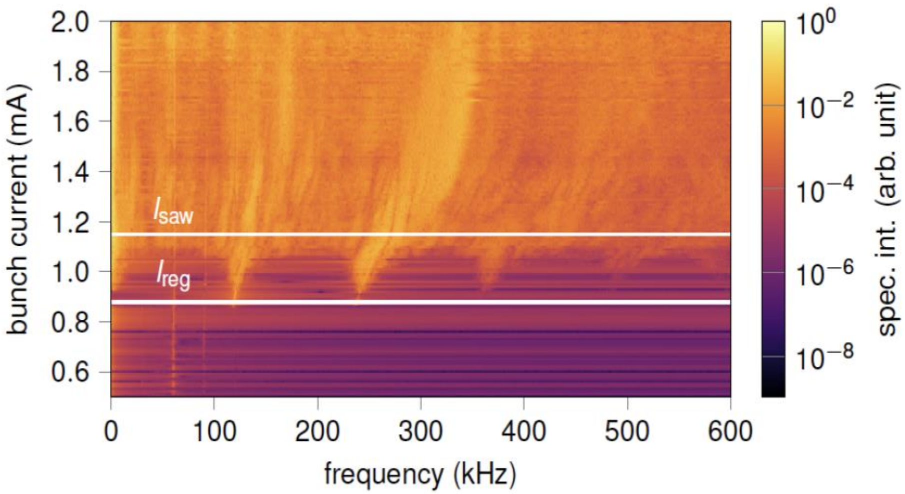


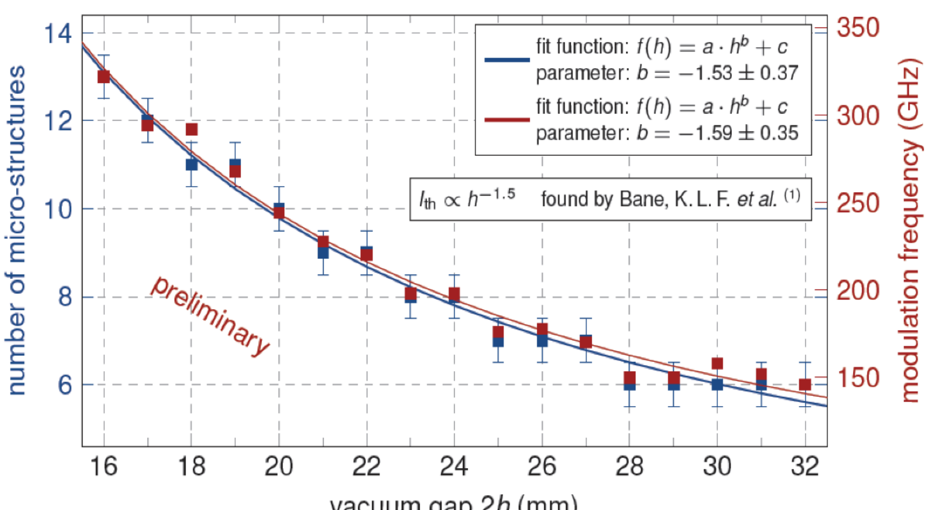


X. Zhang, SLAC

Highlights: Data Analysis

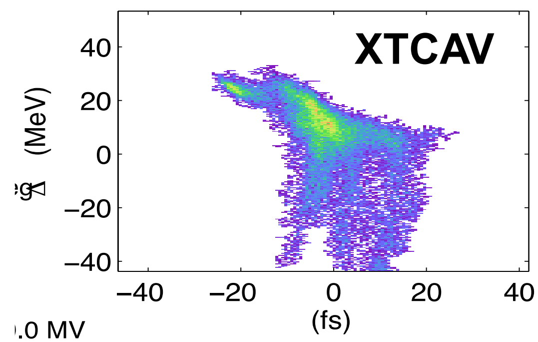
K-means to understand MBI structures

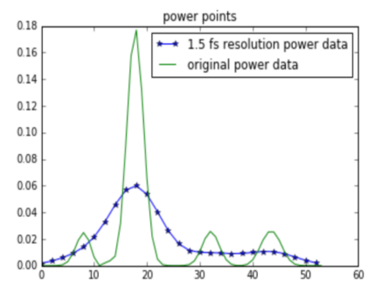




T. Boltz, KIT

Reconstructing FEL Pulses



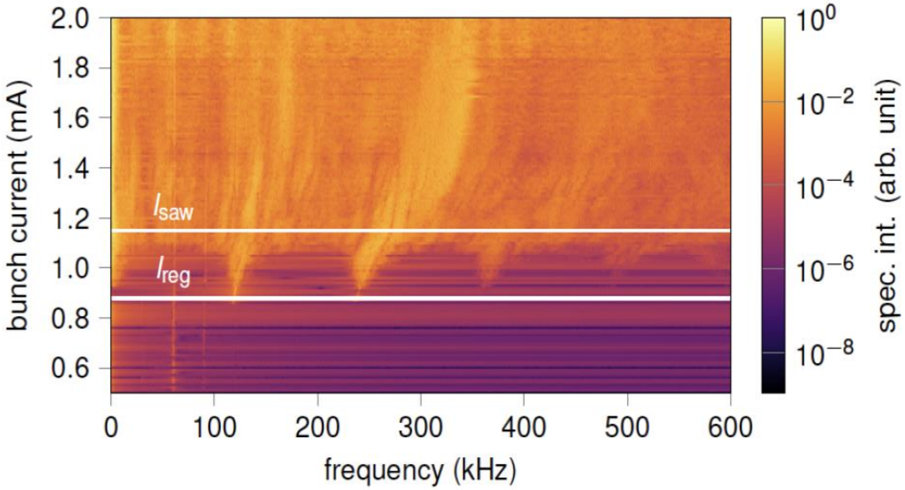


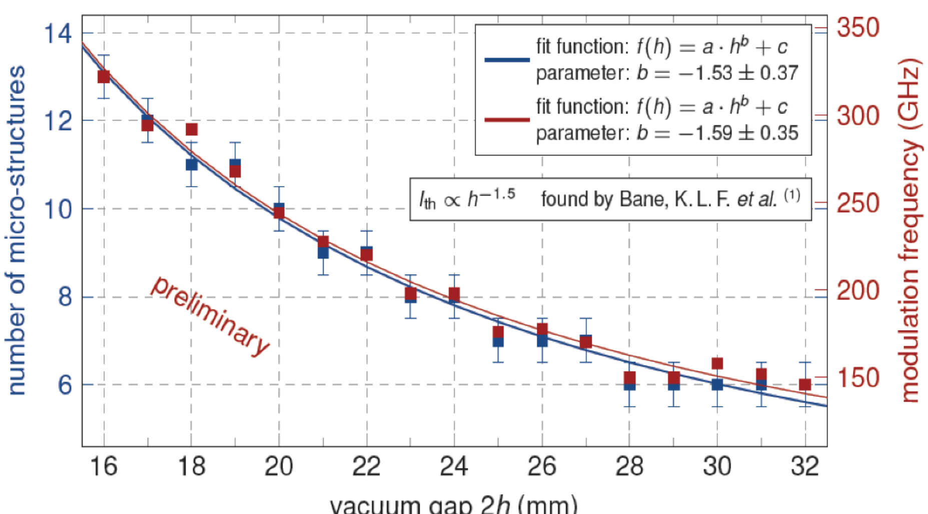
+

X. Zhang, SLAC

Highlights: Data Analysis

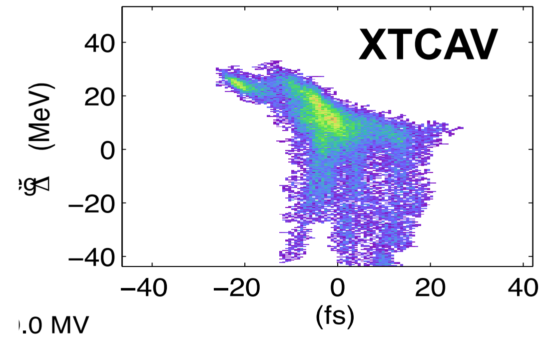
K-means to understand MBI structures

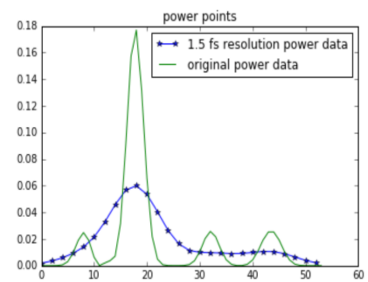




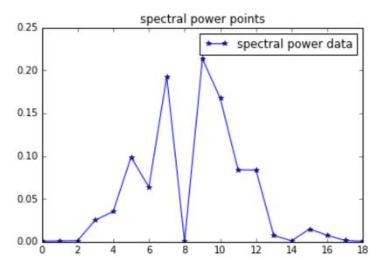
T. Boltz, KIT

Reconstructing FEL Pulses





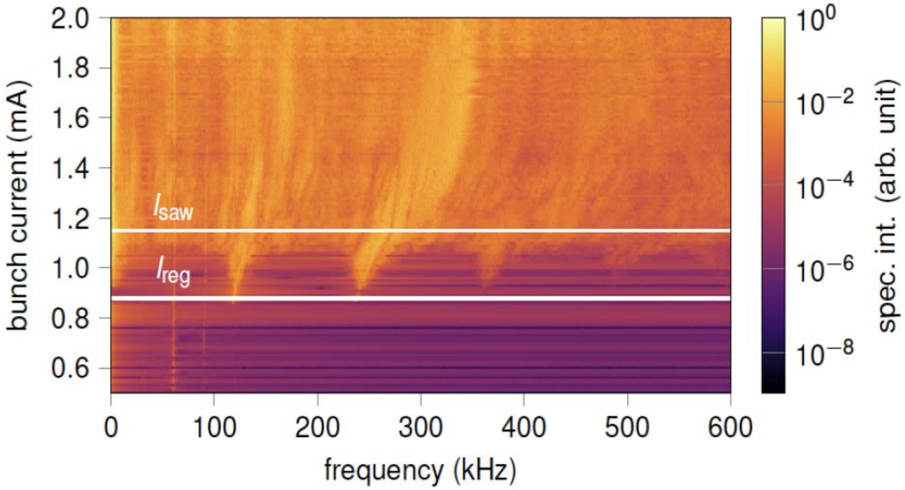
+

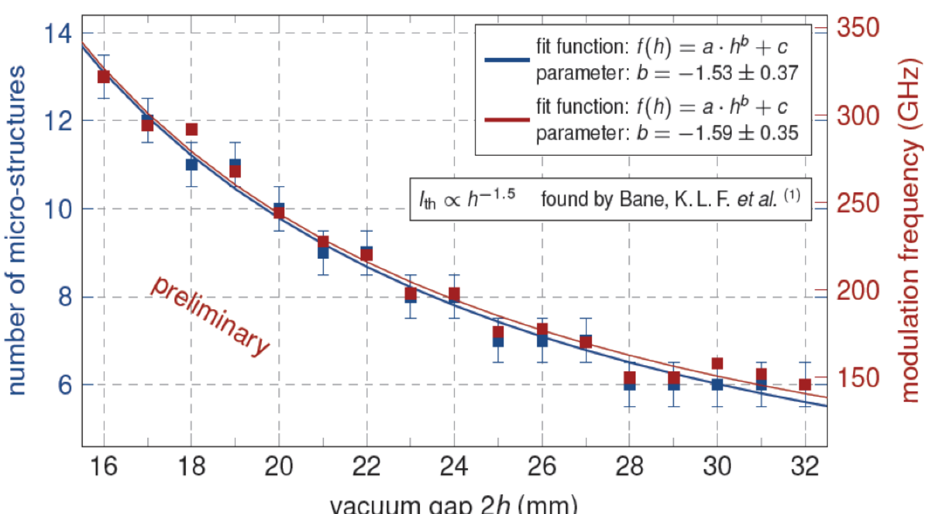


X. Zhang, SLAC

Highlights: Data Analysis

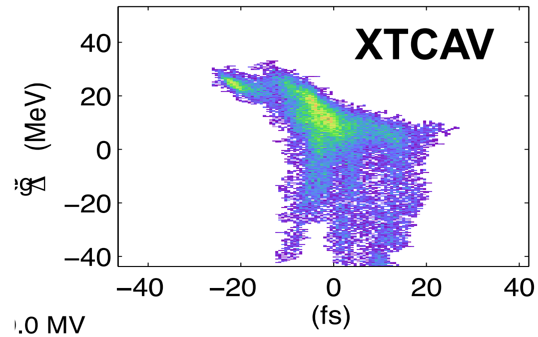
K-means to understand MBI structures



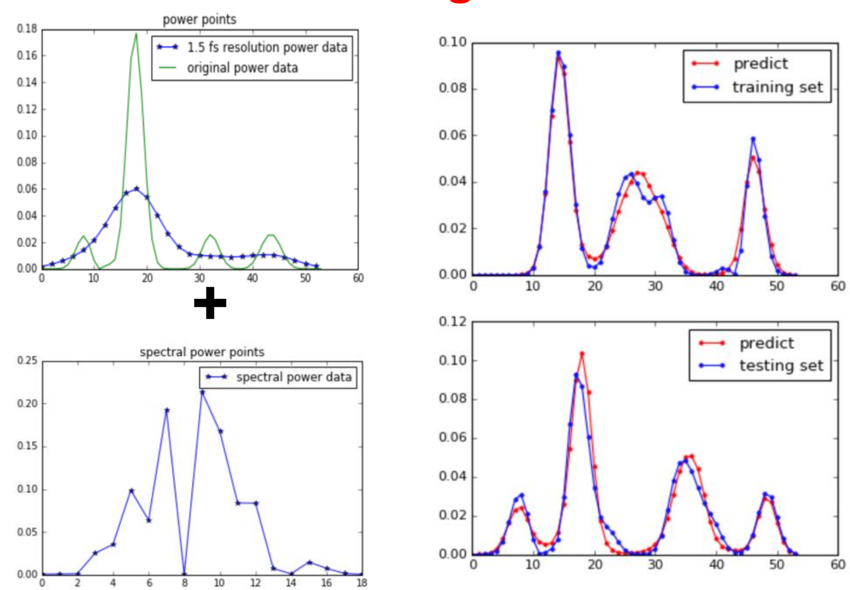


T. Boltz, KIT

Reconstructing FEL Pulses



5 orders of magnitude faster!



X. Zhang, SLAC

Highlights: Data Analysis

General data mining: Can we use ML to learn about our machines?

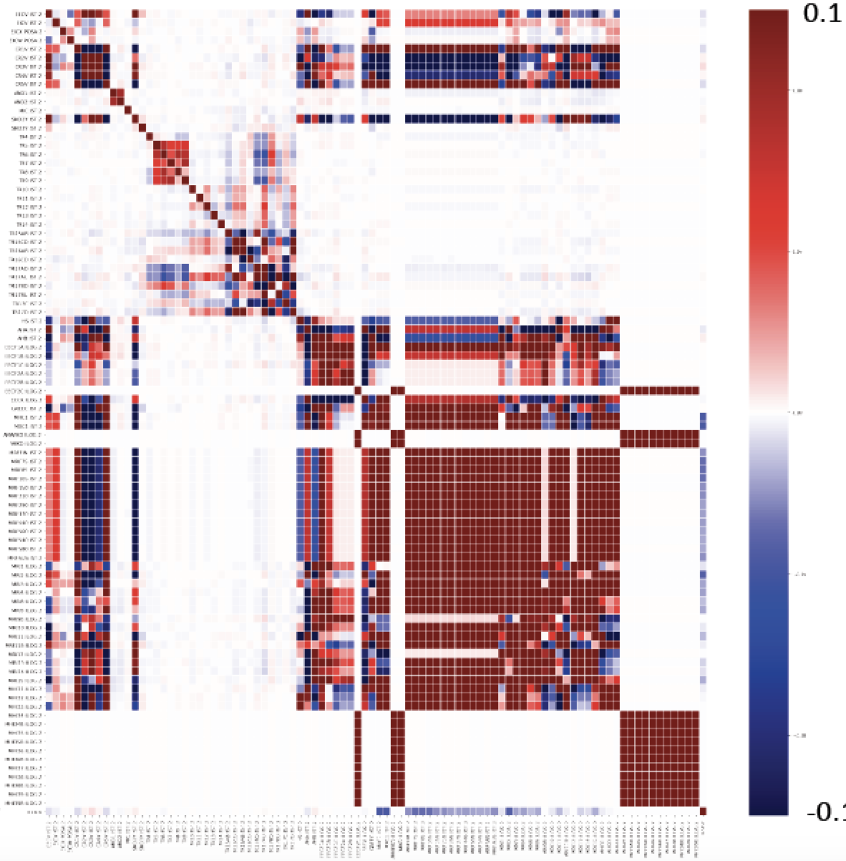
extraction and RF
in ring cyclotron

trim coils
in ring cyclotron

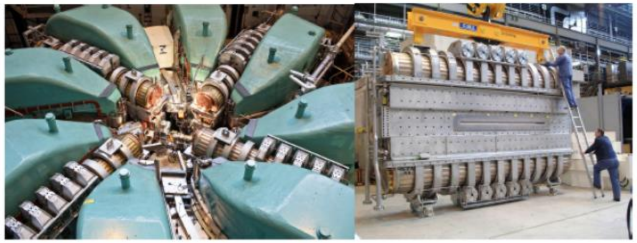
phase and loss
monitors
in ring cyclotron

loss monitors in
targets beamline

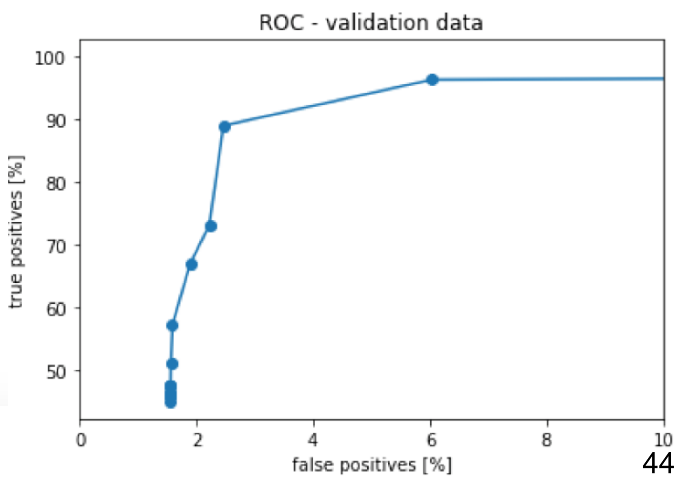
Interlocks →



PSI cyclotron



Predict interlock



J. Snuverink, A. Adelmann

Future directions

Currently writing white paper to summarize first workshop

Future directions

Currently writing white paper to summarize first workshop

Then next workshop: March, 2019 in Switzerland!

Future directions

Thanks for listening and hope to see you at a future ML workshop!