General Experience With Online Optimization

Alexander Scheinker
(ascheink@lanl.gov)

LCLS ML Workshop, February 27, 2018

Motivation

Difficult to control and quickly tune for specific beam properties using traditional model-based approaches

- **Dynamics of intense charged particle bunches (FELs and PWFAs) dominated by:**
  - **Complex collective effects:**
    - **Wakefields**
    - **Space charge**
    - **Coherent synchrotron radiation**
    - **Uncertain and time varying electron bunch distribution off cathode**
  - **Components drift unpredictably with time, misalignments**
  - **Limited diagnostics**

Need for model-independent feedback algorithms for time-varying, uncertain systems.

- **Requirements:**
  - **Tune many coupled parameters simultaneously.**
  - **Can be applied in-hardware as real-time feedback.**
  - **Work with noisy data.**

Online tuning and optimization approaches

Robust conjugate direction search (RCDS) [1]

- Robust to noise
- Model independent
- Tune many parameters simultaneously

Extremum Seeking (ES) [2,3]

- Robust to noise
- Model independent
- Tune many parameters simultaneously
- Automatically adapts to time-variation of the system

[1] X. Huang, J. Corbett, J. Safranek, J. Wu, “An algorithm for online optimization of accelerators”, Nucl. Instr. Methods, A 726 (2013) 77-83.

[2] A. Scheinker, “Extremum Seeking for Stabilization,” PhD Thesis, UCSD, 2012.

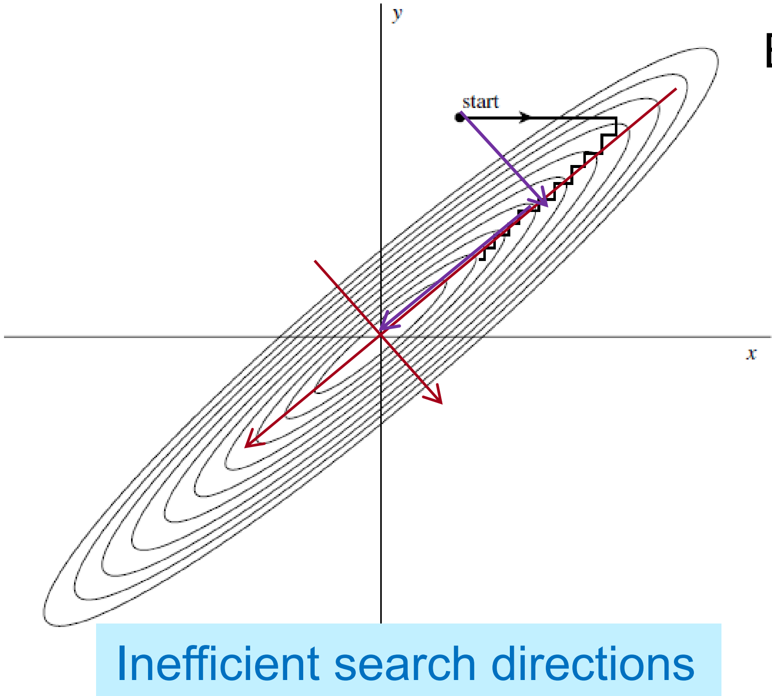
[3] A. Scheinker et al., *Model-free stabilization by extremum seeking*, Springer, 2016.

The development of the RCDS algorithm

- The development was motivated by the need to optimize storage ring nonlinear beam dynamics.
  - Correction of nonlinear dynamics is difficult – lack of direct diagnostics, deterministic method, and even target.
- Robust conjugate direction search (RCDS)* performs iterative search over conjugate directions with a robust (against noise), efficient line (1D) optimizer.
  - The conjugate direction set may be updated with Powell's method.
  - The 1D robust optimizer is designed to deal with noise.

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

Search over conjugate directions



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

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

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

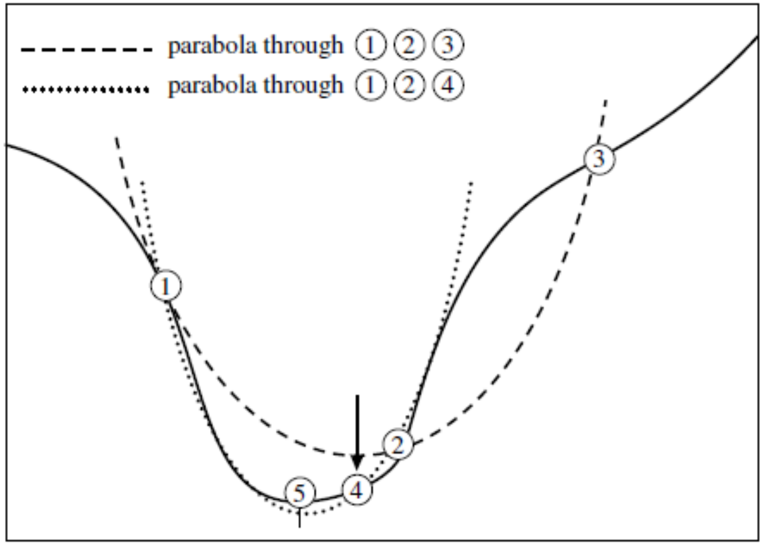
*W.H. Press, et al, Numerical Recipes

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

Anatomy of a line optimizer that is sensitive to noise

Line optimizer – Brent's method

Step 1: Initially bracketing the minimum.
Step 2: Successive interpolation to converge to the minimum.



Inverse quadratic interpolation (figure from Numeric Recipes*.)

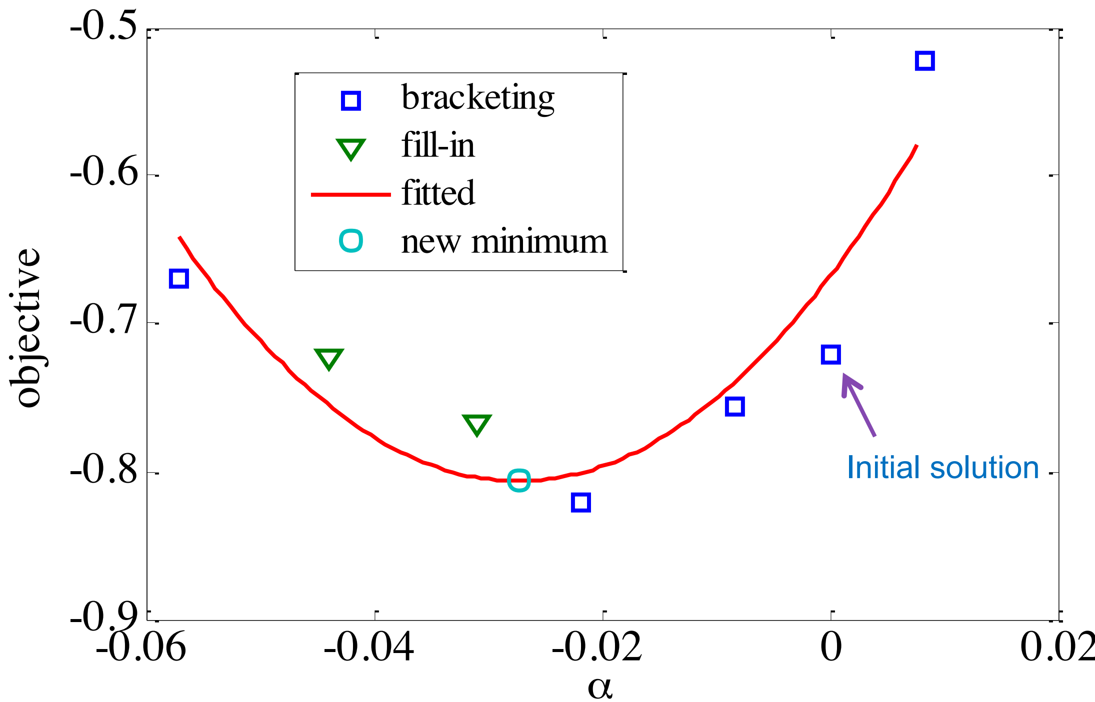
With noise, the comparison of values in both steps can go wrong and the algorithm won't converge.

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

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

The robust 1D optimizer

The robust optimizer is aware of noise in bracketing and uses noise level to filter out outliers. Noise level is detected before optimization.



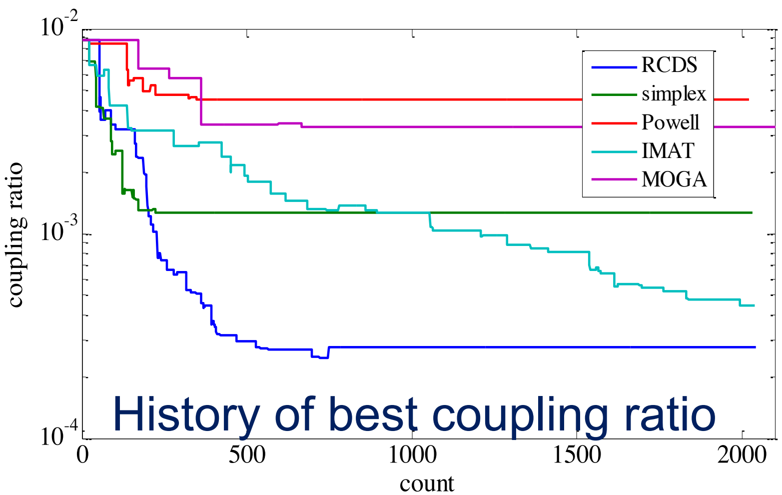
X. Huang et al, Nucl. Instr. Methods, A 726 (2013) 77-83.

Bracketing: step size is increased in the search. Bracket ends are higher than minimum by 3 noise sigma.
Fitting: fill in additional points when necessary to better sample within the bracket and then fit a parabola.

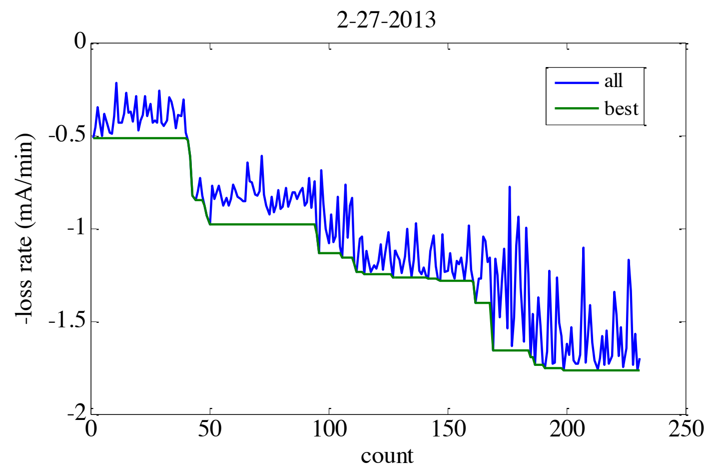
X. Huang (SLAC), Online optimization, ML Workshop 2018

Initial test of RCDS

- Coupling correction (minimization of vertical emittance a storage ring) was used to test the algorithm in both simulation and experiment.
  - Knobs are 13 skew quadrupoles
  - Objective is beam loss over a period of 6 seconds



Simulation



Experiment

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

Los Alamos NATIONAL LABORATORY

Applications of RCDS algorithm

- The RCDS algorithm has been demonstrated on many accelerators.

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

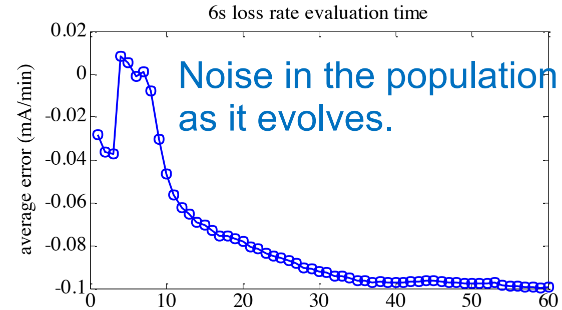
T. Pulampong, et al, IPAC'2017, THPAB151

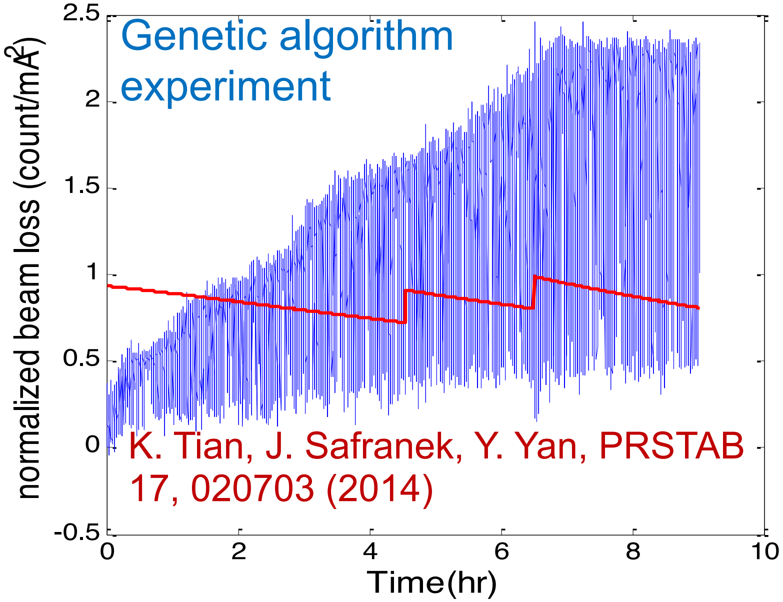
There are many unpublished applications.

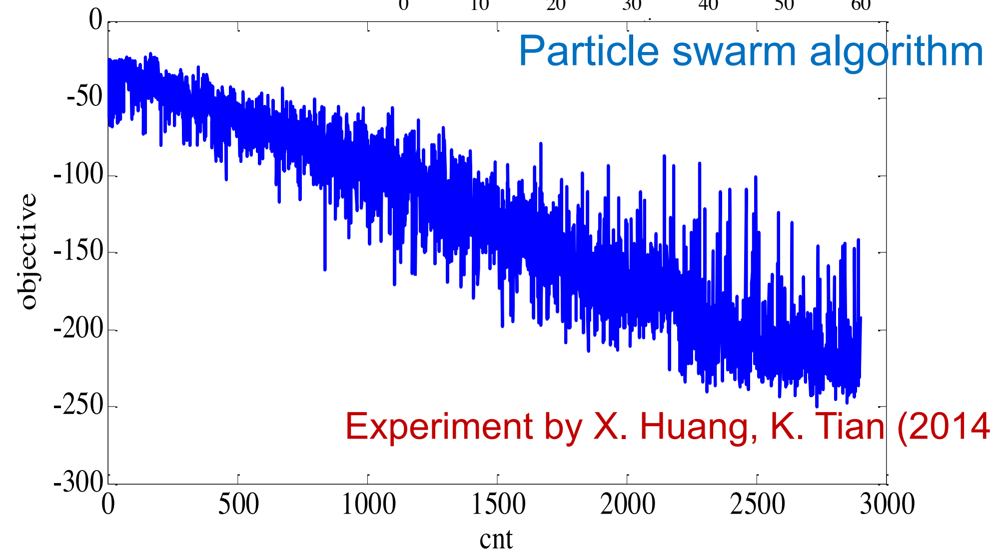
X. Huang (SLAC), Online optimization, ML Workshop 2018

GA vs PSO for online optimization

- Genetic algorithms may not be efficient for online application and suffer bias from noise
- Particle swarm optimization







Using beam loss monitor signal (low noise) as objective. It took **20,000** evaluations.

Same setup as the genetic algorithm experiment. It took **3,000** evaluations.

… while RCDS only took **200** evaluations (see slide 7) for a much noisier setup.

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

RCDS Summary

- Online optimization has advantages in several aspects: low diagnostics demand, no need of model, fast, etc
- At SPEAR3 we have developed the RCDS algorithm and explored other algorithms (GA, PSO, ES).
- The RCDS algorithm has been successful in nonlinear beam dynamics optimization and many applications at other labs.
- Continuing development in online optimization could greatly benefits the community.

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

Extremum Seeking for Optimization and Adaptive Tuning

- **Algorithm overview and simulation studies**
  - Automatic magnet tuning at **LANSCE**
- **Beam energy control**
  - Iterative feed forward for beam loading at **LANSCE**
- **Transverse beam dynamics (x,x',y,y')**
  - Automatic magnet tuning at **SPEAR3**
- **Diagnostics**
  - Non-invasive longitudinal phase space predictions at **FACET**
- **Longitudinal phase space (Δz,ΔE)**
  - Automatic longitudinal phase space tuning at **LCLS**

Motivating Example: Open-loop unstable system with unknown control direction

Consider the scalar system with unknown control direction b(t):

$$\dot{x} = ax + b(t)u$$

Consider the following bounded extremum seeking (ES) controller:

$$u = \sqrt{\alpha\omega}\cos\left(\omega t + kx^2\right)$$

For large ω, the system dynamics are approximately given by:

$$\begin{aligned}\dot{\bar{x}} &= a\bar{x} - k\alpha b^2(t)\bar{x} \\ &= \left(a - k\alpha b^2(t)\right)\bar{x}\end{aligned}$$

This system is stabilized by choosing sufficiently large kα > 0.

$$\max_{t\in[0,T]} |x(t) - \bar{x}(t)| < \delta$$

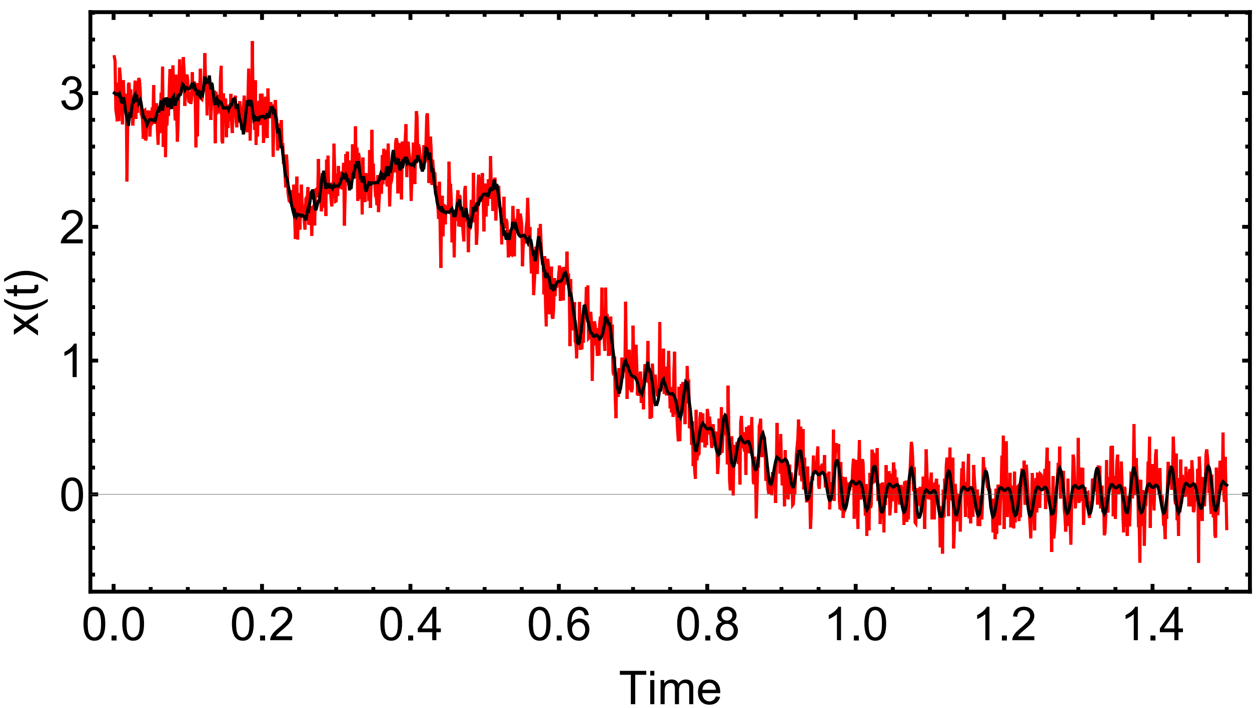
Motivating Example: Open-loop unstable system with unknown control direction

$$\dot{x} = x + b(t)u, \;\; u = \sqrt{\alpha\omega}\cos\left(\omega t + k\hat{x}^2\right), \;\; \hat{x}(t) = x(t) + n(t)$$

$$\dot{x} = x + b(t)\sqrt{\alpha\omega}\cos\left(\omega t + k\hat{x}^2\right)$$

$$\dot{\bar{x}} = \bar{x} - k\alpha b^2(t)\bar{x}$$

$b(t)=\sin(\omega_b t)$, $\omega_b=2\pi\times 10$, $\alpha=2$, $k=10$, $\omega=2\pi\times 50$



Los Alamos NATIONAL LABORATORY

Model Independent Feedback

A. Scheinker, "Extremum Seeking for Stabilization," PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, "Extremum Seeking with Discontinuous Dithers," *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t) = \begin{bmatrix} f_1(x_1, \dots, x_n, p_1, \dots, p_m, t) \\ \vdots \\ f_n(x_1, \dots, x_n, p_1, \dots, p_m, t) \end{bmatrix}$$

$$y = V(\mathbf{x}, t) + n(t)$$

$$p_1(t) = Q_1 \text{ current}$$

$$\vdots$$

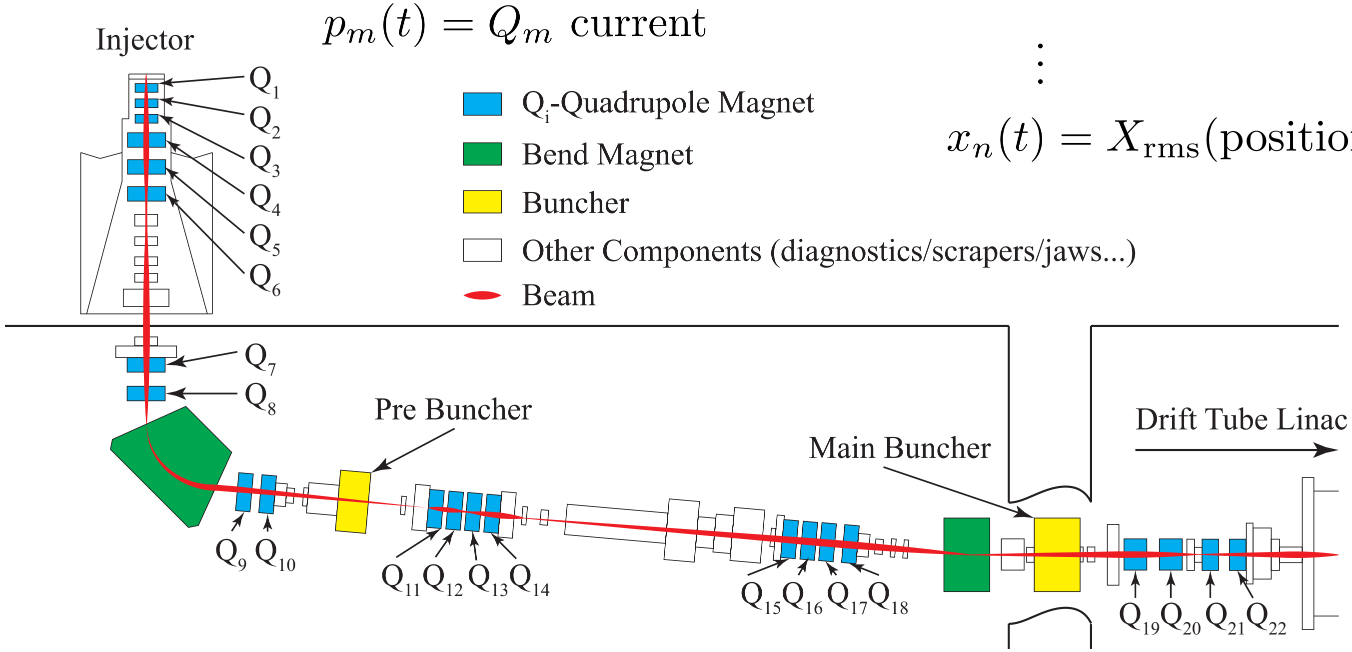
$$p_m(t) = Q_m \text{ current}$$

$$x_1(t) = X_{\text{rms}}(\text{position 1})$$

$$x_2(t) = Y_{\text{rms}}(\text{position 1})$$

$$\vdots$$

$$x_n(t) = X_{\text{rms}}(\text{position n})$$



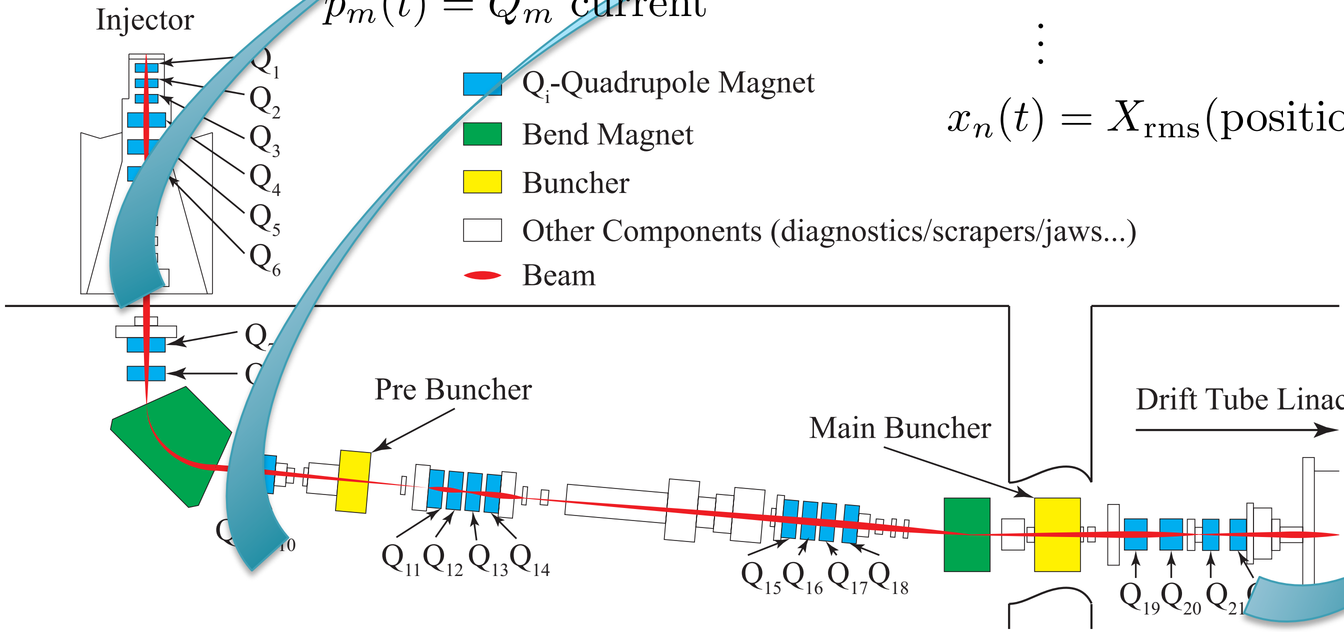
Los Alamos NATIONAL LABORATORY

Model Independent Feedback A. Scheinker, “Extremum Seeking for Stabilization,” PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, “Extremum Seeking with Discontinuous Dithers,” *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t) = \begin{bmatrix} f_1(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \\ \vdots \\ f_n(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \end{bmatrix} \qquad y = V(\mathbf{x}, t) + n(t)$$

$$p_1(t) = Q_1 \text{ current}$$
$$\vdots$$
$$p_m(t) = Q_m \text{ current}$$

$$x_1(t) = X_{\text{rms}}(\text{position 1})$$
$$x_2(t) = Y_{\text{rms}}(\text{position 1})$$
$$\vdots$$
$$x_n(t) = X_{\text{rms}}(\text{position n})$$



Los Alamos NATIONAL LABORATORY

Model Independent Feedback

A. Scheinker, “Extremum Seeking for Stabilization,” PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, “Extremum Seeking with Discontinuous Dithers,” *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t) = \begin{bmatrix} f_1(x_1, \dots, x_n, p_1, \dots, p_m, t) \\ \vdots \\ f_n(x_1, \dots, x_n, p_1, \dots, p_m, t) \end{bmatrix}$$

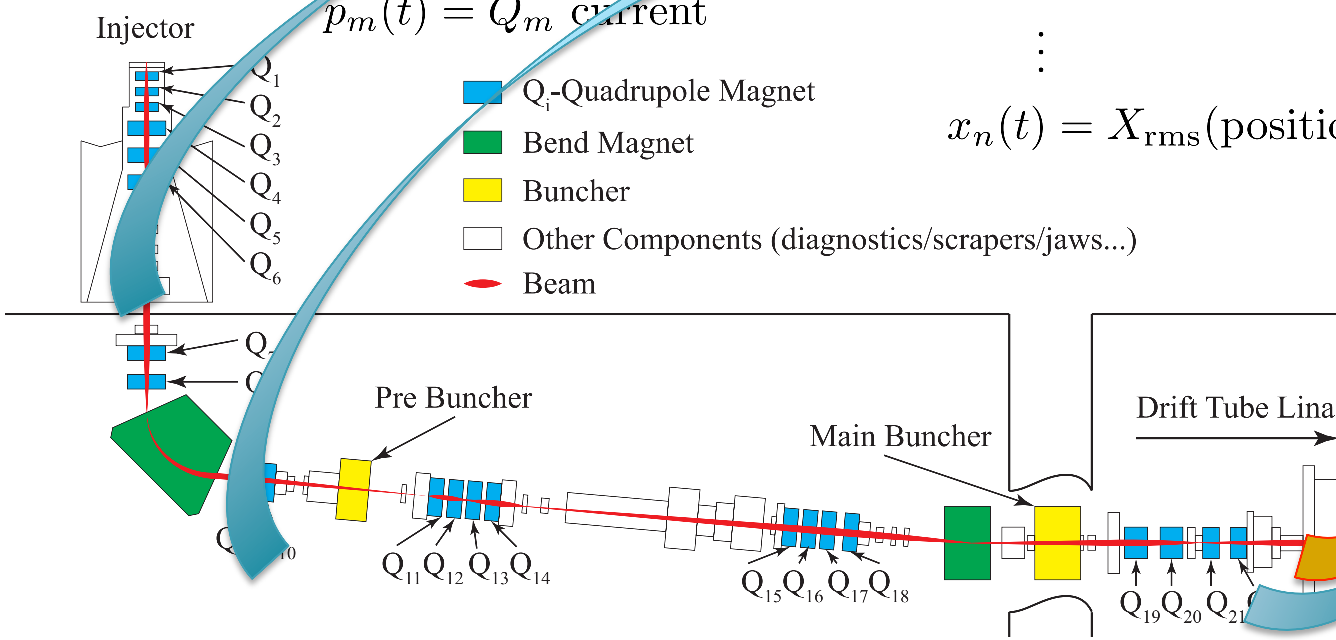
$$\begin{aligned} y &= V(\mathbf{x}, t) + n(t) \\ &= (I(t) - I_0)^2 + n(t) \end{aligned}$$

Surviving beam current at end.

noise

$$p_1(t) = Q_1 \text{ current}$$
$$\vdots$$
$$p_m(t) = Q_m \text{ current}$$

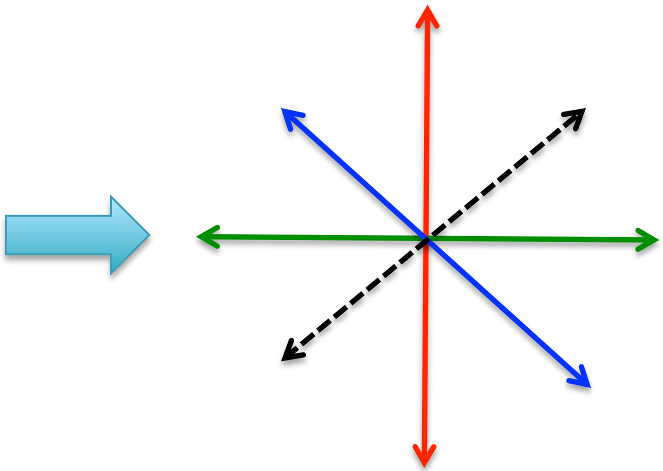
$$x_1(t) = X_{\mathrm{rms}}(\text{position 1})$$
$$x_2(t) = Y_{\mathrm{rms}}(\text{position 1})$$
$$\vdots$$
$$x_n(t) = X_{\mathrm{rms}}(\text{position n})$$



Model Independent Feedback A. Scheinker, "Extremum Seeking for Stabilization," PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, "Extremum Seeking with Discontinuous Dithers," *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t) = \begin{bmatrix} f_1(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \\ \vdots \\ f_n(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \end{bmatrix} \qquad y = V(\mathbf{x}, t) + n(t)$$

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



Dithering with different frequencies makes the parameters "perpendicular" in Hilbert space.

$$\omega_i = \omega r_i, \; r_i \neq r_j \implies \text{for any } t > 0$$

$$\lim_{\omega \to \infty} \langle \cos(\omega_i t), \cos(\omega_j t) \rangle = \lim_{\omega \to \infty} \int_0^t \cos(\omega_i \tau)\cos(\omega_j \tau) d\tau = 0$$

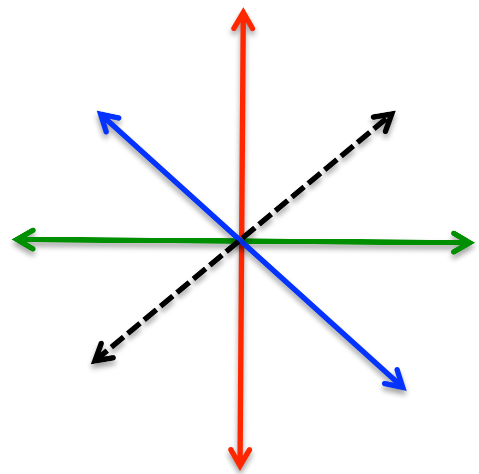
Los Alamos NATIONAL LABORATORY

Model Independent Feedback

A. Scheinker, “Extremum Seeking for Stabilization,” PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, “Extremum Seeking with Discontinuous Dithers,” *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}, t) = \begin{bmatrix} f_1(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \\ \vdots \\ f_n(x_1, \ldots, x_n, p_1, \ldots, p_m, t) \end{bmatrix} \qquad y = V(\mathbf{x}, t) + n(t)$$

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



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

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.

$$\omega_i = \omega r_i, \;\; r_i \neq r_j \implies \text{for any } t > 0$$

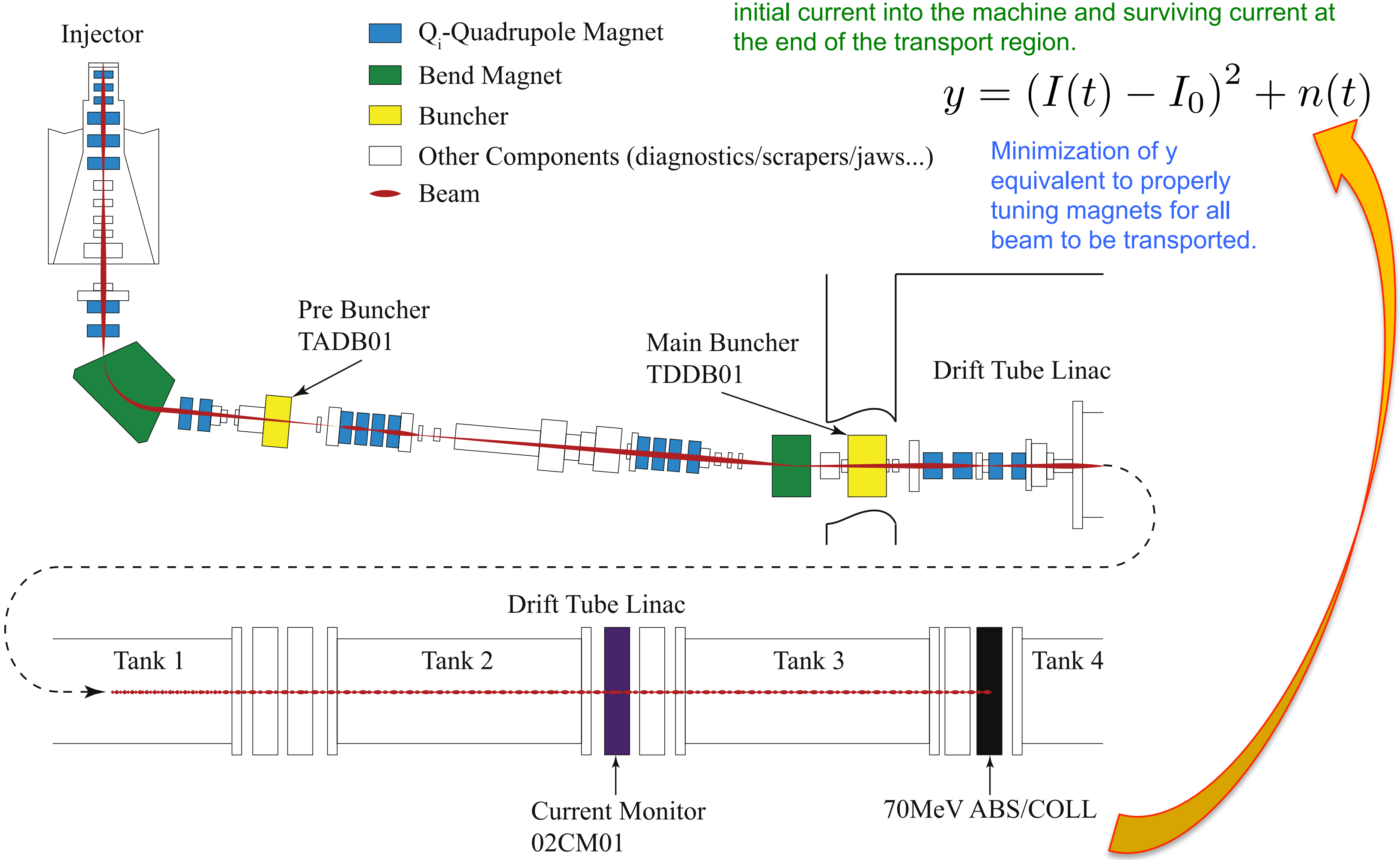
$$\lim_{\omega\to\infty} \langle \cos(\omega_i t), \cos(\omega_j t) \rangle = \lim_{\omega\to\infty} \int_0^t \cos(\omega_i \tau)\cos(\omega_j \tau) d\tau = 0$$

Model Independent Feedback A. Scheinker, “Extremum Seeking for Stabilization,” PhD Thesis, UCSD, 2012. A. Scheinker and D. Scheinker, “Extremum Seeking with Discontinuous Dithers,” *Automatica*, vol. 69, pp. 250-257, 2016. "Constrained extremum seeking stabilization of systems not affine in control." *International Journal of Robust and Nonlinear Control* 28.2, 2018.

Cost is noisy measurement of the difference between initial current into the machine and surviving current at the end of the transport region.

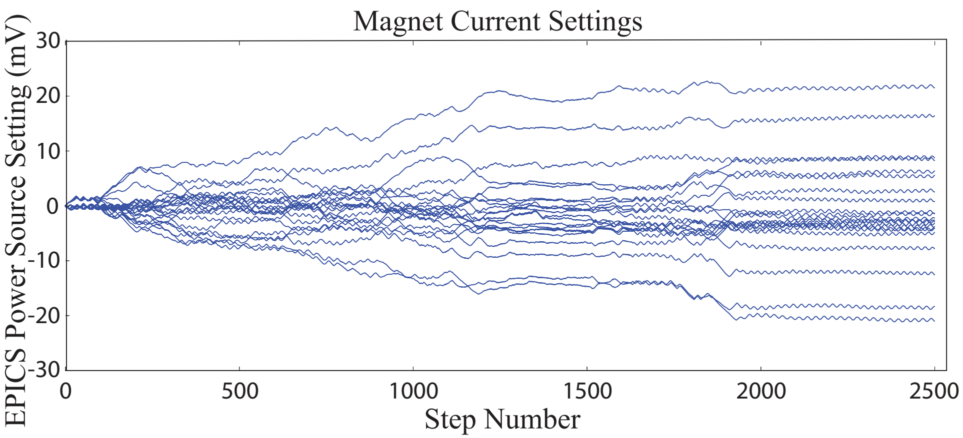
$$y = (I(t) - I_0)^2 + n(t)$$

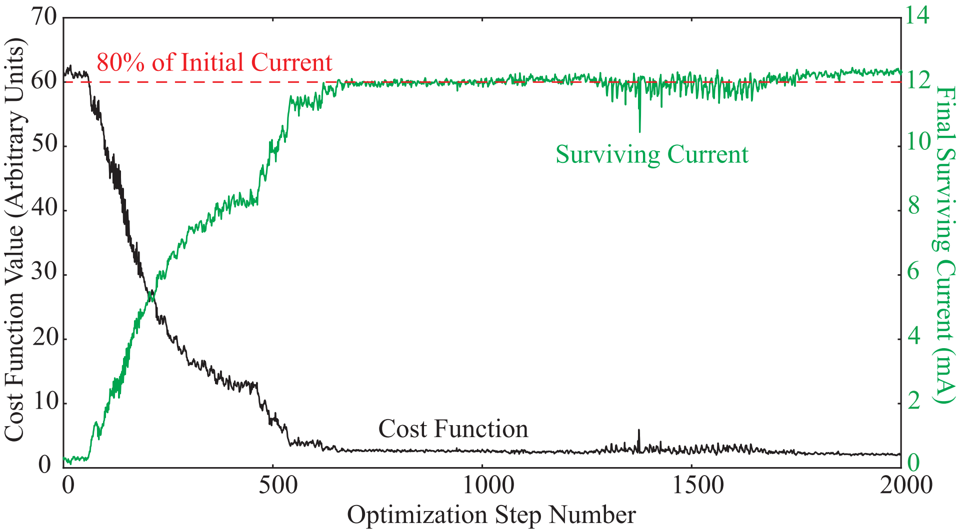
Minimization of y equivalent to properly tuning magnets for all beam to be transported.



LANSCE (simulation) Automated low energy beam transport region magnet tuning.

A. Scheinker, X. Pang, and L. Rybarcyk, “Model-independent particle accelerator tuning,” *Physical Review Accelerators and Beams*, 16, 102803, 2013.





Cost is noisy measurement of the difference between initial current into the machine and surviving current at the end of the transport region.

Minimization of y equivalent to properly tuning magnets for all beam to be transported.

Magnets were all turned off and then the algorithm automatically set them up to maximize the amount of surviving current at the end of the transport region.

80% is the typical maximum value achieved by operators in this beam line.

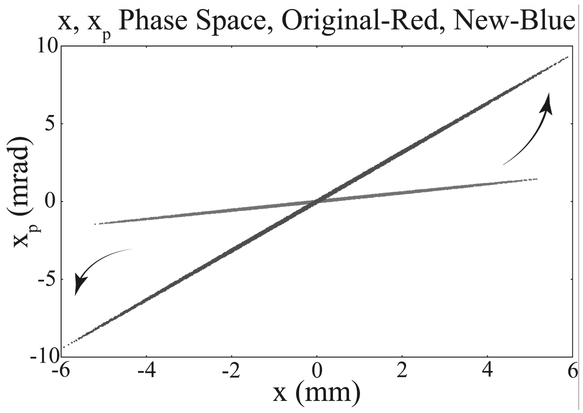
LANSCE (simulation) Automated low energy beam transport region magnet tuning.

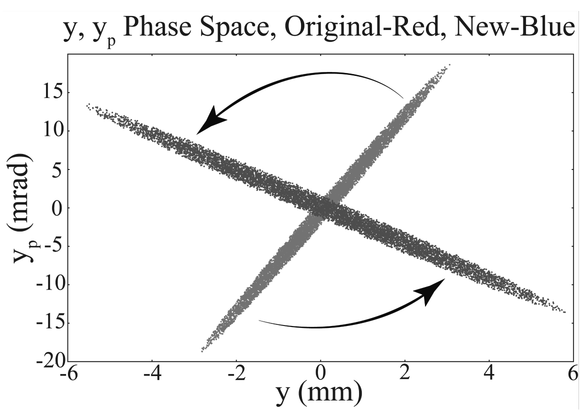
A. Scheinker, X. Pang, and L. Rybarcyk, “Model-independent particle accelerator tuning,” *Physical Review Accelerators and Beams*, 16, 102803, 2013.

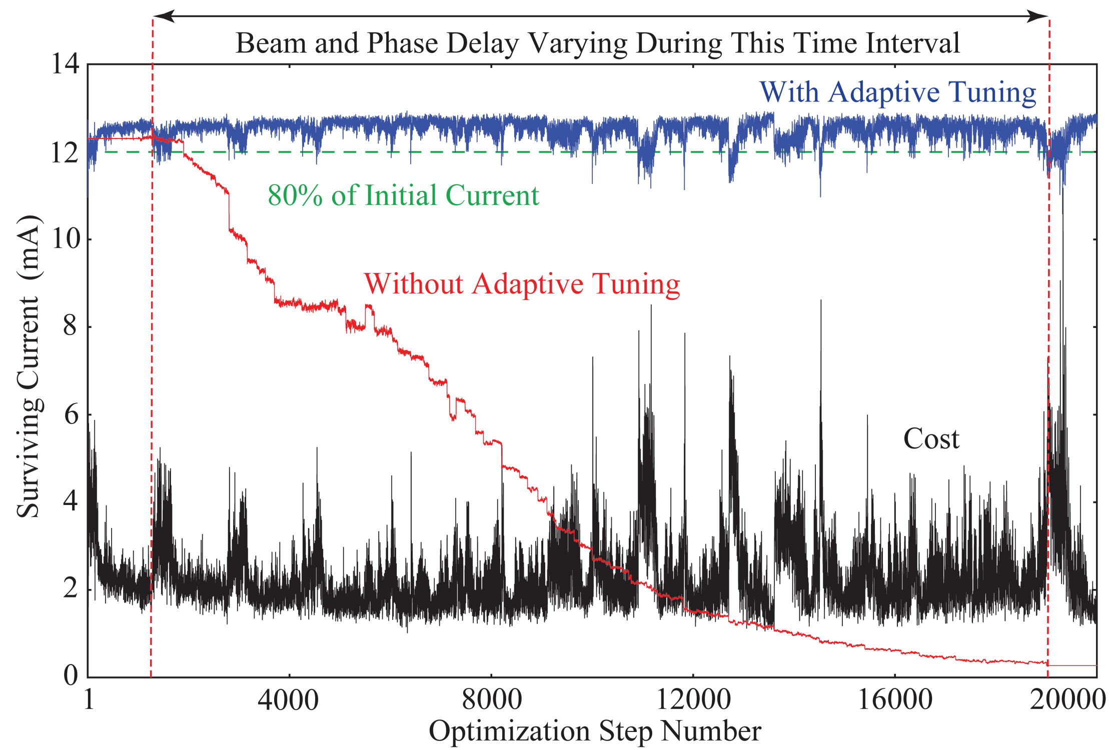
After the magnetic lattice was matched to transport the beam, beam phase space was continuously varied, and arbitrary phase drifts were introduced into the RF buncher cavities.

Without adaptive feedback all beam is quickly lost (red line in figure below).

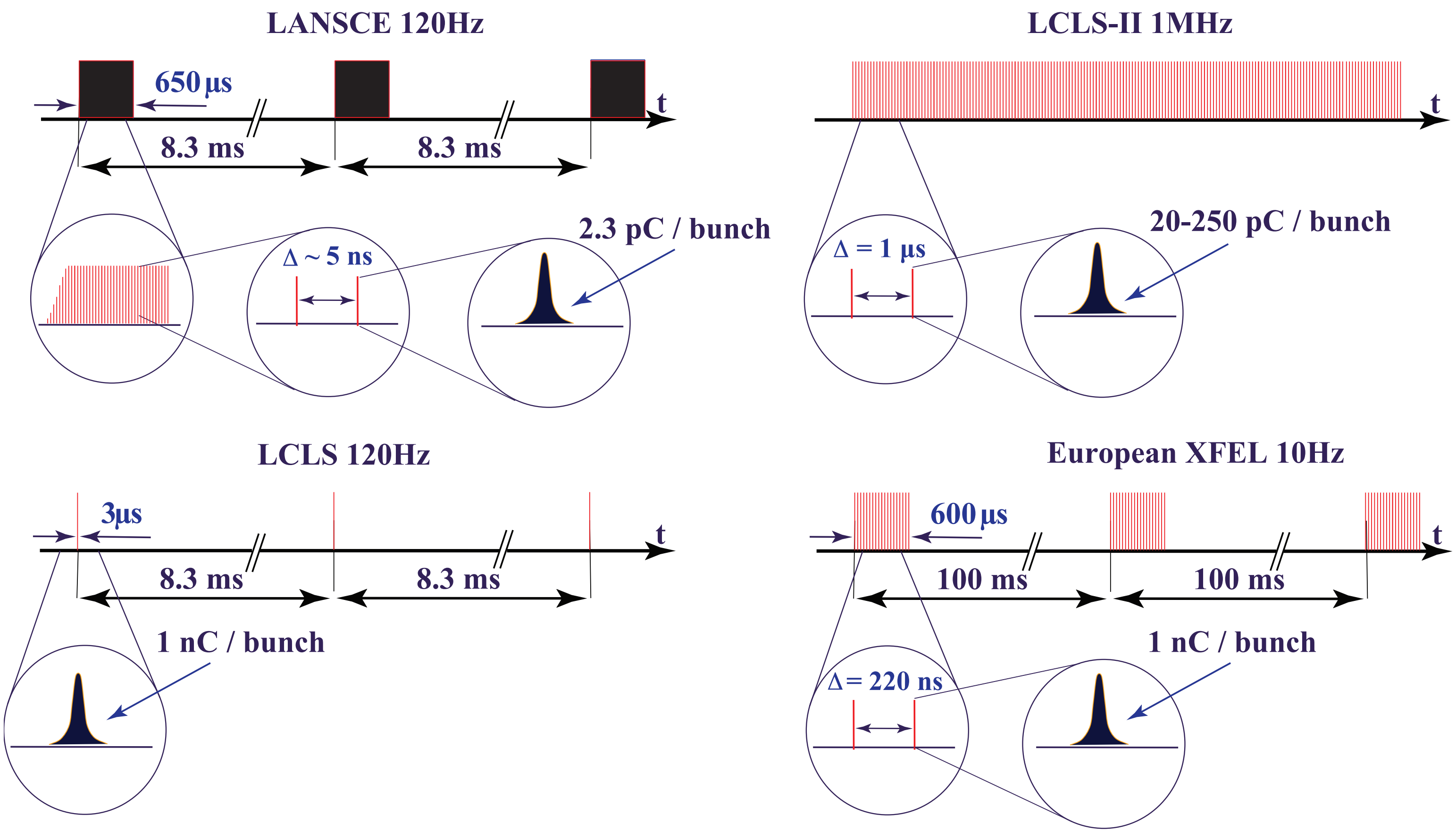
With adaptive tuning the 22 quad magnetic lattice and 2 RF buncher cavities are continuously re-tuned to maintain maximal beam transmission and acceleration.





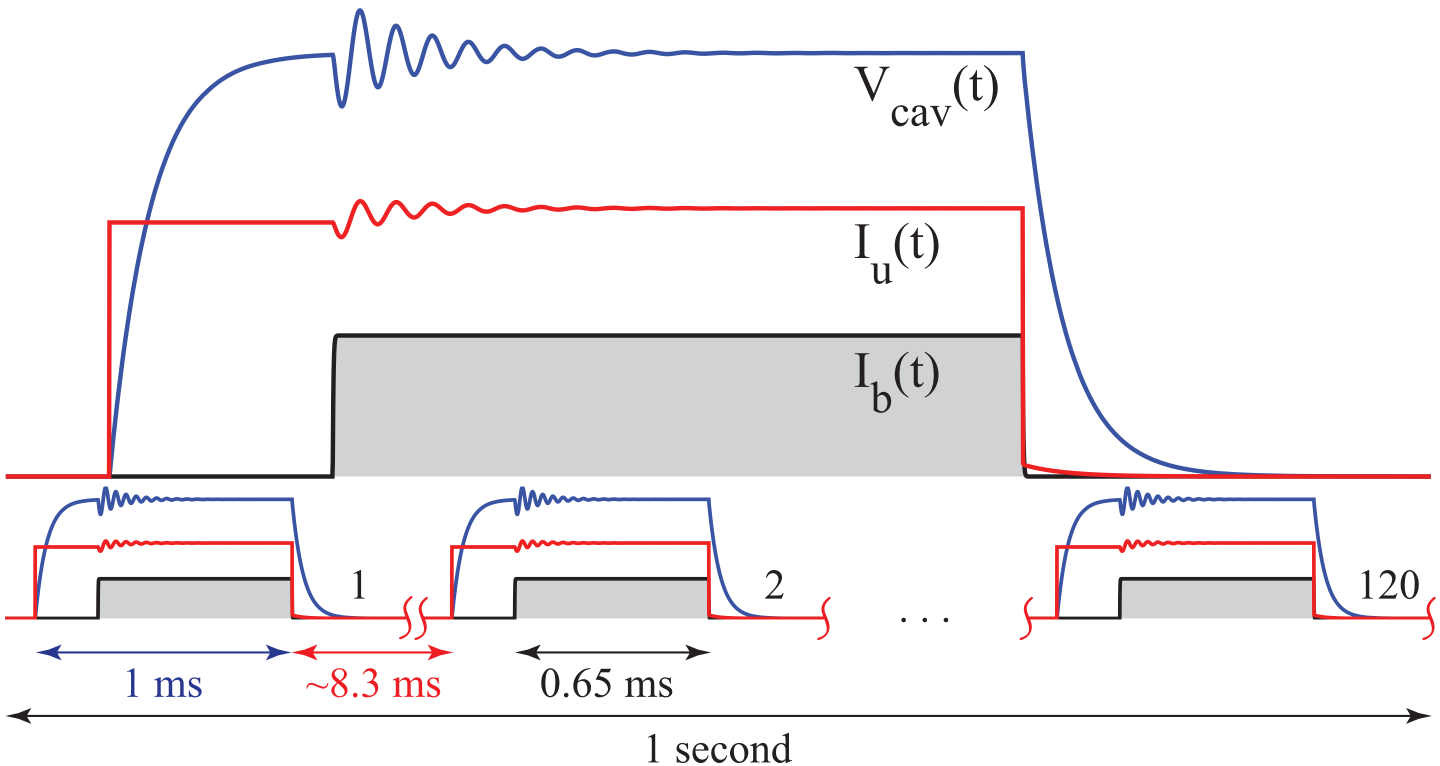


Beam loading



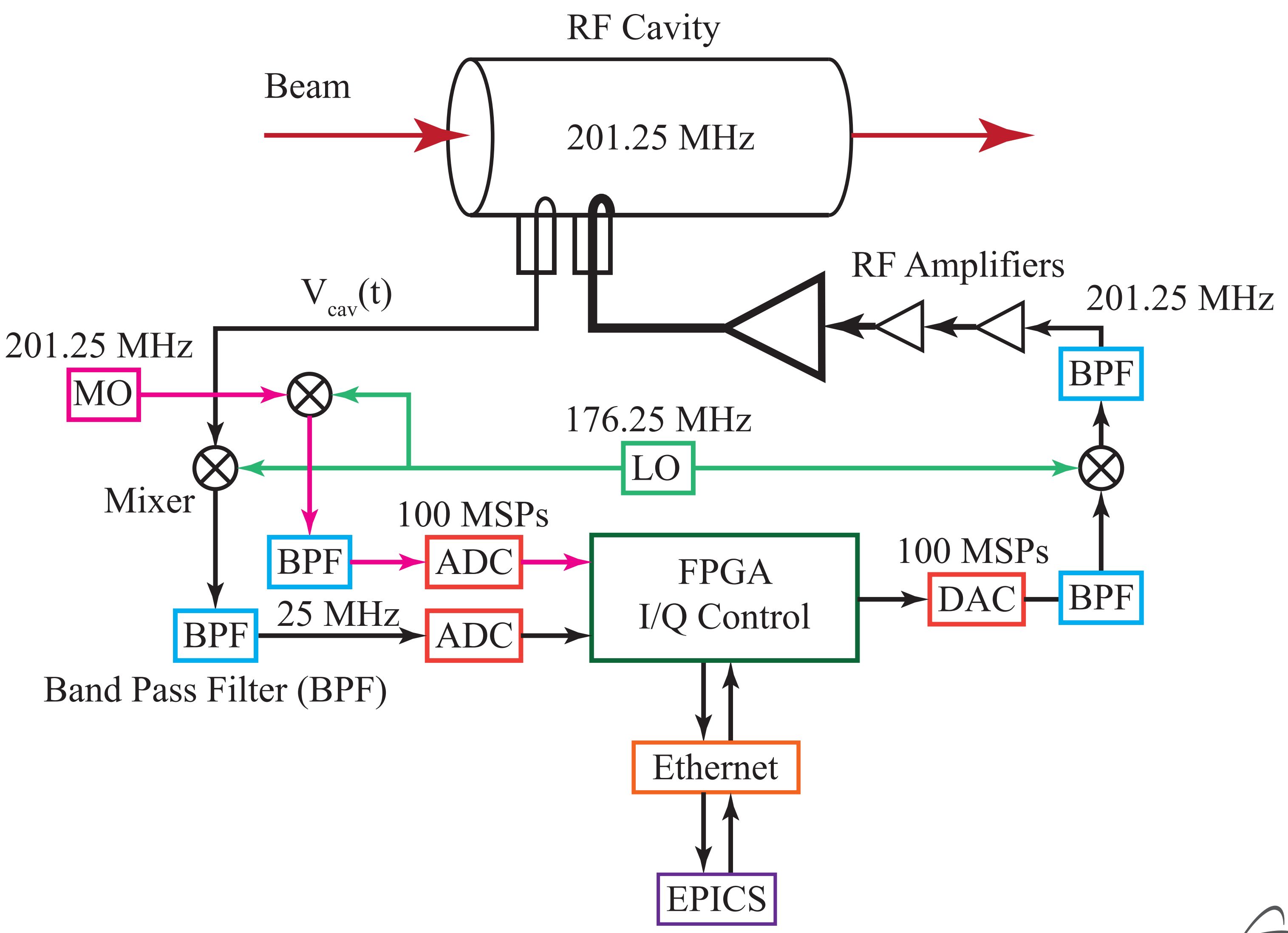
LANSCE Beam Loading

H+ ~400μA average current: 21mA peak current, 625μs pulses @ 30Hz



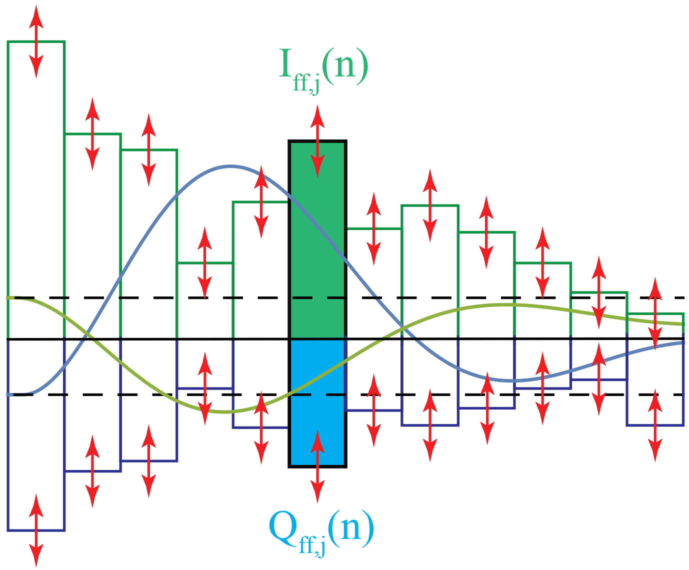
LANSCE: Iterative feed-forward beam loading compensation.

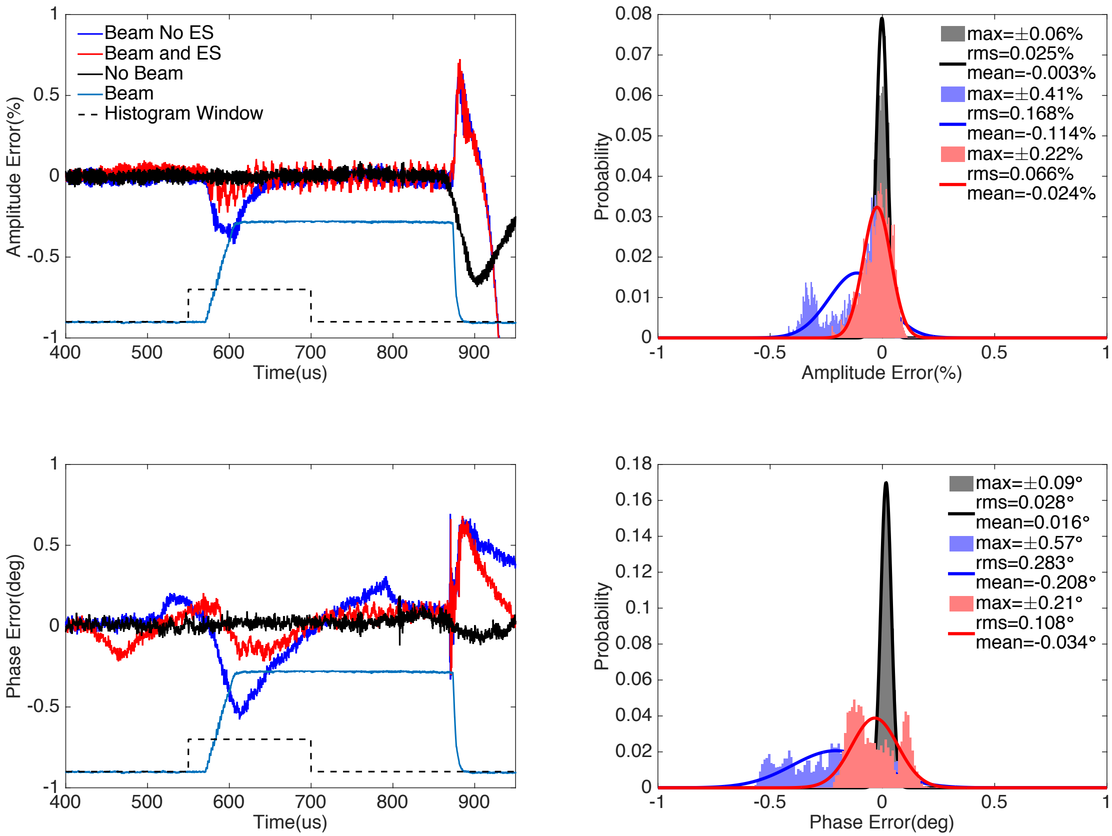
A. Scheinker et al., "Preliminary Study of Advanced LLRF Controls at LANSCE for Beam Loading Compensation in the MaRIE X-FEL." 2016 *North American Particle Accelerator Conf.*



Experimental results at LANSCE, with 12mA proton beam.

High speed FPGAs:
5ns – 10us resolution (1ns soon)

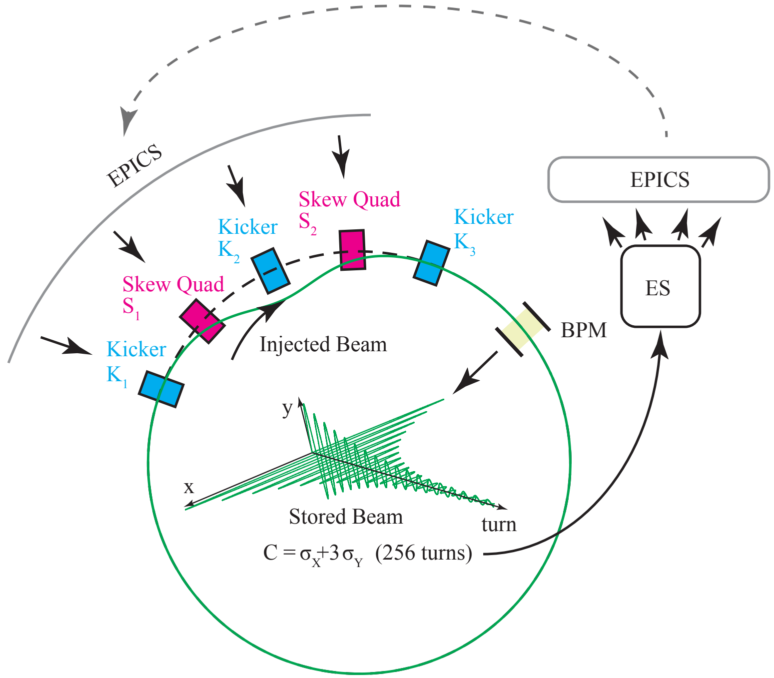


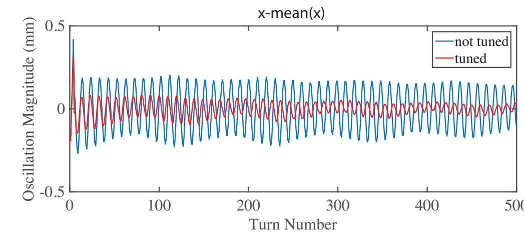


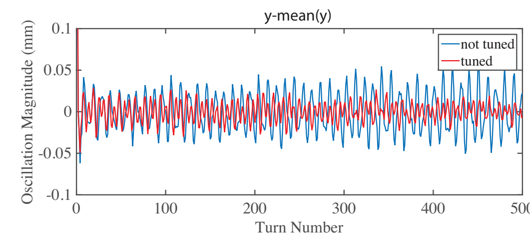
	No Beam	Beam, No ES	Beam & ES
max A error (%)	±0.06	±0.41	±0.22
rms A error (%)	0.025	0.168	0.066
mean A error (%)	-0.003	-0.114	-0.024
max θ error (deg)	±0.09	±0.57	±0.21
rms θ error (deg)	0.028	0.283	0.108
meanθ error (deg)	0.016	-0.208	-0.034

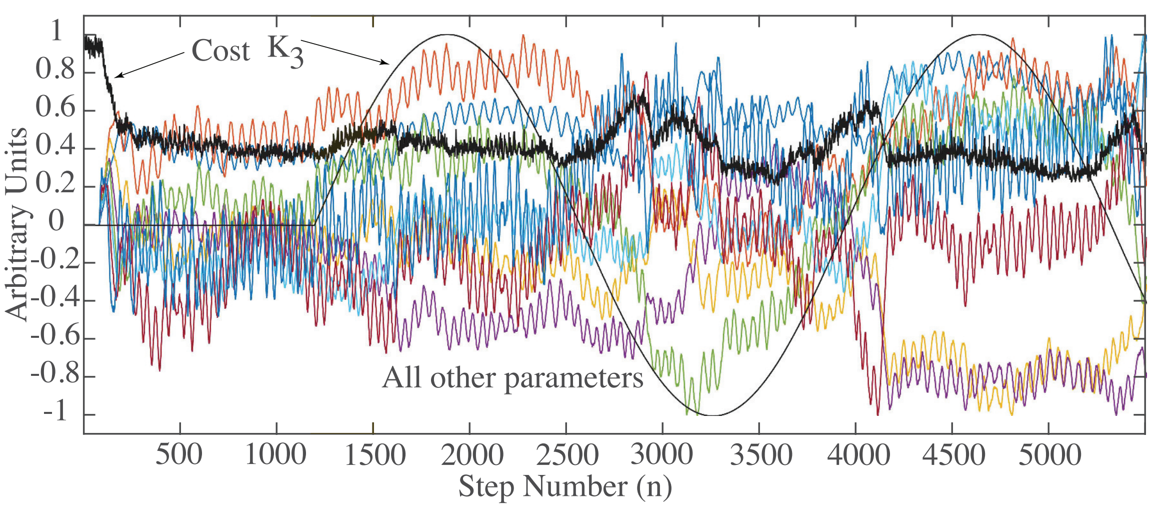
SPEAR3: Beam-based Feedback, Betatron oscillation minimization in time-varying lattice.

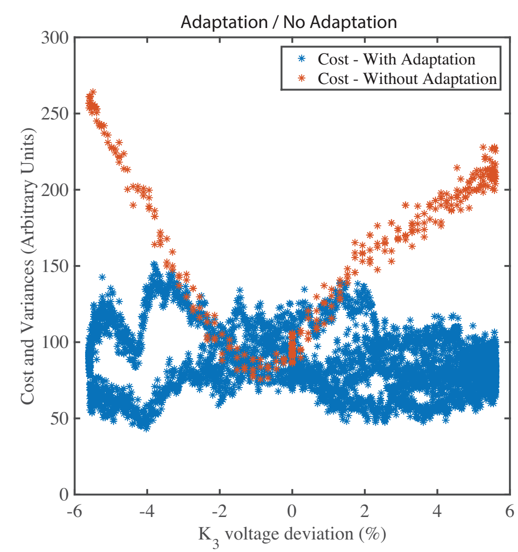
A. Scheinker, X. Huang, and J. Wu, “Minimization of Betatron oscillations of electron beam injected into a time-varying lattice via extremum seeking,” *IEEE Transactions on Control Systems Technology*, under review, 2016.



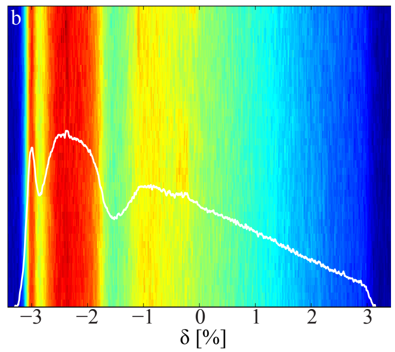


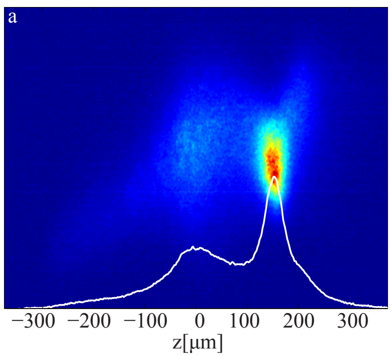


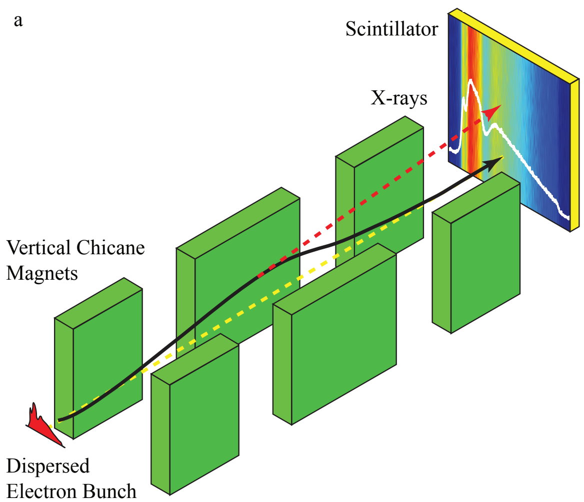


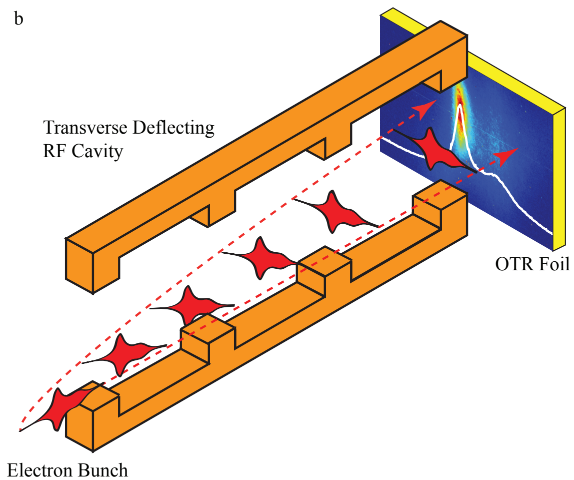


FACET: E-bunch profile prediction based on non-destructive measurements of energy spread spectrum. A. Scheinker and S. Gessner, "Adaptive method for electron bunch profile prediction," *Physical Review Accelerators and Beams*, 18, 102801, 2015.

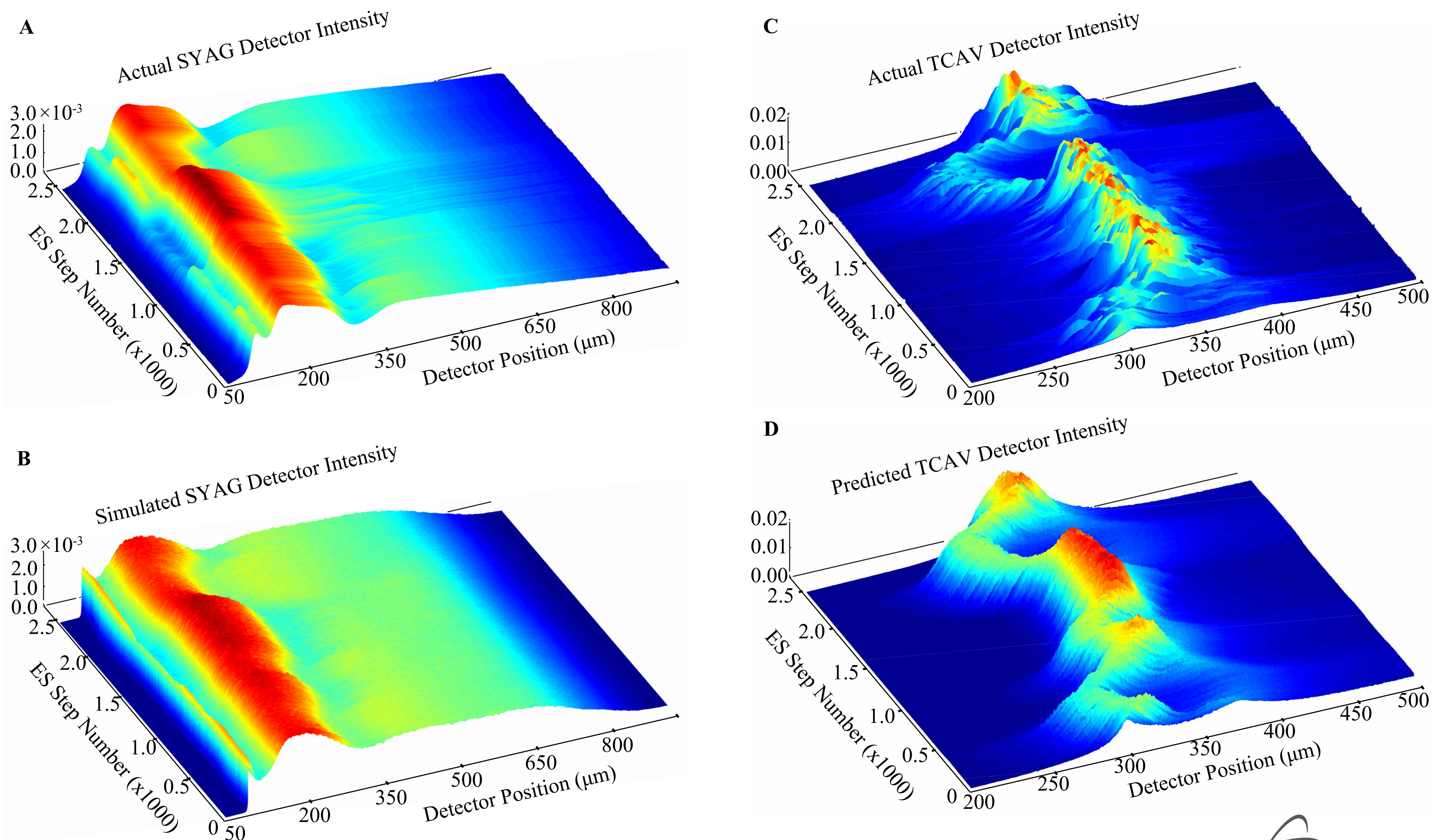


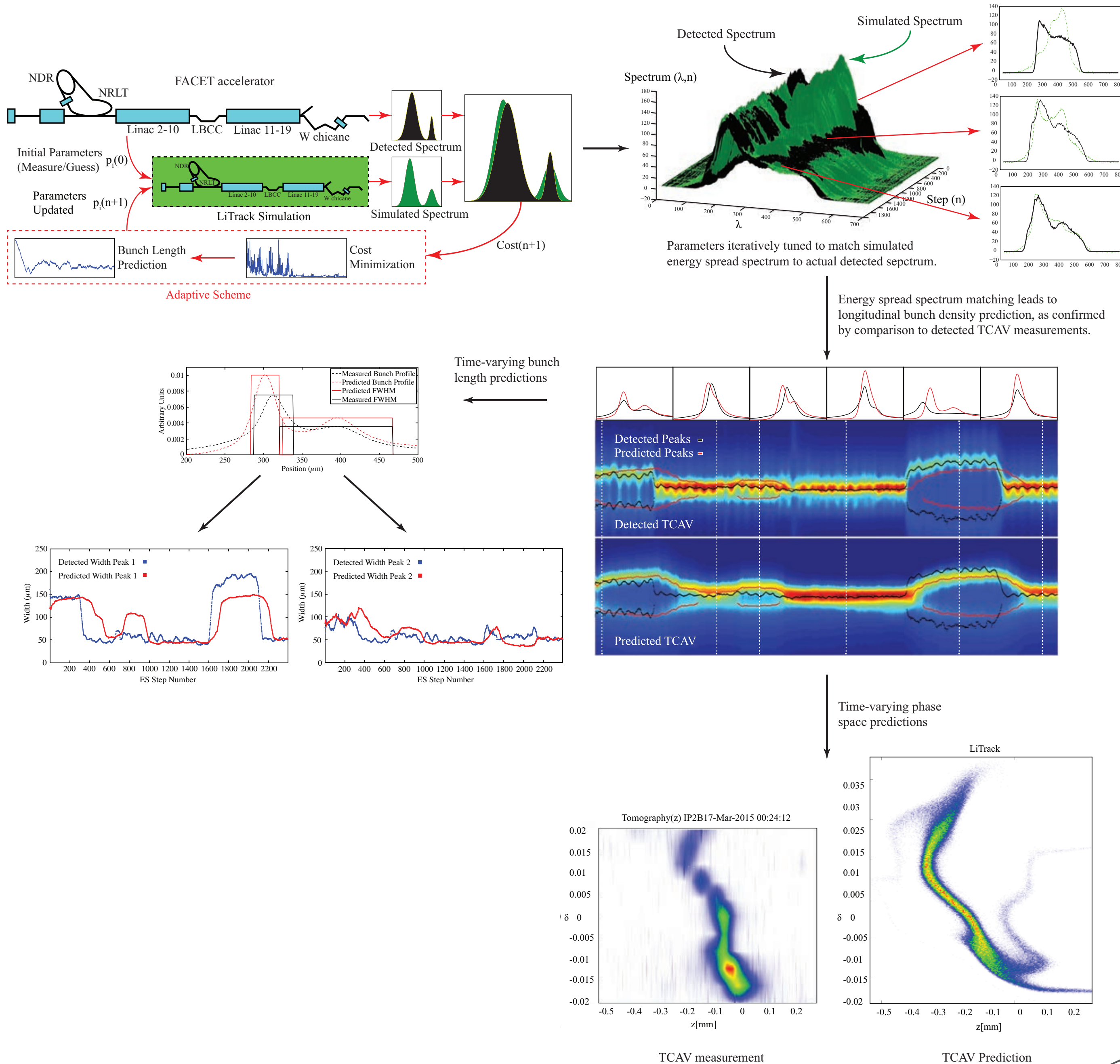






FACET: Tracking time-varying electron bunch profile based on non-destructive measurements of energy spread

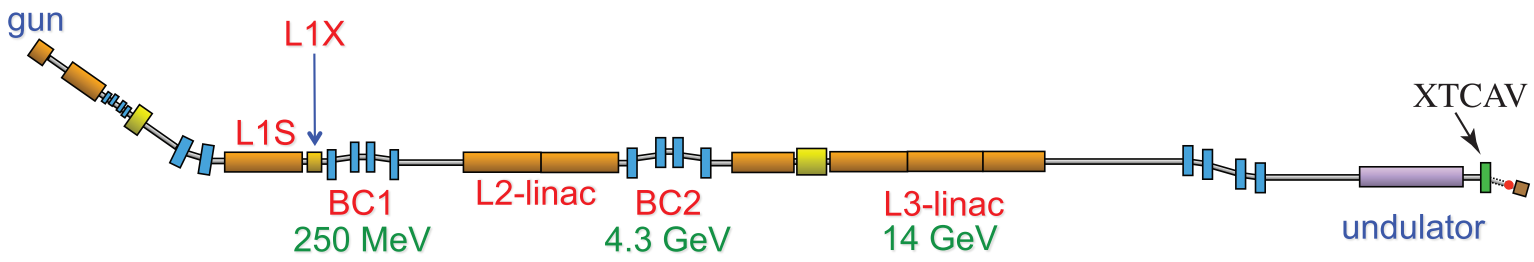


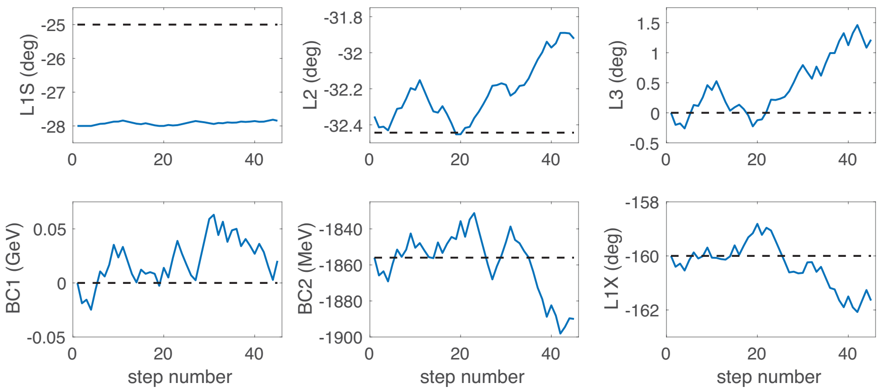


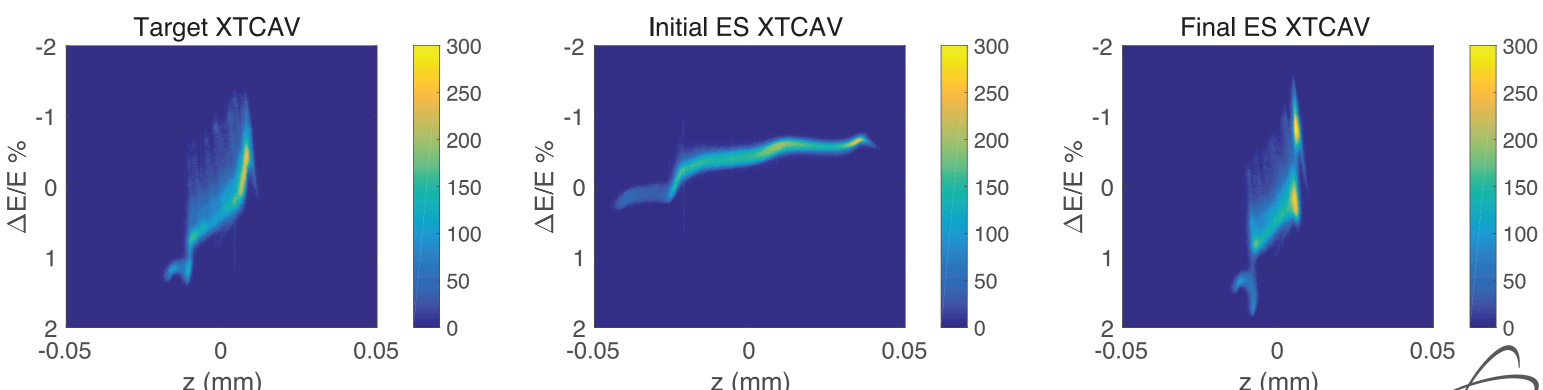
Los Alamos NATIONAL LABORATORY

LCLS automatic longitudinal phase space tuning. A. Scheinker. and D Bohler, “Demonstration of model-independent control of the longitudinal phase space of electron beams in the LCLS with fs resoltuions,” *submitted.*

Motivation: Fast switching/tuning between XFEL experiments, custom current profiles (FACET-II).

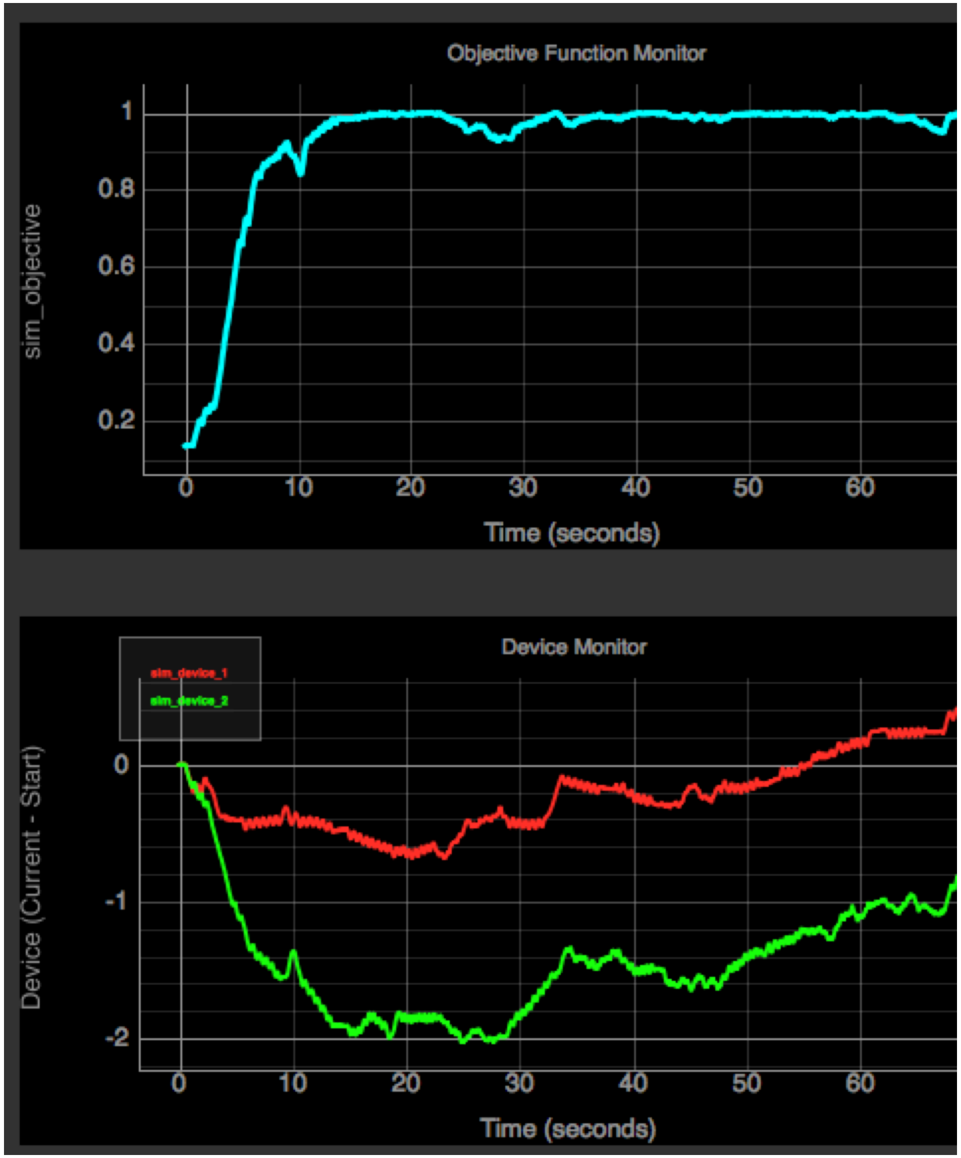






Los Alamos NATIONAL LABORATORY

ES in OCELOT for time-varying system



Adaptive machine learning