



Source Size Stabilization at the ALS using Machine Learning

Simon C. Leemann

ALS Accelerator Physics, ATAP & ALS Divisions, Lawrence Berkeley National Laboratory
October 28, 2020



8th Low Emittance Rings Workshop, October 26-30, 2020, LNFN-INFN, Frascati, Italy



U.S. DEPARTMENT OF
ENERGY

Office of
Science

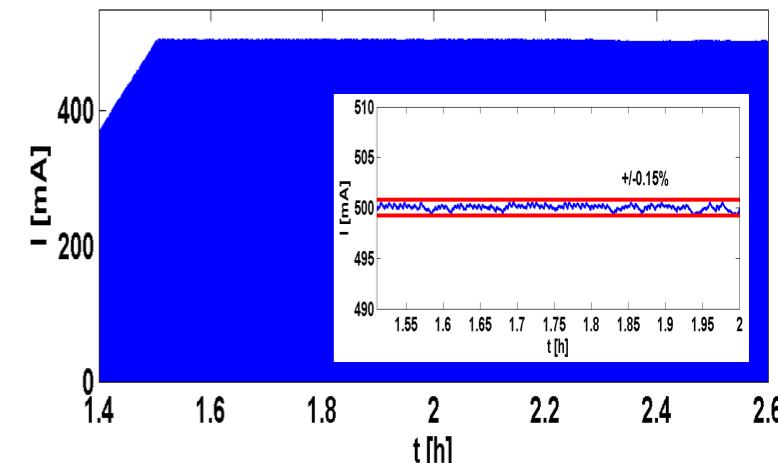


ACCELERATOR TECHNOLOGY &
APPLIED PHYSICS DIVISION **ATAP**

ALS
ADVANCED LIGHT SOURCE

Many Successful Efforts to Stabilize Electron Beams

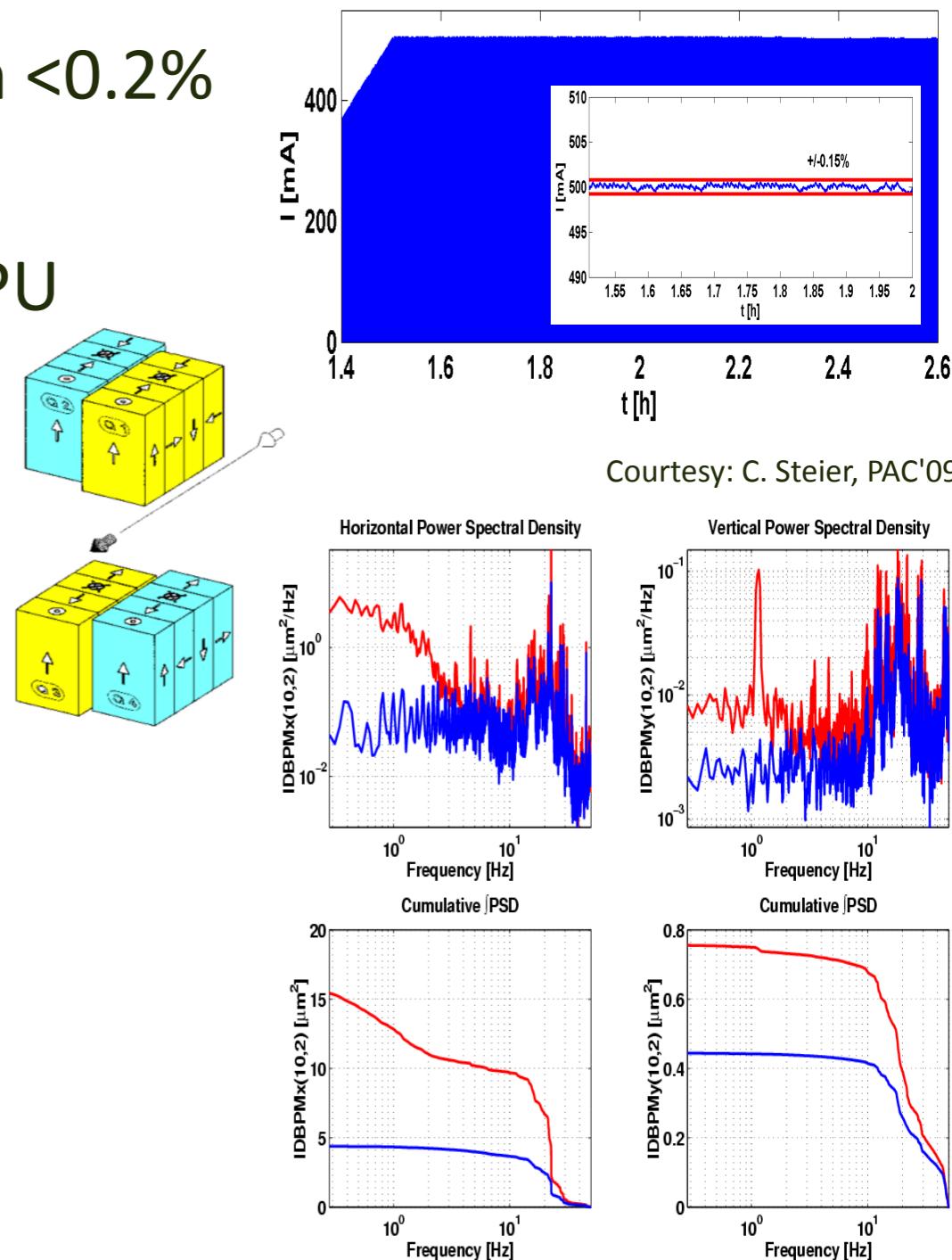
- Top-off keeps ALS stored current variation <0.2%



Courtesy: C. Steier, PAC'09

Many Successful Efforts to Stabilize Electron Beams

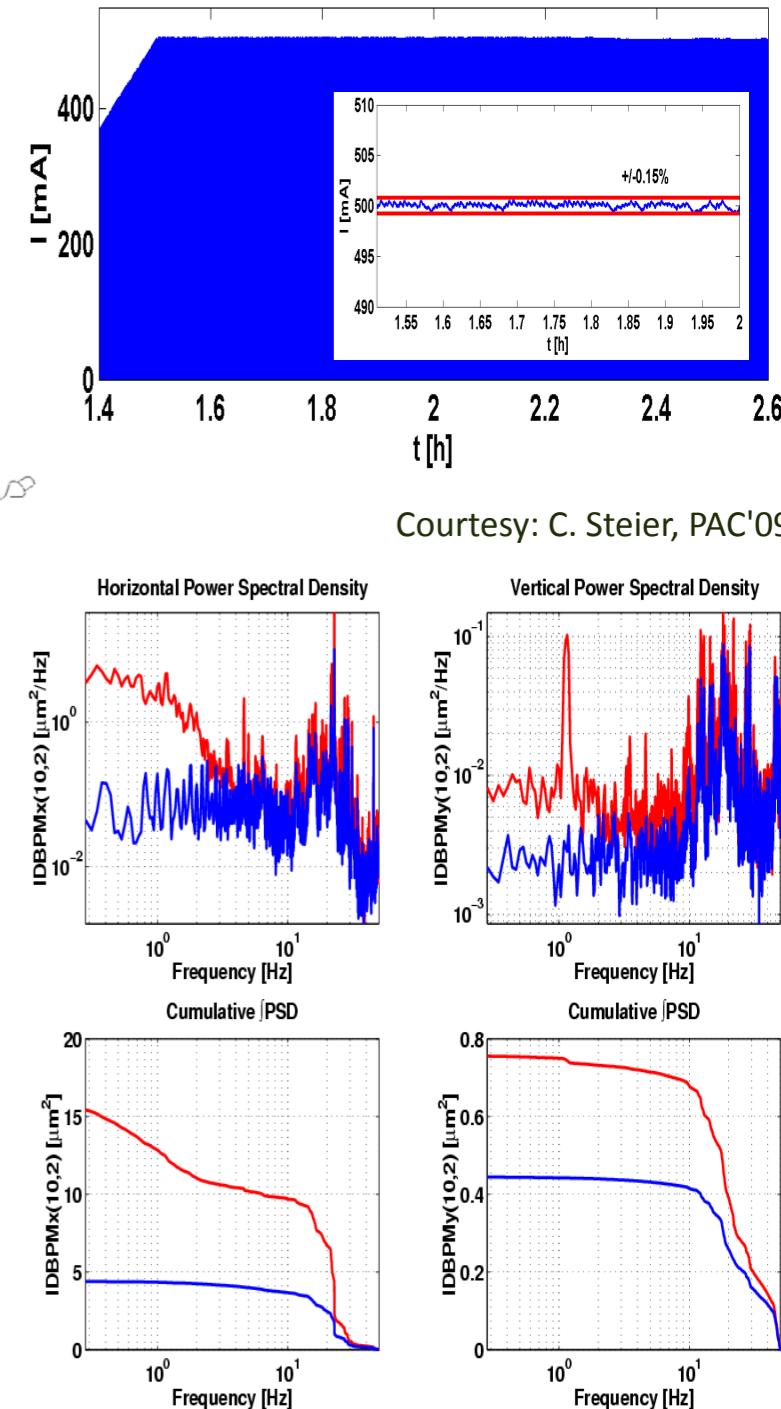
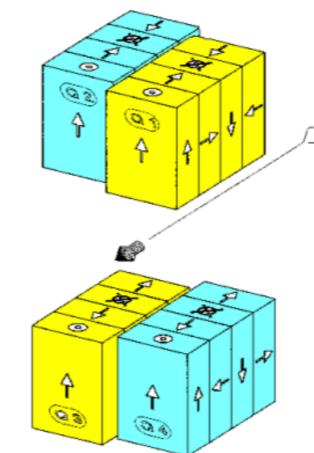
- Top-off keeps ALS stored current variation <0.2%
- At low energy, ALS strongly affected by ID imperfections & continuously changing EPU gaps/phases
 - Orbit feedback and ID feed-forwards stabilize source positions/angles to **sub-micron** level at many tens of Hz



Courtesy: C. Steier, PAC'09

Many Successful Efforts to Stabilize Electron Beams

- Top-off keeps ALS stored current variation <0.2%
- At low energy, ALS strongly affected by ID imperfections & continuously changing EPU gaps/phases
 - **Orbit feedback** and ID feed-forwards stabilize source positions/angles to **sub-micron** level at many tens of Hz
 - **ID feed-forwards** & tune feedback stabilize optics at source points
 - **ID skew feed-forwards** stabilize source size
 - require recording lookup tables (time consuming)
 - tables are imperfect and **machine drifts** over time

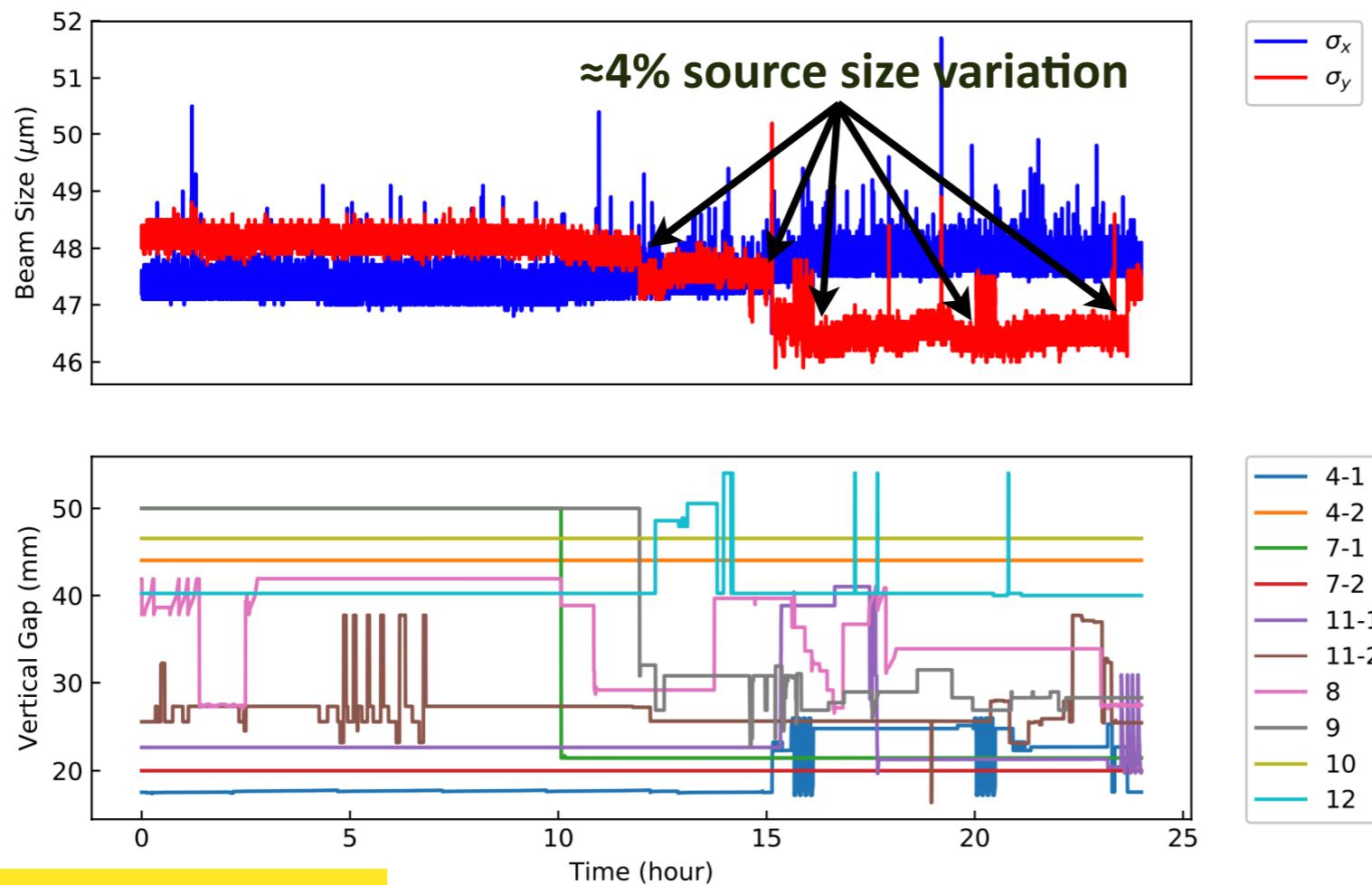


Courtesy: C. Steier, PAC'09

Thermal, Ground, Water Table, etc.

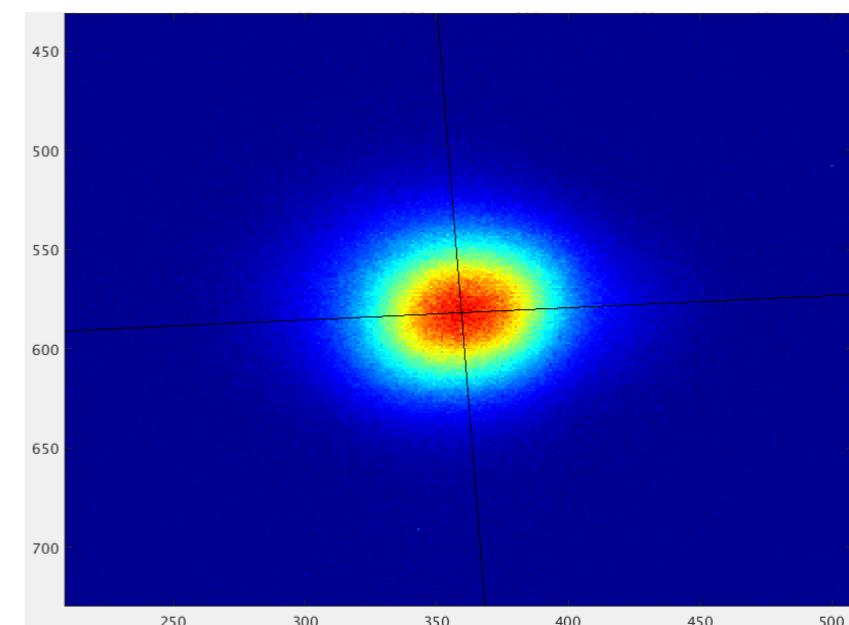
The Problem: Beam Size vs. ID Motion

- Nevertheless, during routine user ops observe vertical source size variations when ID configurations change



PRL 123, 194801 (2019)

ALS Diagnostic Beamline 3.1



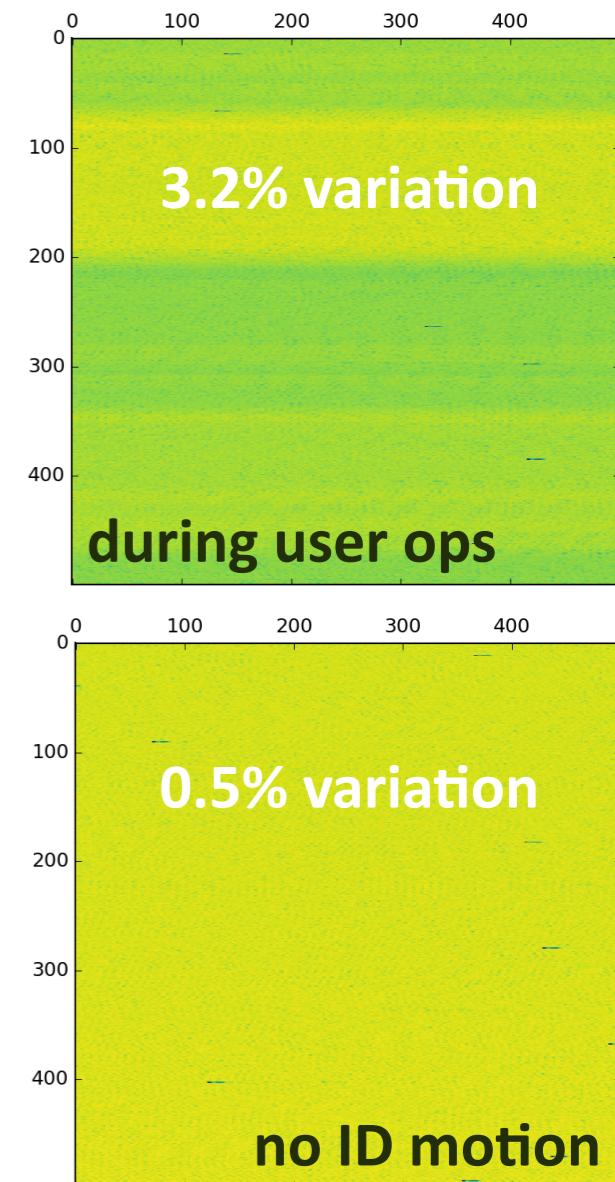
SR from 1st arc dipole ("round beam") → KB mirrors → C filter → 1-3 keV x-rays → LYSO scintillator crystal → visible → CCD

Rev. Sci. Instrum. 67, 3368 (1996)

- Traditionally 3rd-gen. sources considered <10% acceptable, but...

How this Problem Affects Sensitive Experiments

- Vertical source size fluctuations show up as intensity variations at highly sensitive beamlines, such as the STXM at ALS beamline 5.3.2.2
 - STXM zone plate focal length ≈ 1 mm \rightarrow no independent & reliable I_0 measurement
 - Very small spot size in focus (>20 nm \rightarrow scan $>10 \times 10 \mu\text{m}^2$)
 - Fast raster scanning for differential measurements \rightarrow no averaging (≈ 1 ms/pixel, 1 s/line, 6 min/scan)
 - Monochromator plane is H \rightarrow V source size fluctuations directly affect experimental noise floor
- 4th-gen. sources such as ALS-U will be equipped with many more such highly sensitive beamlines: STXM, XPCS, ptychography, etc.



PRL 123, 194801 (2019)

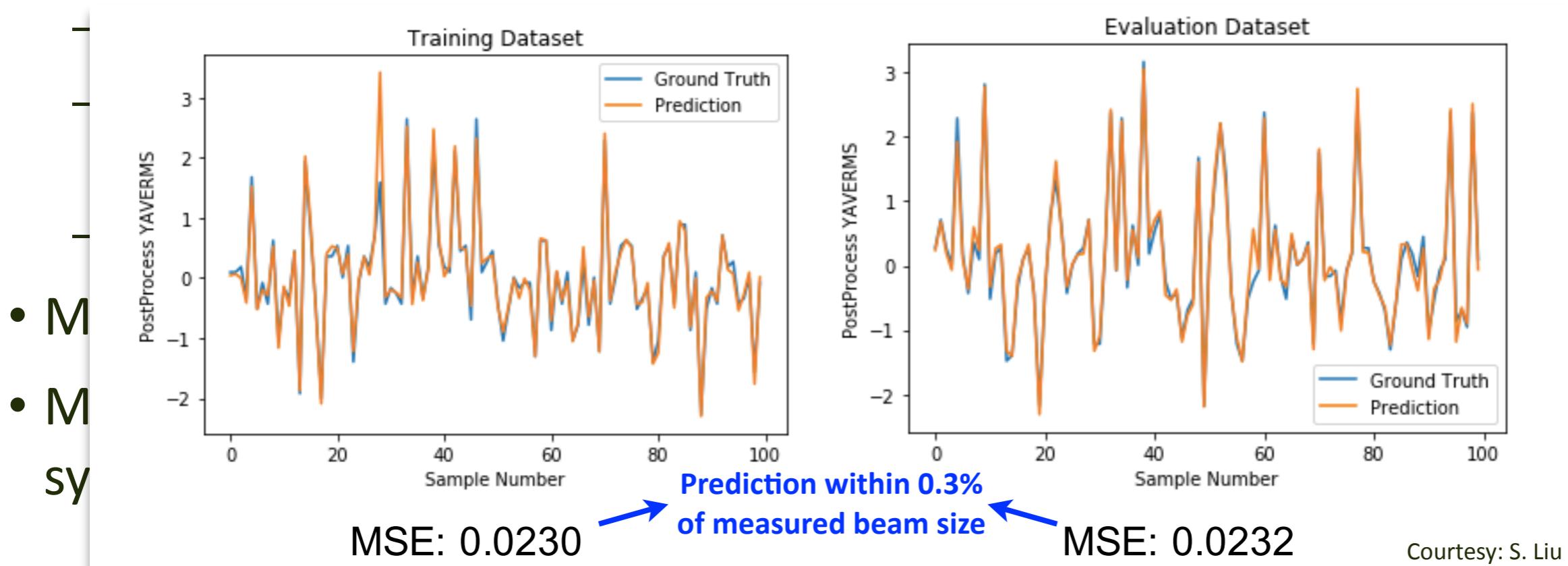


We Need to Solve This Problem at the Source

- Why use **Machine Learning (ML)** to attack this issue?
 - ML can model highly nonlinear processes and is extremely flexible
 - ML does not require a priori understanding underlying physics (e.g. machine drift) → but allows extracting valuable system information a posteriori
 - ML can substantially outperform conventional fitting (polynomial regression)
- ML requires reproducible events → confirmed in experiments
- ML ideally requires large data sets for training → ALS digital control system can provide that

We Need to Solve This Problem at the Source

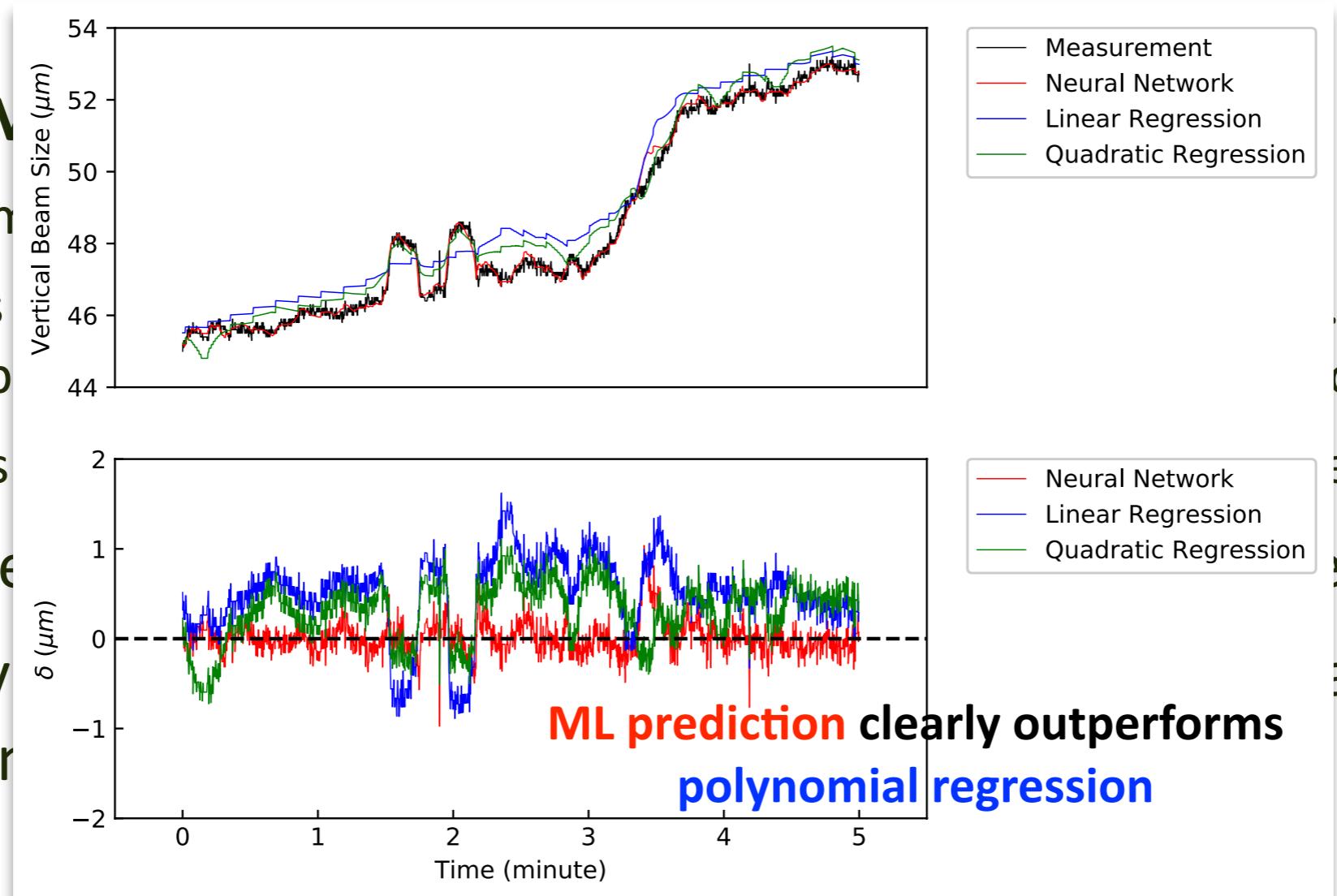
- Why use Machine Learning (ML) to attack this issue?



- First example: offline analysis of user ops data
 - 26 ID parameters ("input") → predict V beam size @ BL3.1 ("output")
 - Recorded 8 Msamples @ 10 Hz → 6 Msamples used for training, 2 Msamples for validation → training took 30 min on powerful GPU

We Need to Solve This Problem at the Source

- Why use ML?
 - ML can handle non-linearities
 - ML does not drift (\rightarrow better control)
 - ML can stabilize
- ML requires training
- ML ideally needs to be part of the system control



- First example: offline analysis of user ops data
 - 26 ID parameters ("input") \rightarrow predict V beam size @ BL3.1 ("output")
 - Recorded 8 Msamples @ 10 Hz \rightarrow 6 Msamples used for training, 2 Msamples for validation \rightarrow training took 30 min on powerful GPU

PRL 123, 194801 (2019)

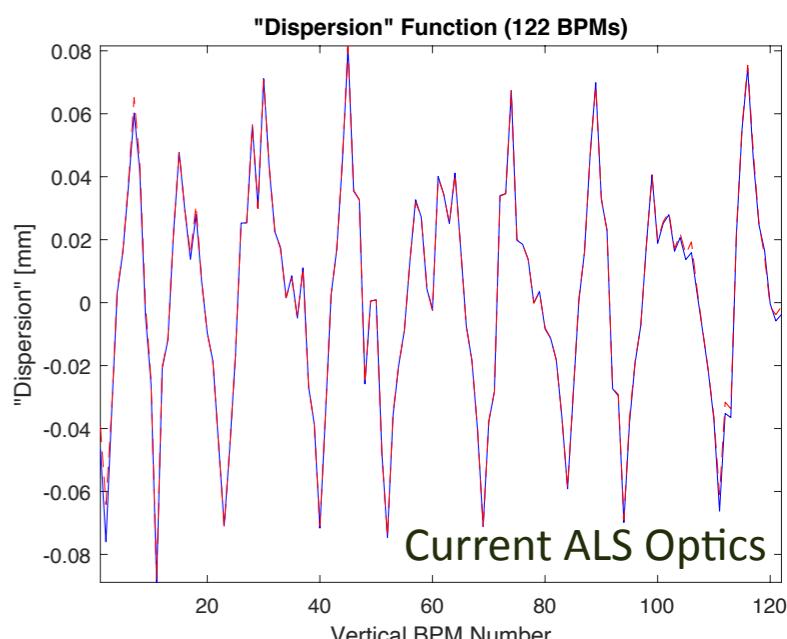
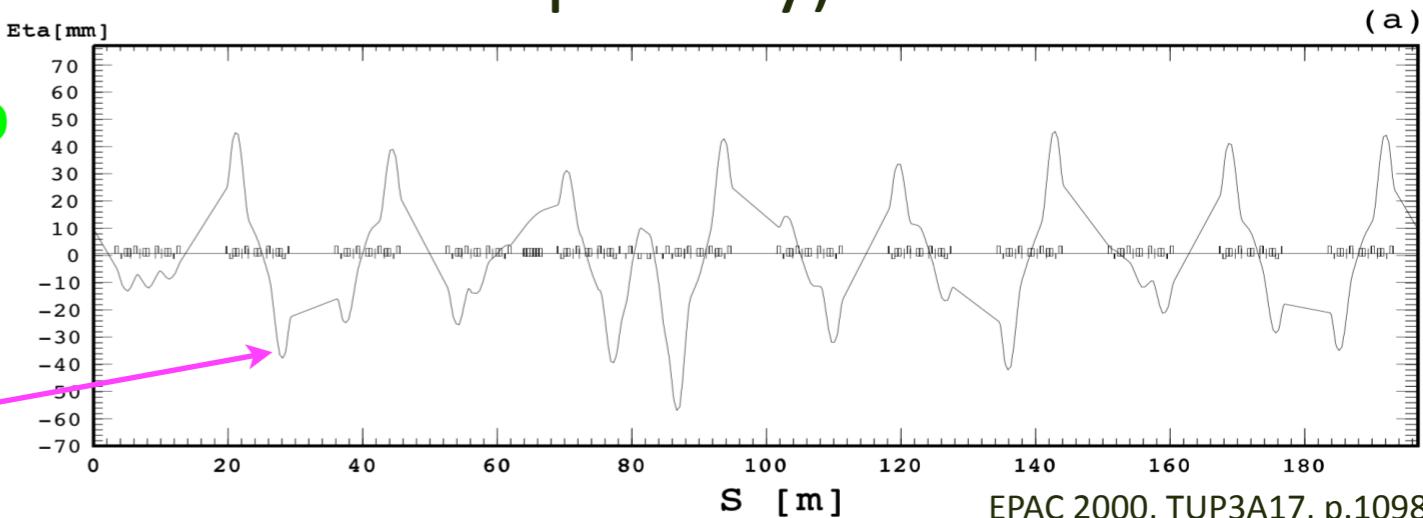
From Prediction to Correction

- Introduced "dispersion wave parameter" (DWP) to modify standard ALS dispersion wave (skew quadrupole excitation pattern) → allows adjusting vertical emittance (global conserved quantity)

$$\vec{K} = \vec{K}_0 + (\chi_0 + \chi) \Delta \vec{K}, \quad \vec{K} \in \mathbb{R}^{16+16}$$

LOCO & Setup DWP Dispersion Wave

SQSF SQSD



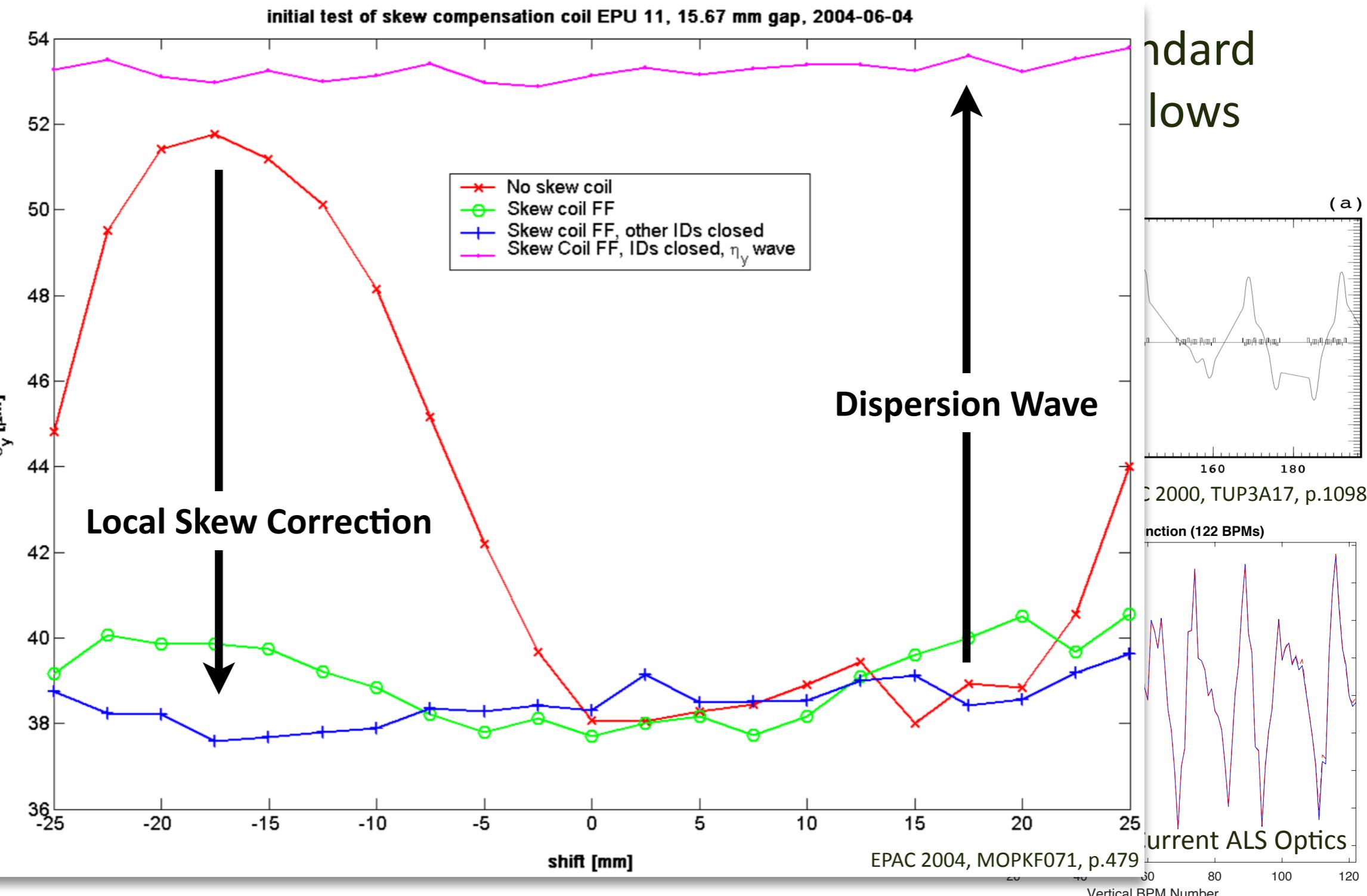
From Prediction to Correction (cont.)

- Introducing ALS dispersion adjustment

$$\vec{K} = \vec{K}_0 +$$

LOCO & Setup

σ_y [μm]



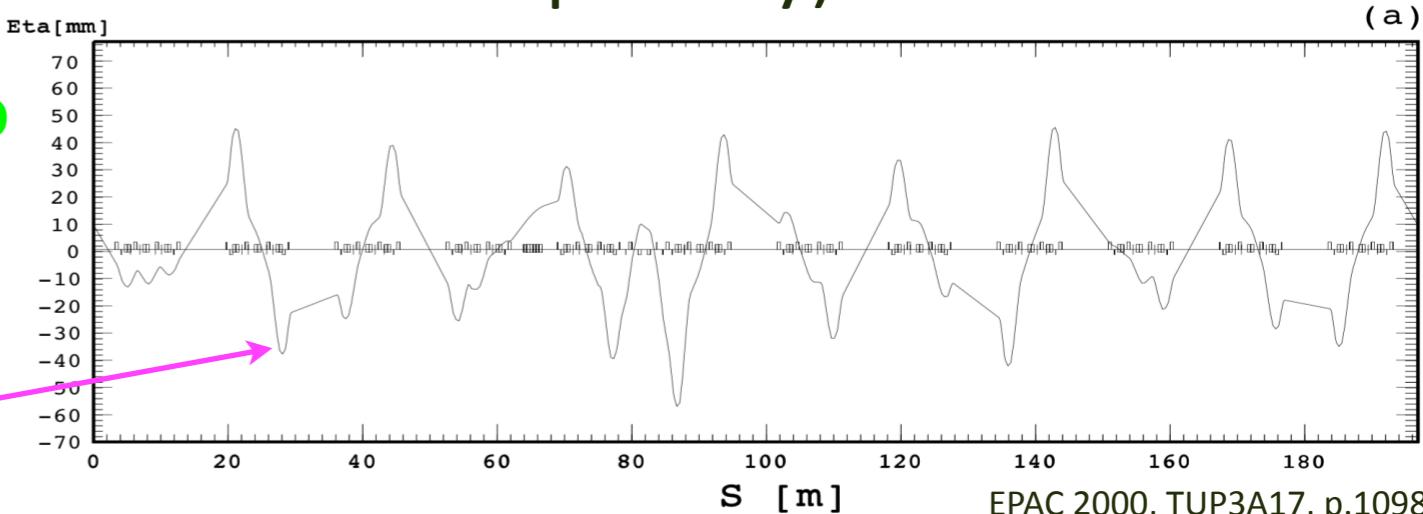
From Prediction to Correction (cont.)

- Introduced "dispersion wave parameter" (DWP) to modify standard ALS dispersion wave (skew quadrupole excitation pattern) → allows adjusting vertical emittance (global conserved quantity)

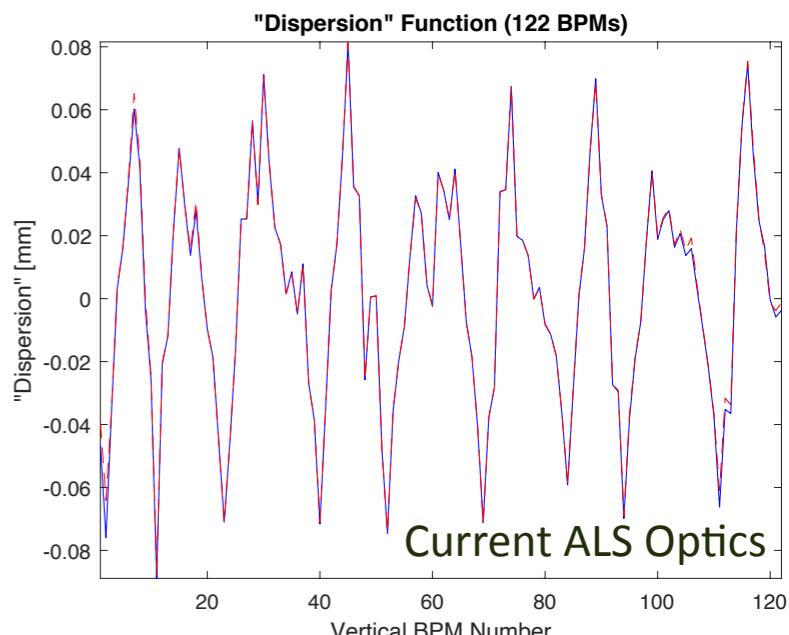
$$\vec{K} = \vec{K}_0 + (\chi_0 + \chi) \Delta \vec{K}, \quad \vec{K} \in \mathbb{R}^{16+16}$$

SQSF SQSD
Dispersion Wave

LOCO & Setup DWP

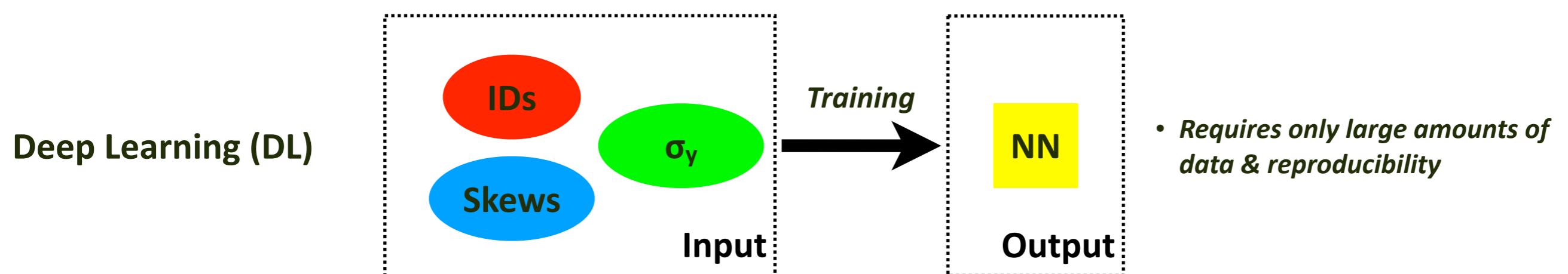


- Observed varying ID configurations affect primarily vertical dispersion $\rightarrow \epsilon_y$
- Can therefore stabilize beam size globally by adjusting DWP



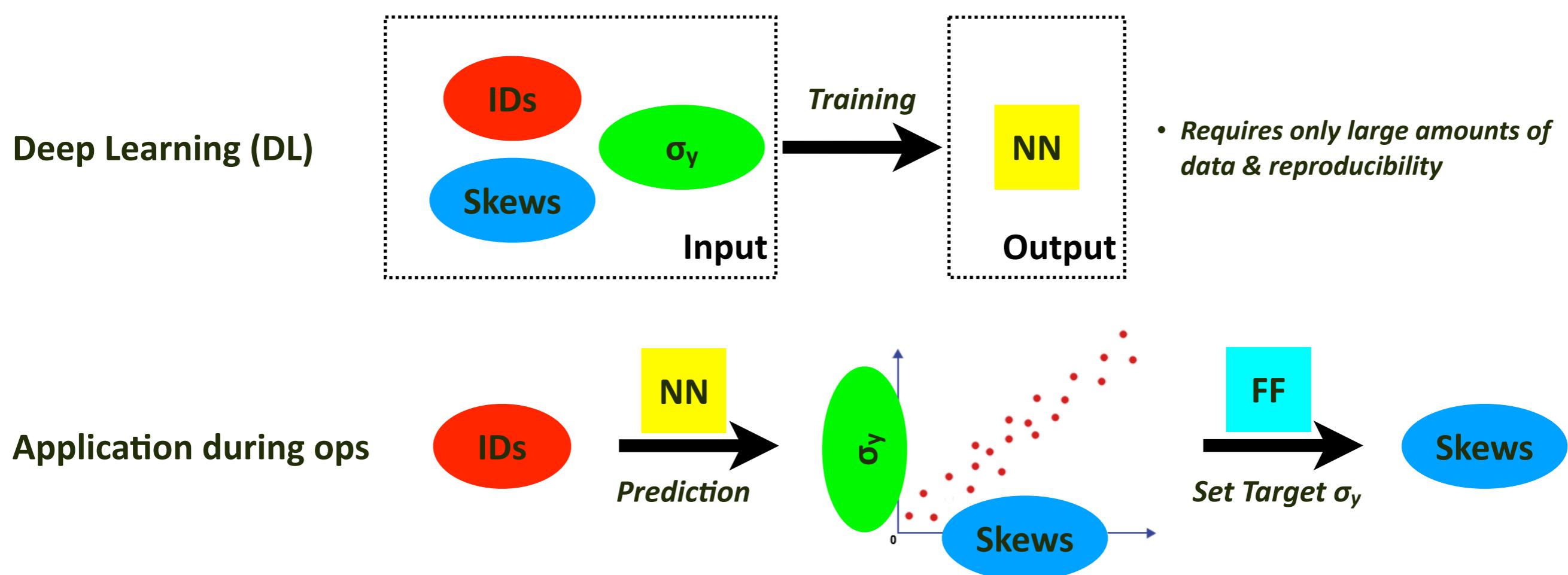
Building a Neural Network-based ID Feed-Forward

- Training: measure beam sizes while scanning DWP & various ID configurations → acquire data at 10 Hz → ***train*** neural network (NN)

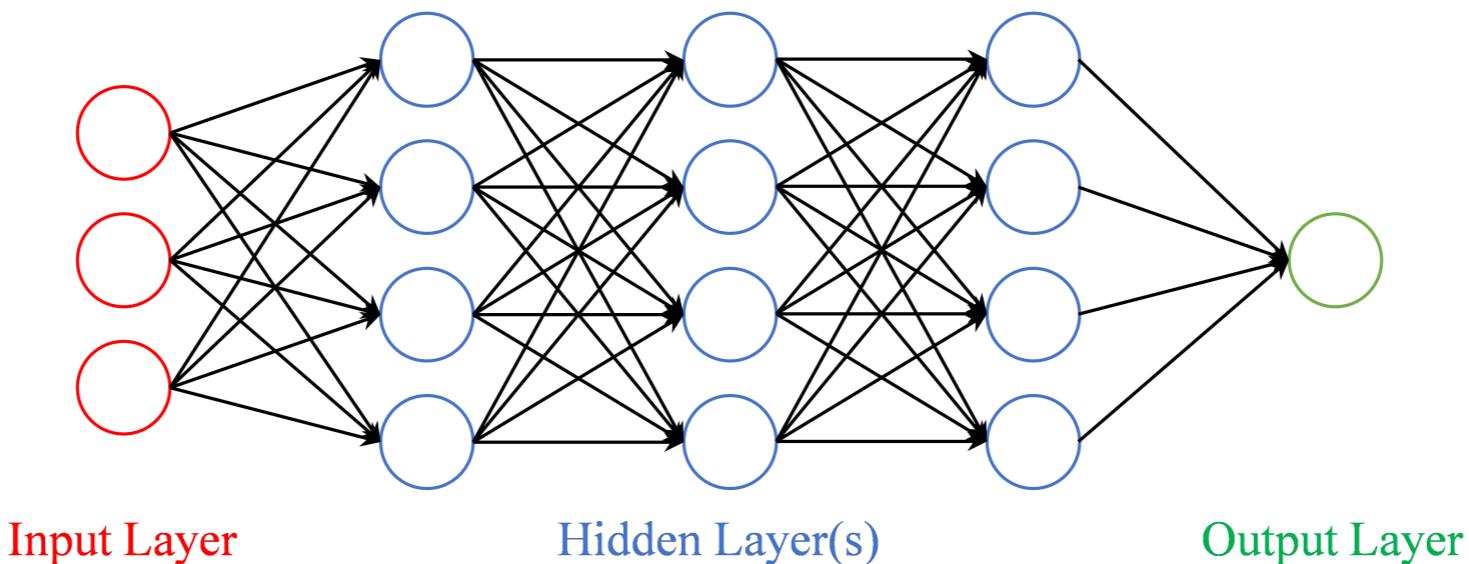


Building a Neural Network-based ID Feed-Forward

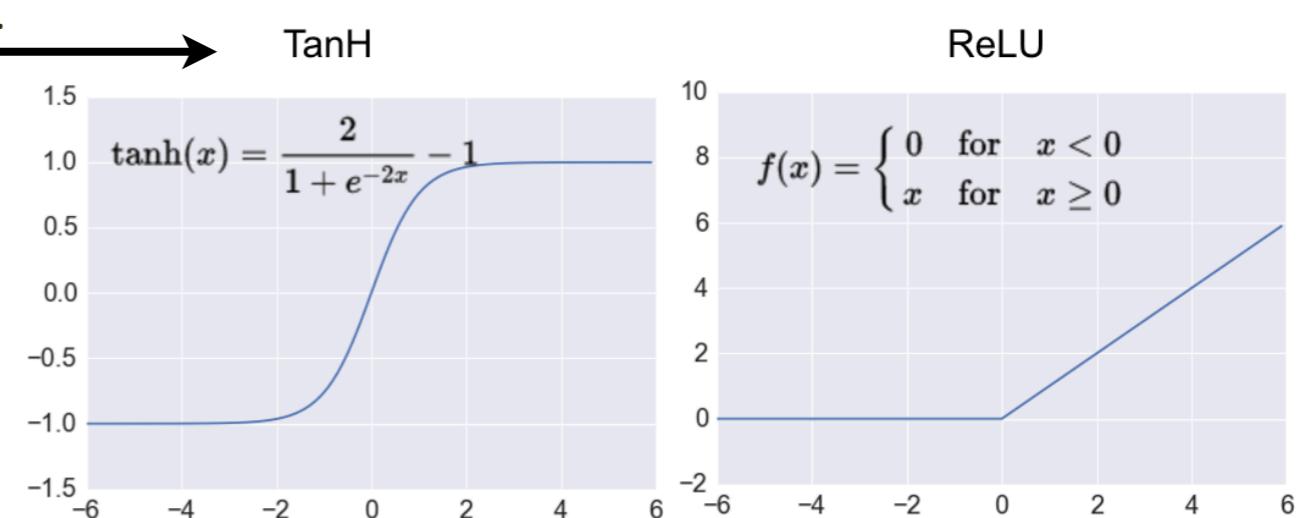
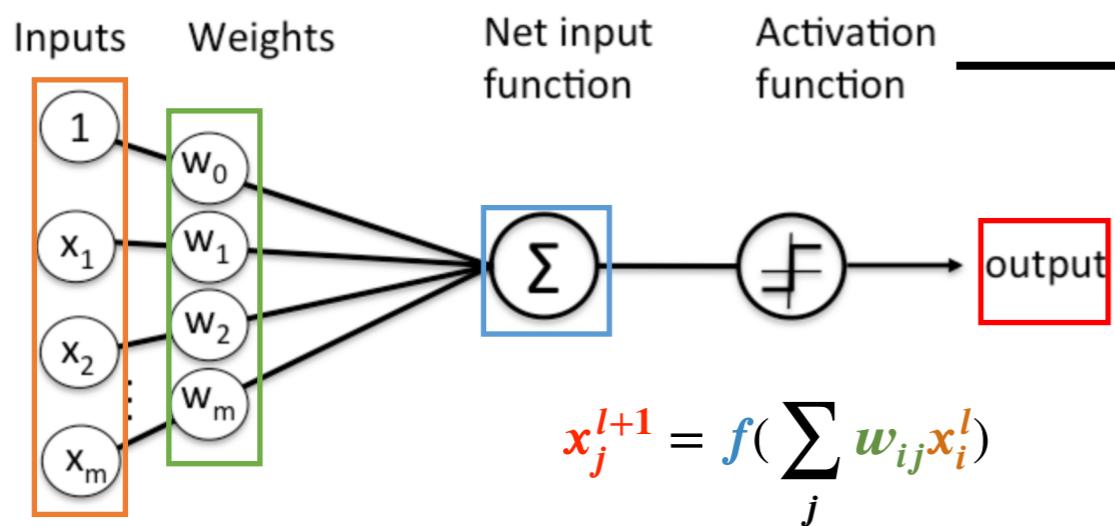
- Training: measure beam sizes while scanning DWP & various ID configurations → acquire data at 10 Hz → ***train*** neural network (NN)
- Resulting NN can give ***prediction*** for DWP required to keep beam size constant for arbitrary ID configurations → run as NN-based ID FF



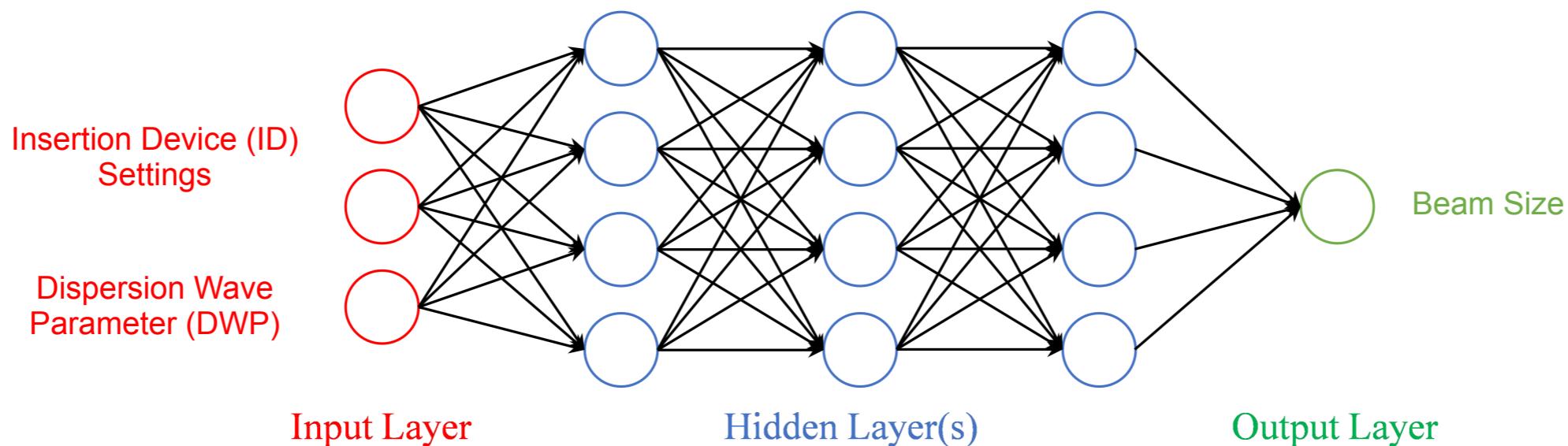
How a Neural Network (NN) Works



Courtesy: S. Liu



Deep Learning: How we Trained the NN



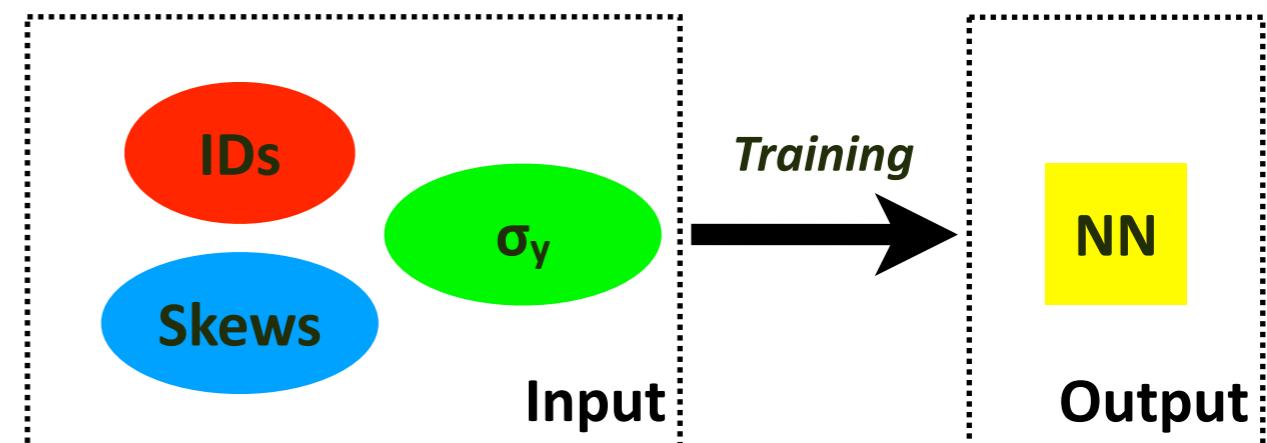
Input Layer: ID settings (22-35 Dimension) and DWP (1 Dimension)

Three Hidden Fully Connected Layers:
128, 64, 32 neurons in each layer

Output Layer: Vertical Beam Size (1 Dimension)

Regularization: L₂ regularizer with $\lambda = 10^{-4}$

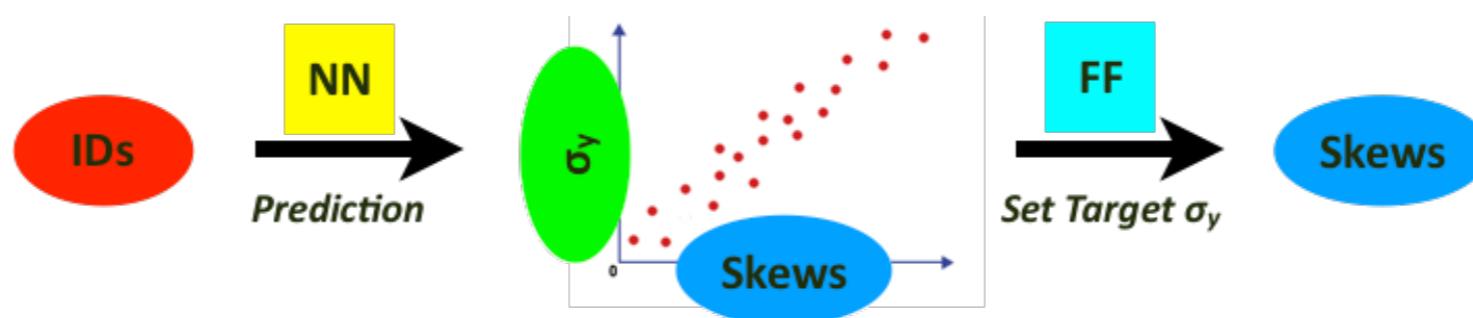
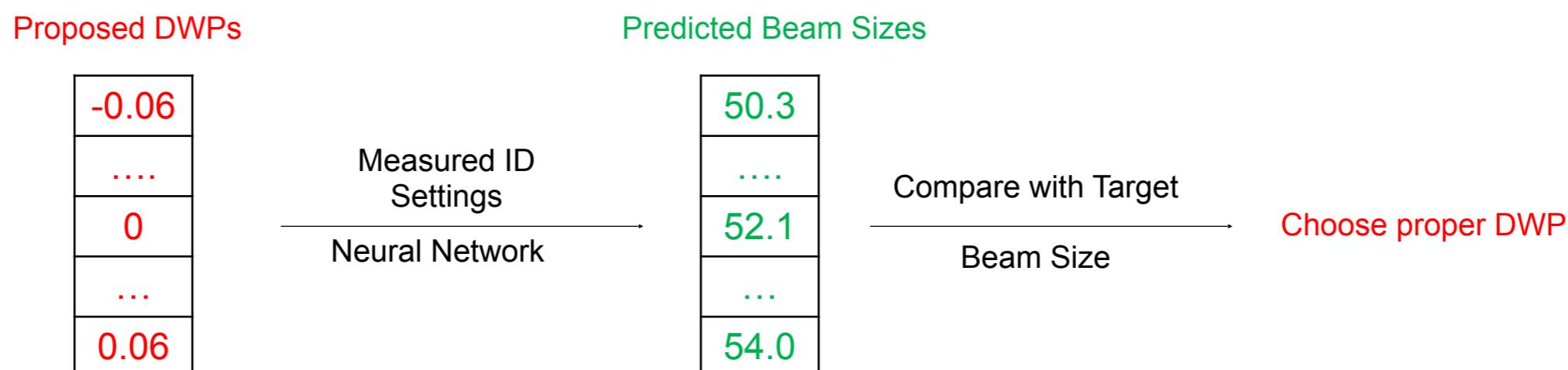
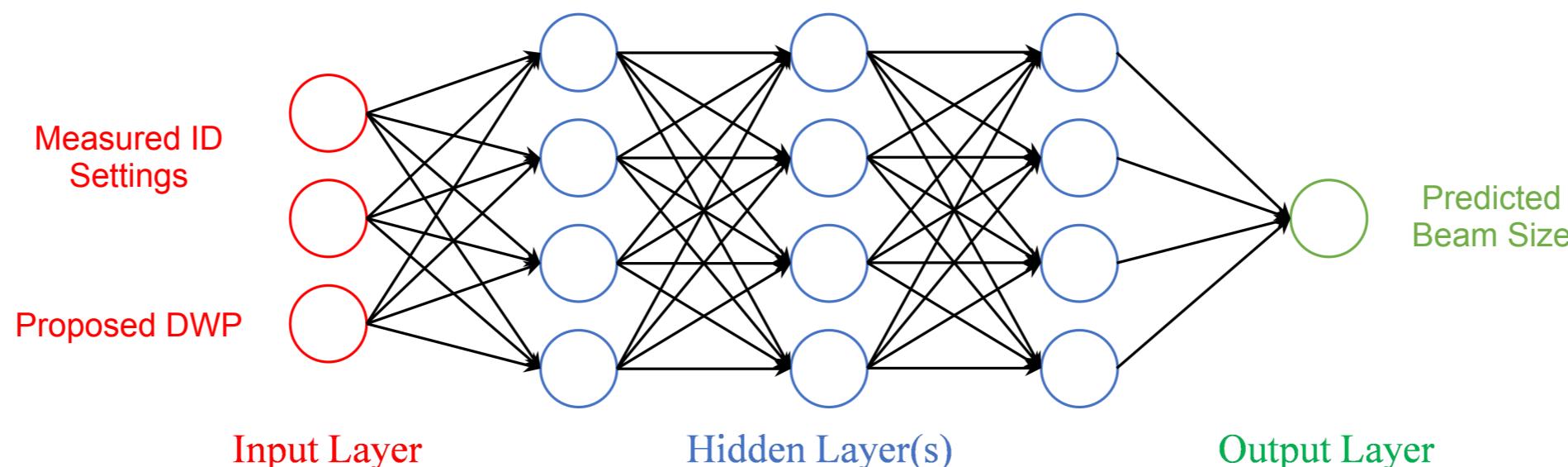
Optimization: Adam Optimizer with learning rate $\alpha = 10^{-3}$



Architecture	Raw Data		With Square Features	
	Training MSE	Evaluation MSE	Training MSE	Evaluation MSE
128-64	0.0265	0.0268	0.0257	0.0260
256-64	0.0243	0.0245	0.0259	0.0262
512-128	0.0243	0.0247	0.0243	0.0247
128-64-32	0.0238	0.0242	0.0243	0.0245
256-128-64	0.0236	0.0240	0.0240	0.0246
256-128-64-32	0.0245	0.0249	0.0245	0.0248

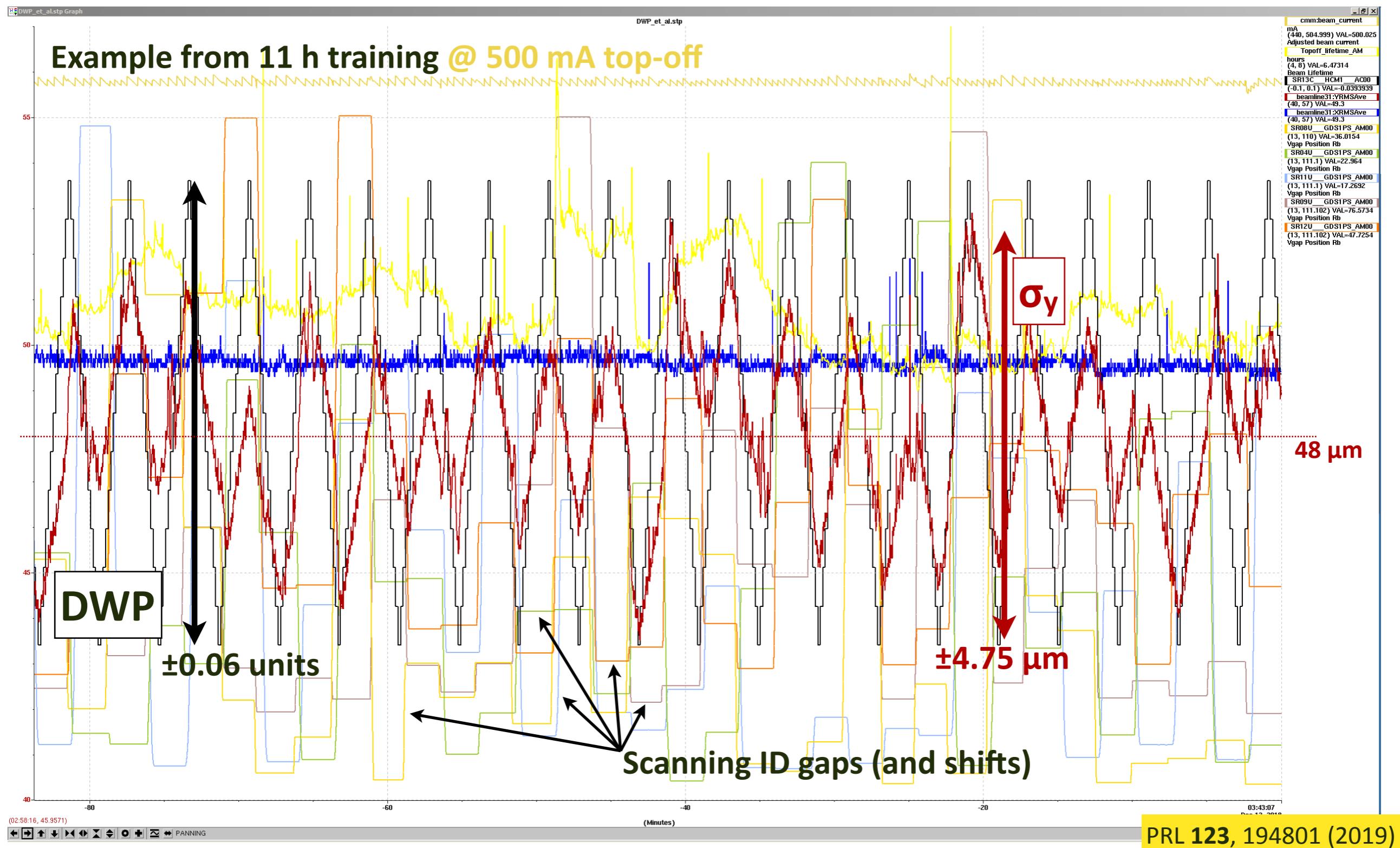
PRL 123, 194801 (2019)

Resulting NN Enables ID Feed-Forward at ≈ 3 Hz

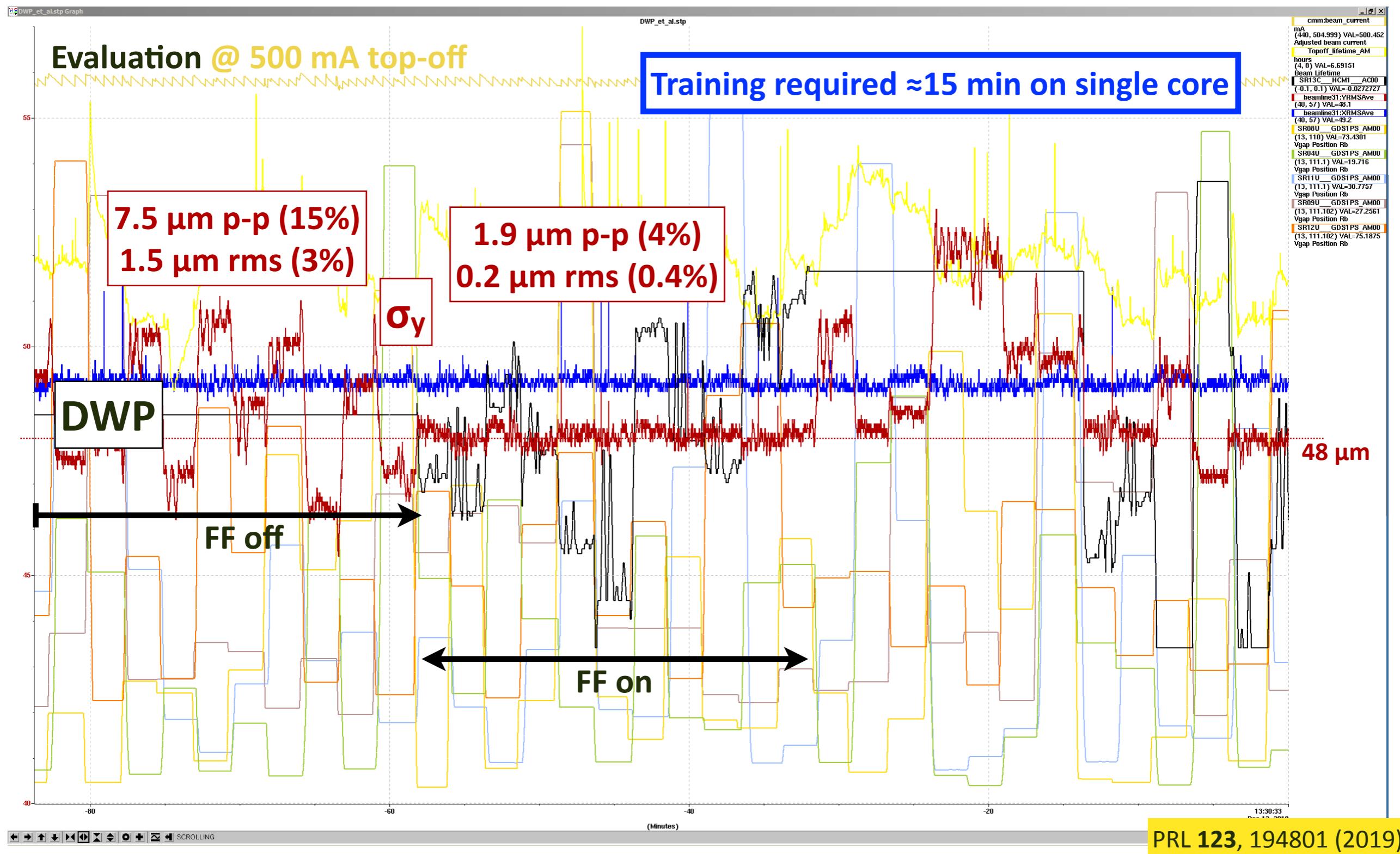


PRL 123, 194801 (2019)

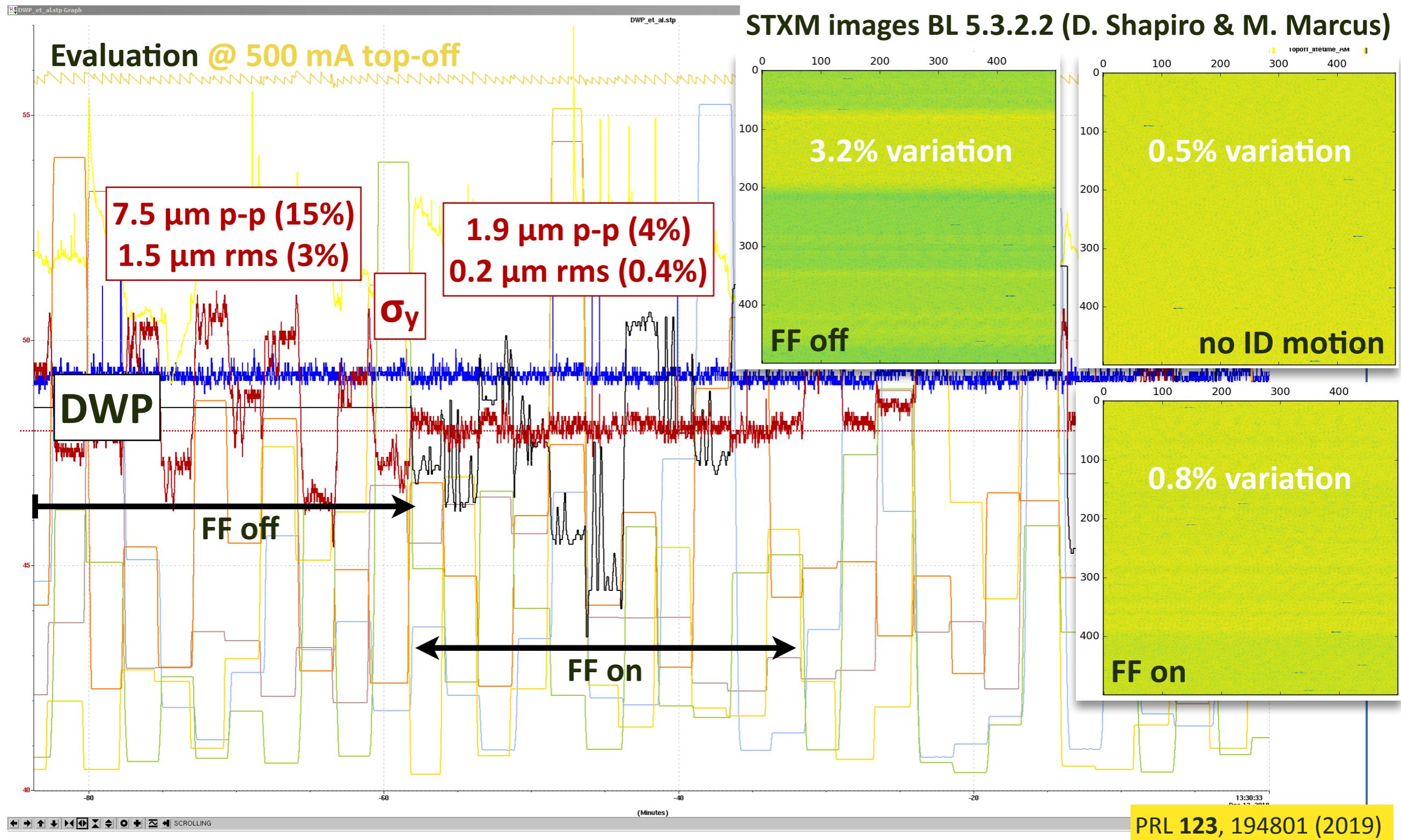
Physics Shift: Data Collection for NN Training



Physics Shift: Running NN-based ID Feed-Forward



Physics Shift: Running NN-based ID Feed-Forward

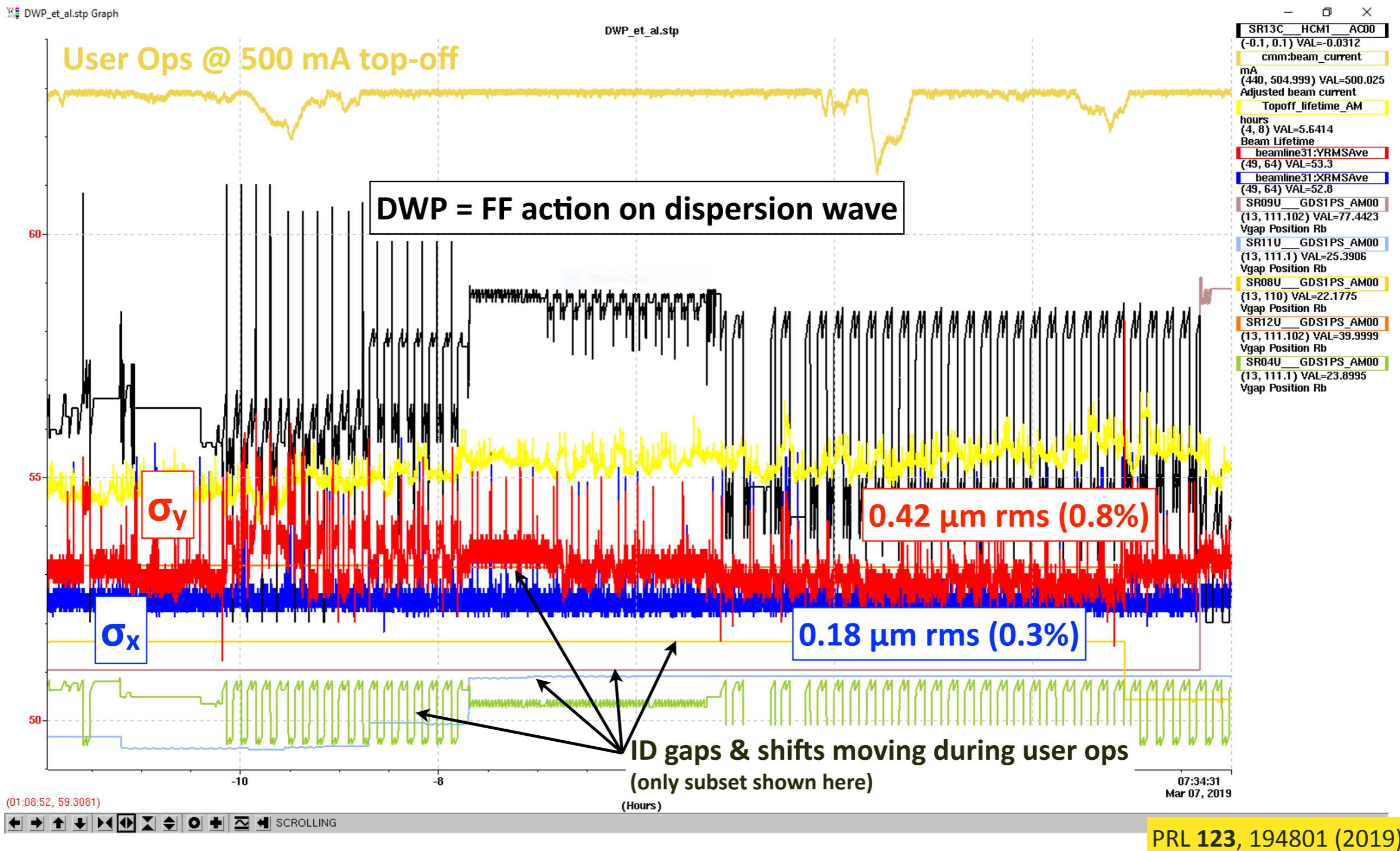


Towards First Experiments During User Ops

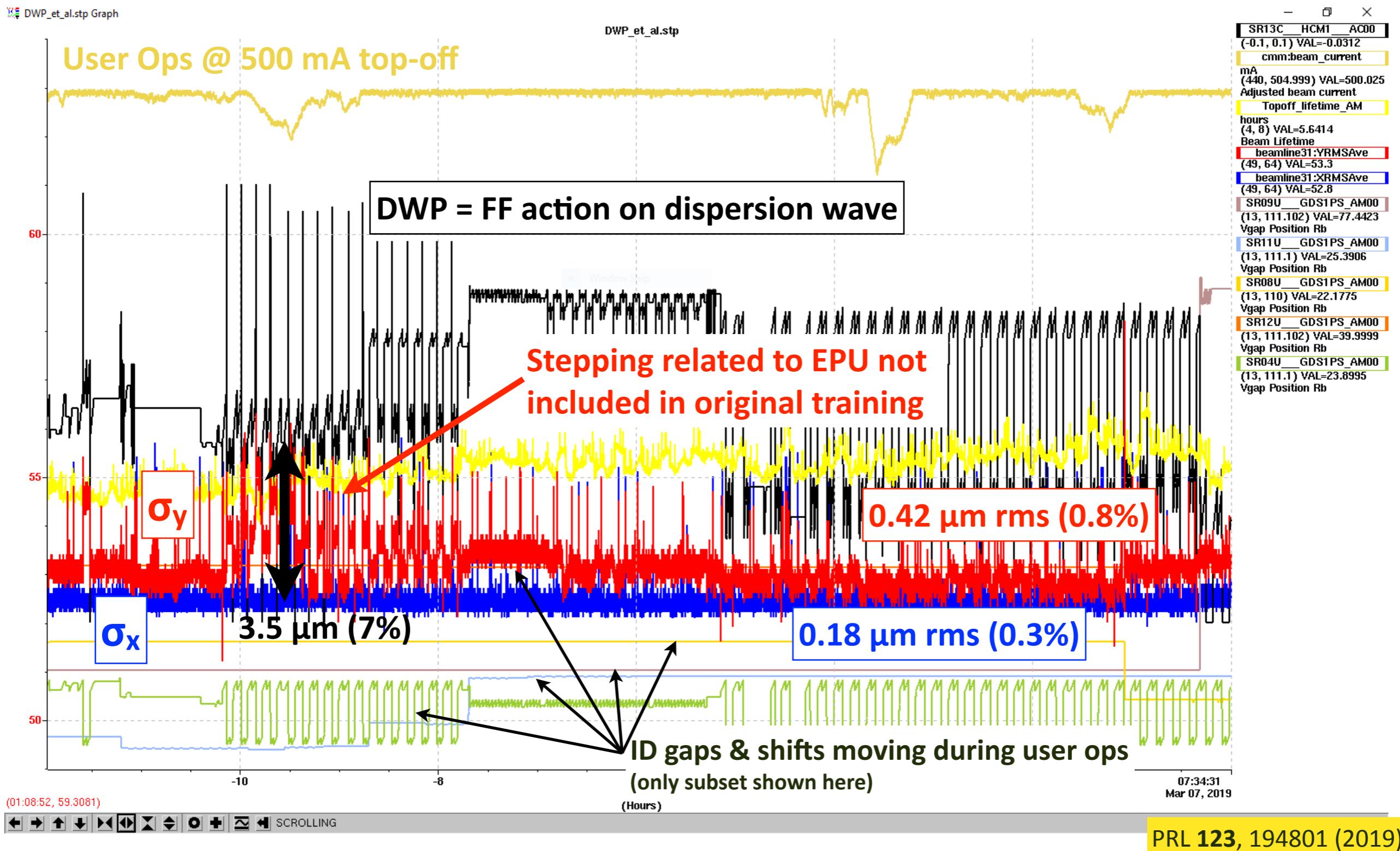
- Use machine shift to acquire training data by scanning operational IDs in a quasi-randomized fashion (favoring operational gap range) → train NN
- Put this NN into FF operation during user ops and evaluate



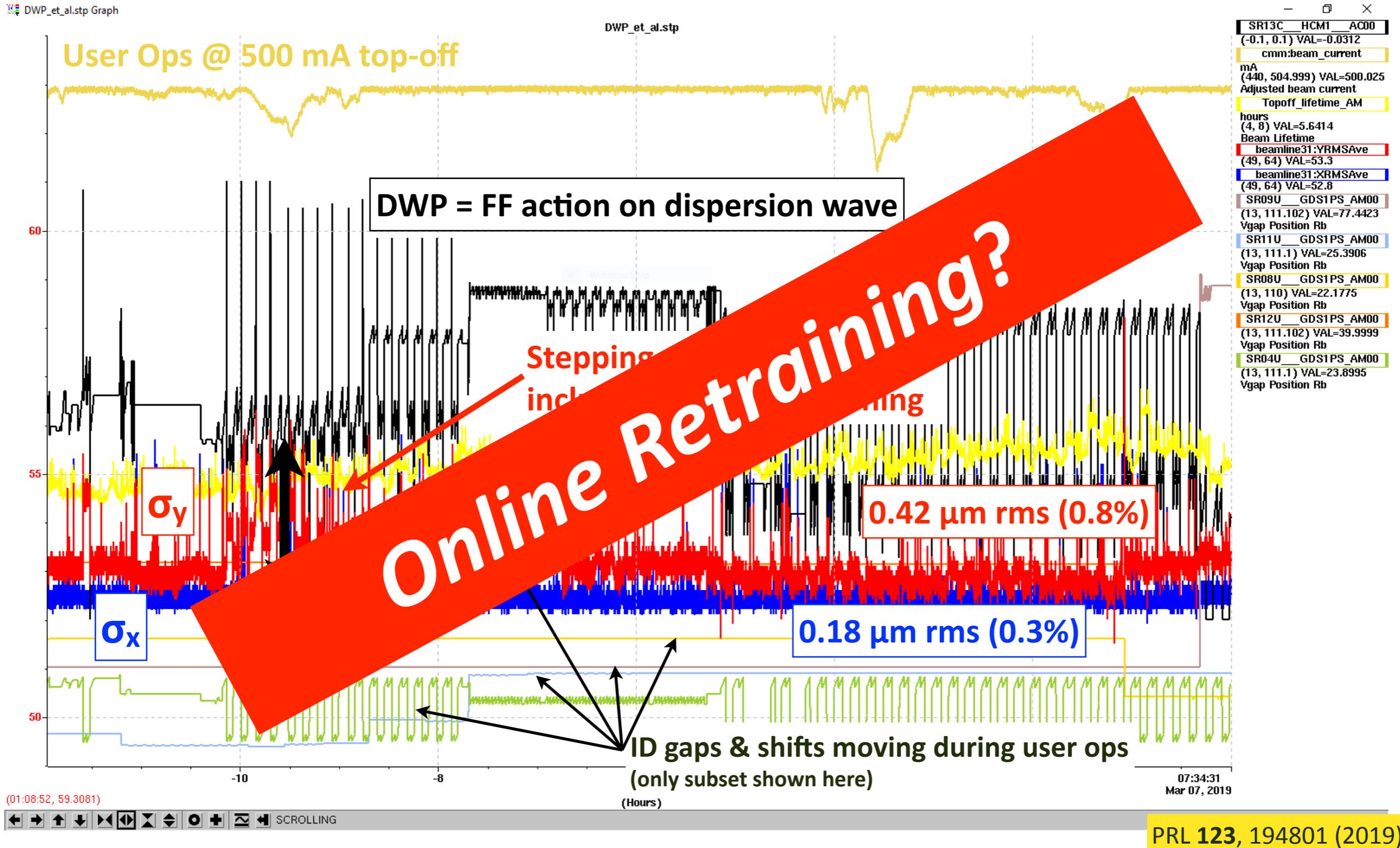
Stabilization Confirmed During First User Ops Trial



Stabilization Confirmed During First User Ops Trial

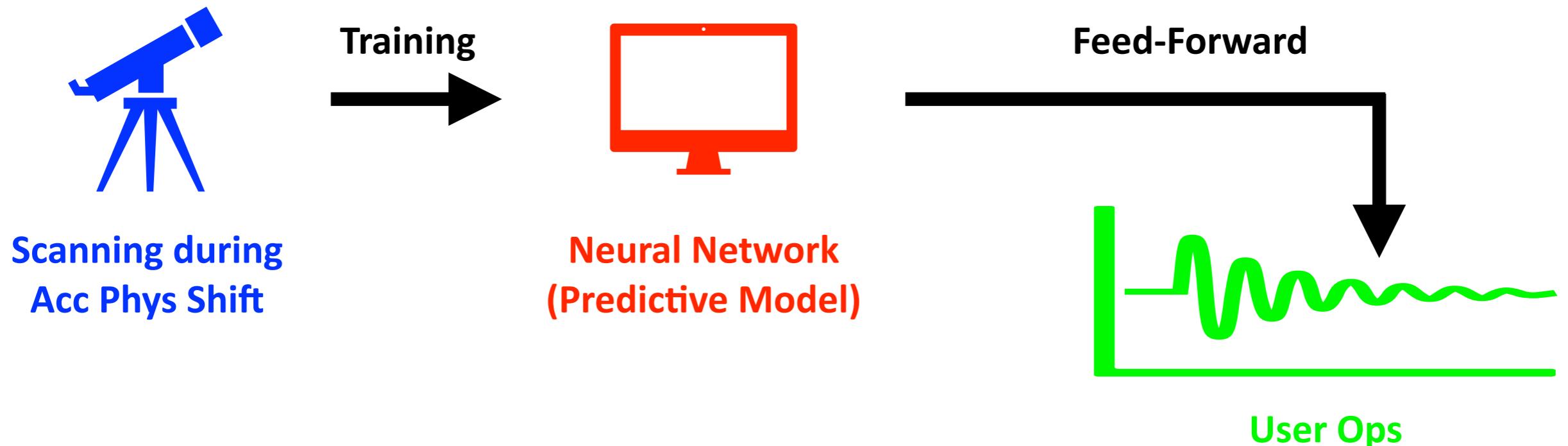


Stabilization Confirmed During First User Ops Trial



Online Retraining: Improve NN with User Ops Data

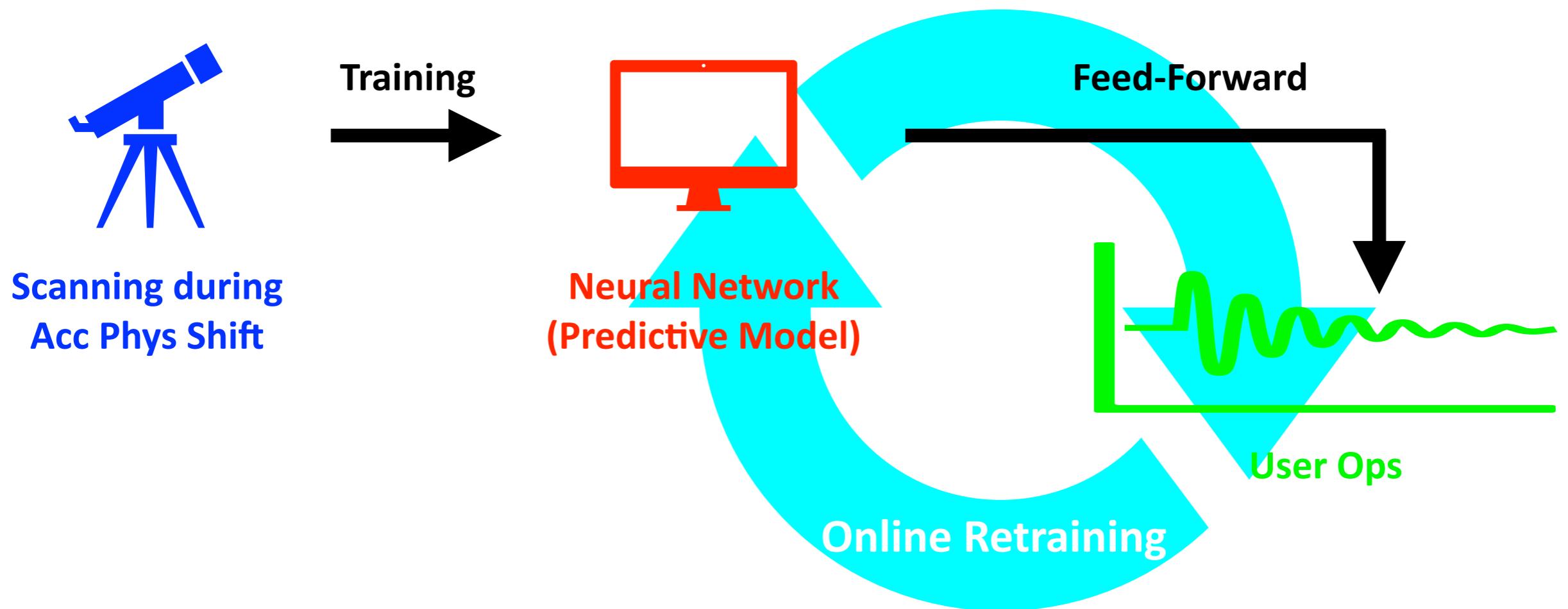
So far: "Conventional" Machine Learning



PRL 123, 194801 (2019)

Online Retraining: Improve NN with User Ops Data

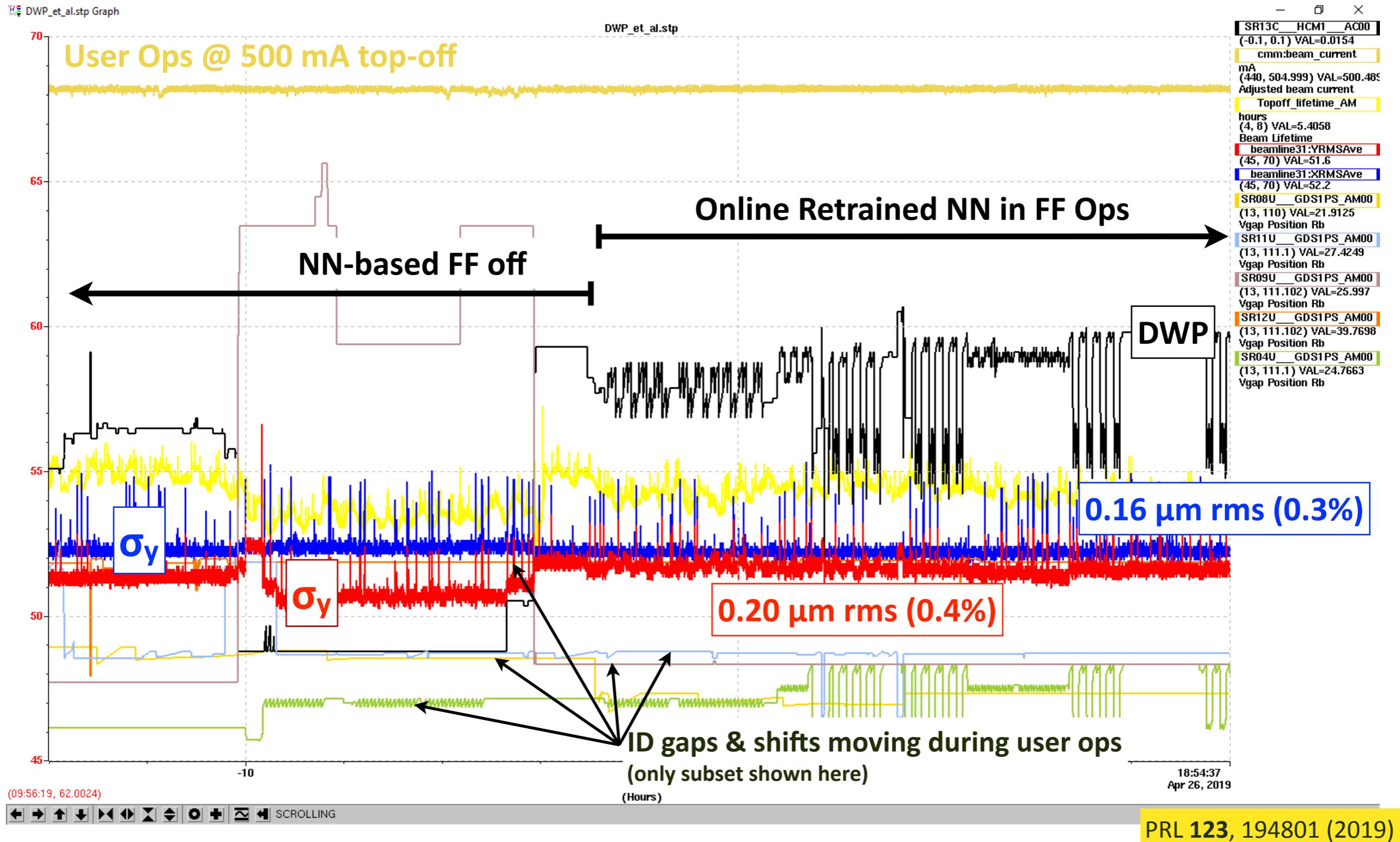
Online Retraining: apply user ops data to improve NN → swap NN used for ID FF on the fly



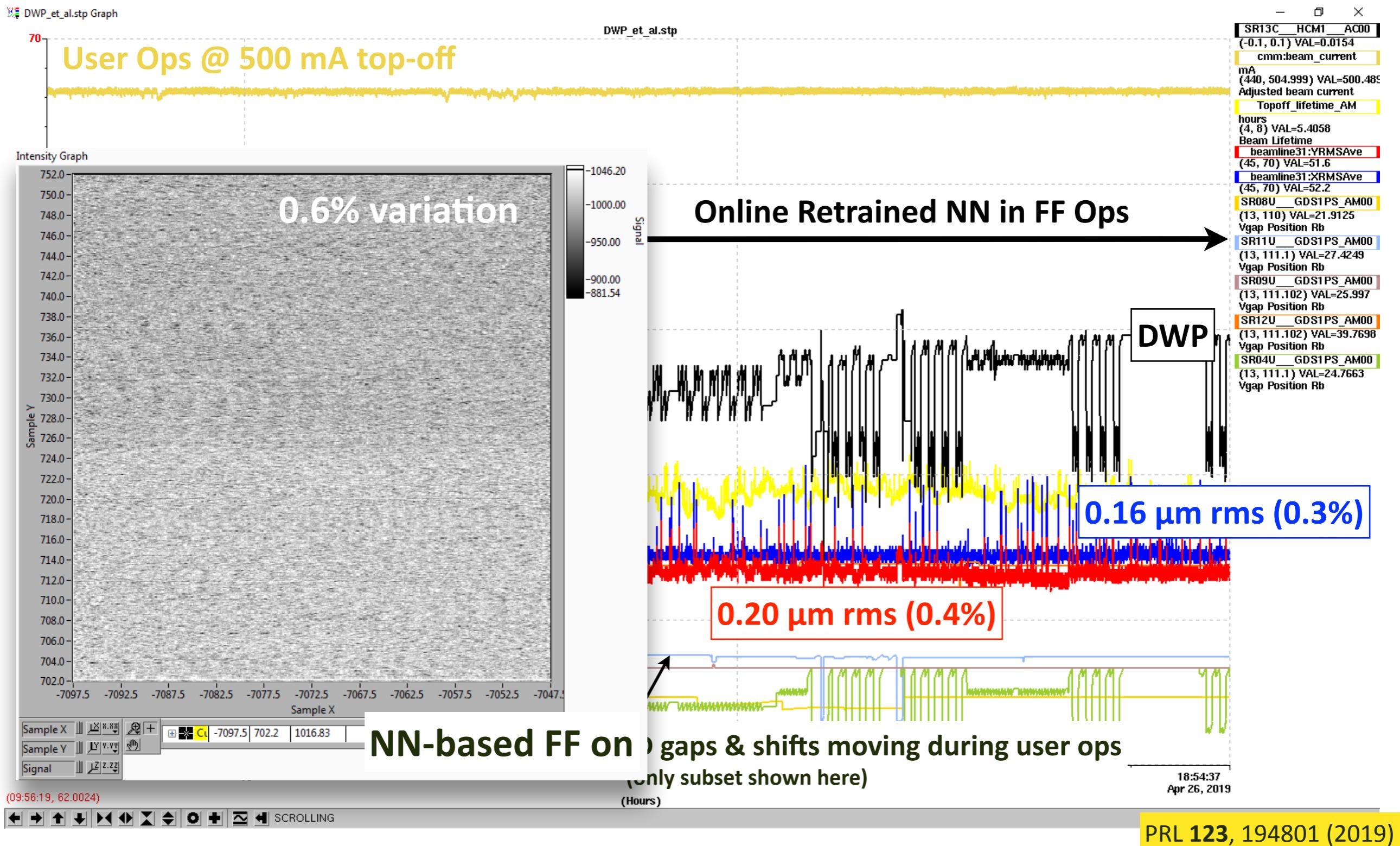
NN can be continuously online retrained during user ops to improve FF performance
(exploiting huge amounts of data acquired during user ops)

PRL 123, 194801 (2019)

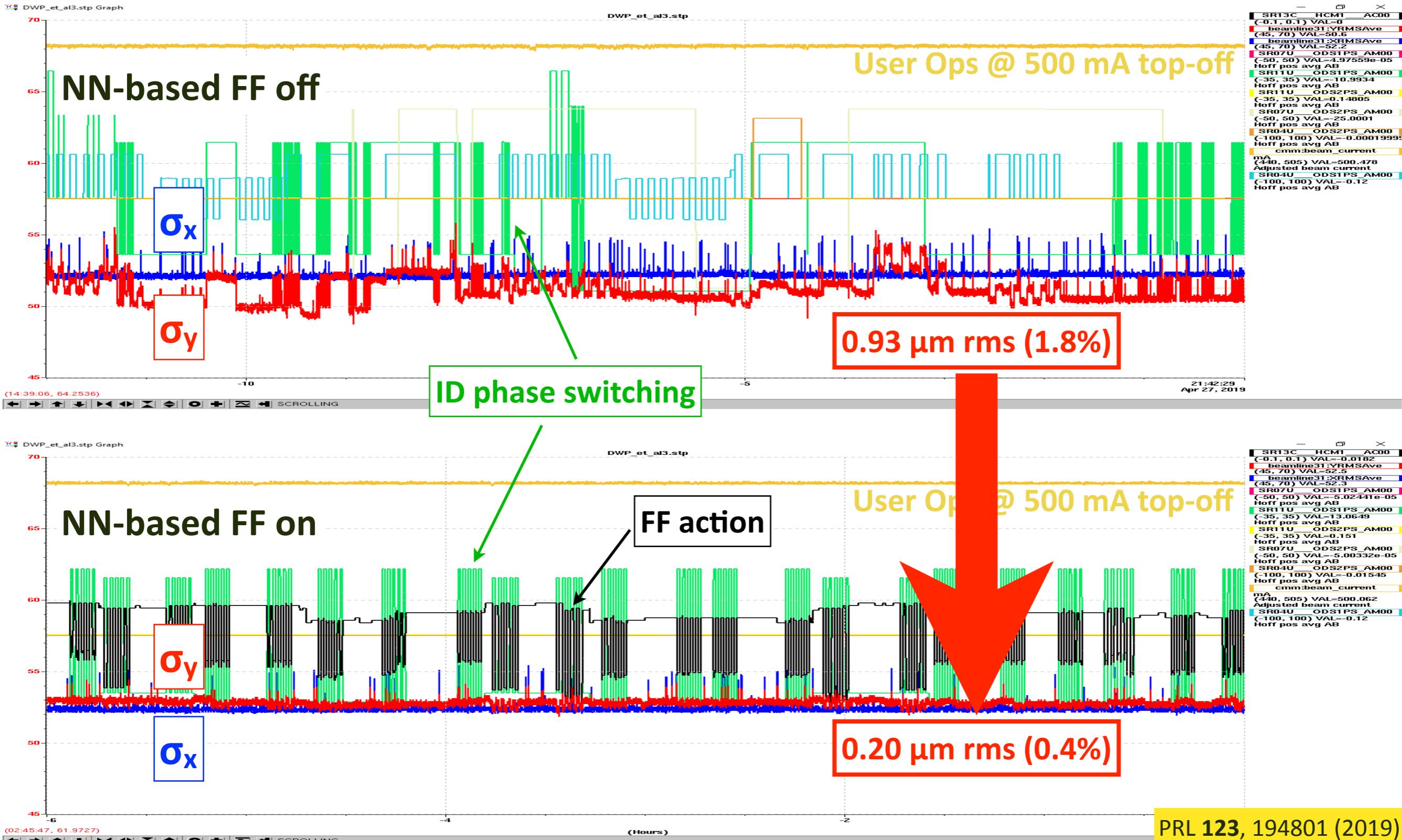
Substantial Improvement After Online Retraining



Substantial Improvement After Online Retraining

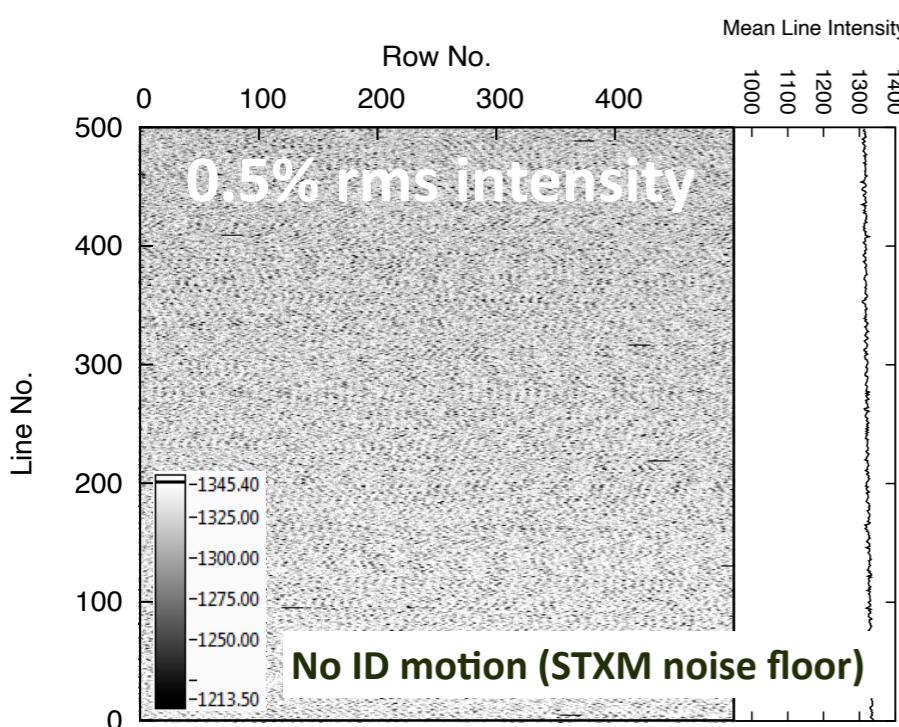


Summary: NN-based FF Off vs. On During User Ops

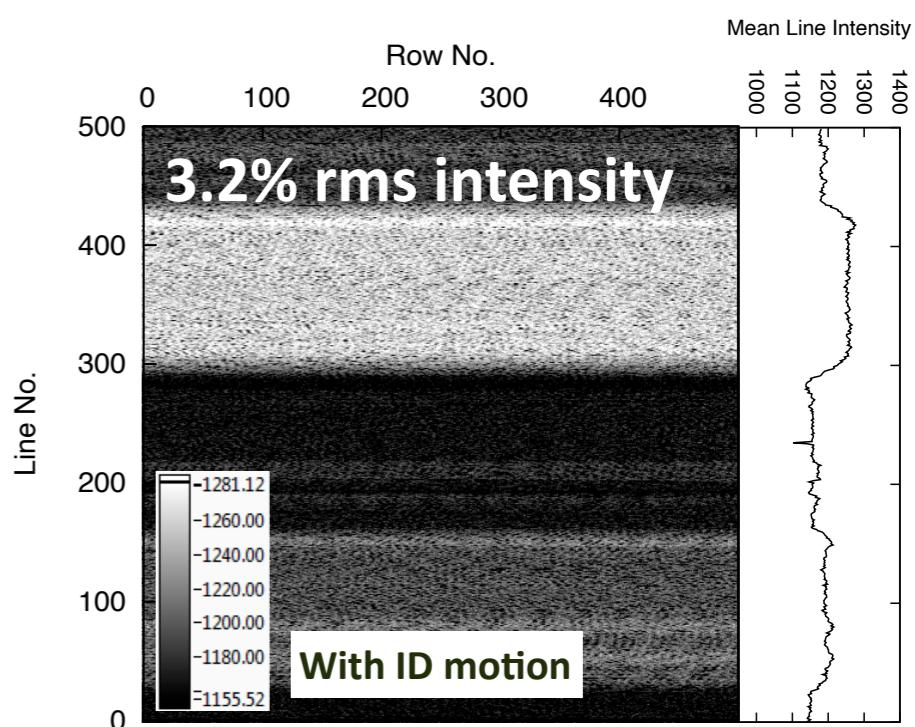


Stabilization Confirmed at Experiment

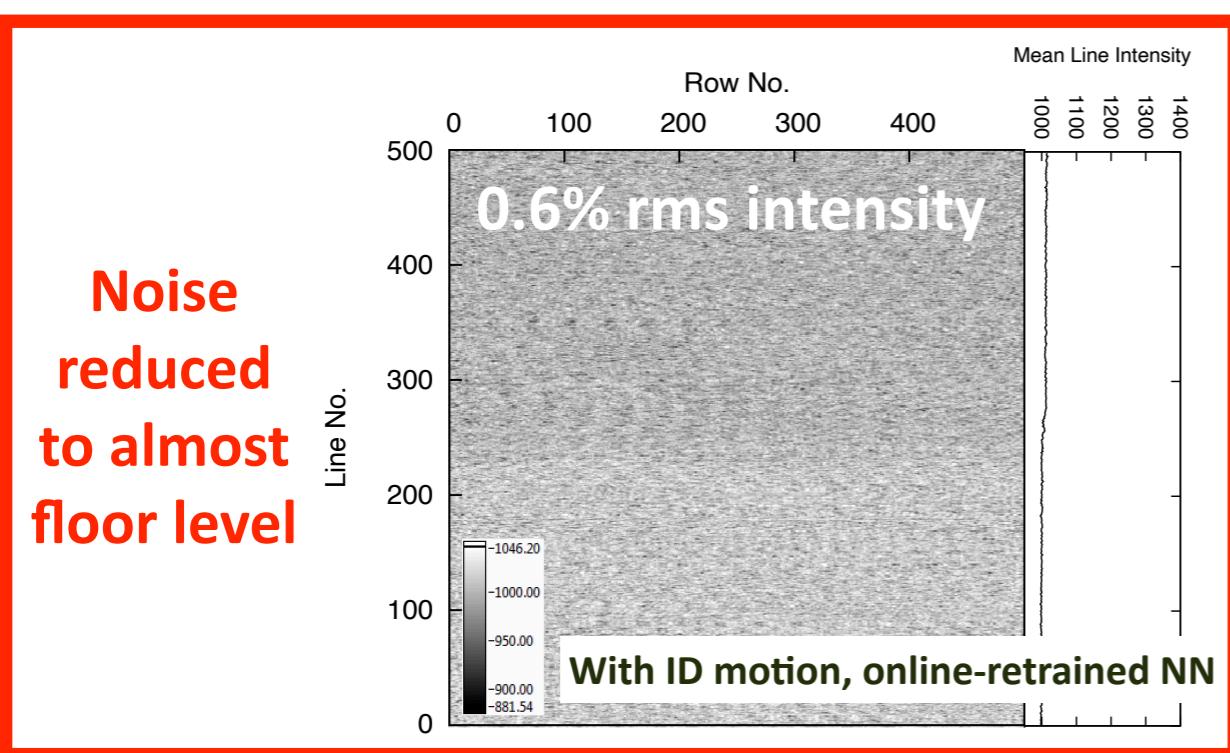
ALS Beamline 5.3.2.2



ID Motion

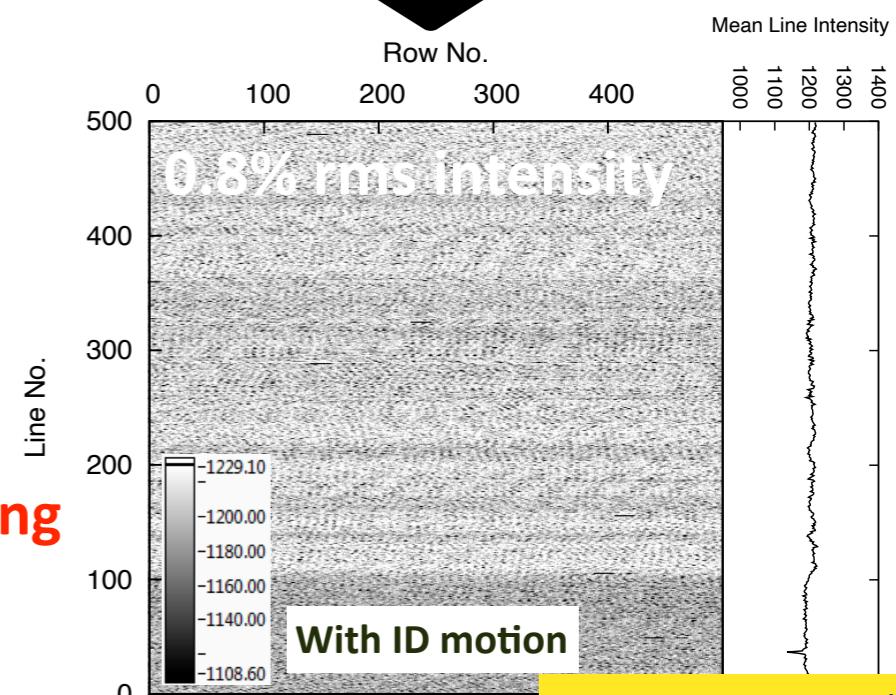


NN-based FF on



Noise reduced to almost floor level

Online Retraining



PRL 123, 194801 (2019)

Thank You!

Questions?

Acknowledgments: Shuai Liu, Hiroshi Nishimura, Matthew A. Marcus, David Shapiro, Changchun Sun, Nathan Melton, Alex Hexemer, Dani Ushizima, Mike Ehrlichman, Gregg Penn, Thorsten Hellert, Yuping Lu, Erik Wallen, Warren Byrne, Fernando Sannibale, Marco Venturini, Andreas Scholl



BERKELEY LAB



U.S. DEPARTMENT OF
ENERGY

Office of
Science

ALS
ADVANCED LIGHT SOURCE