

# Machine Learning at LCLS

*Oct 24, 2016*

D. Ratner, A. Ahmed, T. Cope, J. Duris, S. Ermon, M. Gibbs, T. J. Lane,  
S. Li, T. Maxwell, M. McIntire, M. Mongia, N. Norvell, D. Sanzone, D.  
Schneider, C. Yoon

SLAC National Accelerator Laboratory

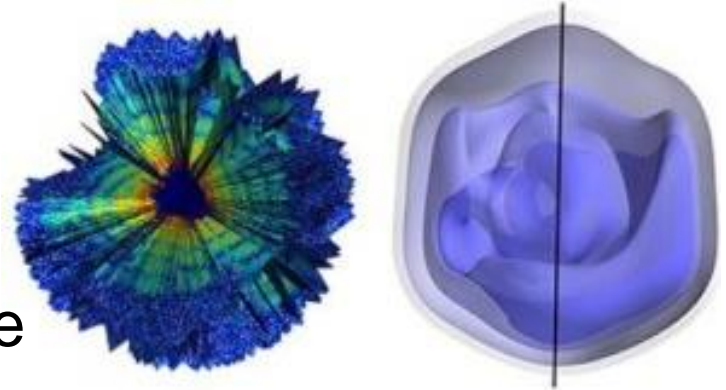
## Big Data comes to Photon Science

### User side:

LCLS: 120 Hz images  $\rightarrow$  15 TB/hour

LCLS-II: 100 kHz  $\rightarrow$  1 PB/hour!

$\rightarrow$  exascale computing initiative

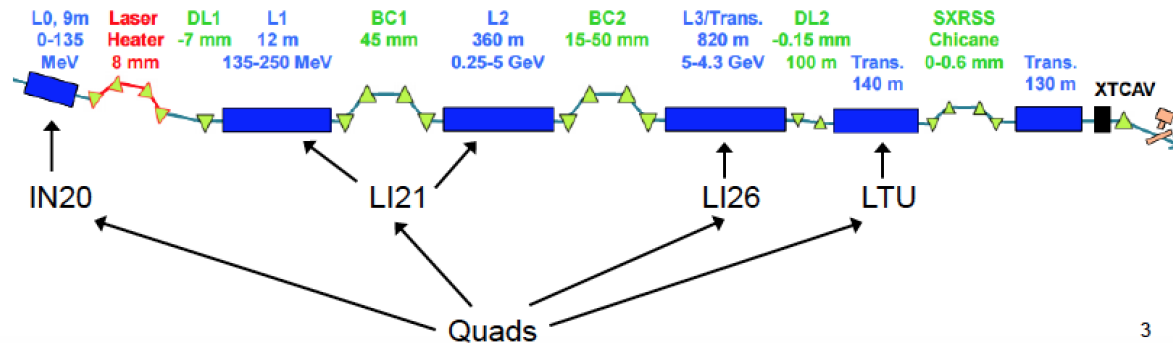


### Machine side:

Archive 200k variables at 1Hz  $\rightarrow$   $10^{12}$  data points so far

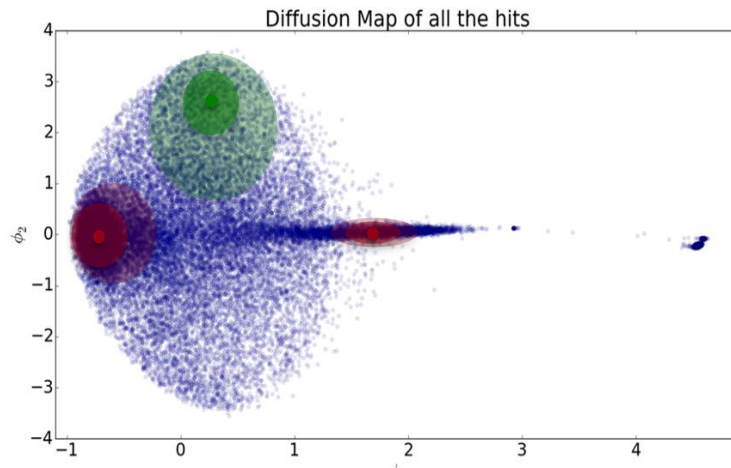
Online optimization of  $\sim 30$  dimensional space

Alarm/anomaly/breakout handling

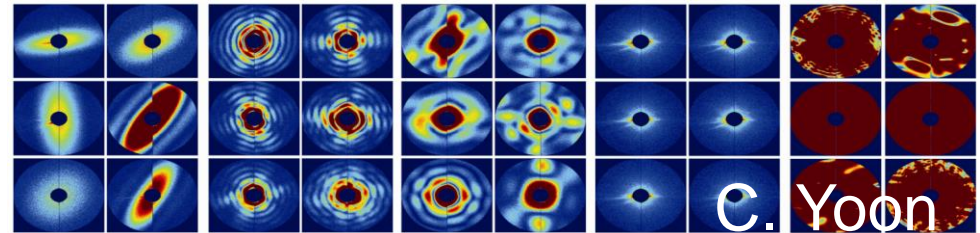


# Computer vision: biological imaging (C. Yoon)

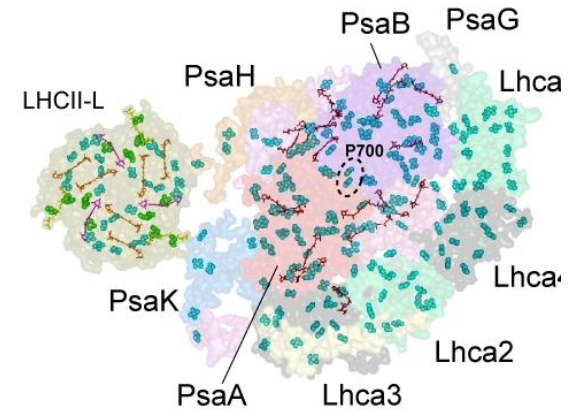
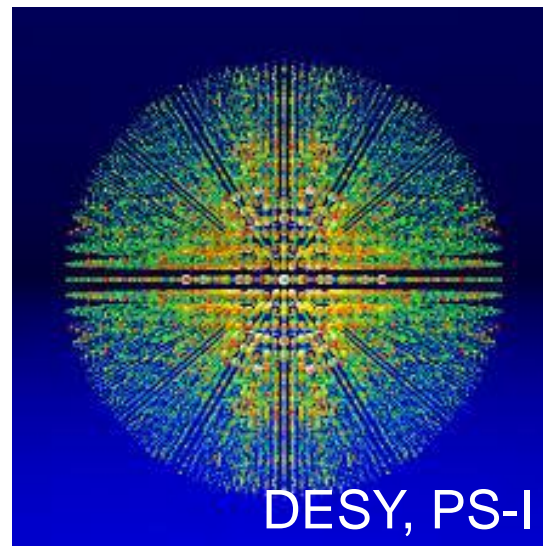
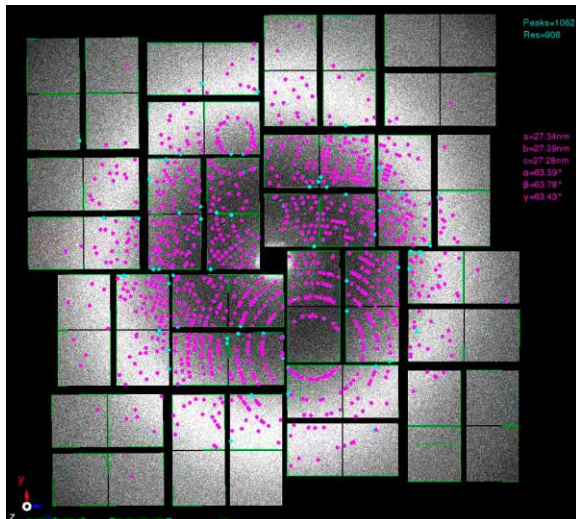
## Classification of single particle images



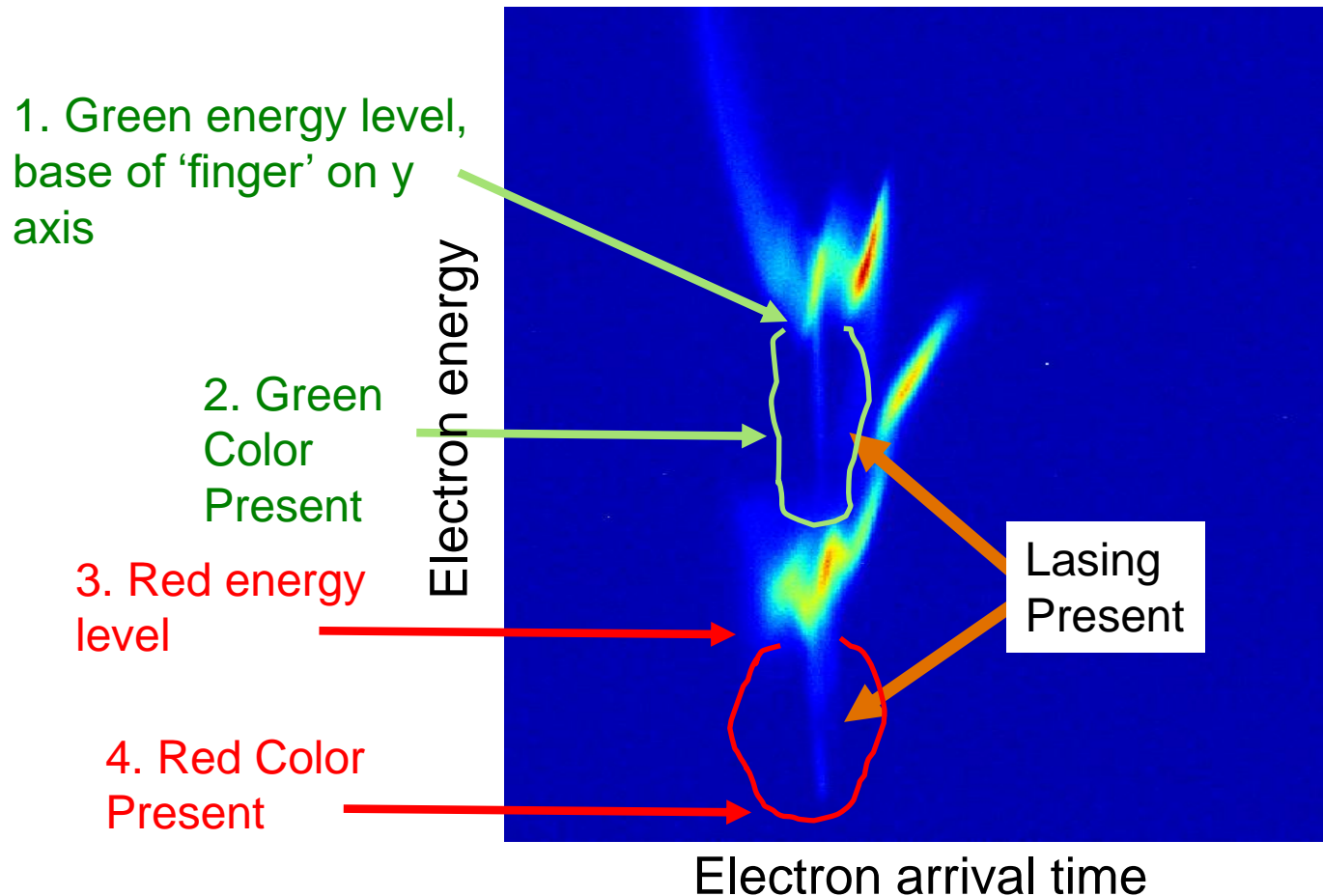
C. Yoon, A. AbuHashem



## Indexing and classification of nano-xtal images (Google Accelerated Science)



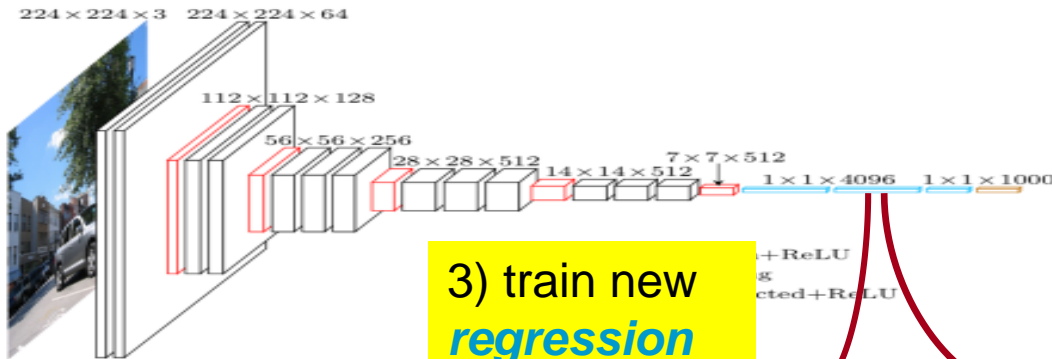
## XTCAV electron diagnostic: best source of X-ray temporal info!



# Computer vision: X-ray/electron beams (D. Schneider)

1) start with fully trained *ImageNet* based convnet

Millions of images

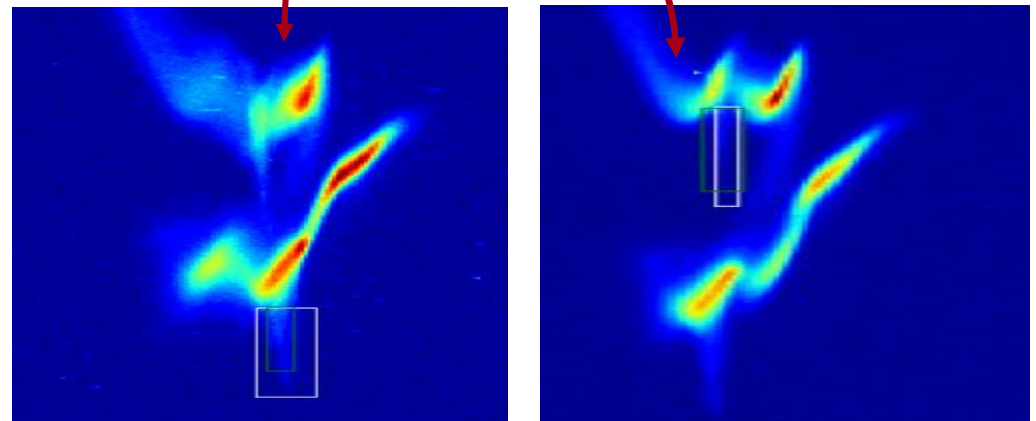
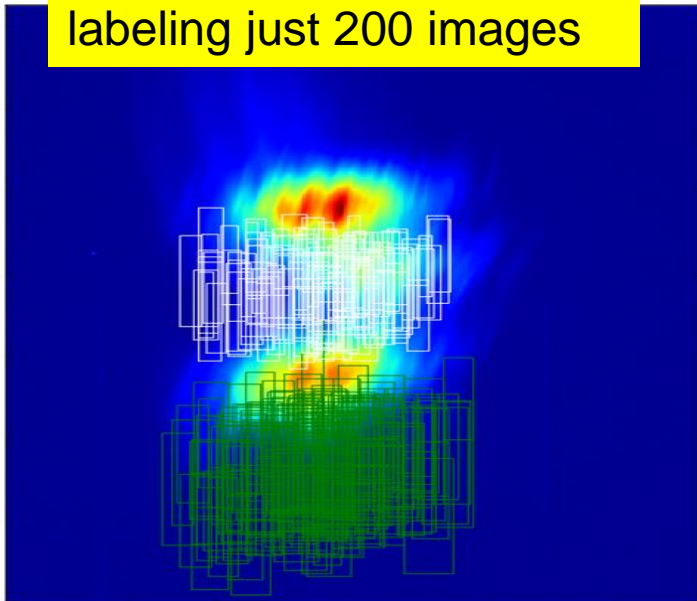


Classifies a 1000 categories - cat, dog, bus, etc

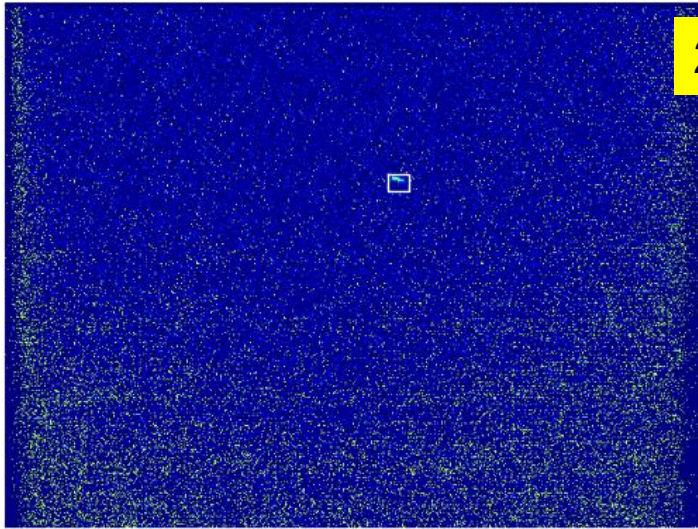
2) spend *only 3 hours* labeling just 200 images

3) train new *regression heads* on 200 images

4) Good results

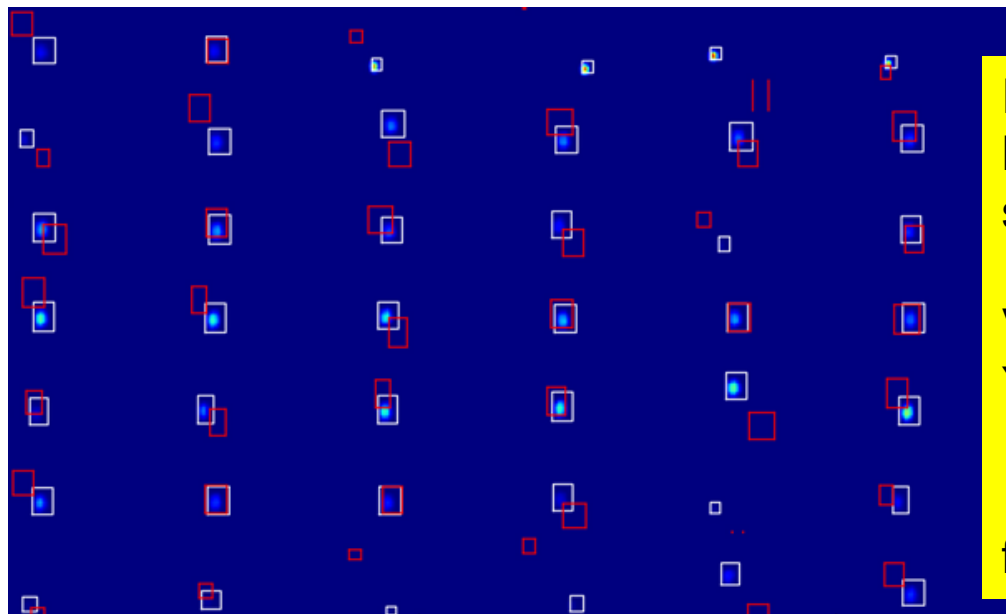
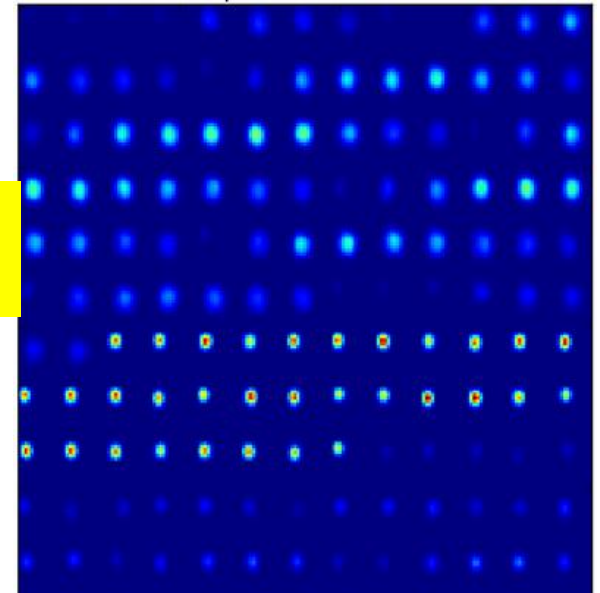


# Computer vision: X-ray/electron beams (D. Schneider)



250 labeled boxes

For VCC – high variation in beam



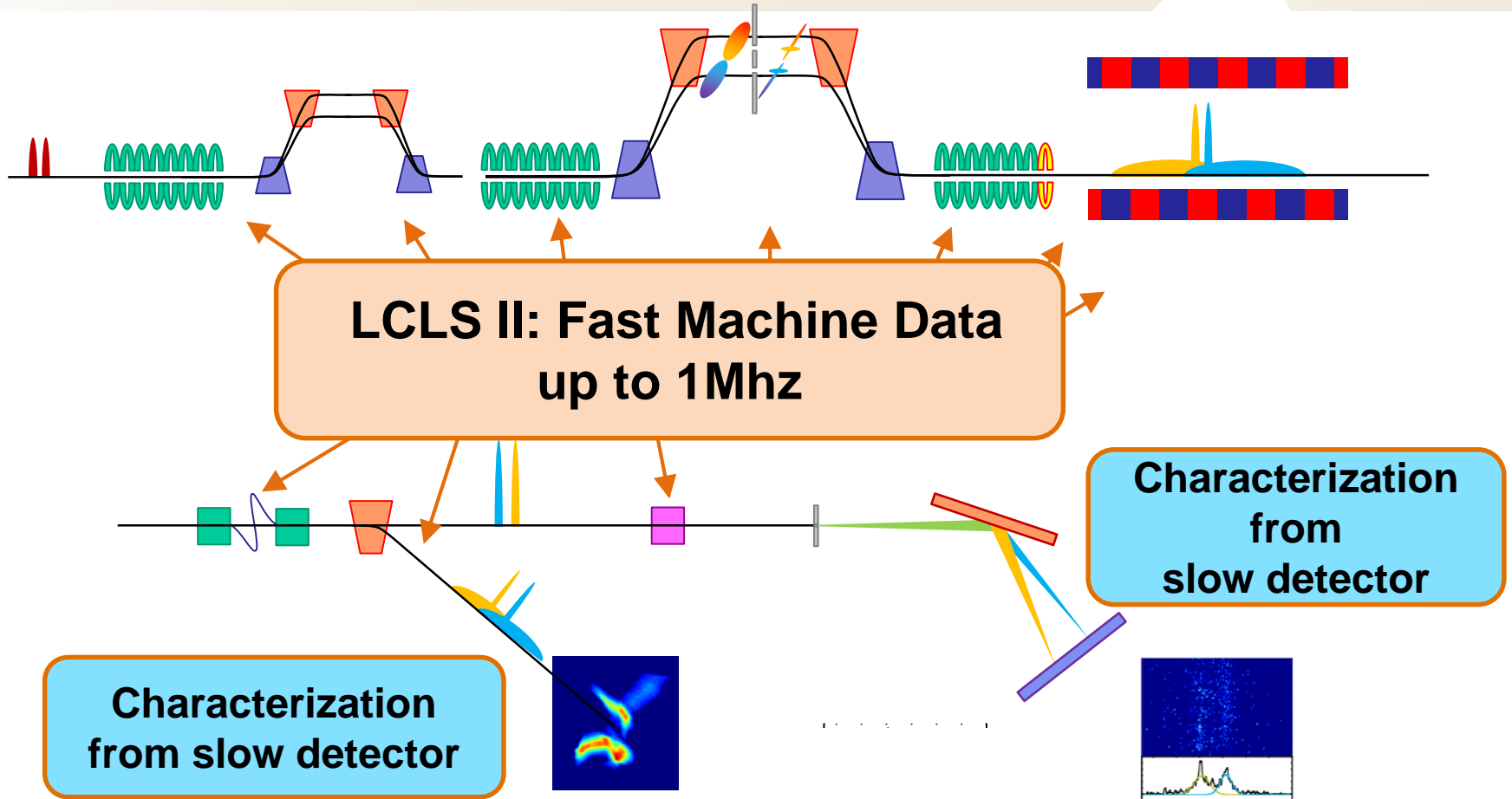
Regression Results:  
Measurement of accuracy: for how many shots is area of intersection/union  $> 0.5$ ?

VCC Screen – only **0.09** (left)

YAG Screen – **0.89** (not shown)

**Critical step:** de-noising preprocessing  
from Abdullah Rashed Ahmed

# How to reconcile MHz beam and 120 Hz diagnostic?

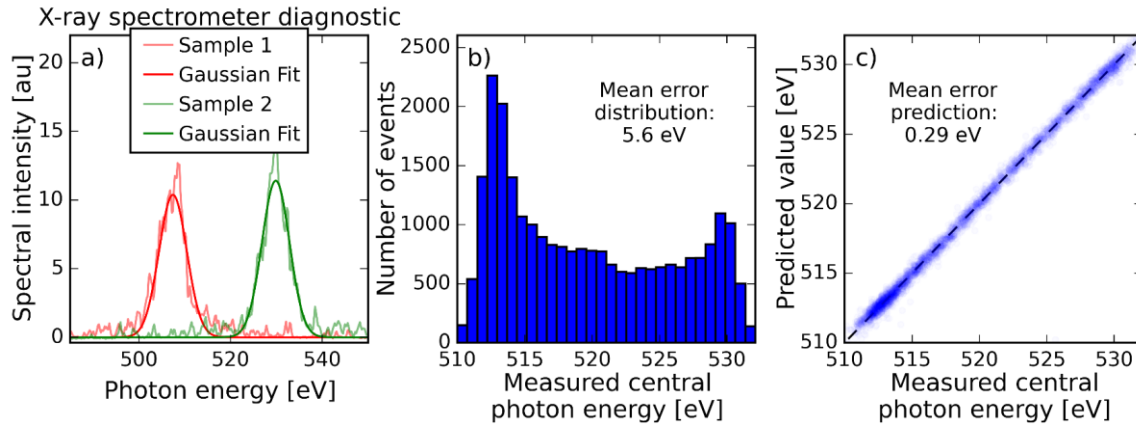


Characterization from slow detector

Characterization from slow detector

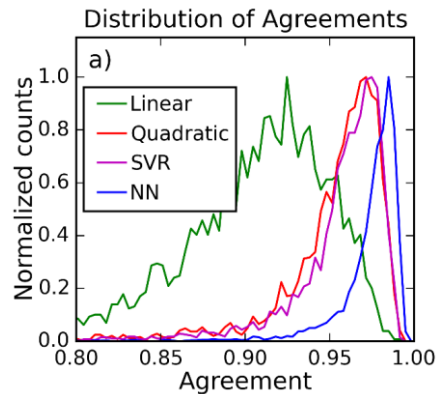
Predict/Interpolate slow characterization with Machine Learning

# How to reconcile MHz beam and 120 Hz diagnostic?

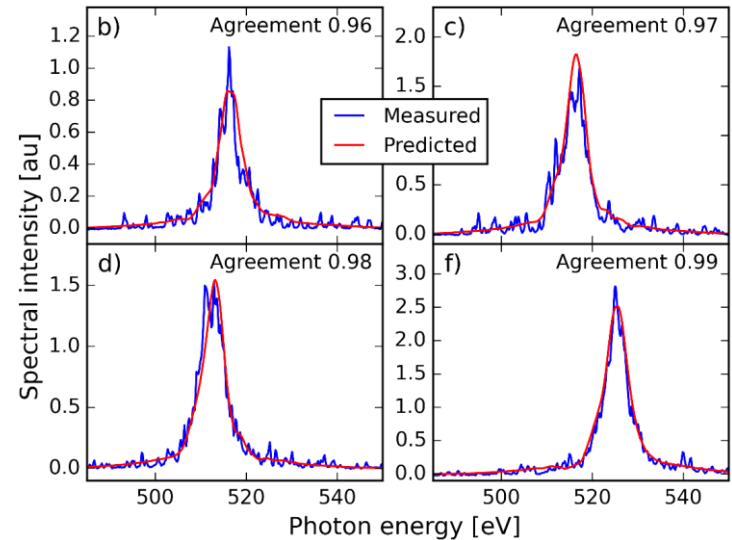


**Photon Energy**  
Prediction error  
smaller than 0.3 eV

**Spectral Shape**  
Agreements better  
than 0.97%

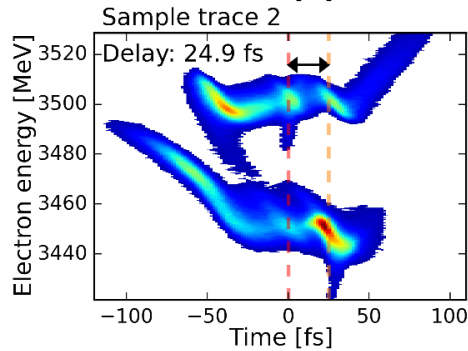
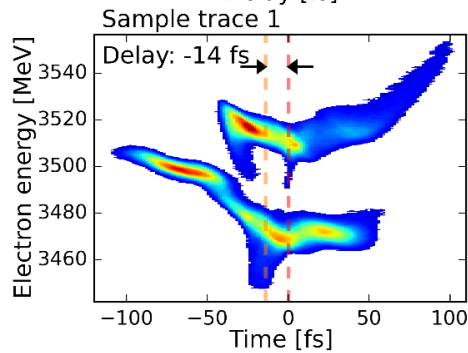
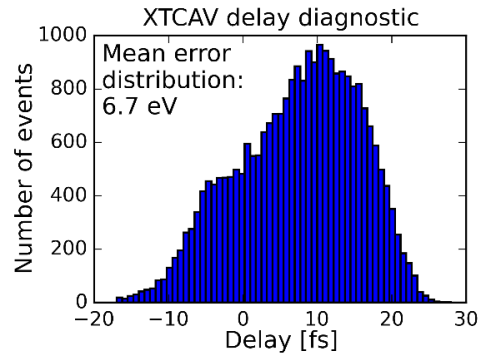


Examples of agreements



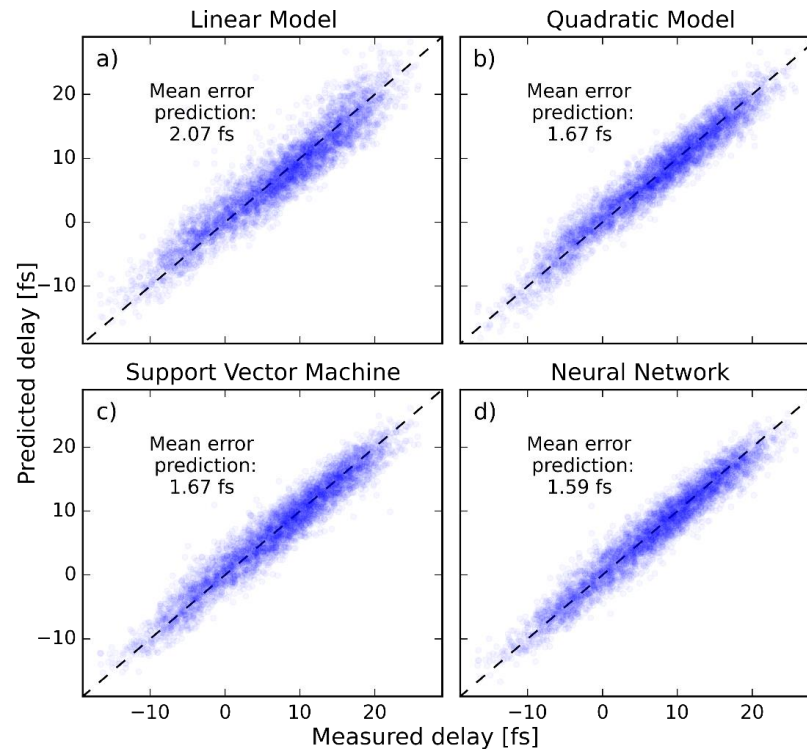


# How to reconcile MHz beam and 120 Hz diagnostic?



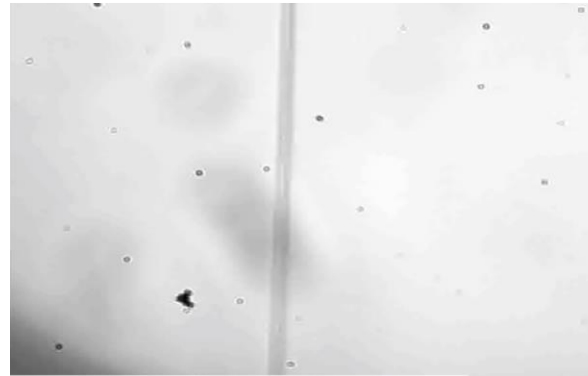
## 2-pulse delays

< 1.57 fs mean error

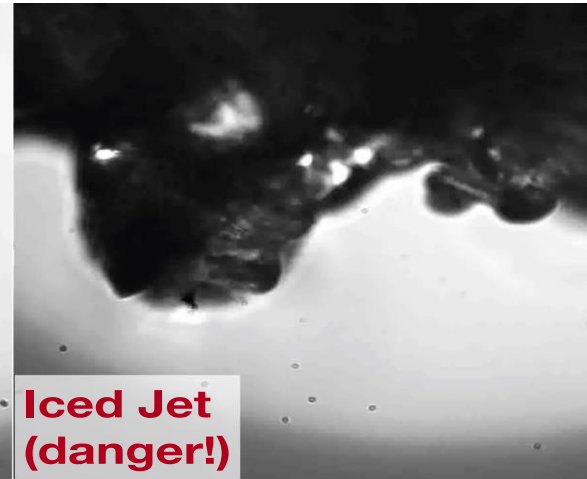


# Anomaly/Breakout detection (T.J. Lane)

**Machine protection:**  
e.g. detecting ice to  
protect the detector

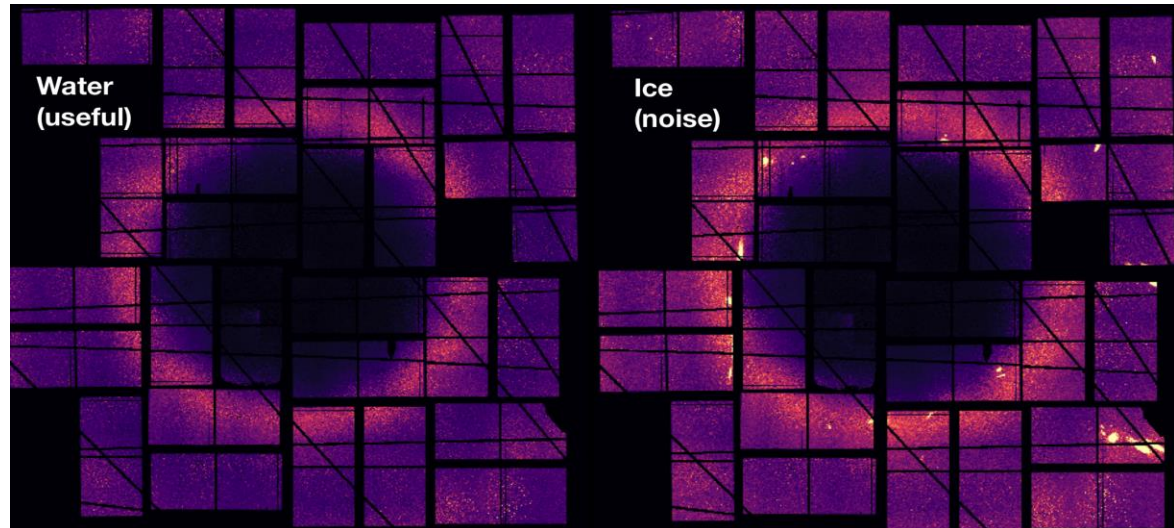


**Normal Jet  
(delivering sample)**



**Iced Jet  
(danger!)**

**Data analysis:**  
e.g. sorting shots

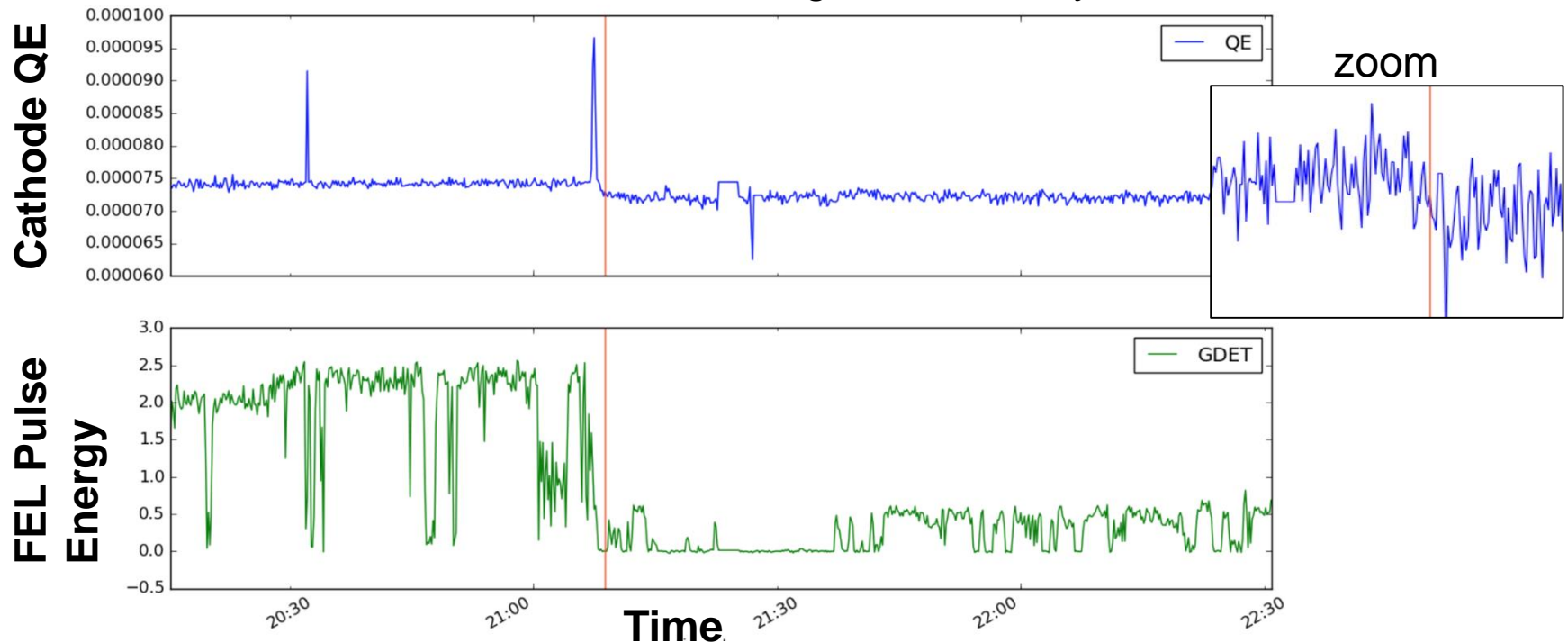


# Anomaly/Breakout Detection

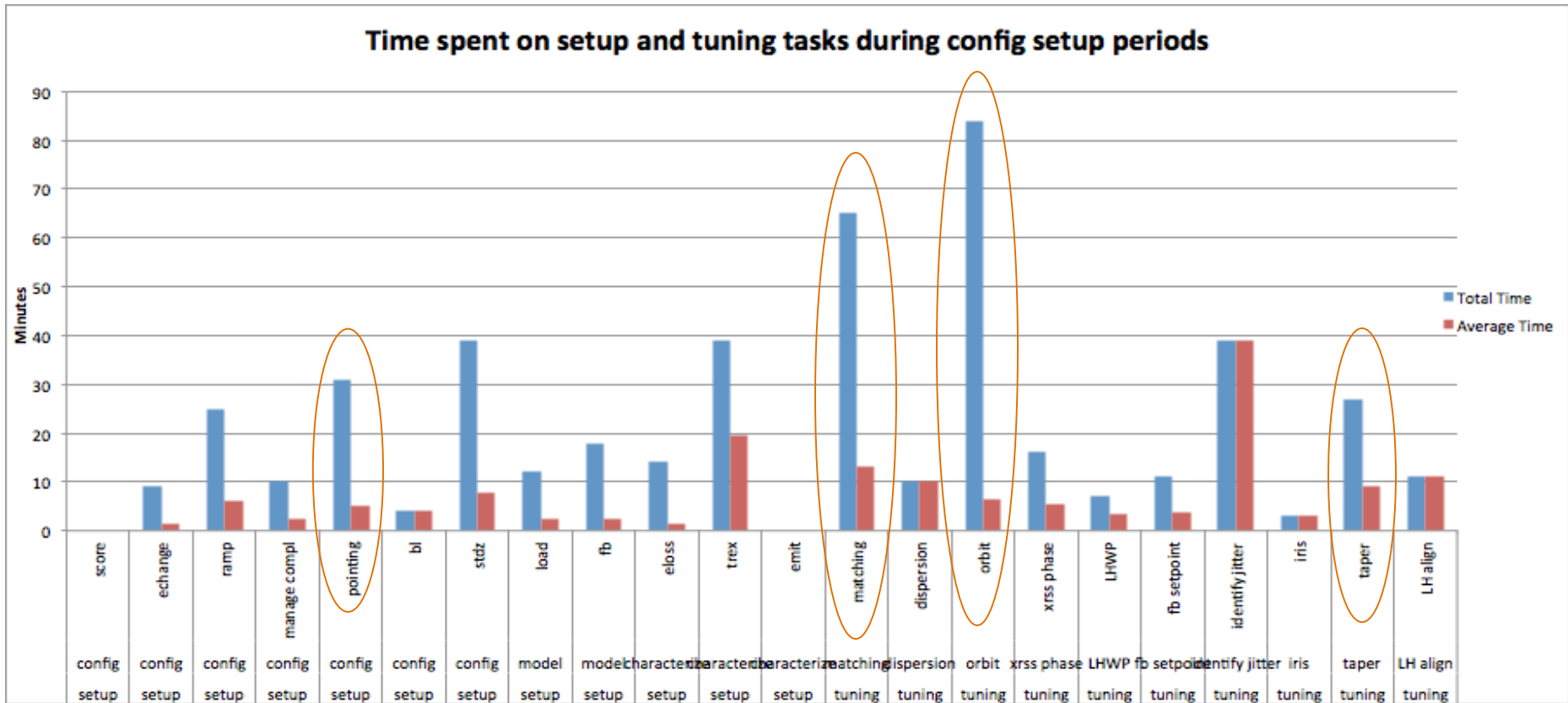
Can we detect if something is broken or about to break?

- 200,000 PVs: no human can keep an eye on all of them
- Signals are complex: simple thresholds cannot work

Cathode QE drop caused hours of downtime. Breakout detection would have found change immediately!



## How to optimize 2km long machine?



2015: 450 hand tuning hours, 250 dedicated!

⇒ Lots of opportunity to speed operations and relieve operator load

# Machine Tuning Automation

## Working with AOSD – Faster tuning, fewer errors

MainWindow <@lcls-srv01>

Choose Charge and Step Through Steps

Desired Charge: 20    Previous    Next

Step Number: 1    Save Current Configuration using SCORE

Parameter Settings and Readback

| Parameter       | Original    | Current     | Action Processed                            |
|-----------------|-------------|-------------|---|
| Mech Shutter    | Open        | Open        | 1 2016-10-07 11:00:31 - I'm all charged up! |
| Charge setpoint | 0.25 nC     | 0.25 nC     |   |
| Charge FB On    | Matlab On   | Matlab On   |   |
| Iris Value      | 1.2 mm      | 1.2 mm      |   |
| BPM Attenuation | 0.25 nC     | 0.25 nC     |   |
| Laser           | Coherent #1 | Coherent #1 |   |
| Laser Phase     | 0.0 deg     | 0.0 deg     |   |
| Gun Phase       | 0.0 deg     | 0.0 deg     |   |
| Finished!       | Nope        | Nope        |   |

**Injector setup procedures**

Charge Change GUI Initialized

Optimization Scan Panel    Scan Setup Panel

|    | PVs             | Saved Value     | Current Value   | Active                              |
|----|-----------------|-----------------|-----------------|-------------------------------------|
| 1  | QUAD:LTU1:62... | -93.23365       | -93.23365       | <input type="checkbox"/>            |
| 2  | QUAD:LTU1:64... | 86.19807        | 86.19807        | <input type="checkbox"/>            |
| 3  | QUAD:LTU1:66... | -81.05878       | -81.05878       | <input type="checkbox"/>            |
| 4  | QUAD:LTU1:68... | 61.30179        | 61.30179        | <input type="checkbox"/>            |
| 5  | QUAD:LI26:20... | 7.19764113071   | 7.19764113071   | <input type="checkbox"/>            |
| 6  | QUAD:LI26:30... | -15.3553777167  | -15.3553777167  | <input type="checkbox"/>            |
| 7  | QUAD:LI26:40... | 21.384          | 21.384          | <input type="checkbox"/>            |
| 8  | QUAD:LI26:50... | -8.22554056197  | -8.22554056197  | <input type="checkbox"/>            |
| 9  | QUAD:LI26:60... | 18.5904298714   | 18.5904298714   | <input type="checkbox"/>            |
| 10 | QUAD:LI26:70... | -13.8490479091  | -13.8490479091  | <input type="checkbox"/>            |
| 11 | QUAD:LI26:80... | 17.4268804015   | 17.4268804015   | <input type="checkbox"/>            |
| 12 | QUAD:LI26:90... | -15.1786363158  | -15.1786363158  | <input type="checkbox"/>            |
| 13 | QUAD:LI21:22... | 0.44626         | 0.44626         | <input type="checkbox"/>            |
| 14 | QUAD:LI21:25... | -0.5662         | -0.5662         | <input type="checkbox"/>            |
| 15 | QUAD:LI24:74... | -0.314594375    | -0.314594375    | <input type="checkbox"/>            |
| 16 | QUAD:LI24:86... | 0.437890625     | 0.437890625     | <input type="checkbox"/>            |
| 17 | QUAD:LTU1:44... | -0.846989375    | -0.846989375    | <input type="checkbox"/>            |
| 18 | QUAD:LTU1:46... | -2.96971        | -2.96971        | <input type="checkbox"/>            |
| 19 | QUAD:LI21:20... | -5.60147        | -5.53939858216  | <input checked="" type="checkbox"/> |
| 20 | QUAD:LI21:21... | 4.74459         | 4.67726546309   | <input checked="" type="checkbox"/> |
| 21 | QUAD:LI21:27... | -7.09875        | -6.94980214166  | <input checked="" type="checkbox"/> |
| 22 | QUAD:LI21:27... | 8.17251         | 9.37881217554   | <input checked="" type="checkbox"/> |
| 23 | PHYS:ACR0:OC... | -0.000146939... | -0.000146939... | <input type="checkbox"/>            |
| 24 | PHYS:ACR0:OC... | 0.000214041...  | 0.000214041...  | <input type="checkbox"/>            |
| 25 | PHYS:ACR0:OC... | 0.000214041...  | 0.000214041...  | <input type="checkbox"/>            |

**FEL optimization**

Middle click a PV then the table to add your favorite device!

Update reference    Reset All    Check    Uncheck

Start scan    Logbook    GP 2D Heatmap    Help/Docs

Objective Function Monitor

Device Monitor

**Fast global steering**

Steer Dump (1-1)

Restore Dump

Ready to 1-1 Steer Dump

Regold L3,U28    Stop/Enable L3,U28

Show Ref    Get Reference Orbit    Auto Scale Y Axis

Diff Orbit    Fix Y Axis    +/-mm    L

Help    Relaunch GUI    Pointopia GUI    Logbook LCLS

Mirror and Camera

Current Destination: Hard Line    Current Camera: P3H

Gas Attenuator: Off    Transmission (%): 100

Gas Attenuator    10% Transmission    Set/Restore 10 Hz    Check Positions

Off    On    Set    Rest    Set    Rest    Measure

Reference Positions

X 500    Y 2350

Update Ref Pos's

Measured Positions

X 406    Y 2333

Repoint Manually

Xi 0    Yi 0

Xf 94    Yf 8.5

Repoint(Xi, Yi, Xf, Yf)

**FEL pointing**

01-Oct-2016 01:08:09 Finished Repointing

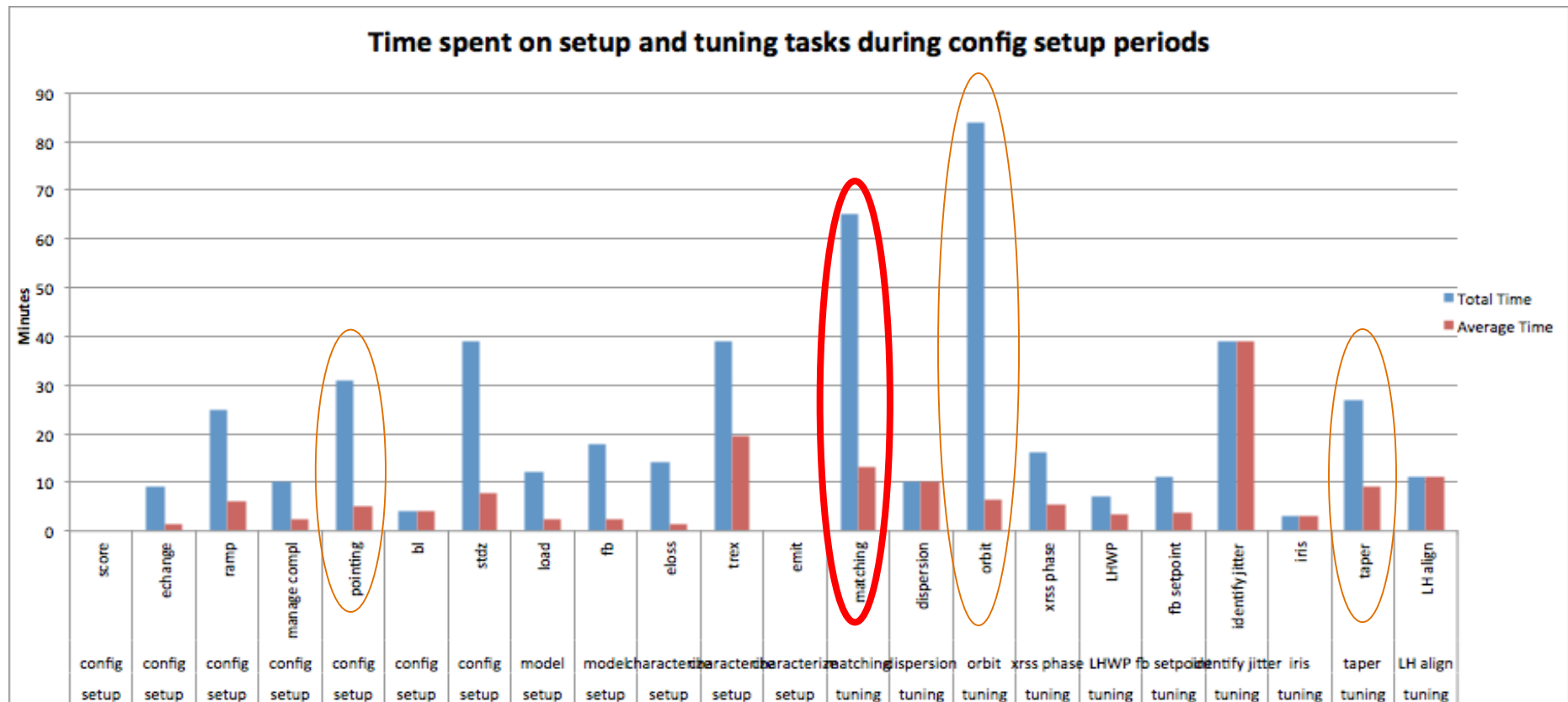
01-Oct-2016 01:08:40 Repointing

01-Oct-2016 01:04:07 Pulling P3H

01-Oct-2016 01:03:59 Averaging 50 shots

01-Oct-2016 01:03:57 Putting in P3H

## How to optimize 2km long machine?

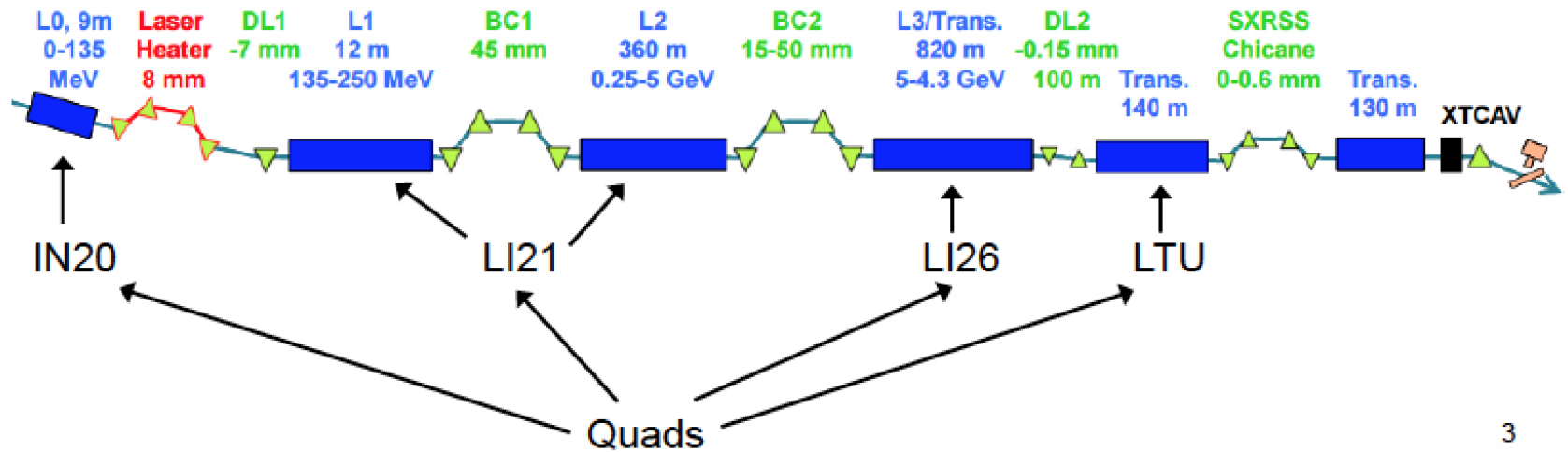


2015: 450 hand tuning hours, 250 dedicated!

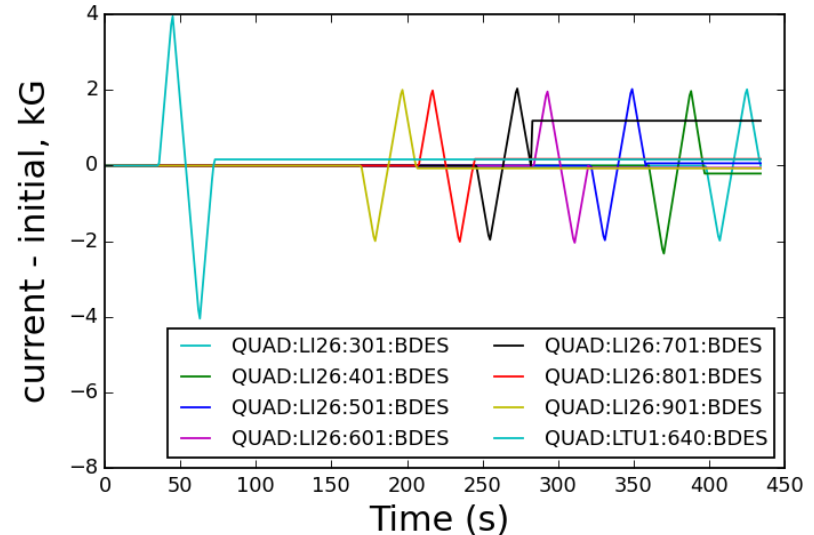
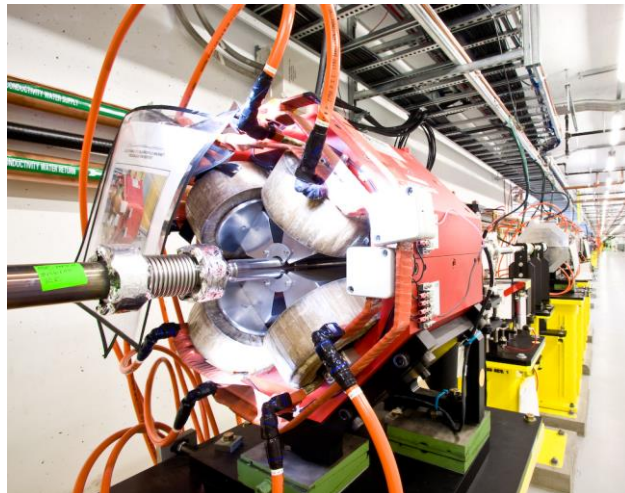
⇒ Lots of opportunity to speed operations and relieve operator load

# Online optimization

## Online optimization of quadrupole magnets



3

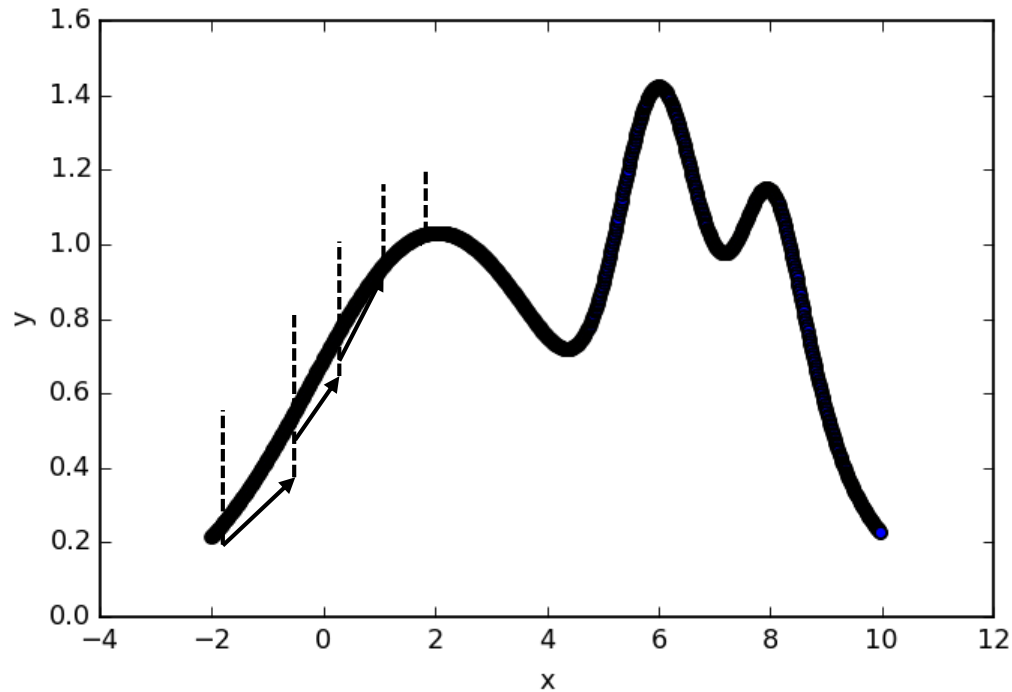


# Online optimization

Tried several optimization approaches:

→ Gradient/simplex methods

(Nelder-Mead in general use)

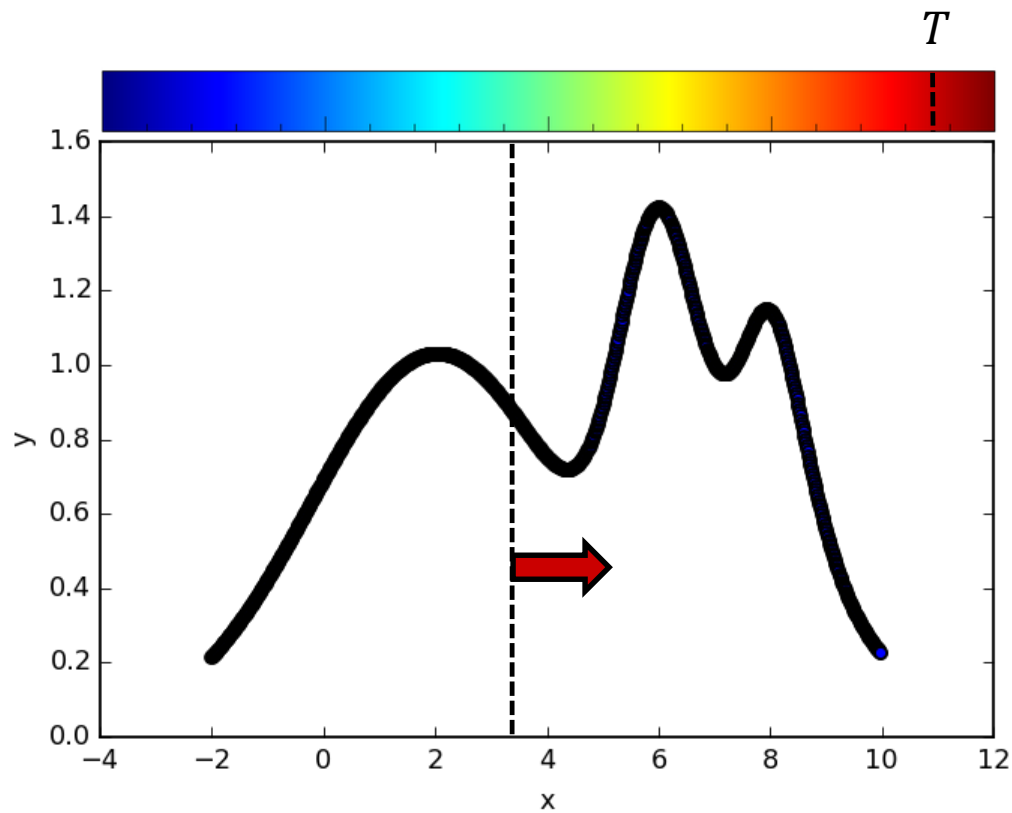




# Online optimization

Still many optimizers to try:

→ Simulated annealing, genetic algorithms, etc.



# Bayesian optimization

Bayesian approach: introduce probabilistic model

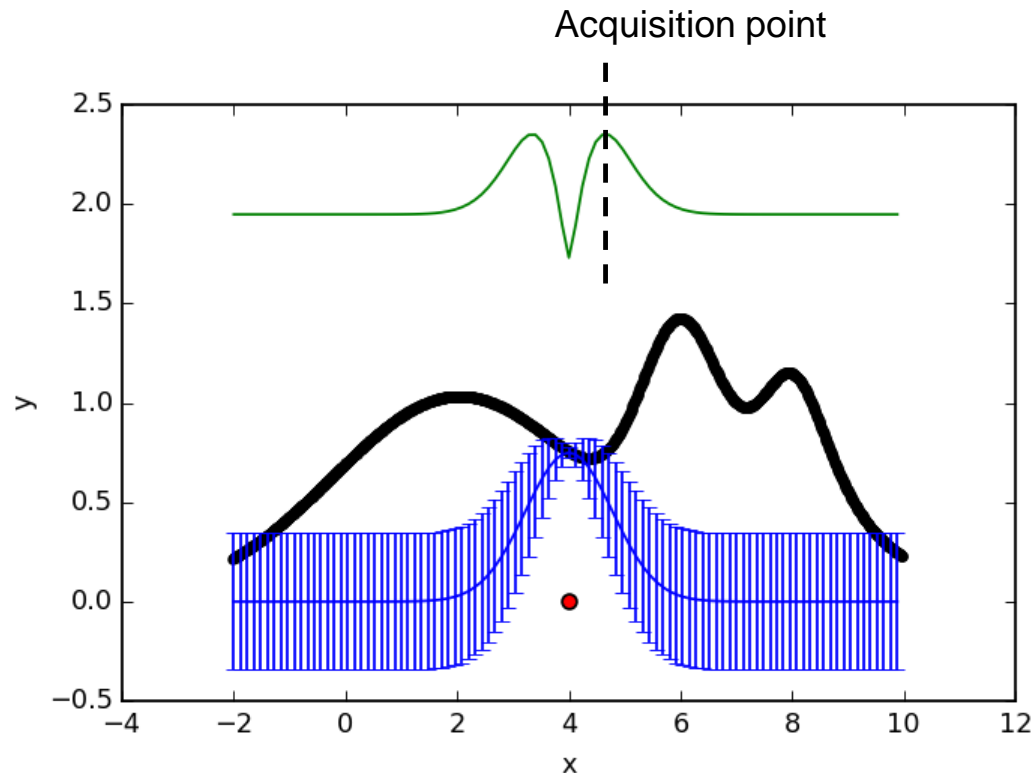
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

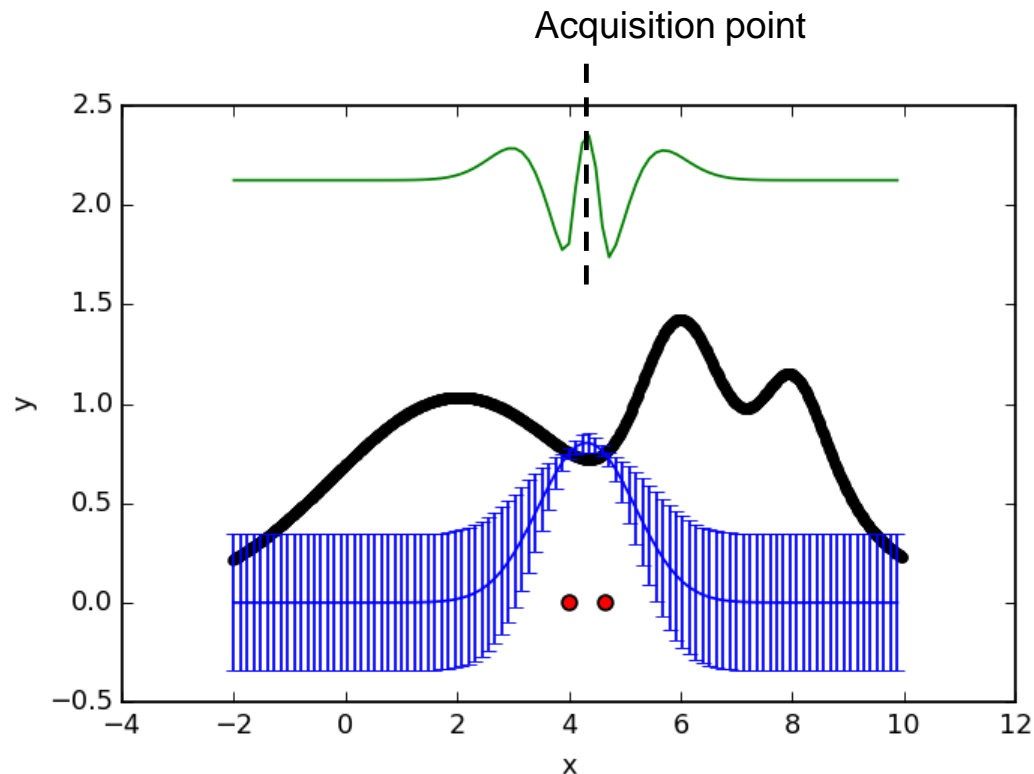
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

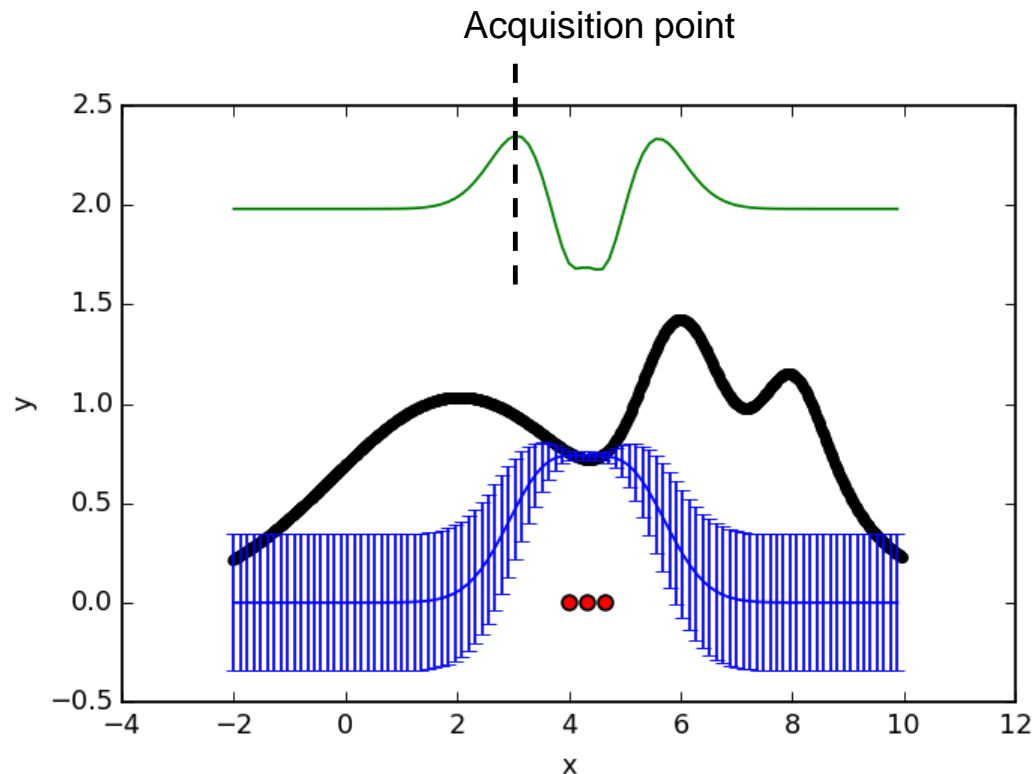
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

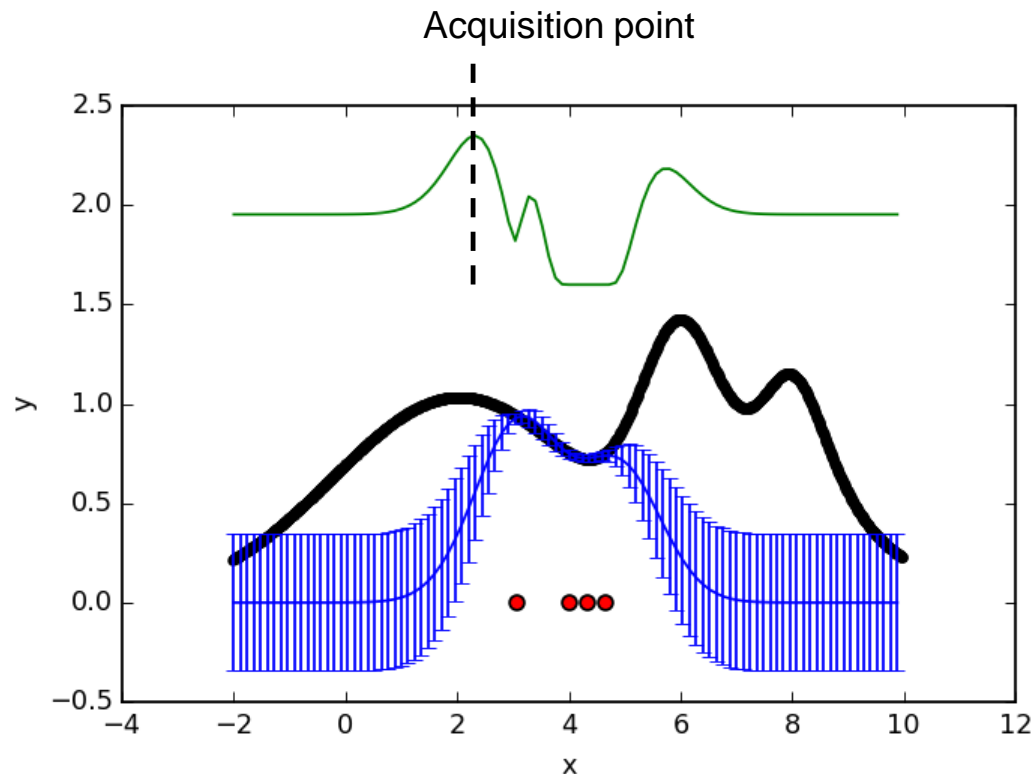
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

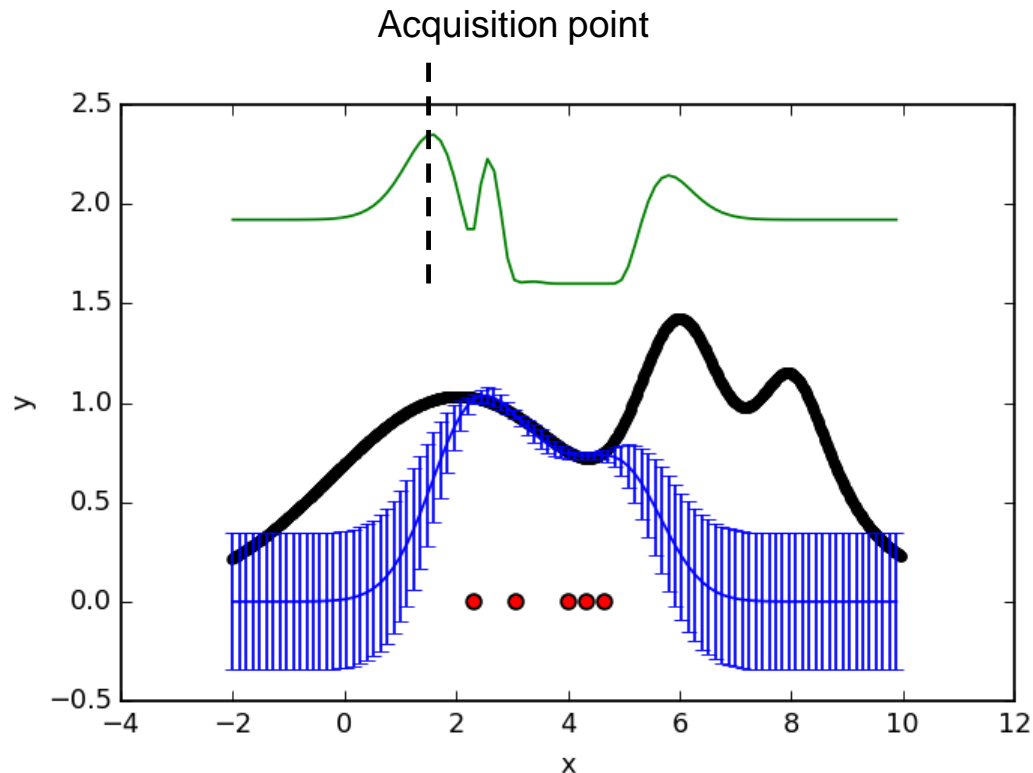
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

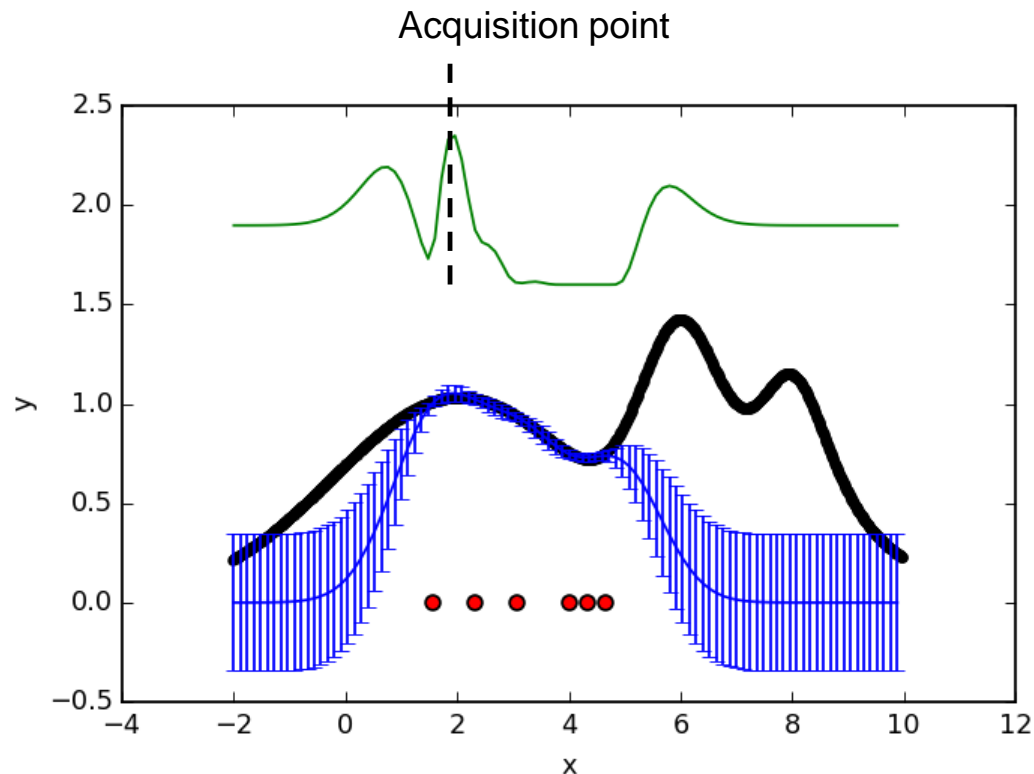
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

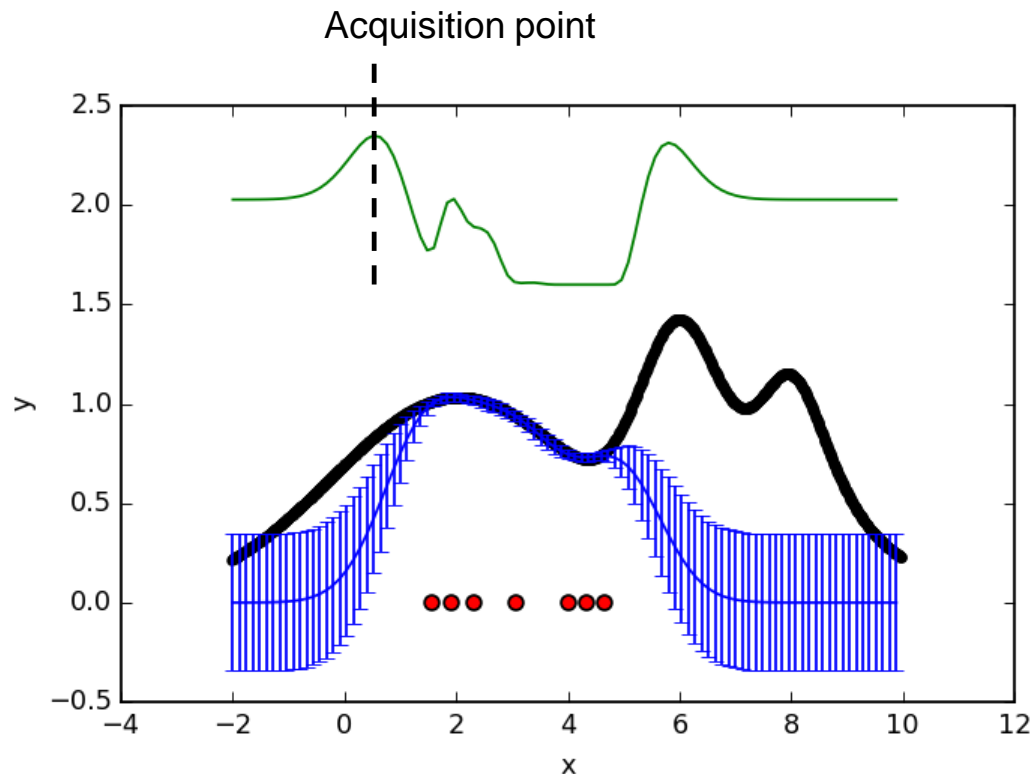
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior





# Bayesian optimization

Add probabilistic model

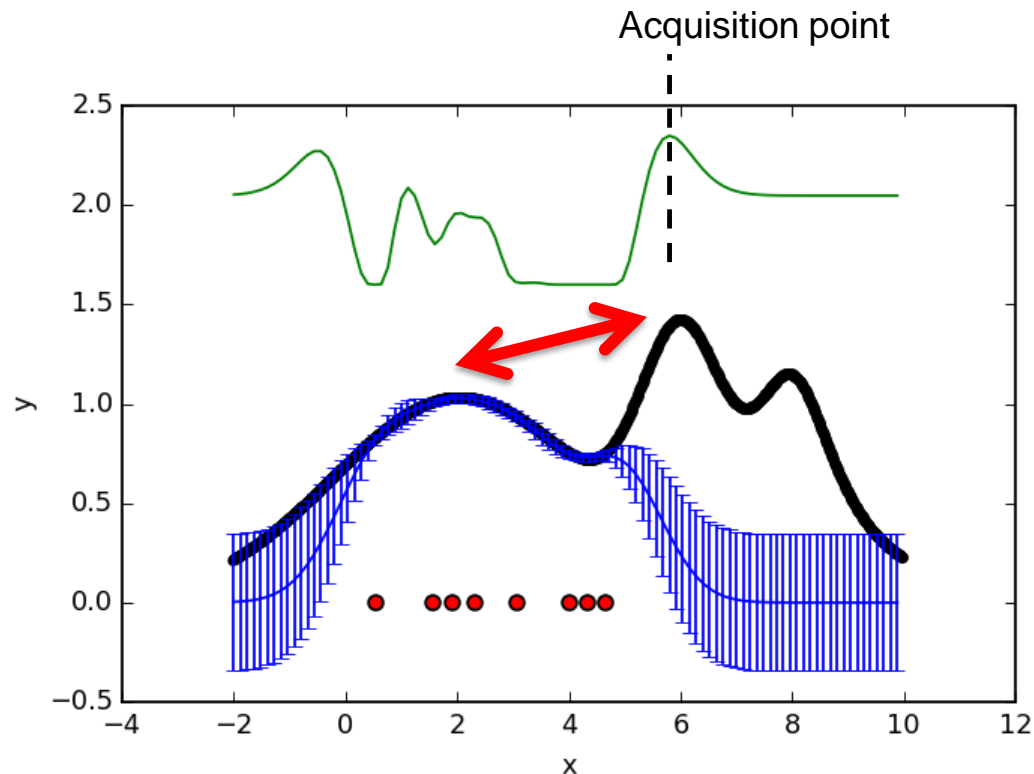
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

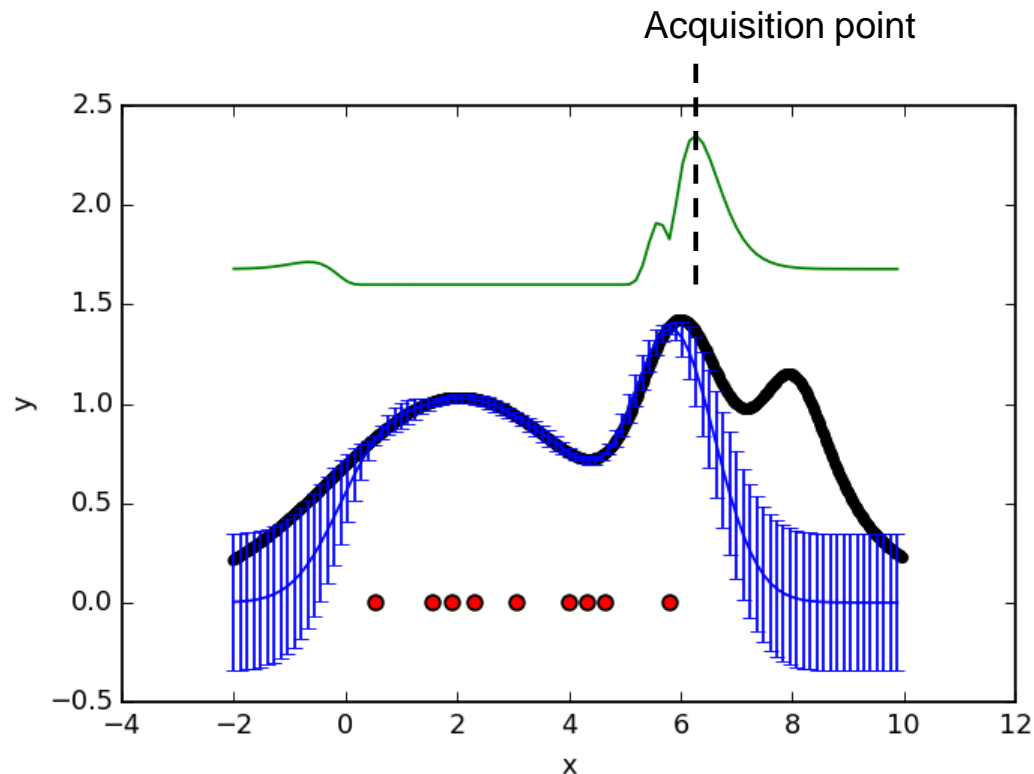
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

Posterior



# Bayesian optimization

Add probabilistic model

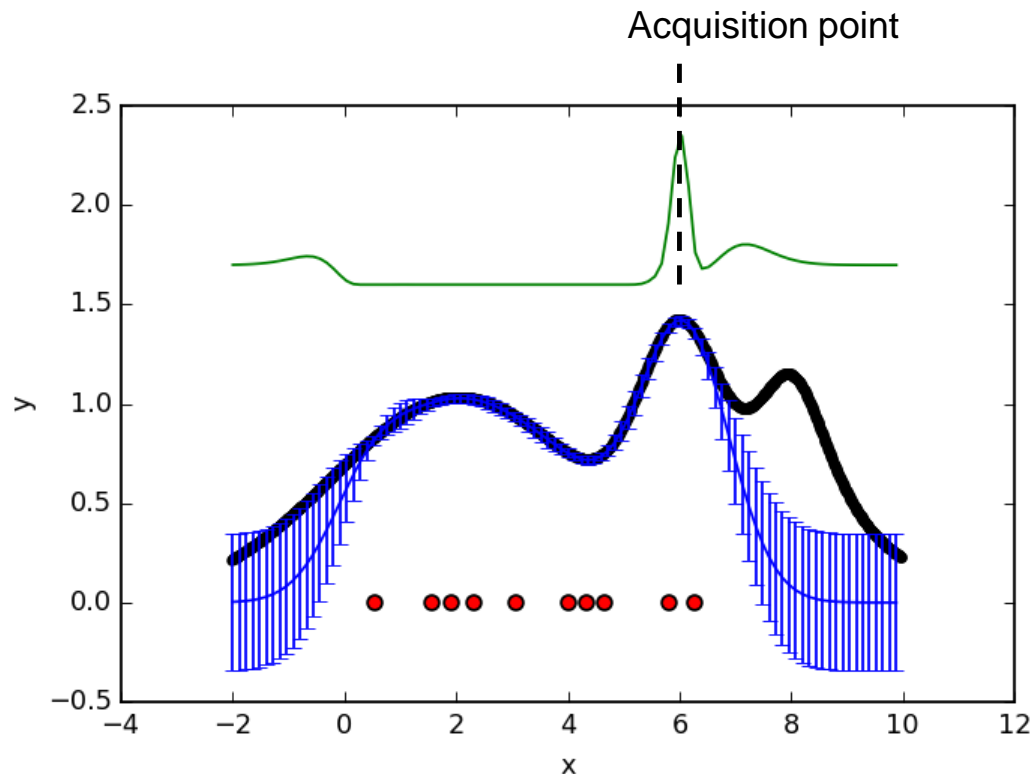
→ create acquisition function

→ more efficient search of high dimensional space

Acquisition function

Ground truth

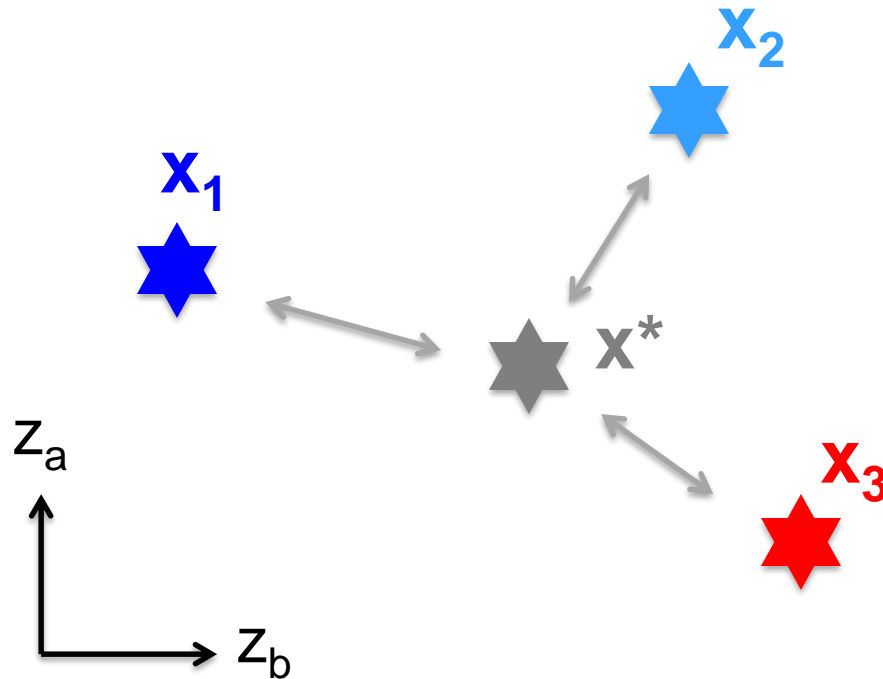
Posterior



# Gaussian Process Optimizer

Gaussian process: instance based learning method

Covariance function:  $k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$

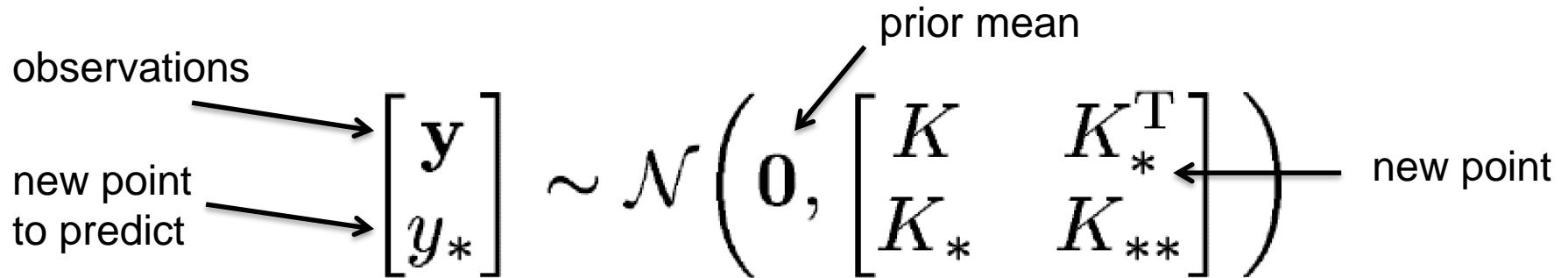


$$y^* = f(x^*) \\ = f(z_a, z_b, z_a^* z_b, \dots)$$

# Gaussian Process Optimizer

Gaussian process: instance based learning method

Covariance function:  $k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$

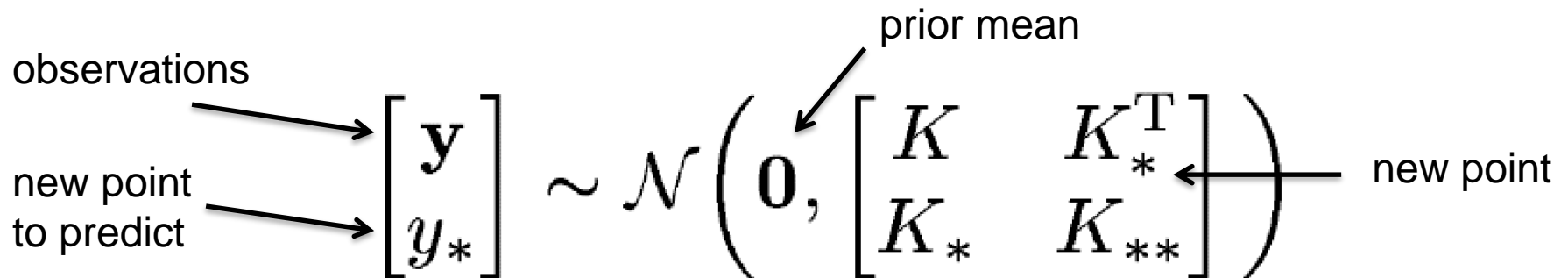


$$K = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{bmatrix} \quad \begin{aligned} K_* &= [k(x_*, x_1) \cdots k(x_*, x_n)] \\ K_{**} &= k(x_*, x_*) \end{aligned}$$

# Gaussian Process Optimizer

Gaussian process: instance based learning method

Covariance function:  $k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$



Prediction of new point:  $\bar{y}_* = K_* K^{-1} \mathbf{y}$

Variance of new point:  $\text{var}(y_*) = K_{**} - K_* K^{-1} K_*^T$

# Gaussian Process Optimizer

Gaussian process: instance based learning method

Covariance function:  $k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$

observations

new point to predict

prior mean

new point

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N} \left( \mathbf{0}, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix} \right)$$

Acquisition function:

$$EI(x^*) = \int_{\tilde{y}}^{\infty} (y^* - \tilde{y}) P(y^* | x^*) dy^*$$

best observed point

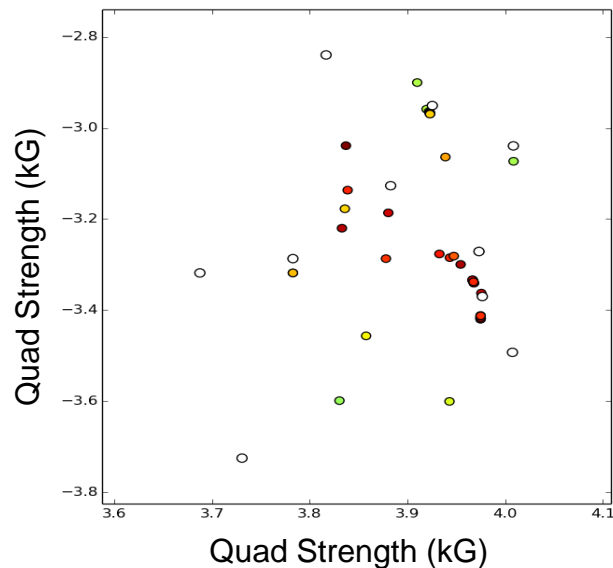
# Gaussian Process Optimizer

Gaussian process: instance based learning method

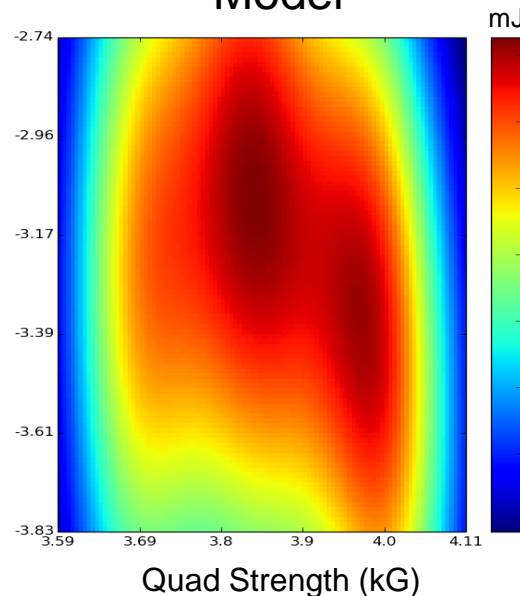
Similarity function:  $k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$

Acquisition function:  $EI(x^*) = \int_{\tilde{y}}^{\infty} (y^* - \tilde{y}) P(y^* | x^*) dy^*$

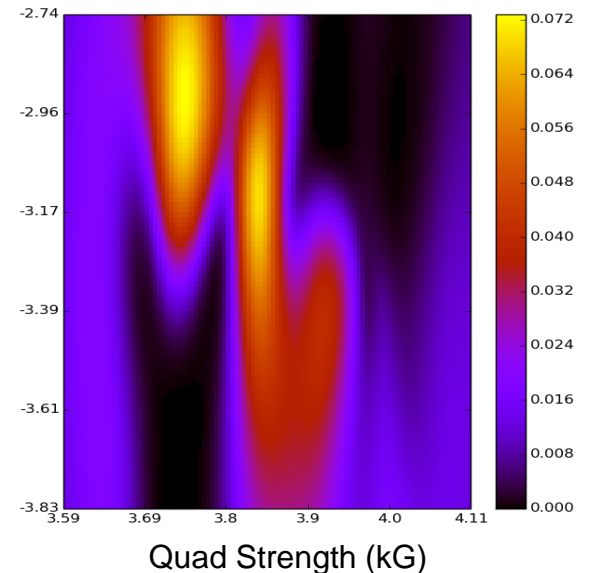
Observations



Model

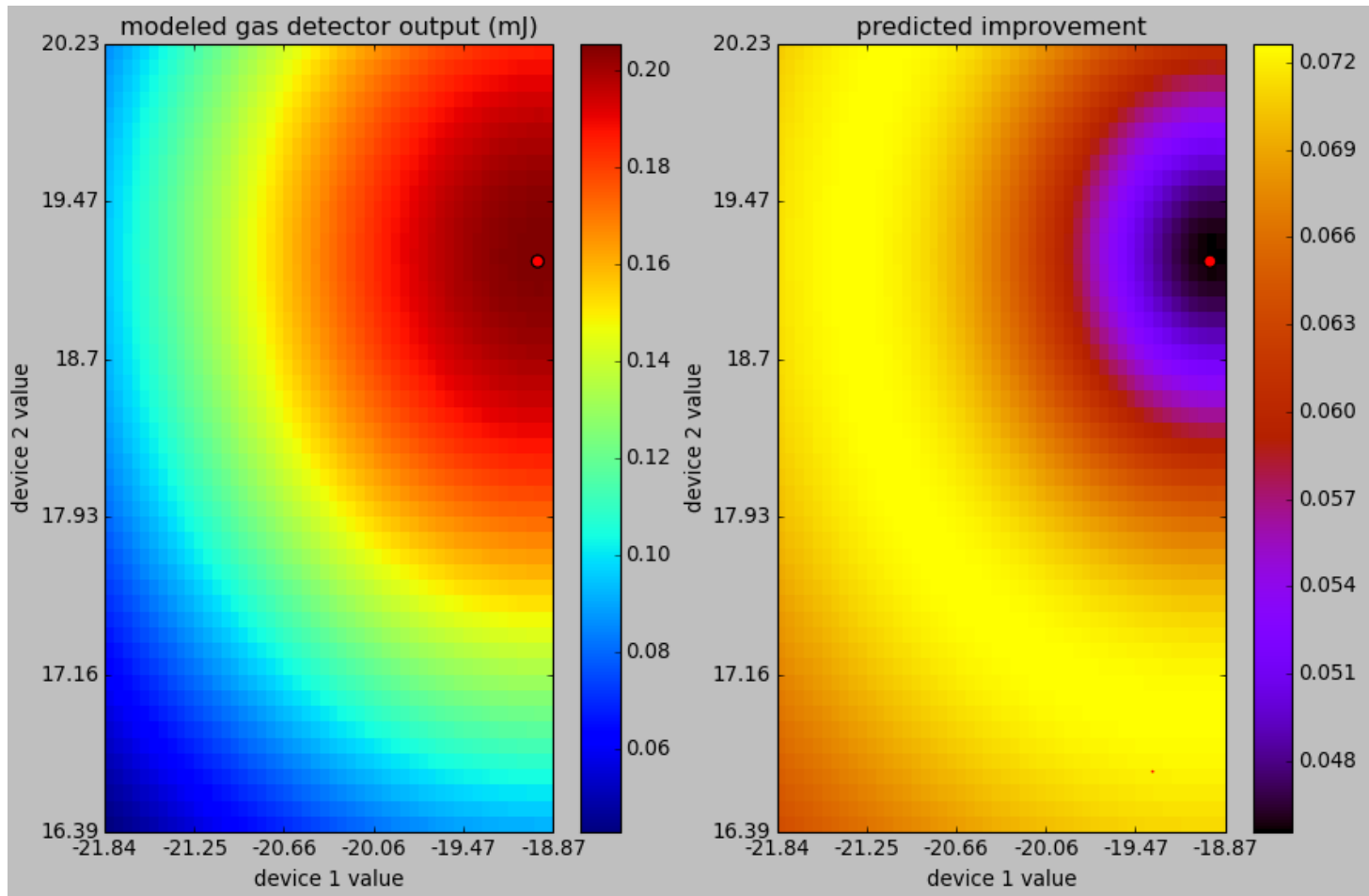


Expected Improvement

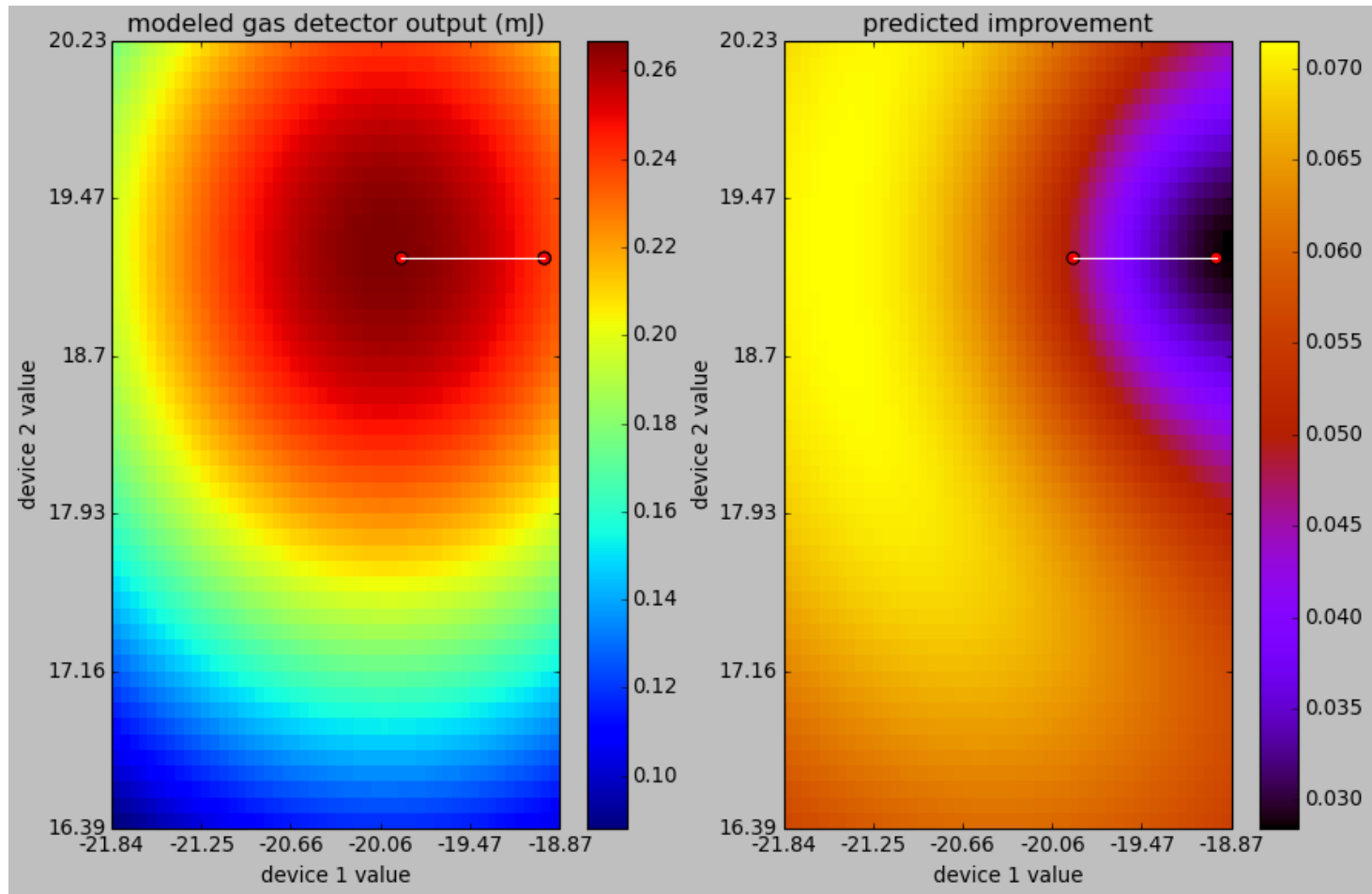




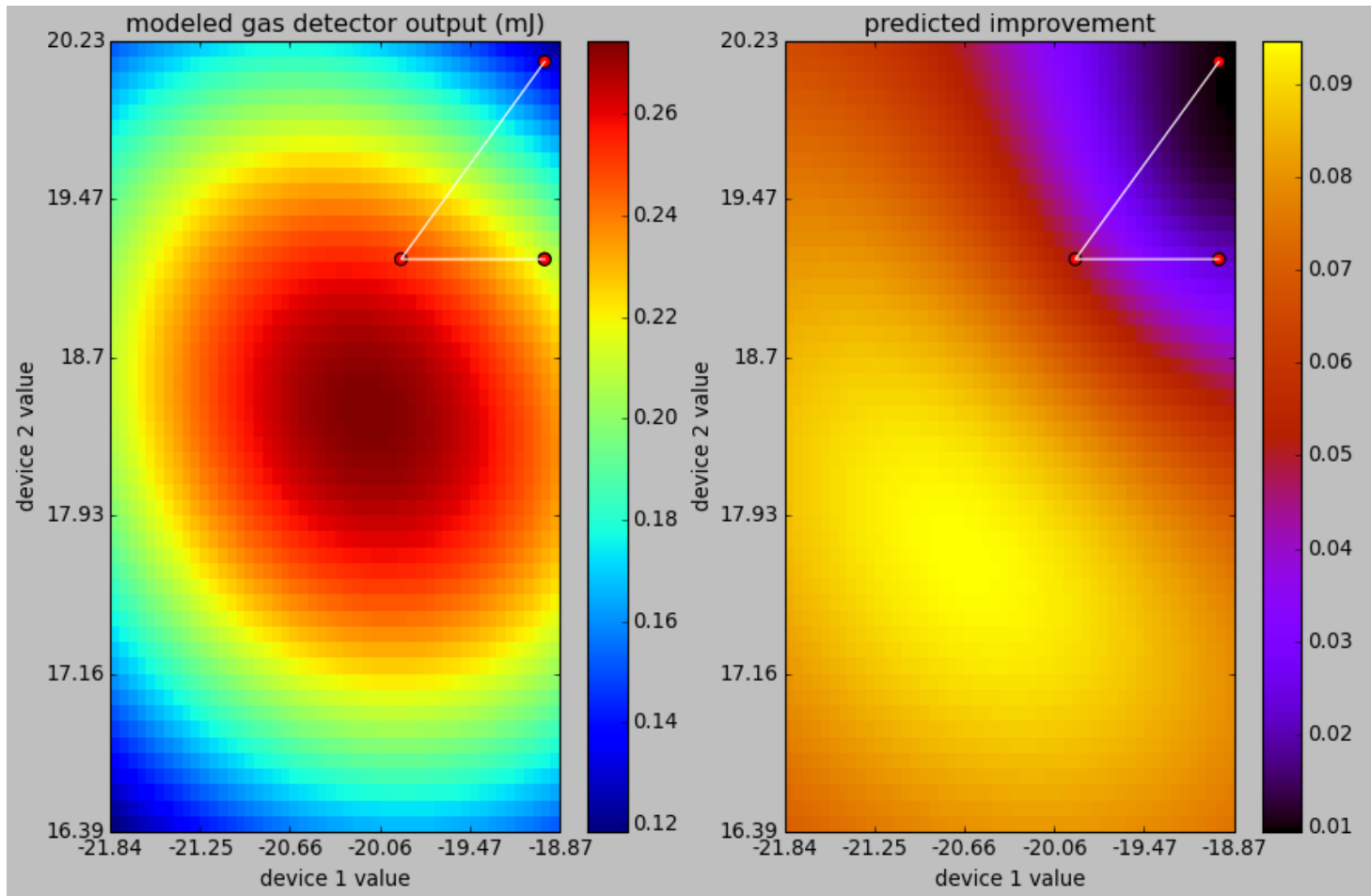
# Point 2



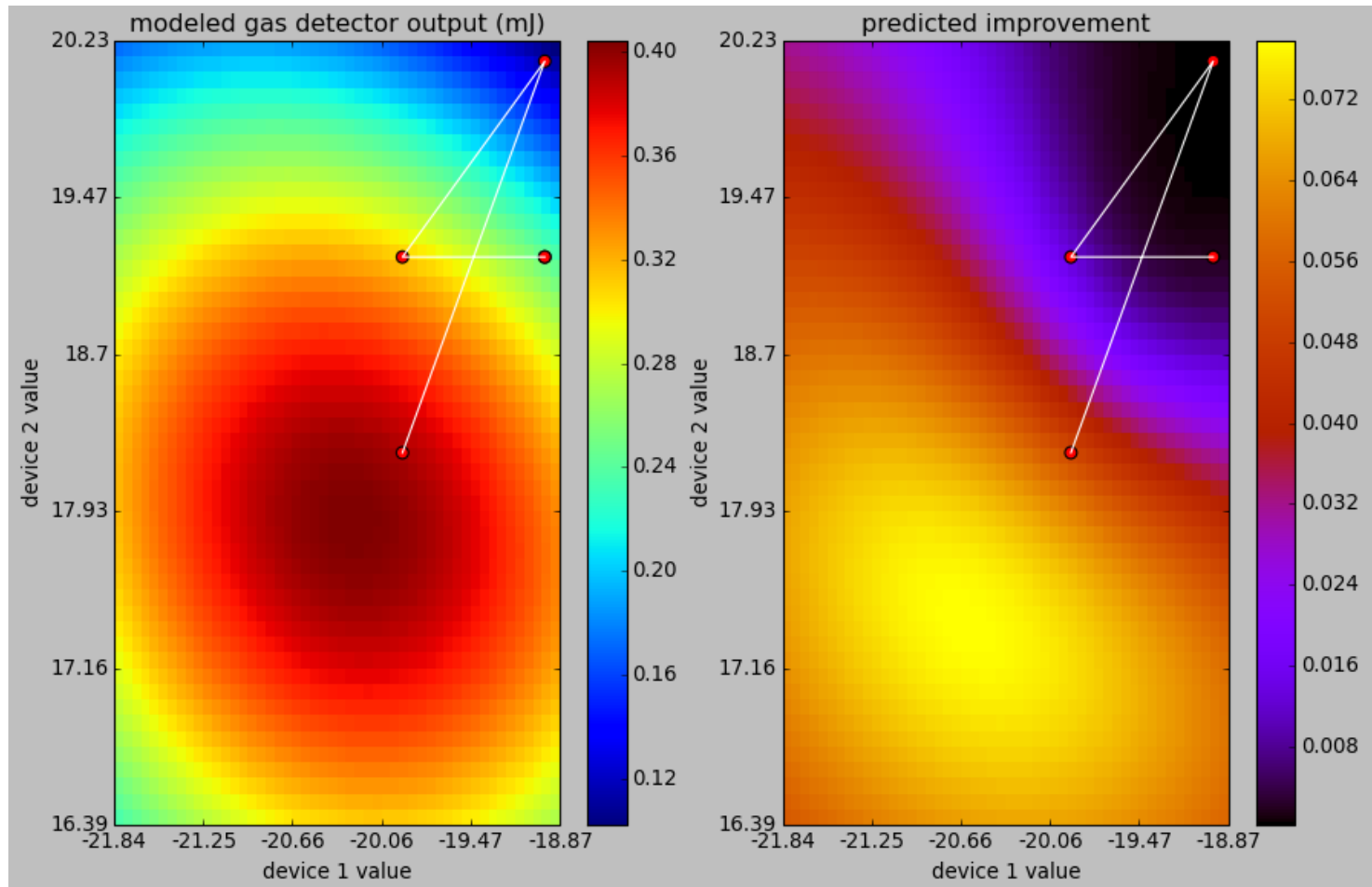
# Point 3



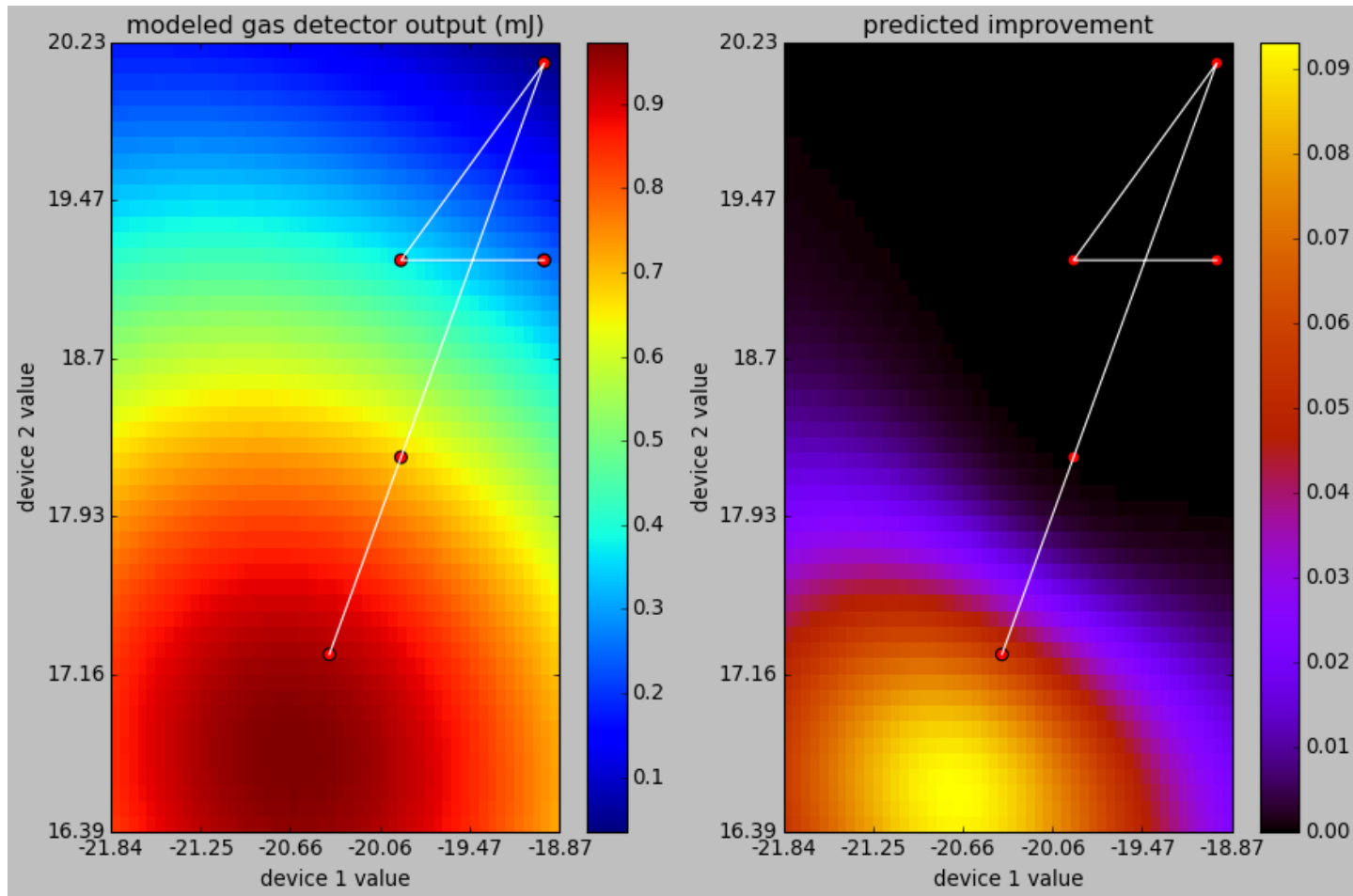
# Point 4



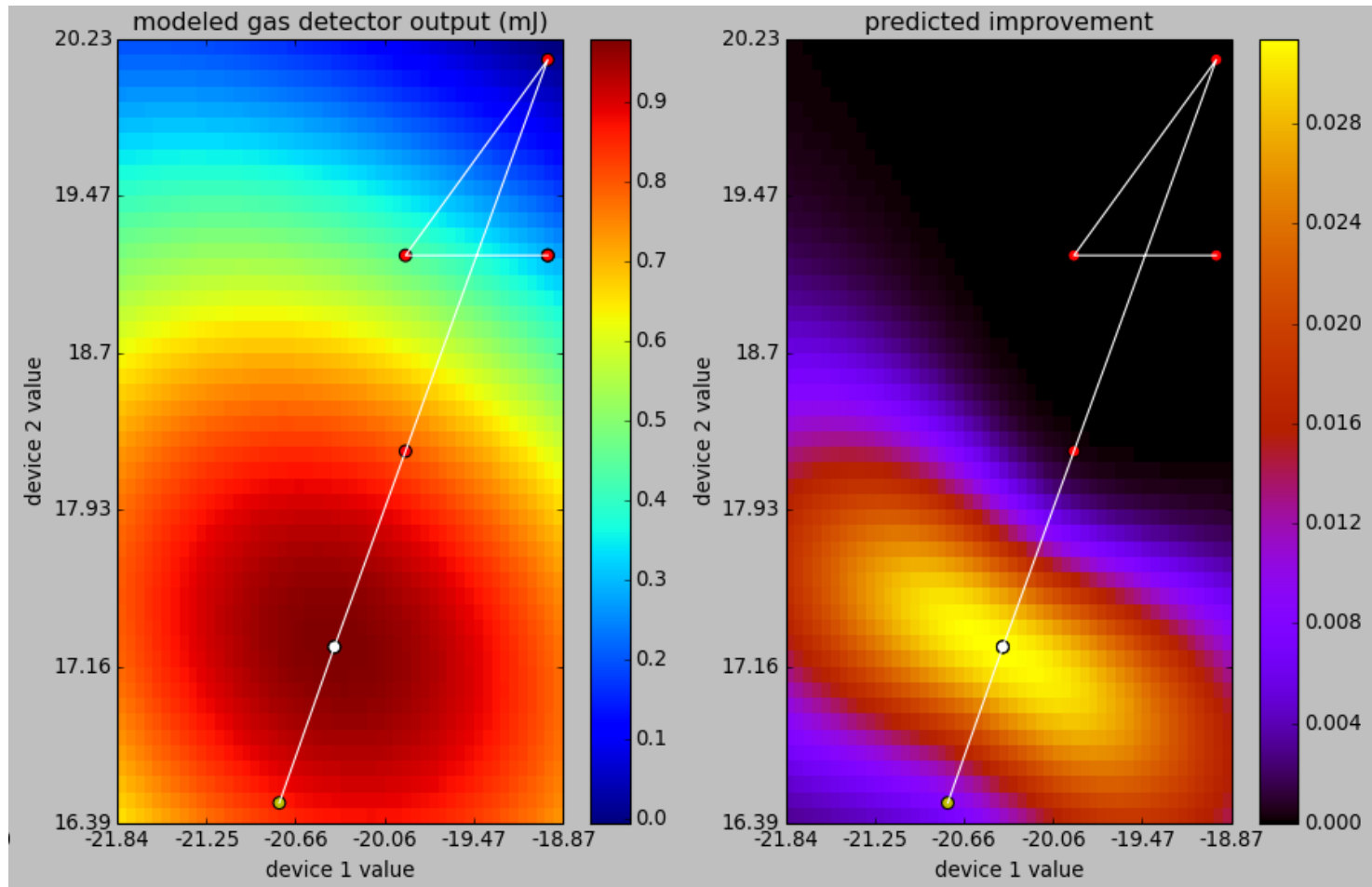
# Point 5



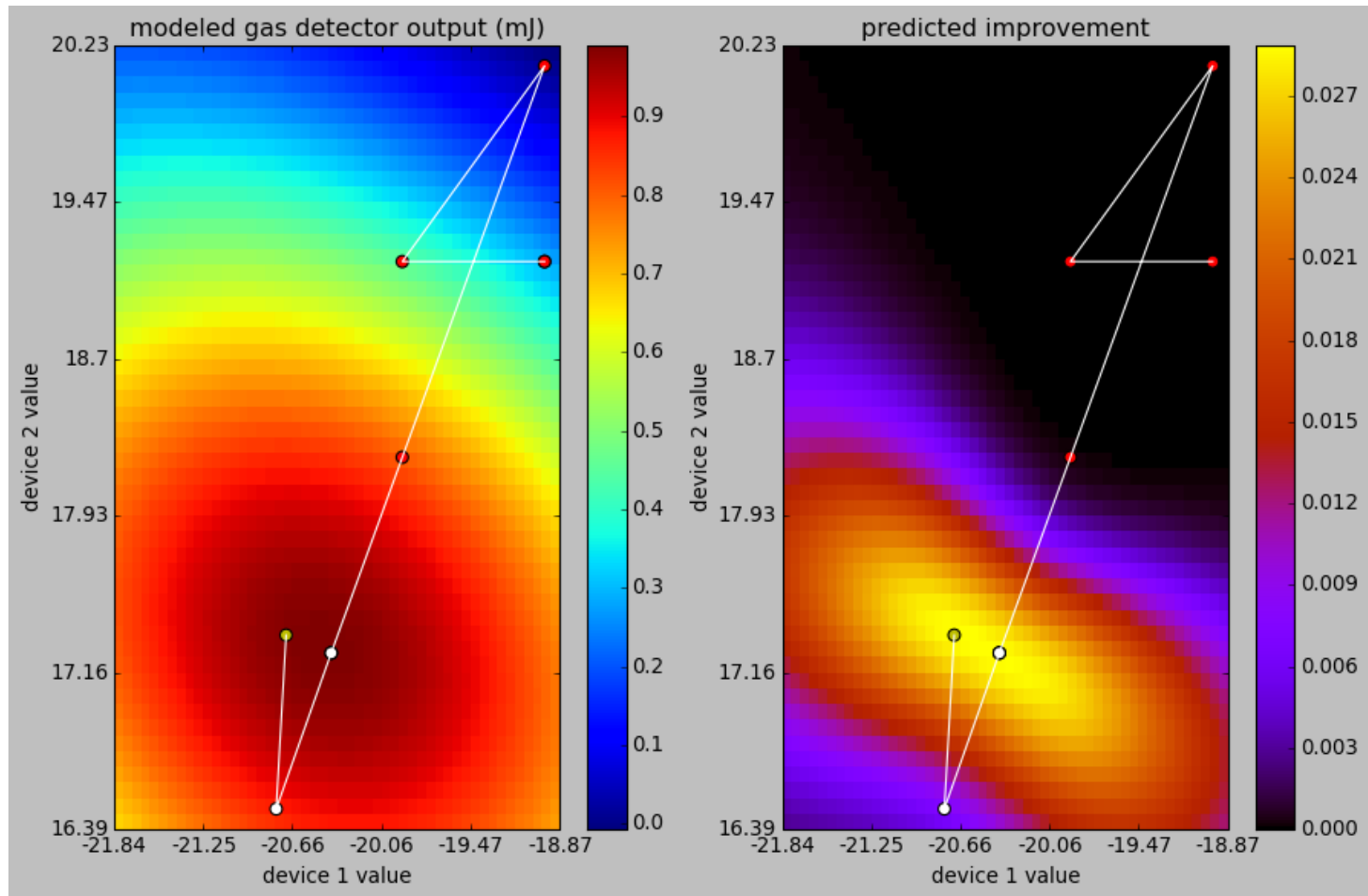
# Point 6



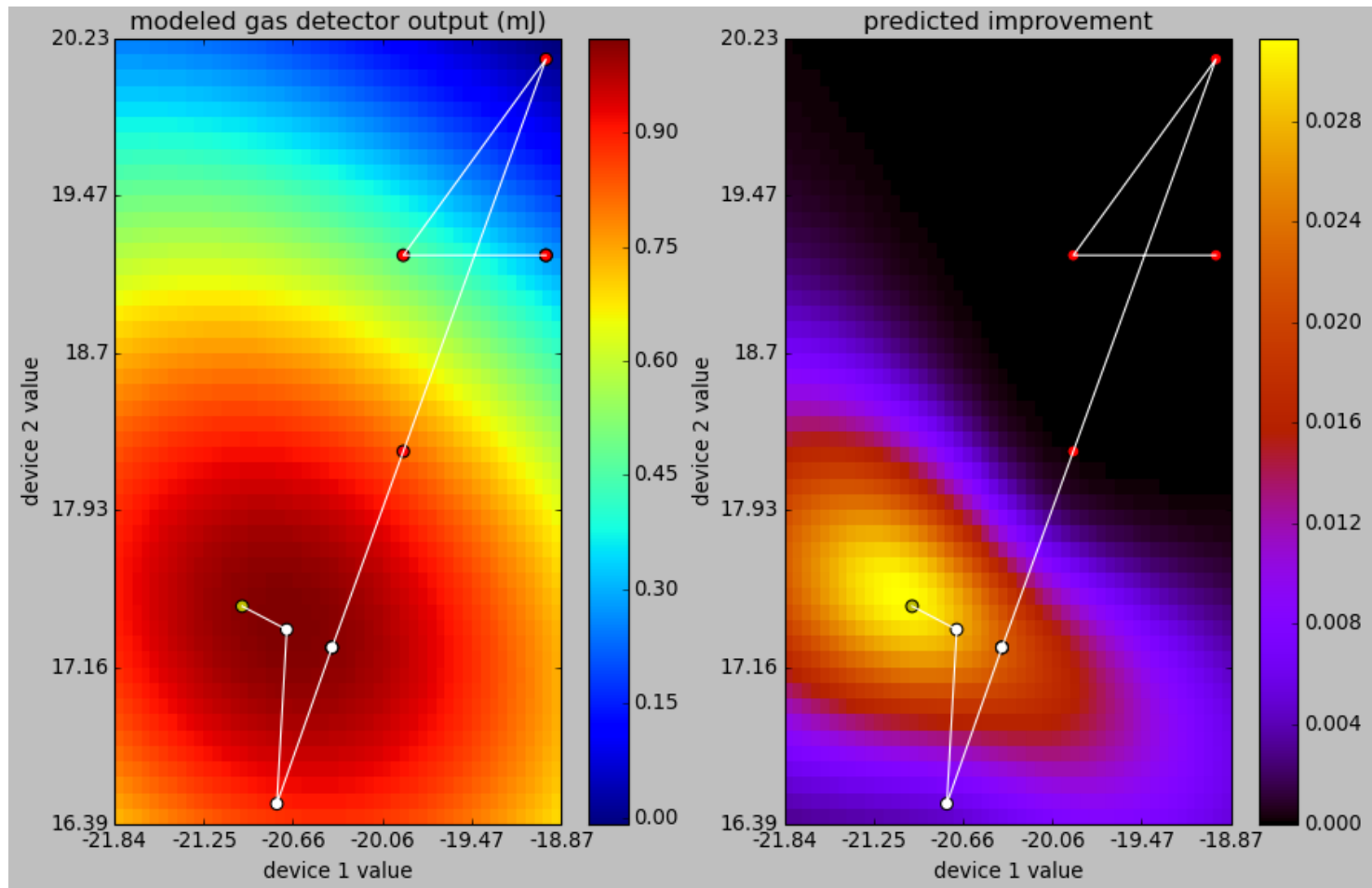
# Point 9



# Point 10

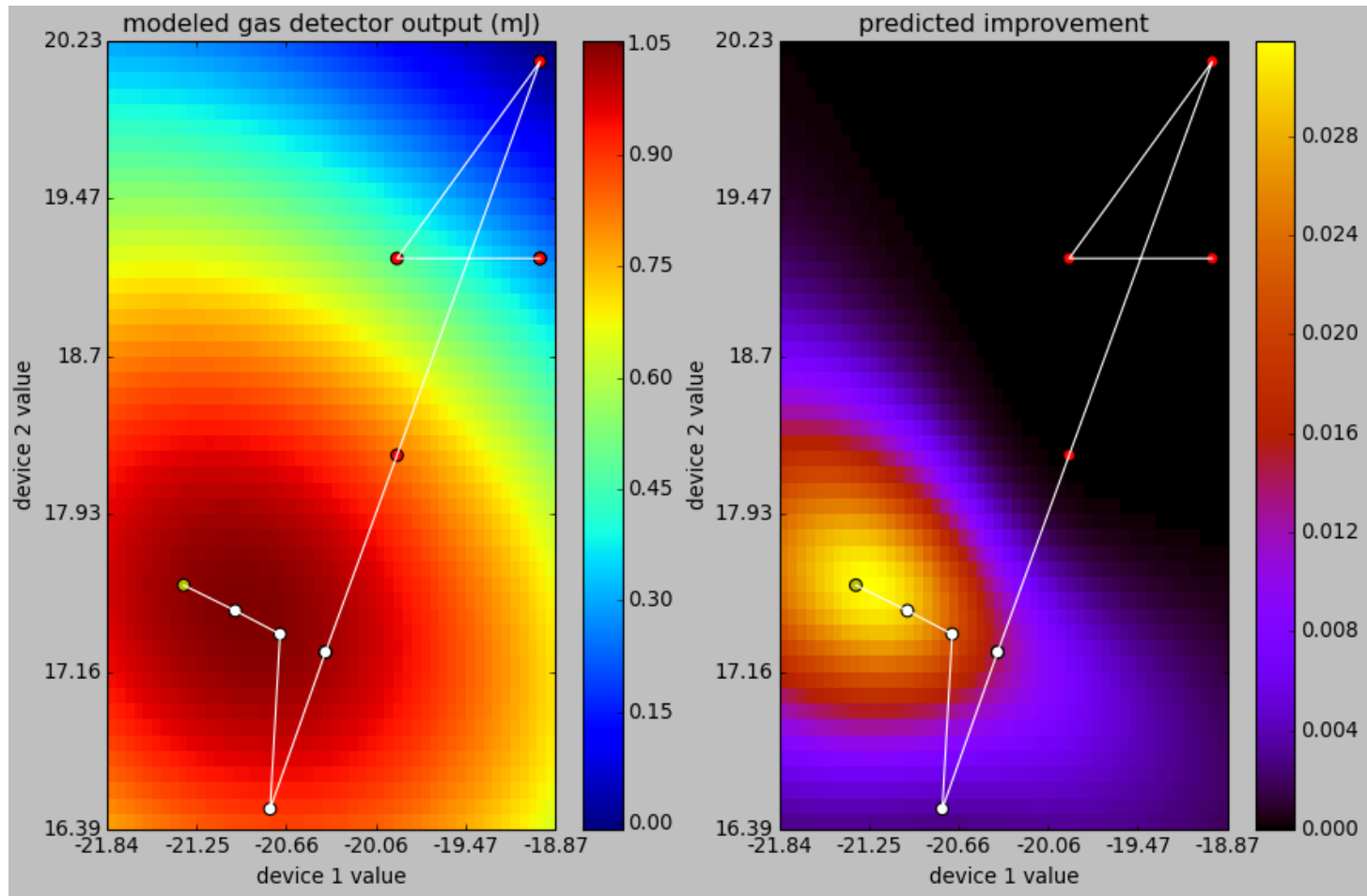


# Point 11

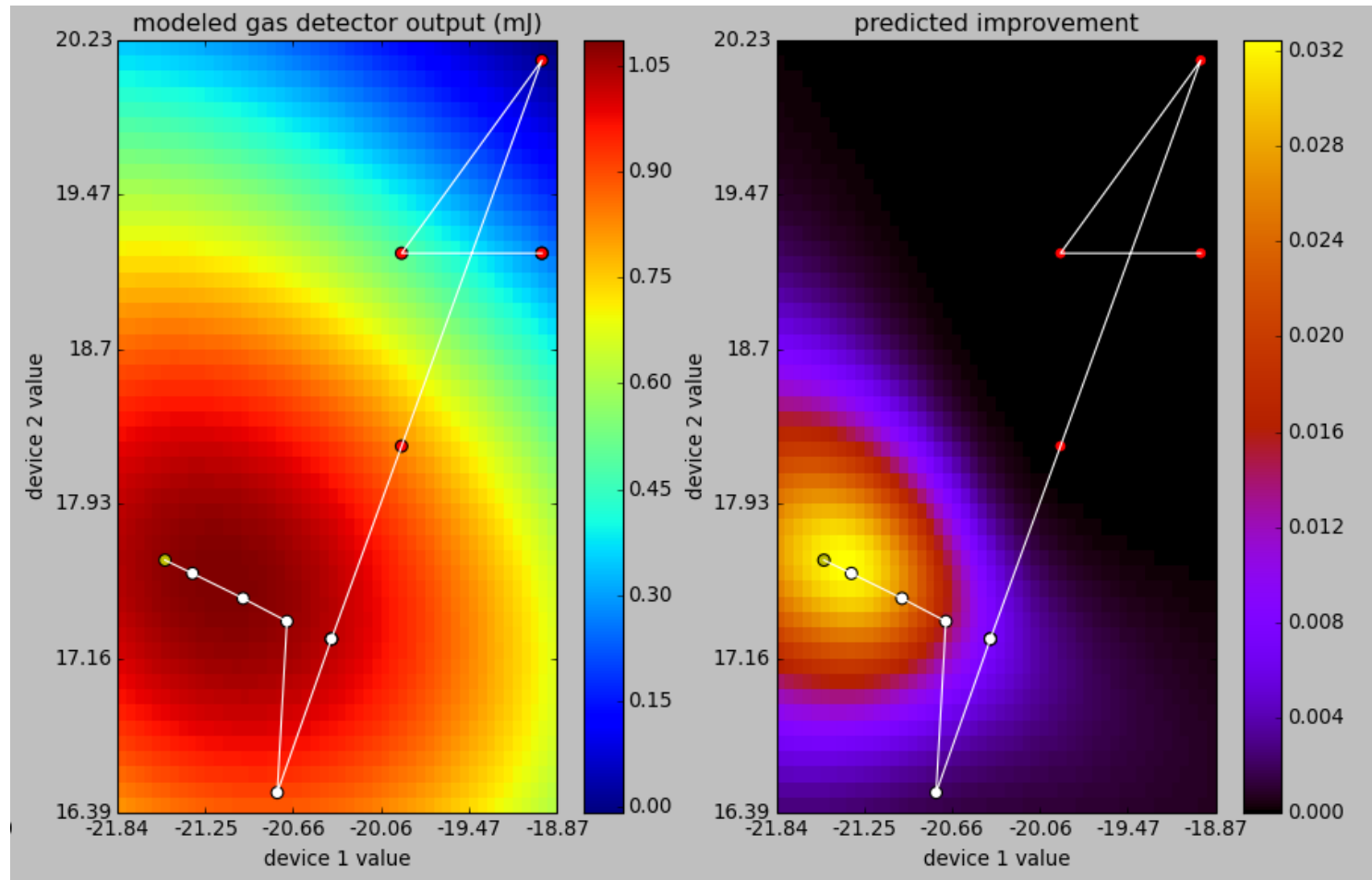




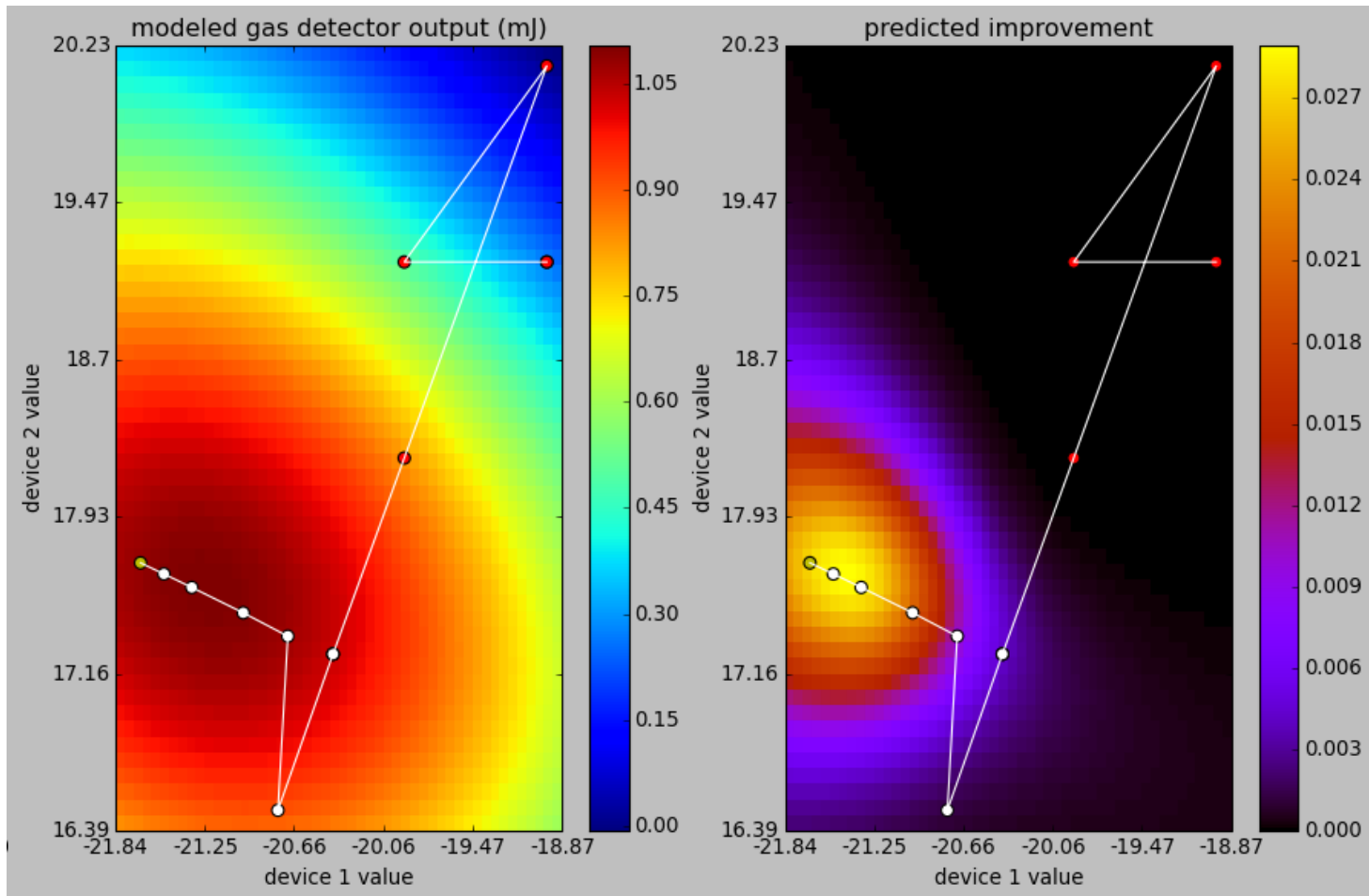
# Point 12



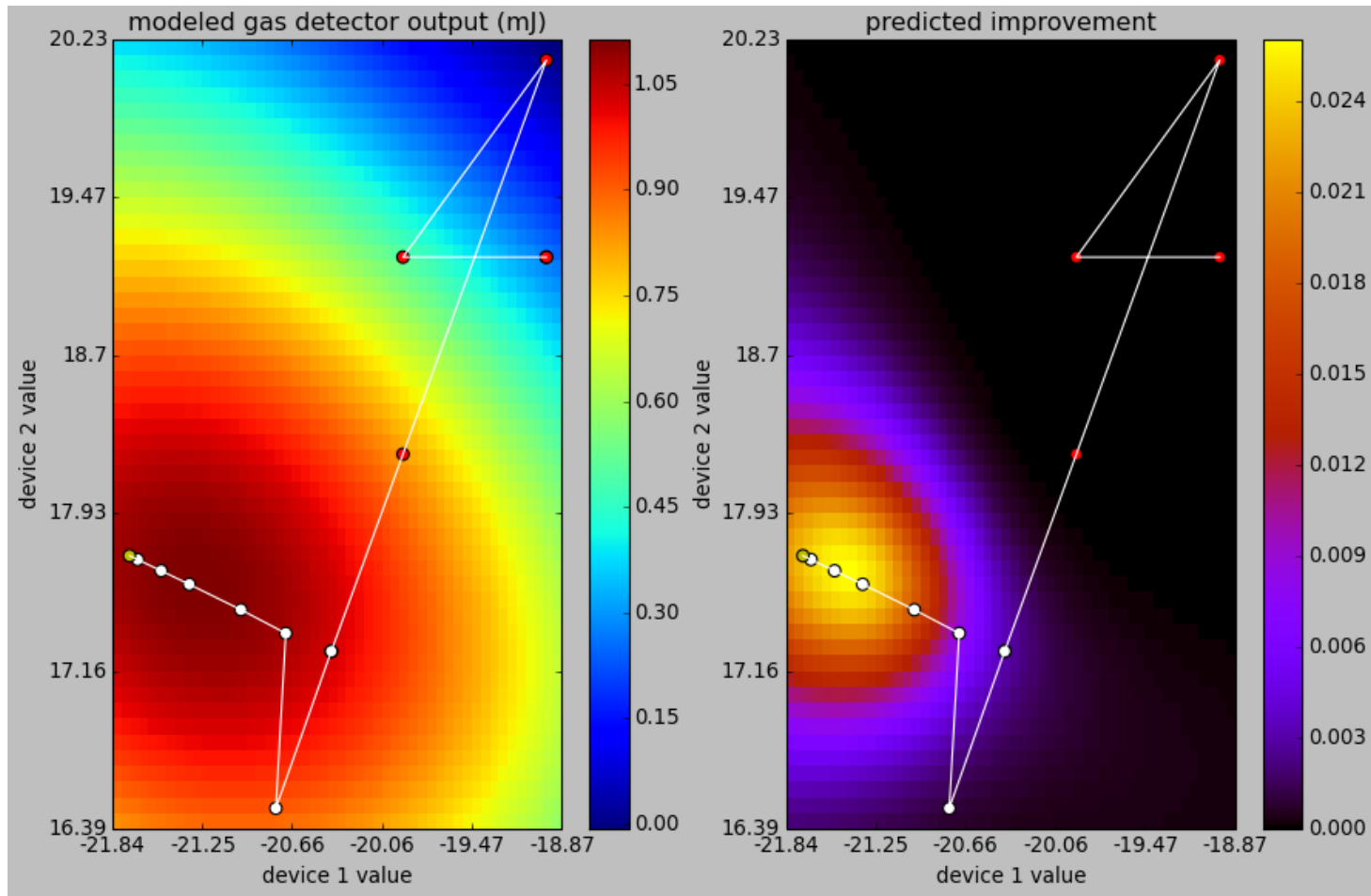
# Point 13



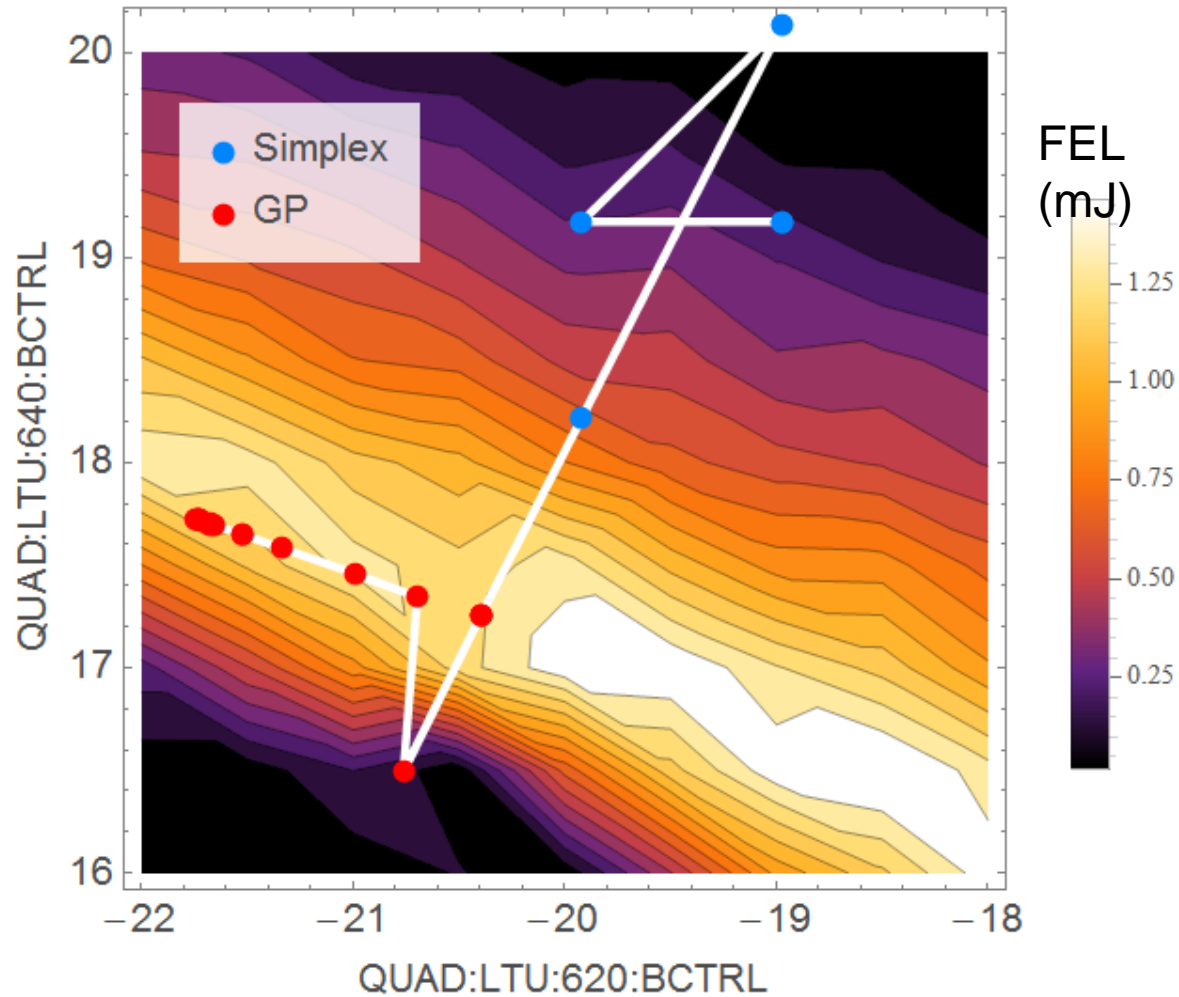
# Point 14



# Point 15



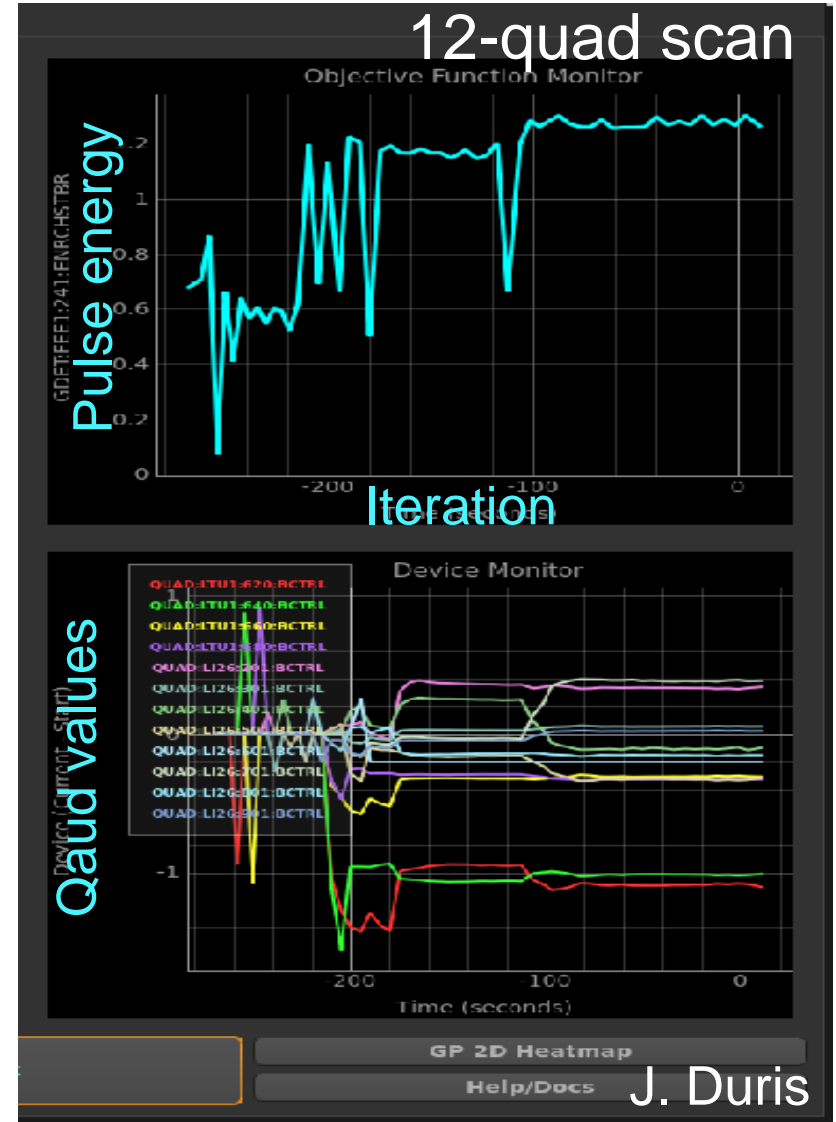
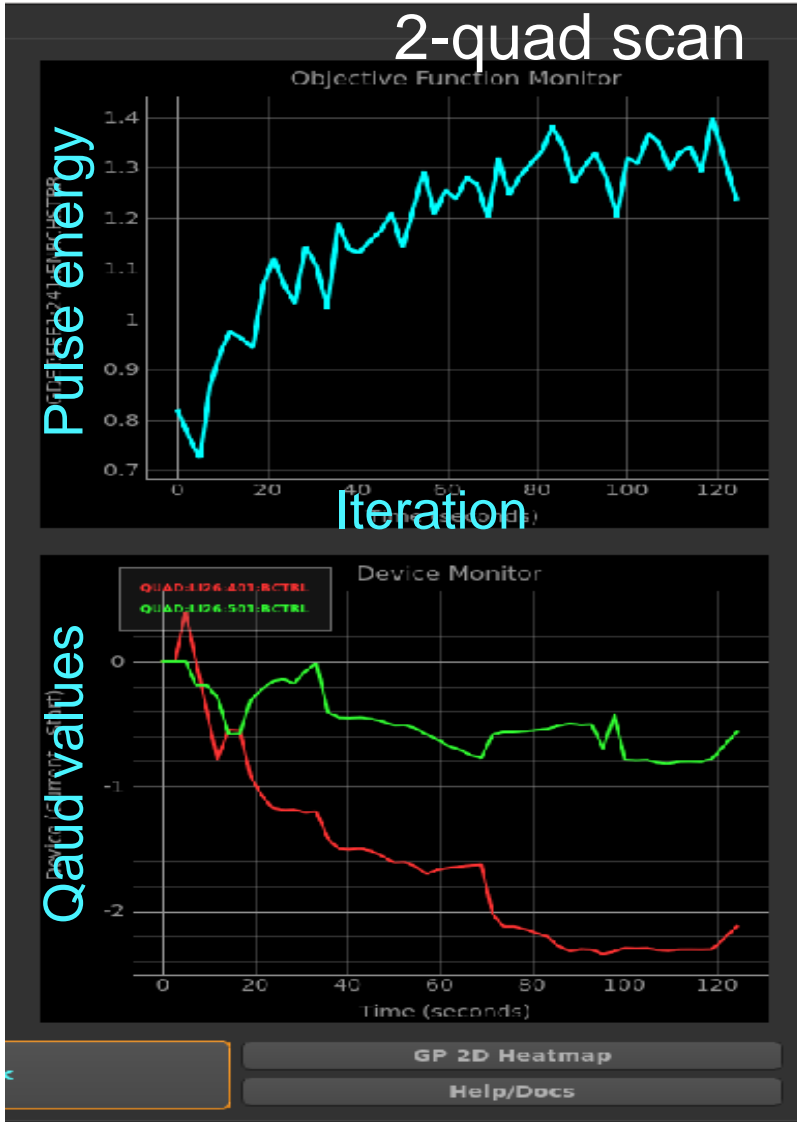
# 2-quad raster scan + Ocelot path



OcelotScan-2016-09-21-185122.mat

CorrelationPlot-QUAD\_LTU1\_620\_BCTRL-2016-09-21-185628.mat

# Recent results



# Machine Tuning Automation

Performance summary:  
Already as good as best human operators!  
Still many improvements to come...

|  | Hand tuning     | Simplex Ocelot   | GP Ocelot       |
|--|-----------------|------------------|-----------------|
| All runs                                     | 25 (30) $\mu$ J | 72 (50) $\mu$ J  | 95 (51) $\mu$ J |
| Infrequent tuning (>20 mins since last tune) | 39 (38) $\mu$ J | 128 (90) $\mu$ J | 38 (62) $\mu$ J |

~4x better than hand tuning

~3x better than hand tuning

Means (and medians in parentheses) FEL change during tune

# Machine Tuning Automation

**FY16 end** – Focused on completing most frequent tasks w/ fast ROI:

Time savings = Est. 103 min / wk

Goal = 210 min / wk

**To date 49% of goal**

(Evaluating actual integrated savings thru Dec.)

| Procedure                | Tune time (m) |       |
|--------------------------|---------------|-------|
|                          | Past          | Now   |
| Injector Tune            | 180           | < 120 |
| Global Steering          | 6             | < 1.5 |
| Und. Pointing            | 7             | 3     |
| Global Quad Optimization | 20            | 7     |

Man hours

Algorithms

**FY17 plan** – (LFD, AOSD, EED)

- Further code standardization
- Completion of more involved A.I.'s: *XTCAV, true emittance measure/model, E change management, still-faster inj. tuning*
- Plans for LCLS device integration
- Extend machine-agnostic code to add'l SLAC accelerators



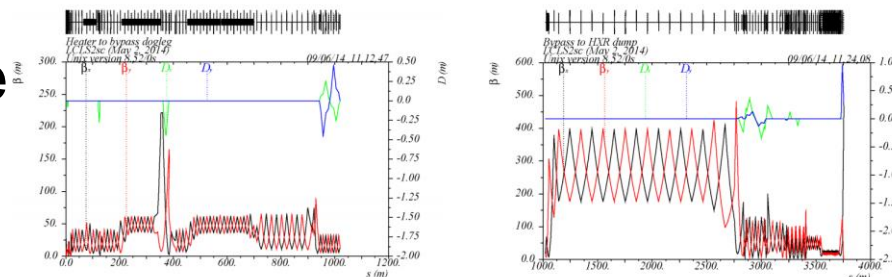
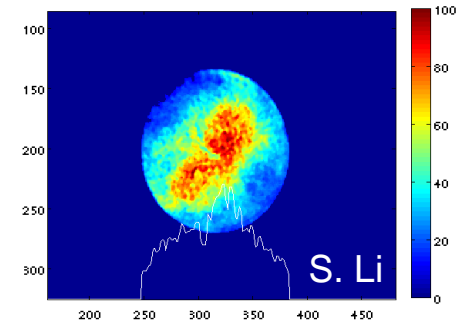
# Comparison of FEL changes for different tunes

## Future directions:

1. Use ground truth to fit hyper-parameters
2. Use archive/ground truth to introduce prior-mean
3. Expand to more complicated optimization problems (laser profiles, multi-objective functions, etc.)
4. Incorporate physical parameters into the model (i.e. fit physical models, not blind tuning parameters.)

$$k(x_1, x_2) = \theta e^{-(x_1 - x_2)^T \Lambda (x_1 - x_2)}$$

$$\begin{bmatrix} \mathbf{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K & K_*^T \\ K_* & K_{**} \end{bmatrix}\right)$$



# Comparison of FEL changes for different tunes

Future directions:

Hoping to develop international collaborations on shared online tuning algorithm for accelerators!

## Thanks for your attention!

Big thanks to people who did this work: A. Ahmed, T. Cope, J. Duris, S. Ermon, M. Gibbs, T. J. Lane, S. Li, T. Maxwell, M. McIntire, M. Mongia, N. Norvell, D. Sanzone, D. Schneider, C. Yoon