Machine Learning for an X-ray Laser

Oct. 21, 2018
SLAC National Accelerator Laboratory
Recent Machine Learning Highlights at LCLS

**Optimization**
- Model independent (simplex, gradient, RCDS, ES, etc.)
- Reinforcement learning
- Bayesian optimization

**Data analysis**
- Computer vision (conv-nets)
- Compressed sensing (convex opt.)

**Surrogate models**
- Generative adversarial nets

**Statistical analysis**
- True Wigner Matrix
- NN prediction

**Modeled FEL vs quads**

**Measured FEL vs quads**

**XTCAV**

**Computer vision**

**Inverse problems**

**Computer vision (conv-nets)**

**Compressed sensing (convex opt.)**
Optimization

Online tuning:
- Twice daily, ~500 of hours/year
- A single task, quadrupole tuning, required 1 hour/day
Treat optimization like a game: FEL power is the score

LCLS Experiment: 5.5 KeV Self-seeding FEL, Zig-zag doubles power from continuous profile

Juhao Wu
Optimization: Bayesian optimization

Model-based optimization

Advantage 1: Balance “exploitation vs. exploration”

⇒ Find global maximum

Gradient optimizer

Bayesian optimizer

Acquisition function

Ground truth

Posterior

Joe Duris
Optimization: Bayesian optimization

Model-based optimization

Advantage 2: Existence of model enables use of physics, archived data

Gaussian process: instance based learning method

Kernel function:

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

FEL vs quads

Joe Duris
Optimization: Bayesian optimization

Gaussian process: instance based learning method

Example: tuning quadrupoles from pure noise
Surrogate Models: Accelerator models

High fidelity physics simulations are remarkable:

LCLS microbunching instability

...but also slow. (e.g. hours on NERSC)

How can we best support design of a new machine?

LCLS-II simulations, Y. Ding
Surrogate Models: Accelerator models

- Predict XTCAV image from other diagnostic output or upstream machine settings to create a non-destructive virtual diagnostic
- Simulation + neural network results match well for FACET-II (see left)
- Small study with LCLS machine data and XTCAV images (scan of L1S phase and BC2 peak current at 13.4 GeV)

Emma, Edelen, et al. in preparation
Surrogate Models: FEL simulations

What happens if you turn around your trained network?

Deep dreaming of dogs

Generative adversarial network (GAN)

Style transfer

Simulation

Discriminator

Generator

Fake features

Fake label

Real

Features

Label

Genesis:
~1000 cpu-sec

GAN (neural net):
~0.001 gpu-sec

Gatys, et al.

X. Ren
Best measurements of X-ray beam come from electrons

Users require:

1. High resolution X-ray power profile shot-by-shot
2. Full (phase and amplitude) reconstruction of X-ray pulses

Current analysis algorithm has ~5 fs resolution

➔ Use computer vision to improve algorithm, speed up reconstruction

Combine spectra and high resolution power to reconstruct full FEL pulse

➔ Neural network speeds up to beam rate for users

A. Edelen, X. Zhang, X. Ren
Data Analysis: Pulse reconstruction

Computer vision

Classification + Localization

Object Detection

Aphex34 https://commons.wikimedia.org/w/index.php?curid=45679374

D. Mendez
Data Analysis: Pulse reconstruction

XTCAV Analysis

Before lasing

After lasing

XTCAV

Real and Predicted Power Profile

CNN prediction
Full beam reconstruction

- Combine time and frequency-domain measurements to retrieve field
- Iterative optimization methods used historically, but slow (minutes to hours)
- Neural network approach: do the optimization once (hours to days) then each example is fast (minutes)
Data Analysis: Statistical methods for data analysis

Ghost Imaging / Single Pixel Camera

Riddle: How can I take a picture with a spectrometer?

Answer: Have a friend with a flashlight
Data Analysis: Statistical methods for data analysis

Ghost Imaging / Single Pixel Camera

\[ x^* = \arg\min_x \left( \|Ax - B\|^2 + \lambda_2 \|x\|^2 + \lambda_1 \sum_j |x_j| \right) \]
subject to \( x_j \geq 0 \)

Images, \( I \)

Random patterns

\( B = [ 5037, 4783, 4891, 5940, \ldots ] \)

Target reconstruction

Target
Data Analysis: Statistical methods for data analysis

Example application: photocathode quantum efficiency

Don’t need DMD: exploit natural variation, jitter of drive laser
Conclusion

Summary:
X-ray FELs are complex, challenging machines. We need all the computational help we can get!

Applications include:
1. **Online tuning:** transverse matching, longitudinal phase space, X-ray spectrum, emittance minimization, etc.)
2. **Surrogate modeling:** efficient machine design, user support, predictive control
3. **Data analysis:** X-ray pulse reconstructions, electron parameters, user experiments
Thanks for your attention!

And thanks to the people who did the work:

Gaussian process: instance based learning method

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

\[
\begin{pmatrix}
\mathbf{y} \\
y_*
\end{pmatrix}
\sim \mathcal{N}
\begin{pmatrix}
\mathbf{0}, \\
K \\
K_* \\
K_{**}
\end{pmatrix}
\]

\[
K =
\begin{bmatrix}
k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\
k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\
\vdots & \vdots & \ddots & \vdots \\
k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n)
\end{bmatrix}
\]

\[
K_* =
\begin{bmatrix}
k(x_*, x_1) & \cdots & k(x_*, x_n)
\end{bmatrix}
\]

\[
K_{**} = k(x_*, x_*)
\]

taken from M. Ebner, GP for Regression
Gaussian Process Optimizer

Gaussian process: instance based learning method

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

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

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

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

taken from M. Ebner, GP for Regression
Gaussian Process Optimizer

Gaussian process: instance based learning method

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

\[
\begin{bmatrix}
  y \\
  y^*_n
\end{bmatrix} \sim \mathcal{N}
\begin{pmatrix}
  0, \\
  K \\
  K^*_n
\end{pmatrix}
\]

Acquisition function:
\[
UCB(x^*) = \mu(x^*) + \sqrt{\nu \tau(t)} \sigma(x^*)
\]
\[
\tau(t) = 2 \log(t^{d/2} + 2\pi^2 / 3\delta), \quad 0 < \delta < 1, \quad 0 < \nu
\]
Data Analysis: Statistical methods for data analysis

Experimental Results

Target

Transmission at camera

Ground truth

ADMM reconstruction

DMD

Cathode laser

Cathode

target

Siqi Li