Phase Space Reconstruction from Accelerator Beam Measurements Using Neural Networks and Differentiable Simulations

Ryan Roussel
rroussel@slac.stanford.edu
Manipulating Beams in Phase Space

How do we measure particle beam distributions in 6D phase space?

\[ \rho(x, p_x, y, p_y, z, \delta) \]
Determining the Beam Distribution in Phase Space

How do we determine the distribution $\rho(x, p_x, y, p_y, z, \delta)$?

**Tomographic methods**

- Gao Q. et al, PRAB 2018

**Projection methods**

- Penco. G. et al IPAC10, 2010
- Power. J. et al PAC07, 2007

**Specialized diagnostics**

- Gao Q. et al, PRAB 2018
Simple Tomography Analysis

\[ \sigma_x^2 = (1 + dlk)^2 \sigma_{11} + 2(1 + dlk)\sigma_{12} + d^2 \sigma_{22} \]

\[ \Sigma_x = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \quad \varepsilon = \sqrt{\det(\Sigma)} = \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \]

Rotate the beam at a diagnostic screen in phase space by scanning focusing strength and measure the beamsize.
Filtered Backprojection Tomography

Rotate phase space as before, but reconstruct the distribution from 1D projections

\[
\rho(x, p_x) = \int_0^\pi \lambda^\dagger_\phi(\xi) \, d\phi \mid_{\xi = x \cos \phi + p_x \sin \phi}
\]

Yakimenko V. et al PR STAB, 2003
Maximum Entropy Tomography (MENT)

Rotate phase space as before, but reconstruct the distribution from 1D projections + **maximize the beam distribution entropy**

\[
\rho^* = \arg \min \{-H(\rho) + \lambda f(\rho)\}
\]

Note: \( H \propto \log(\varepsilon) \)

Hock K. and Ibison M., JINST, 2013
Inferring Beam Distributions Using Optimization

Represent beam distribution with principal component analysis (PCA) and optimize weights to fit experimental measurements.
Tomography Challenges

- Algebraic methods don’t scale to higher dimensions (> 2D) without specialized procedures / diagnostics / assumptions
- Numerical reconstruction costs increase for high resolution reconstructions / higher dimensions
- Few algorithms account for particle-particle interactions (space charge) or collective effects (wakefields, synchrotron radiation)

Wolski, Andrzej, et al. arXiv, 2022
Hermann B. et al., PRAB, 2021
Wong J. et al., PRAB, 2022
Optimization Strategies for Inference

Can you calculate gradients easily?

- Yes
  - Gradient descent (SGD, Adam etc.)
    - Scales to >10k parameters (ML training)

- No
  - Black box optimization algorithms
    - Scales poorly with input dimension if not unimodal

- Analytical models
- Simulations
- Experiments
Optimization Strategies for Inference

Can you calculate gradients easily?

- Yes
  - Gradient descent (SGD, Adam etc.)
    - Scales to >10k parameters (ML training)
- No
  - Black box optimization algorithms
    - Scales poorly with input dimension if not unimodal

- Analytical models
  - Simulations

Go ahead, try it with simplex…
Differentiable Simulations

Keep track of derivative information during every calculation step.

Enables **gradient based optimization** of model error with respect to all free parameters using the chain rule.

Easily optimize models with >10k free parameters.
Differentiable Simulations

Keep track of derivative information during every calculation step.

Enables gradient based optimization of model error with respect to all free parameters using the chain rule.

Easily optimize models with >10k free parameters.
Want to parameterize 6D phase space distributions with a function that is **flexible** and learnable.

\[
X \sim \mathcal{N}(0, I)
\]

\[
g(x; \theta_t) : \mathbb{R}^6 \mapsto \mathbb{R}^6
\]

\[
Y = g(X; \theta_t)
\]

**Fully connected NN with \sim O(1k) parameters**
Phase Space Reconstruction Pipeline

Base Particle Distribution

$X \sim \mathcal{N}(0, I)$

Neural Network Parameterized Transform

$g(x; \theta_t) : \mathbb{R}^6 \rightarrow \mathbb{R}^6$

Proposed Initial Particle Distribution

$Y = g(X; \theta_t)$
Phase Space Reconstruction Pipeline

Base Particle Distribution

$X \sim \mathcal{N}(0, I)$

Neural Network Parameterized Transform

$g(x; \theta_t) : \mathbb{R}^6 \rightarrow \mathbb{R}^6$

Proposed Initial Particle Distribution

$Y = g(X; \theta_t)$

Differentiable Accelerator Simulations

$Z_n = f(Y; K_n)$

Simulated Screen Images

$Q_n^{(i,j)} = \text{KDE}(Z_n)$

Experimental Screen Images

$R_n^{(i,j)}$
Phase Space Reconstruction Pipeline

Base Particle Distribution

\[ X \sim \mathcal{N}(0, I) \]

Neural Network Parameterized Transform

\[ g(x; \theta_t) : \mathbb{R}^6 \rightarrow \mathbb{R}^6 \]

Proposed Initial Particle Distribution

\[ Y = g(X; \theta_t) \]

Differentiable Accelerator Simulations

\[ n = 2 \]

\[ n = 1 \]

\[ K_n \]

\[ Z_n = f(Y; K_n) \]

Simulated Screen Images

\[ Q_{n(i,j)}^{(i,j)} = \text{KDE}(Z_n) \]

Experimental Screen Images

\[ R_n^{(i,j)} \]

Gradient calculation

\[ \theta_t = \theta_{t-1} - h(\nabla_{\theta_t} l) \]

Optimization Step

Loss Function

\[ l = -\log \left[ (2\pi e)^{3/2} \varepsilon_{6D} \right] + \lambda \sum_{n,i,j} R_n^{(i,j)} \log \left( \frac{R_n^{(i,j)}}{Q_n^{(i,j)}} \right) \]

Initial Distribution Entropy

Image Divergence

Constraint Penalty
Phase Space Reconstruction Pipeline

Base Particle Distribution: $X \sim \mathcal{N}(0, I)$

Neural Network Parameterized Transform: $g(x; \theta_t): \mathbb{R}^6 \mapsto \mathbb{R}^6$

Proposed Initial Particle Distribution: $Y = g(X; \theta_t)$

Differentiable Accelerator Simulations:

$Z_n = f(Y; K_n)$

$K_n$

$Q_n^{(i,j)} = \text{KDE}(Z_n)$

Simulated Screen Images

$n = 2$

$n = 1$

Experimental Screen Images

$n = 2$

$n = 1$

Reconstructed Initial Distribution: $Y^* = g(X; \theta^*)$

$\theta^* = \arg \min_{\theta} l$

Optimization Step: $\theta_t = \theta_{t-1} - h(\nabla_{\theta} l)$

Gradient calculation

Loss Function:

$l = -\log \left( \frac{(2\pi e)^3}{6D} \right) + \lambda \sum_{n,i,j} R_n^{(i,j)} \log \left( \frac{R_n^{(i,j)}}{Q_n^{(i,j)}} \right)$

Initial Distribution Entropy

Image Divergence

Constraint Penalty
Maximum Entropy Loss Function

\[ l = - \log \left( (2\pi e)^3 \varepsilon_{6D} \right) + \lambda \sum_{n,i,j} R^{(i,j)}_n \left| \log \left( \frac{R^{(i,j)}_n}{Q^{(i,j)}_n} \right) \right| \]

Initial Distribution Entropy

Image Divergence Constraint Penalty

No evidence

Weak evidence

Strong evidence
Synthetic Example

Synthetic beam distribution in simulation

Screen images

\[ K \]

\[ Z = f(Y; K) \]
Synthetic Example Reconstruction
Measuring Model Uncertainty

Create a snapshot ensemble to measure uncertainty by cycling the learning rate

Huang G. et al., ICLR 2017
Measuring Model Uncertainty and Convergence

Loss Function

\[ l = -\log \left[ \frac{1}{2\pi e} \varepsilon_{6D} \right] + \lambda \sum_{n,i,j} R_{n}^{(i,j)} \log \left( \frac{R_{n}^{(i,j)}}{Q_{n}^{(i,j)}} \right) \]

- Initial Distribution Entropy
- Image Divergence
- Constraint Penalty

No information

Some information

Lots of information
Measuring Model Uncertainty and Convergence

Ground truth

Reconstruction
Tomography Example from AWA

Drive beamline

- Photoinjector
- Linac
- Focusing Quadrupole
- Screen
- Camera

Images showing various tomography data with different $k$ values:
- $k = -3.4 \text{ m}^{-2}$
- $k = -2.5 \text{ m}^{-2}$
- $k = -1.5 \text{ m}^{-2}$
- $k = -0.49 \text{ m}^{-2}$
- $k = 0.49 \text{ m}^{-2}$
- $k = 1.5 \text{ m}^{-2}$
- $k = 2.5 \text{ m}^{-2}$
AWA Reconstruction Results

Red border denotes test samples
AWA Reconstruction Results

(a) Graph showing the relationship between $\sigma^2$ and $k$ with various markers for $\sigma^2_{x,\text{real}}$, $\sigma^2_{y,\text{real}}$, $\sigma^2_{x,\text{pred}}$, and $\sigma^2_{y,\text{pred}}$. Inset images illustrate experimental (exp.) and predicted (pred.) results.

(b) - (g) Images depicting $x$, $y$, $p_x$, and $p_y$ measurements for different cases.
Conclusions

- We can create detailed reconstructions of beam phase spaces from simple tomographic accelerator measurements without special diagnostics.
- Reconstructions from differentiable simulations are not limited by analytical tractability, number of free parameters.
- Theoretically we are only limited by model detail and accuracy, need further investment in differentiable simulations.
- Need to expand our idea of what can be used as a diagnostic.

Details [https://arxiv.org/abs/2209.04505](https://arxiv.org/abs/2209.04505)
Thanks!

SLAC
- Auralee Edelen
- Chris Mayes
- Daniel Ratner

UChicago
- Juan Pablo Gonzalez-Aguilera

Argonne Wakefield Accelerator
- Seongyeol Kim
- John Power
- Eric Wisniewski

Questions?
Other Applications of Differentiable Simulations

High dimensional optimization + model calibration

Hysteresis modeling + ML models

JP Gonzalez-Aguilera, UChicago

Hessian matrix determination

Fitting quadrupole offsets from beam measurements

Roussel, R., et al. PRL (2022)