Transverse 2D phase-space tomography using Beam Position Monitor data of kicked beams

Kilean Hwang, Chad Mitchell, Robert Ryne

Lawrence Berkeley National Laboratory

May 2021

Abstract

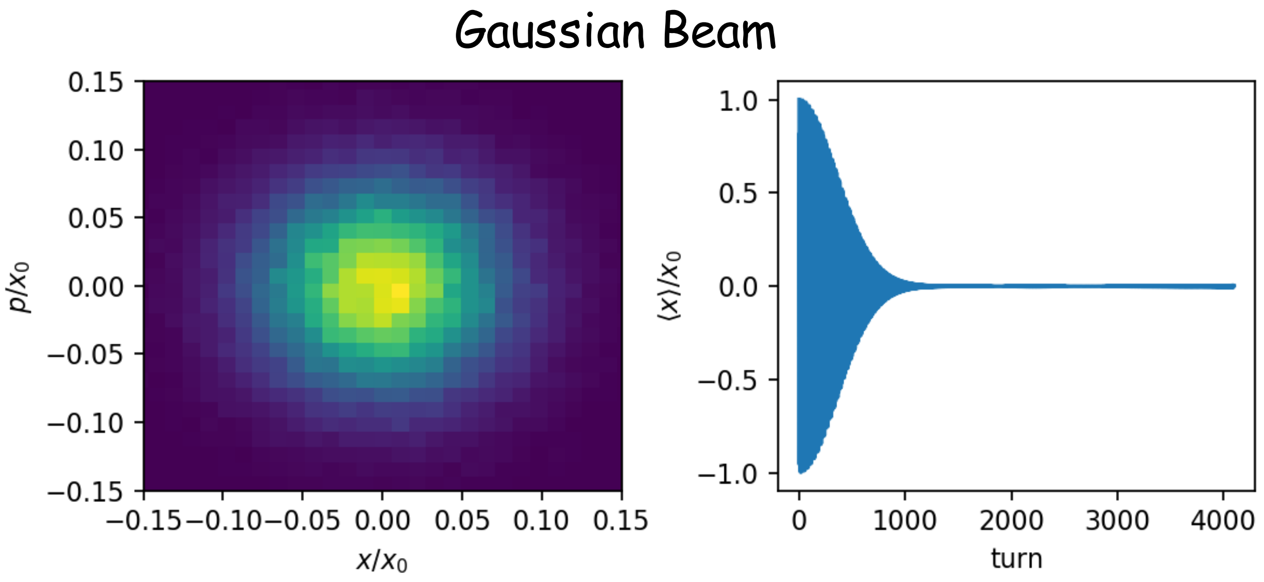
The time-series Beam Position Monitor (BPM) data of kicked beam is a function of lattice parameters and beam parameters including phase-space density. The decoherence model using the first-order detuning parameter has an exact solution when the beam is Gaussian. We parameterize the beam phase-space density by multiple Gaussian kernels of different weights, means, and sizes to formulate the inverse problem for 2D phase-space tomography. Numerical optimization and Bayesian inference are used to infer the beam density and uncertainty.

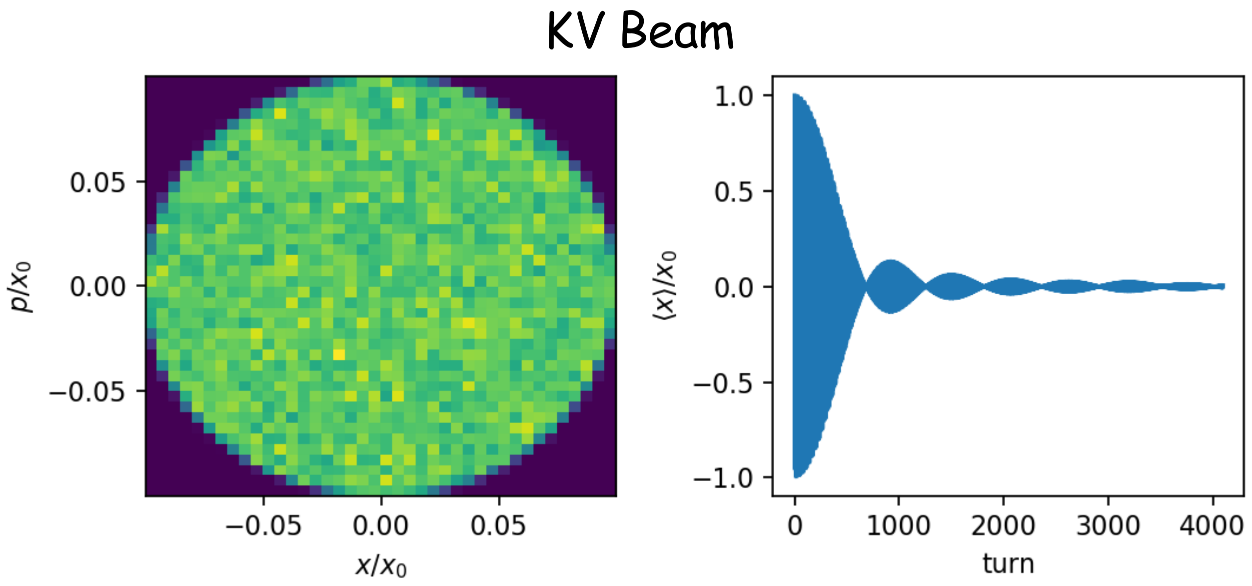
Overview

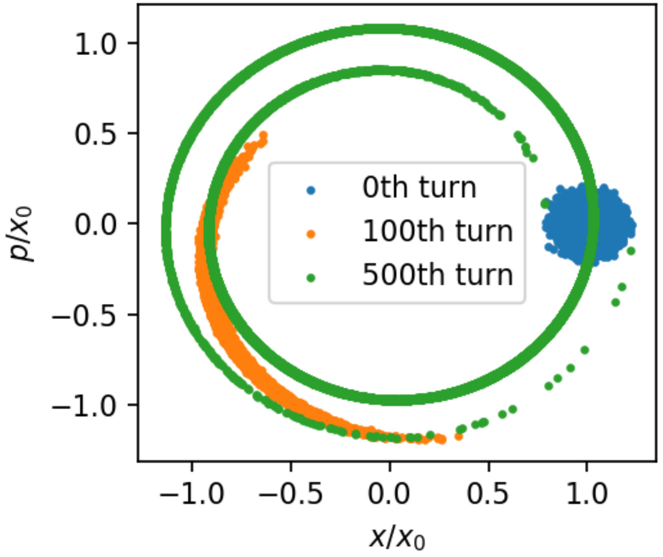
- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

Introduction

The BPM data of a kicked beam can be modeled as a function of the linear and nonlinear optics parameters, and the beam phase-space density







Decoherence BPM data model

We model BPM data by

$$\begin{aligned} \langle X \rangle_t &= \Re \langle X - iP \rangle_t \\ &= \Re \int (X - iP)\, e^{i\omega t} \rho_{X,P}(X - X_0, P - P_0)\, dX dP \end{aligned}$$

Assuming slowly varying frequency in the scale of the beam size, the frequency is modeled by

$$\omega(\Delta I) = \mu_0 + \mu_1 \Delta I$$

Exact solution for Gaussian beam

Under the slowly varying frequency assumption: $\omega(\Delta I) = \mu_0 + \mu_1 \Delta I$,

an exact solution for Gaussian beam exists

$$\langle X \rangle_t = \frac{X_0\left(1-\tau^2\right)+2P_0\tau}{\left(1+\tau^2\right)^2}\exp\left[-\frac{I_0}{\epsilon}\frac{\tau^2}{1+\tau^2}\right]\cos\Psi - \frac{2X_0\tau - P_0\left(1-\tau^2\right)}{\left(1+\tau^2\right)^2}\exp\left[-\frac{I_0}{\epsilon}\frac{\tau^2}{1+\tau^2}\right]\sin\Psi$$

where $\tau \equiv \epsilon\mu_1 t$ and $\Psi \equiv \mu_0 t - \frac{(I_0/\epsilon)\,\tau^3}{1+\tau^2}$

General distribution

For general distribution, one need an approximation to extract meaningful expression for the centroid decoherence motion. As the centroid decoherence is due to the phase-mixing, it is also advantageous to work on frequency domain. Define DFT like function

$$G(k) = \frac{2}{\sqrt{2\beta I_0}} \sum_{t=0}^{T} \langle X \rangle_t \left[e^{-ikt}\right] - \cos\theta$$

where θ is the phase-space angle of the initial kick such that

$$X_0 - iP_0 = \sqrt{2I_0\beta} \exp\left(-i\theta\right)$$

Marginalized probability density

One can show that (in the limit of large initial offset compared to the initial beam emittance: i.e. $I_0 \gg \epsilon$) the marginalized 1D (in the direction of initial offset) probability density of the 2D probability density of phase-space can be written in terms of the DFT like function of the centroid data

$$\rho_\theta(x) \quad \simeq \quad \sqrt{2\beta I_0}\frac{|\mu_1|}{\pi}\Re\left[G\left(x\mu_1\sqrt{2\beta I_0}+\mu_0\right)e^{i\theta}\right]$$

This suggest that if we have multiple kicks of different angles θ, we can reconstruct the 2D phase-space.

Gaussian Kernel Density Model

However, the large initial offset ($I_0 \gg \epsilon$) requirement can be tight due to the physical beam pipe aperture.

Recalling that we have exact solution for gaussian beam, we parameterize the 2D beam density model using multiple Gaussian kernels. Each gaussian kernels have 4 parameters: (1) weight, (2,3) relative locations *(x, p)* from the beam center, (4) emittance.

$$\rho_{X,P}(X,P) = \sum K_i(X,P) \qquad K_i(X,P) \equiv w_i \frac{1}{2\pi\epsilon_i} e^{-\frac{(X-X_i)^2+(P-P_i)^2}{2\epsilon_i}}$$

Then, the centroid data of the whole beam is the linear sum of each gaussian kernel contributions. This translates our problem as parameter fitting on the inverse problem.

Curse of Dimensionality

However, the large number of gaussian kernels for good resolution poses the task of the global parameter fitting of the inverse problem very difficult due to the curse of dimensionality.

Bayesian Approach (Brief Sketch)

Our solution is to use the Bayesian approach with the prior from the parameter fitting on the leading order model:

$$\rho_\theta(x) \simeq \sqrt{2\beta I_0}\frac{|\mu_1|}{\pi}\Re\left[G\left(x\mu_1\sqrt{2\beta I_0}+\mu_0\right)e^{i\theta}\right]$$

The leading order model helps us to initialize and fit the parameters approximately close to the true solution. Once we fit the parameters on the leading order model, we can build posterior using

$$\langle X\rangle_t = \Re\int (X-iP)\,e^{i\omega t}\rho_{X,P}(X-X_0,P-P_0)\,dXdP$$

Then perform the local optimization to maximize posterior.

Overview

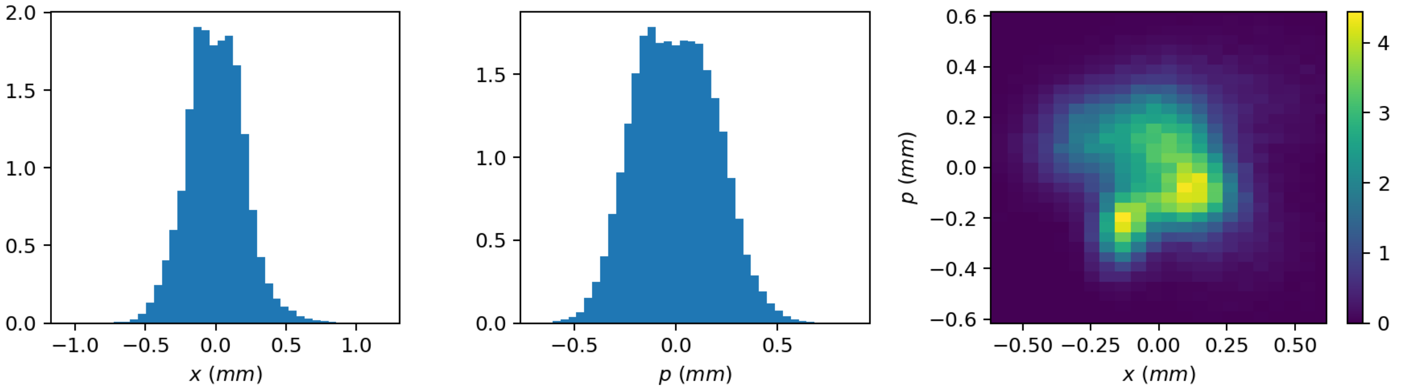
- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

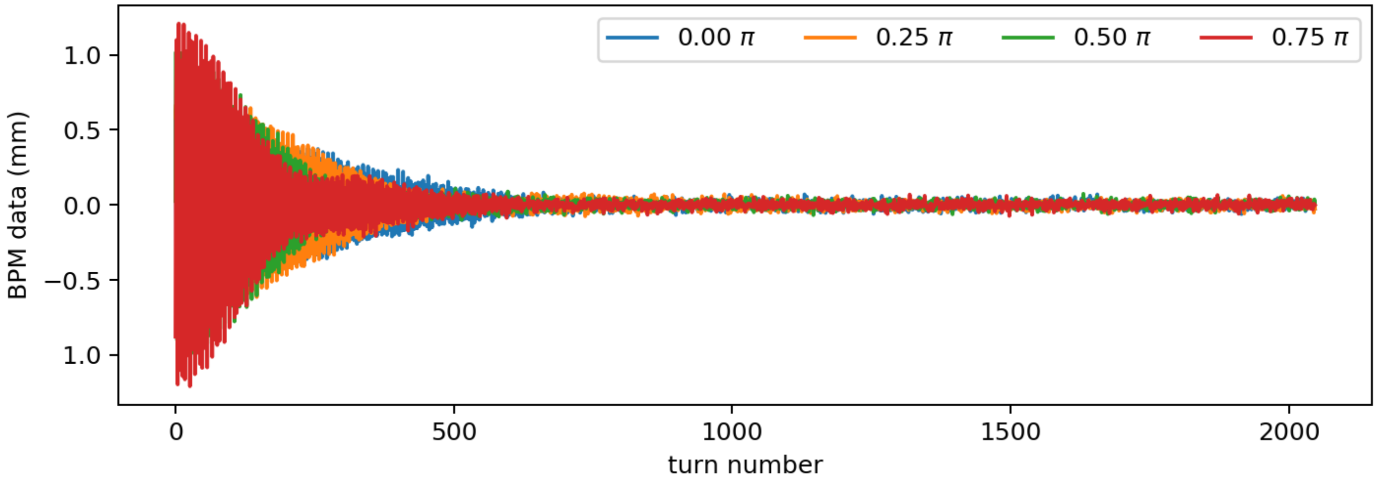
Virtual beam centroid signals

We randomly generate beam density and prepare 4 virtual BPM data using 4 different initial kicks using the following frequency model:

$$\omega(I) \quad = \quad \omega_0 + \omega_1 I + \omega_2 \frac{I^2}{2}$$

The initial beam emittance is 2 nm. The 4 kick strengths are 3,4,5,6 times of the beam emittance and the kick angles equally spaced from 0 to π. We also added virtual noise of RMS size 20 μm





Overview

- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

Construct prior mean: (1) frequencies

Recall

$$\rho_\theta(x) \simeq \sqrt{2\beta I_0}\frac{|\mu_1|}{\pi}\Re\left[G\left(x\mu_1\sqrt{2\beta I_0}+\mu_0\right)e^{i\theta}\right]$$

and notice that the frequency parameter μ_0 is not coupled with any other parameters.

This allow us to optimize the frequencies using the following condition (on the 4 BPM data)

$$\int_{-\infty}^{\infty} x\rho_\theta(x)\,dx = 0$$

Construct prior mean: (2) initial kick strengths and angles

Once we optimize the frequencies μ_0, we further optimize the nonlinear detuning parameter, initial kick strengths, angles and betatron function $\mu_1, I_0, \theta, \beta$ using the following condition.

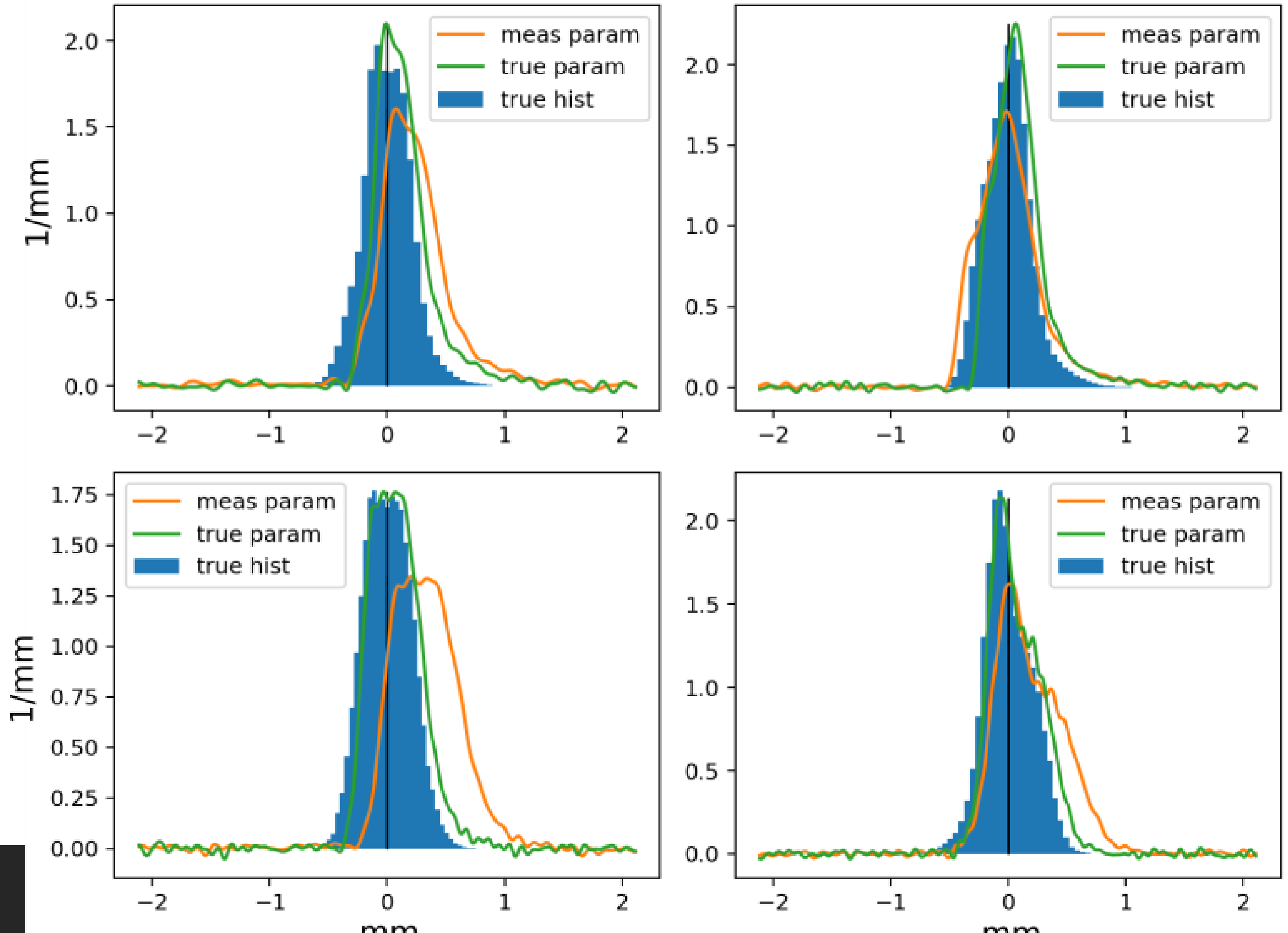
$$\int_{-\infty}^{\infty} \rho_\theta(x)\, dx = 1$$

Here the nonlinear detuning parameter μ_1 is determined from least square fit of the frequencies μ_0 over the initial kicks strengths I_0. Although we have less number of constraints compared to the parameters to fit, we experienced that when the initial guess is close to the ground truth, the local optimization often works.

Construct prior mean

DFT like function using roughly estimated parameters:

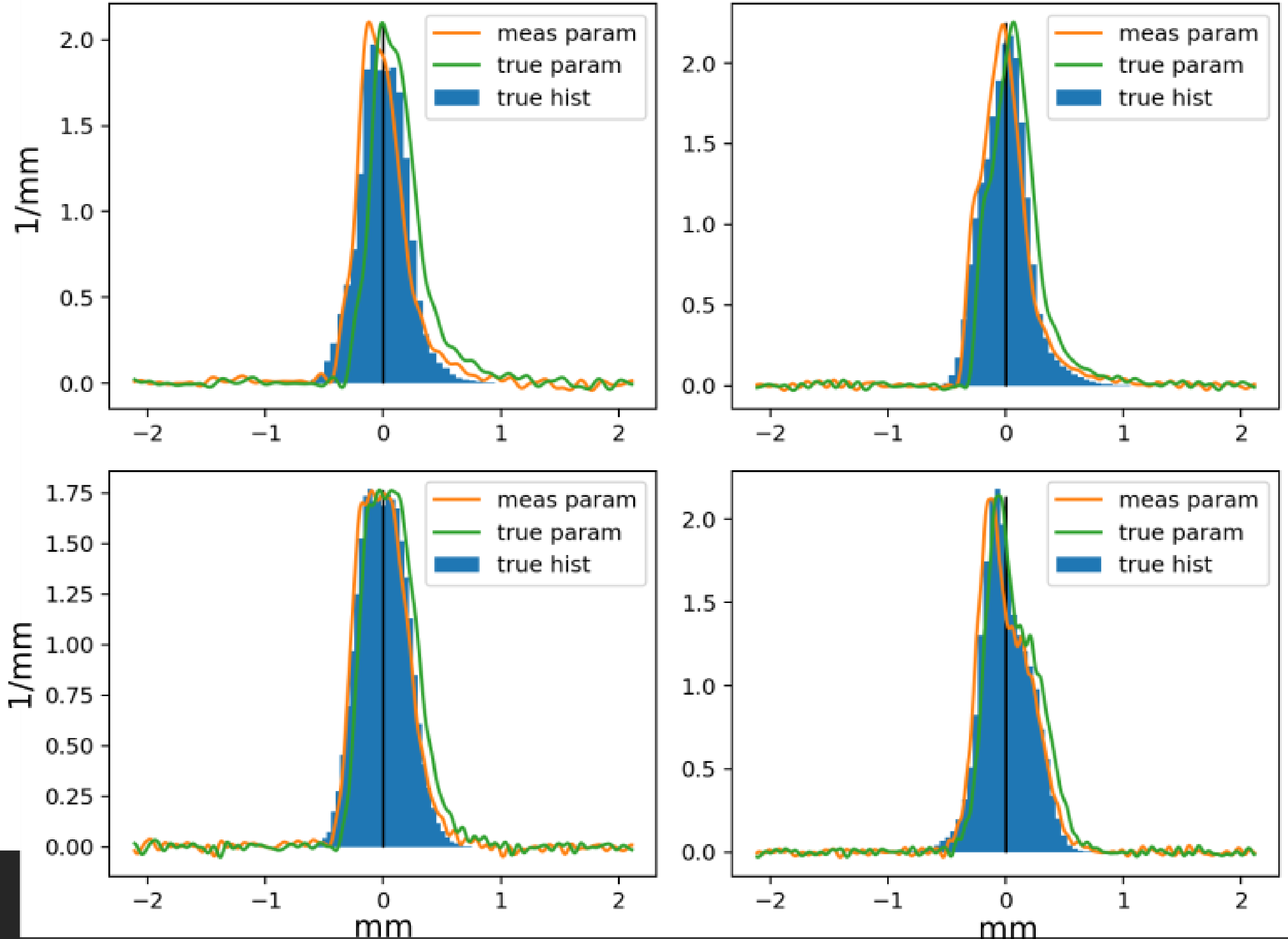
$$\mu_0,\ \mu_1,\ I_0,\ \theta,\ \beta$$



Construct prior mean

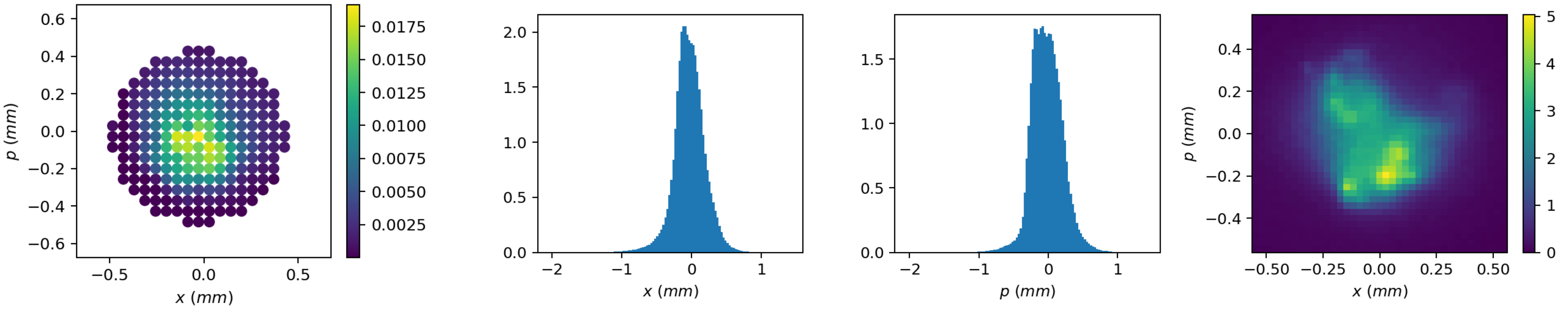
DFT like function after optimization of the following parameters:

$\mu_0,\ \mu_1,\ I_0,\ \theta,\ \beta$



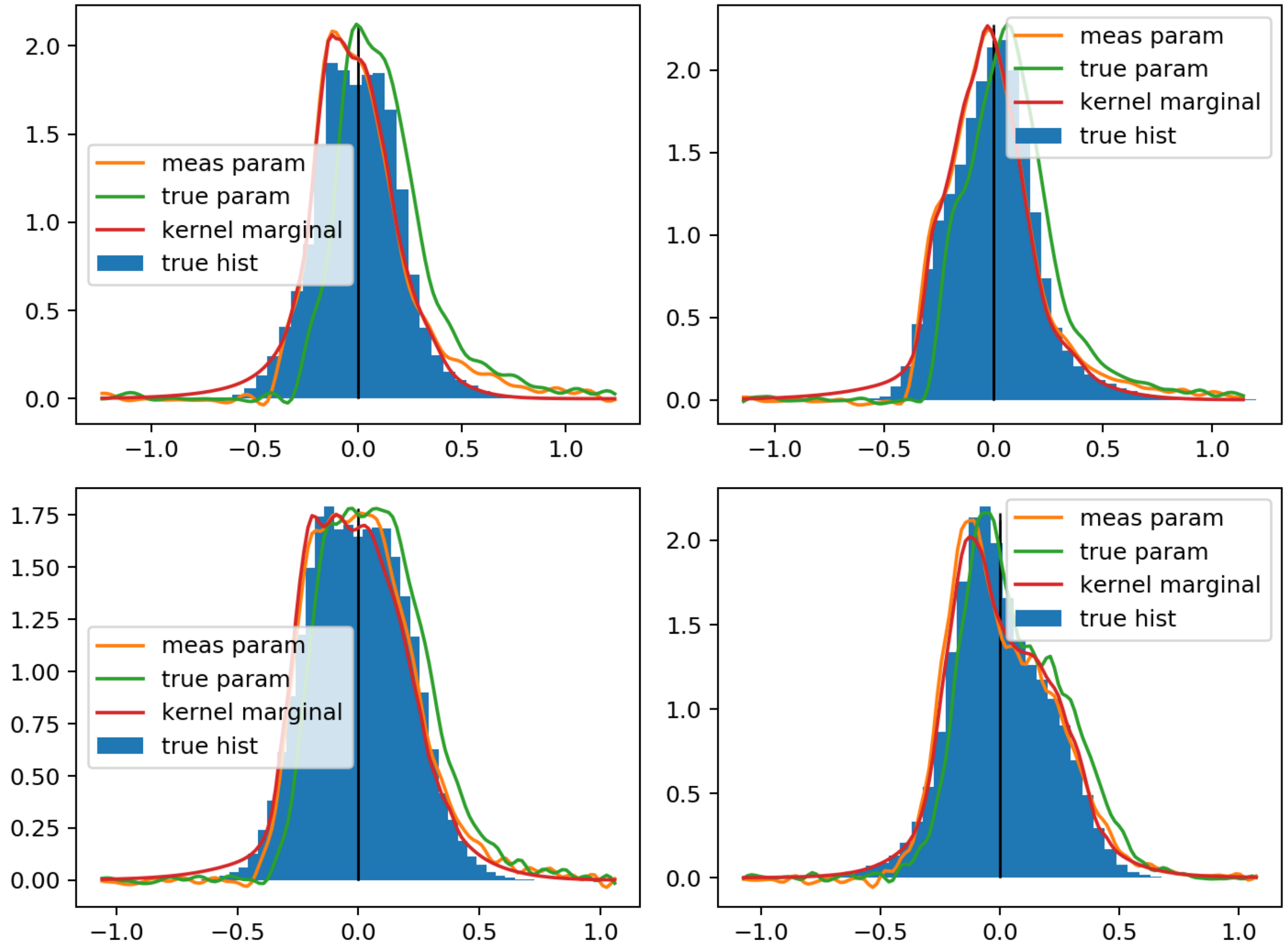
Construct prior mean: (3) Gaussian kernels

Once we have fixed $\mu_0,\ \mu_1,\ I_0,\ \theta,\ \beta$ based on the leading order theory, we continue to fit the parameters gaussian kernels on the leading order theory.



Construct prior mean: (3) Gaussian kernels

DFT like function using gaussian kernels well fitted on the DFT like function using the virtual BPM data



Construct prior

Once we have prior mean, we construct the prior using normal distribution. We use independent normal of the parameters $\mu_0, \mu_1, I_0, \theta, \beta$ with the standard deviation from our belief. For example, since we are building prior from the leading order theory that is in the limit of $I_0 \gg \epsilon$, we choose smaller standard deviation for smaller ϵ/I_0. In other words, we have more confidence on our belief for larger initial kick.

As for the gaussian kernel parameters, we do not construct prior as it is hard to choose the standard deviation in a reasonable away

Overview

- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

Construct posterior mean

Now we construct posterior mean by model fitting on

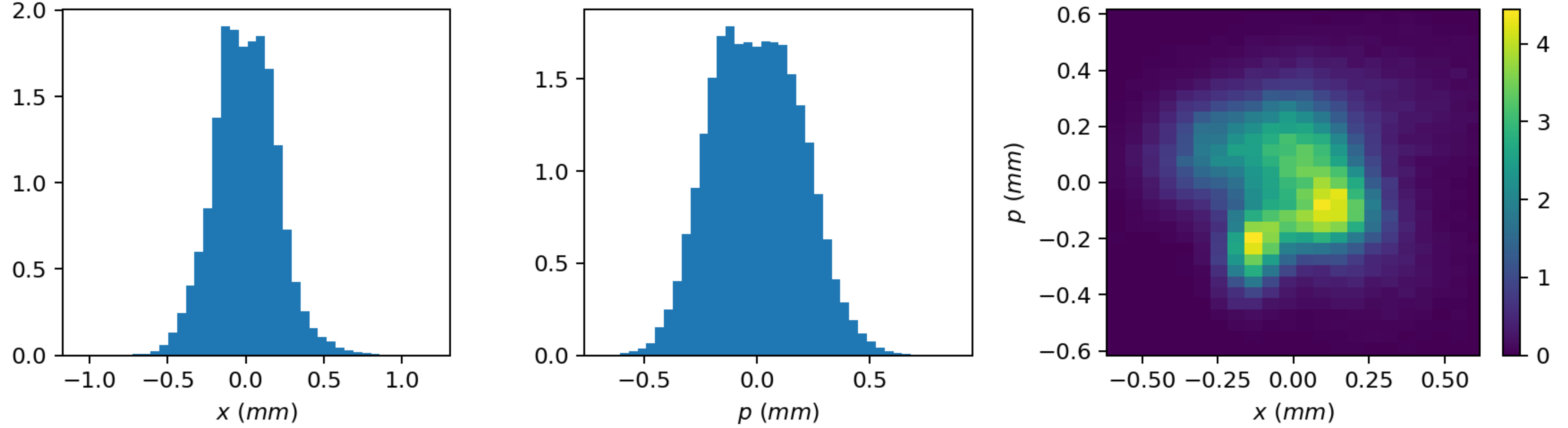
$$\langle X \rangle_t = \Re \int (X - iP)\, e^{i\omega t} \rho_{X,P} (X - X_0, P - P_0)\, dX dP$$

Specifically, we construct our likelihood by

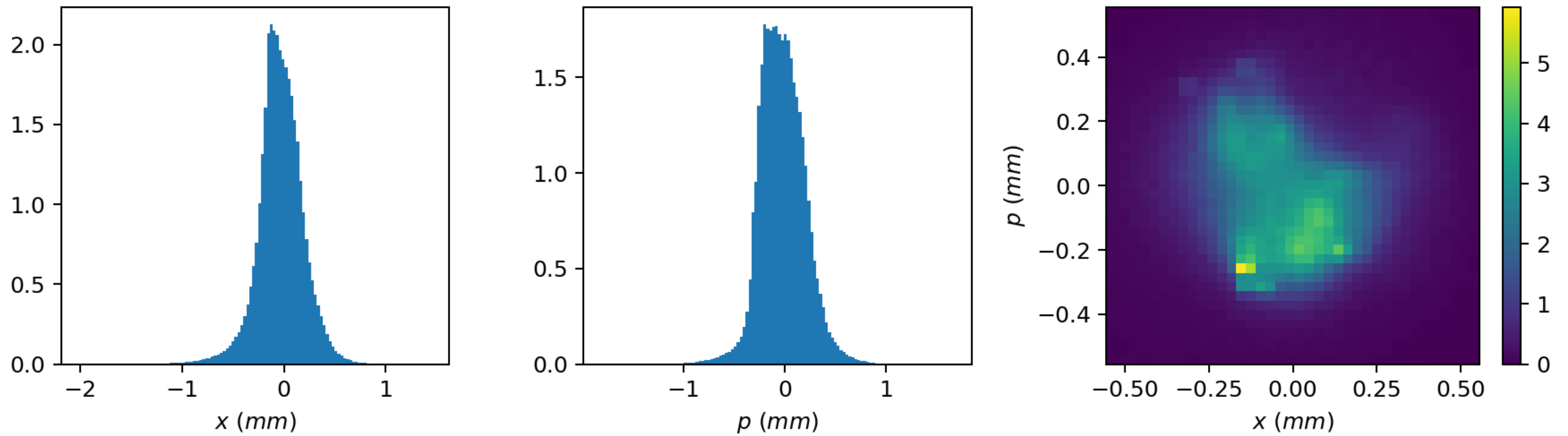
$$P\left(\xi | \langle X \rangle_{t,BPM}\right) = \Pi_{k=1}^{K} \frac{1}{\sqrt{2\pi}\sigma_{BPM}} \exp\left(-\sum_{t=0}^{T-1} \frac{\left(\langle X \rangle_{t,BPM_k} - \langle X \rangle_{t,model_k}\right)^2}{2\sigma_{BPM}^2 T}\right)$$

Where k is the index for each kick and ξ represent the model parameters including the parameter quantifying model error and data noise σ_{BPM} .

Ground truth



maximum a posteriori

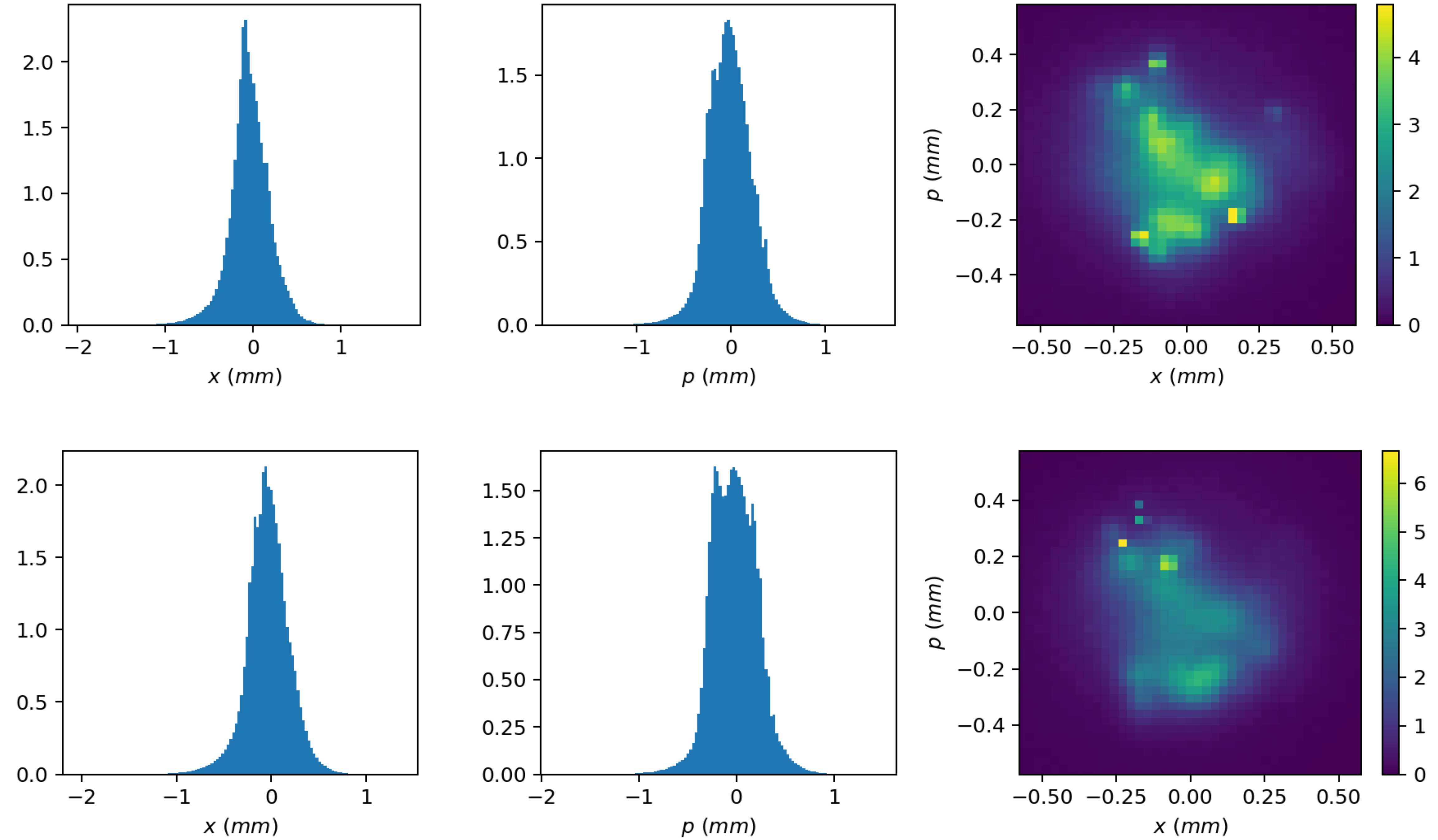


Overview

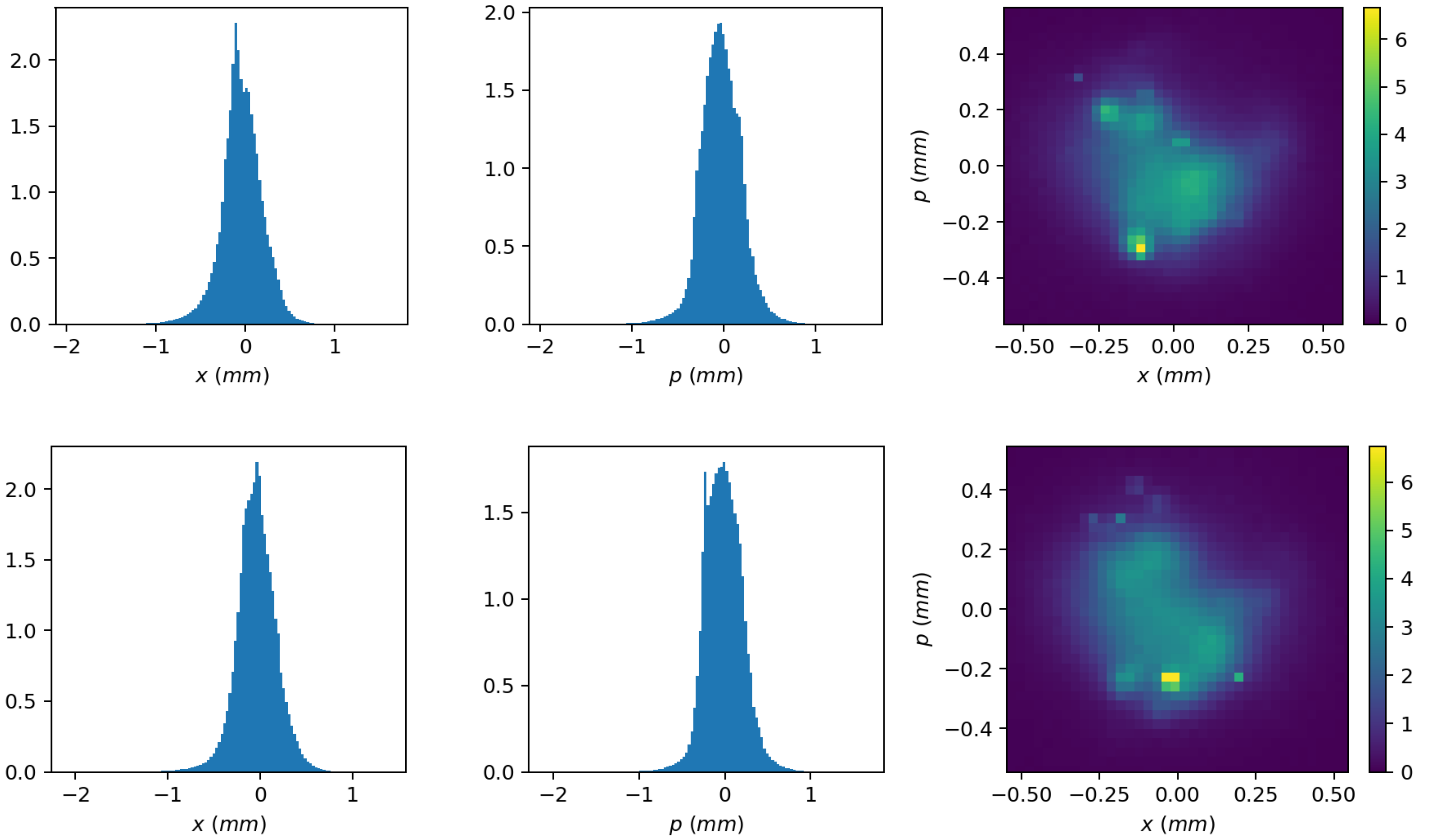
- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

Samples within credential interval

So far we have done point estimate. Since our prior and likelihood are modeled by gaussian distribution, the posterior is also gaussian. This helps us to sample from posterior with known credential level without relying on MCMC (Markov chain Monte Carlo) that can be very computationally heavy for convergency with so many parameters. Here are few samples within 95% confidence level.



x (mm)
p (mm)
x (mm)
p (mm)
x (mm)
p (mm)
x (mm)
p (mm)



x (mm)
p (mm)
p (mm)
x (mm)
x (mm)
p (mm)
p (mm)
x (mm)

Overview

- Problem and strategy overview
- Detail procedure for the proof of concept
  - Generate virtual beam centroid signal with multiple initial kicks
  - Construct prior
  - Construct posterior
  - Uncertainty
- Conclusion and Remarks

Conclusions and Remarks

We performed proof of principle of the 2D phase-space tomography using beam centroid data of multiple kicked beam. For better resolution, we need more kicks in different angles.

The Bayesian approach helped us to avoid curse-of-dimensionality problem through prior belief construction that could be done using the leading order theory and local minimization. Then the maximum a posterior estimation also could be done using local minimization. It also helped us to sample from posterior (to visualize uncertainty) without relying on MCMC that is practically impossible with so many parameters to infer.

Next, we will work on real BPM data