

TRANSVERSE 2D PHASE-SPACE TOMOGRAPHY USING BEAM POSITION MONITOR DATA OF KICKED BEAMS*

Kilean Hwang[†], Chad Mitchell, Robert Ryne
Lawrence Berkeley National Laboratory, Berkeley, USA

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.

INTRODUCTION

The BPM data of a kicked beam is a function of the linear and nonlinear optics parameters [1], and the beam phase-space density. Ignoring nonlinear normalizing map, we model BPM data by

$$\begin{aligned} \langle X \rangle_t &= \Re \langle X - iP \rangle_t \\ &= \Re \int (X - iP) e^{i\omega t} \rho_{X,P}(X - X_0, P - P_0) dXdP, \end{aligned} \quad (1)$$

where X and P are normal conjugate variables

$$\begin{aligned} X(t) &= \sqrt{2\beta I} \cos(\omega t + \phi), \\ P(t) &= -\sqrt{2\beta I} \sin(\omega t + \phi), \end{aligned}$$

and X_0 and P_0 are initial kicks (or offsets). Figures 1 and 2 illustrate the sensitivity of the BPM data on initial beam phase-space density.

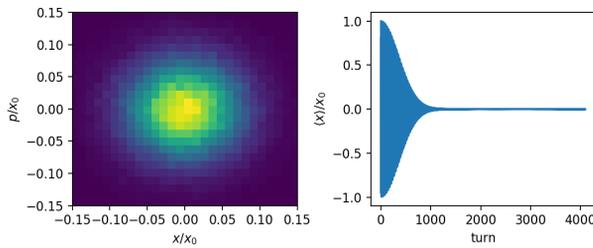


Figure 1: Kicked beam centroid data of initial Gaussian beam.

It is convenient to introduce θ such that

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

where $I_0 = (X_0^2 + P_0^2)/2$ is the action corresponding to the initial beam centroid offset. Assuming slowly varying

* Work supported by the Director of the Office of Science of the US Department of Energy under Contract no. DE-AC02-05CH11231

[†] kilean@lbl.gov

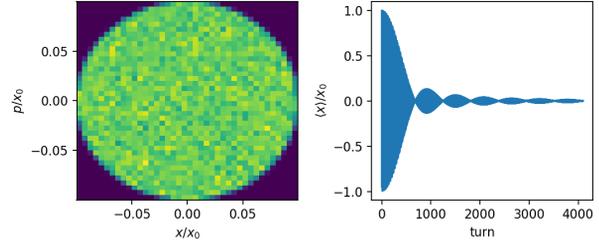


Figure 2: Kicked beam centroid data of initial uniform disk beam.

frequency in the scale of the beam size, the frequency can be modeled by

$$\omega(\Delta I) = \mu_0 + \mu_1 \Delta I, \quad (3)$$

where $\Delta I \equiv I - I_0$.

For Gaussian initial beam distribution, an exact solution exists:

$$\begin{aligned} \langle X \rangle_t &= \frac{X_0(1 - \tau^2) + 2P_0\tau}{(1 + \tau^2)^2} \exp\left[-\frac{I_0}{\epsilon} \frac{\tau^2}{1 + \tau^2}\right] \cos \Psi \\ &\quad - \frac{2X_0\tau - P_0(1 - \tau^2)}{(1 + \tau^2)^2} \exp\left[-\frac{I_0}{\epsilon} \frac{\tau^2}{1 + \tau^2}\right] \sin \Psi, \end{aligned} \quad (4)$$

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

2D PHASE-SPACE TOMOGRAPHY STRATEGY

Marginal Distribution

For general initial beam distribution, one needs 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 the frequency domain. We define the following function in the frequency domain:

$$G(k) = \frac{2}{\sqrt{2\beta I_0}} \sum_{t=0}^T \langle X \rangle_t [e^{-ikt}] - \cos \theta. \quad (5)$$

In the limit of large initial offset compared to the initial beam size $I_0 \gg \epsilon$, it can be shown that [2]

$$\rho_\theta(x) \approx \sqrt{2\beta I_0} \frac{|\mu_1|}{\pi} \Re \left[G\left(x\mu_1\sqrt{2\beta I_0} + \mu_0\right) e^{i\theta} \right]. \quad (6)$$

This suggests that if we have multiple kicks of different angles θ , we can reconstruct the 2D phase-space. And more kicks of different angles increase angular resolution.

Gaussian Kernel Density Model

However, the large initial offset $I_0 \gg \epsilon$ condition can be tight due to the physical beam pipe aperture. Recalling that we have the exact solution for the Gaussian beam, we parameterize the 2D beam density model using multiple Gaussian kernels. Each Gaussian kernels have weight, relative locations (x, p) from the beam center, and emittance parameters.

$$\rho_{X,P}(X, P) = \sum_i K_i(X, P), \quad (7)$$

$$K_i(X, P) \equiv w_i \frac{1}{2\pi\epsilon_i} e^{-\frac{(X-X_i)^2 + (P-P_i)^2}{2\epsilon_i}}. \quad (8)$$

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.

Bayesian Approach

However, a 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. We solve this problem using the Bayesian approach [3,4] with the prior mean from the parameter fitting on the leading order model described in Eq. (6). The leading order model helps us to initialize and fit the parameters approximately close to the true solution as long as the initial kicks are larger than the initial emittance. Once we fit the parameters on the leading order model, we can build the posterior using the following likelihood model:

$$P(\xi | \langle X \rangle_{t,BPM}) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi}\sigma_{BPM}} \exp\left(-\sum_{i=0}^T \frac{(\langle X \rangle_{t,BPM_k} - \langle X \rangle_{t,model_k})^2}{2\sigma_{BPM}^2(T+1)}\right), \quad (9)$$

where k is the index for each kick, and ξ is the set of model parameters including σ_{BPM} which is a parameter quantifying model error and data noise. The prior serves not only as a regularisation for the optimization but also an initial starting point for the maximum a posteriori (MAP) estimation using a local optimization algorithm. It also serves the optimization stable preventing the optimizer explore large parameter space not limited by the prior.

PROOF OF PRINCIPLE

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:

$$\omega(I) = \omega_0 + \omega_1 I + \omega_2 \frac{I^2}{2}, \quad (10)$$

and we ignore the effect by the nonlinear normalization map. The initial beam emittance is chosen to be $2nm$. The 4 kick strengths are 3,4,5 and 6 times the beam emittance and the

kick angles are equally spaced from 0 to π . We also added virtual noise of RMS size $20 \mu m$. The ground truth beam distribution and the BPM signals of 4 different kicks are shown in Figs. 3 and 4.

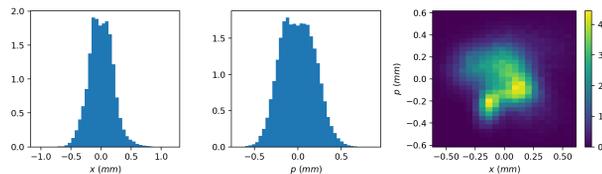


Figure 3: Ground truth beam density is randomly generated.

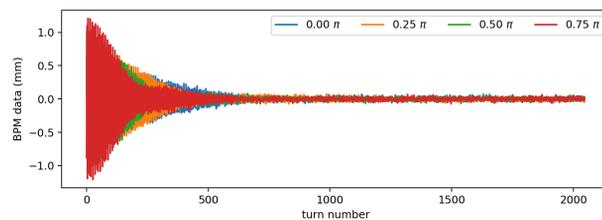


Figure 4: Ground truth BPM signal of various kick angles.

Contraction of Prior

Note that, in Eq. (5) and Eq. (6), the frequencies μ_0 are not coupled with any other parameters. This allows us to optimize the frequencies using the following conditions (on the 4 BPM data).

$$\int_{-\infty}^{\infty} x \rho_{\theta}(x) dx = 0. \quad (11)$$

We further optimize the initial kick strengths I_0 , angles θ , and betatron function β , using the following conditions (on the 4 BPM data).

$$\int_{-\infty}^{\infty} \rho_{\theta}(x) dx = 1. \quad (12)$$

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

Once the prior mean of μ_0 , I_0 , θ , and β are fixed, we fit the gaussian kernel parameters so that the marginal distribution equals to the estimation based on Eq. (6). The prior mean location of Gaussian kernels and weights are shown in Fig. 5

The estimated marginal distribution using Eq. (6), before and after fit are shown in Figs. 6 and 7 respectively. Once we have the prior mean, we construct the prior using independent normal distributions on μ_0 , I_0 , θ , and β with the standard deviation from our belief. For example, since we are building prior from the leading order theory that is justified in the limit of $I_0 \gg \epsilon$, we choose a smaller standard

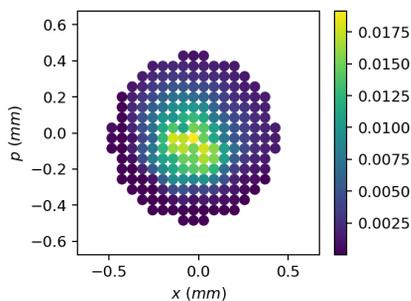


Figure 5: The locations of each Gaussian kernel colored based on weights. The deviation from the grid points of each kernel is not visible in this scale of the plot.

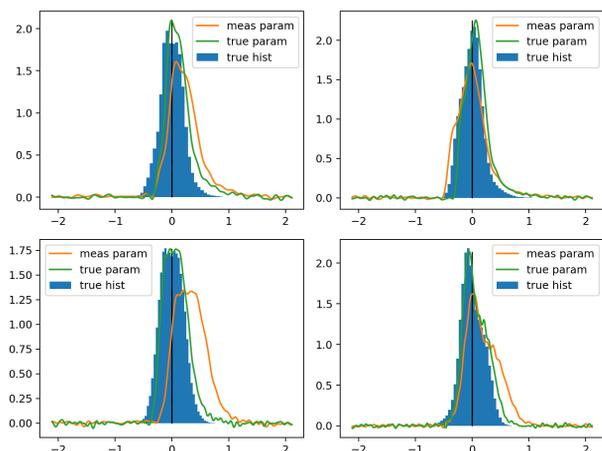


Figure 6: Estimated marginal distribution in 4 different kick direction using Eq. (6) and the estimated parameters before prior mean fitting (orange) and ground truth parameters (green). The ground truth histogram is also shown in blue. The horizontal unit is (mm) and the vertical unit is (1/mm).

deviation for smaller ϵ/I_0 . In other words, we have more confidence in our belief for a larger initial kick. As for the gaussian kernel parameters, we do not construct prior as it is complicated.

Construct Posterior

Now we construct the posterior mean by maximizing the posterior model Eq. (9). The distribution at MAP is shown in Fig. 8.

UNCERTAINTY

So far we have done a 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. One sample from 0.95 confidence interval is shown in Fig. 9.

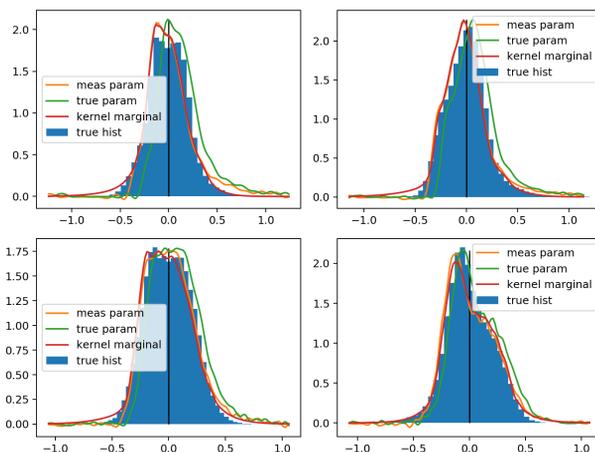


Figure 7: Estimated marginal distribution in 4 different kick direction using Eq. (6) and the parameters after prior mean fitting (orange) and ground truth parameters (green). The ground truth histogram from the gaussian kernel model is shown in red. The marginal distribution from the gaussian kernel model is shown in red. The horizontal unit is (mm) and the vertical unit is (1/mm).

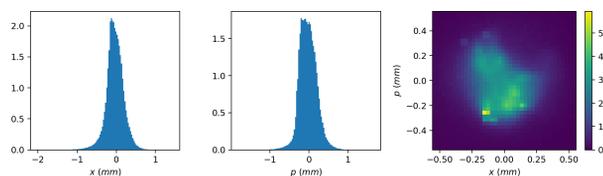


Figure 8: marginal and full 2D histogram of the Gaussian kernel model at MAP. A highly bright point in the 2D histogram can be a numerical artifact of the binning resolution.

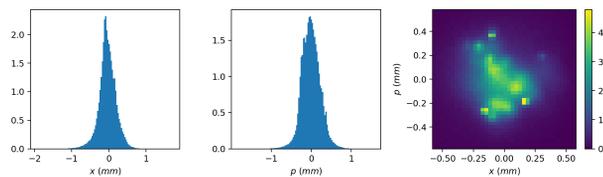


Figure 9: marginal and full 2D histogram of the Gaussian kernel model within 0.95 confidence interval. A highly bright point in the 2D histogram can be a numerical artifact of the binning resolution.

CONCLUSION

We performed proof of principle of the 2D phase-space tomography using beam centroid data of multiple kicked beams. For better resolution, we need more kicks from different angles. The Bayesian approach helped us to avoid the curse-of-dimensionality problem through prior belief construction that could be done using the leading order theory and local minimization. Then the MAP estimation also could be done using local minimization. It also helped us to sample from the posterior (to visualize uncertainty) without relying on MCMC that is practically impossible with so many parameters to infer.

REFERENCES

- [1] R. E. Meller, *et. al.*, “Decoherence of Kicked Beams”, Report number: SSC-N-360, May 1987.
<https://inspirehep.net/literature/248833>
- [2] C. Mitchell and K; Hwang “Thoughts on Using Nonlinear Decoherence for Electron Beam Phase Space Characterization”, IOTA collaboration meeting, LBL, CA, USA, Dec. 2018.
https://indico.fnal.gov/event/24029/attachments/46768/56157/IOTA_NonlinearDecoherence_ElectronPS_17Dec2018.pdf
- [3] Y. Li, R. Rainer, and W. Cheng “Bayesian approach for linear optics correction”, *Phys. Rev. Accel. Beams*, vol. 2, p. 012804, Jan. 2019.
doi:10.1103/PhysRevAccelBeams.22.012804
- [4] Y. Hao, Y. Li, M. Balcewicz, L. Neufcourt, and W. Cheng, “Reconstruction of Storage Ring’s Linear Optics with Bayesian Inference”, Feb. 2019. arXiv:1902.11157