EFFICIENT FIVE-DIMENSIONAL BEAM SIGMA MATRIX DETERMINATION USING DIFFERENTIABLE SIMULATION*

C. Xu†[1], M. Borland[1], L. Emery[1], O. Mohsen[1], Y. Sun[1], K. P. Wootton[1], R. Roussel[2]
[1]Argonne National Laboratory, Lemont, IL, USA
[2]SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Abstract

Precise reconstruction of the beam sigma matrix is critical for transport-line modeling and injection optimization. Our earlier work demonstrated that the differentiable simulation framework Cheetah enables gradient-based recovery of the 5×5 non-temporal sigma matrix at the APS BTS transport line using quadrupole scans. In this paper, we extend the method to improve efficiency and robustness. We introduce a generalized formulation that incorporates multi-screen measurements, providing increased stability in realistic lattice configurations. The differentiable-tracking approach yields physically consistent reconstructions while remaining computationally scalable. These developments form a practical framework for sigma-matrix determination in complex transport lines and support future real-time model calibration and tuning at the APS and similar facilities.

INTRODUCTION

Accurate knowledge of the incoming beam sigma matrix is essential for transport-line modeling, matching, and injection tuning. When performing quadrupole scans in regions with nonzero dispersion, the measured beam moments also contain information about momentum-dependent correlations, allowing reconstruction of the non-temporal 5×5 beam sigma matrix [1]. The downstream beam moments can be determined by linearly solving for the unknown upstream sigma matrix using first-order transport matrices [1, 2]. Despite its efficiency and transparency, this approach's accuracy degrades when the transport line cannot be represented well by a first-order model. Strong chromatic effects, nonlinear fields, and lattice-model errors can produce discrepancies between the measured moments and the linear prediction. These discrepancies may be reduced with subsequent nonlinear fitting or black-box optimization, but the resulting problem is high-dimensional and non-convex, often requiring many tracking evaluations.

Differentiable beam-dynamics simulation provides a practical way to improve the efficiency of general parameter optimization tasks while retaining physics fidelity. Recent codes such as Cheetah [3, 4], Bmad-Julia [5], and JuTrack [6] use auto-differentiation (AD) to propagate derivatives through tracking calculations. This capability has enabled gradient-based high-dimensional beam reconstruction, including generative phase space reconstruction (GPSR) methods for six-dimensional phase-space recovery [7, 8]. In our previous work, we proposed and implemented a Cheetah-based approach for reconstructing the 5 × 5 sigma matrix of the Advanced Photon Source (APS) booster-to-storage-ring (BTS) transport line from quadrupole scan measurements [9]. Although the method showed promising results in simulations and initial measurements, its sensitivity to initial conditions and computational cost limited its application.

In this paper, we extend that method to improve both efficiency and robustness. We formulate the sigma-matrix reconstruction directly in terms of a differentiable, physically constrained covariance parameterization. The method is generalized to include measurements from multiple screens, which adds independent constraints to improve the robustness and reduce the data requirements of real-world measurements. The resulting differentiable-tracking approach preserves physical consistency, remains computationally scalable, and provides a practical path toward online sigma-matrix determination for model calibration and tuning.

GRADIENT-BASED OPTIMIZATION OF BEAM SIGMA MATRIX

We briefly summarize the workflow for reconstructing the incoming beam sigma matrix from quadrupole scan measurements using differentiable simulation. The workflow is outlined in Fig. 1.

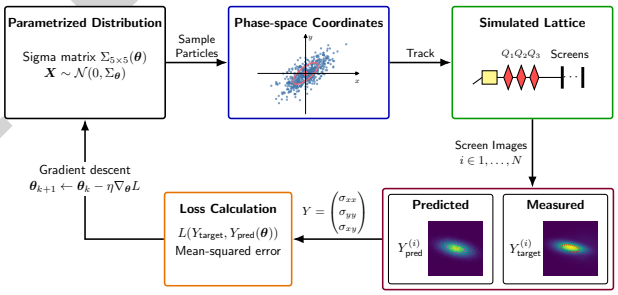

Figure 1: Workflow for the beam sigma matrix reconstruction with differentiable simulation.

The unknown incoming beam distribution is represented by the covariance matrix Σ of the five non-temporal phase-space coordinates (x, p_x, y, p_y, δ). The covariance can be parametrized by a lower-triangular Cholesky factor,

$$\Sigma_{\text{in}}(\boldsymbol{\theta}) = L_{\boldsymbol{\theta}} L_{\boldsymbol{\theta}}^{\top}, \quad \boldsymbol{\theta} \in \mathbb{R}^{15}, \tag{1}$$

which results in 15 degrees of freedom corresponding to the independent elements of the symmetric sigma matrix. The diagonal elements of $L_{\boldsymbol{\theta}}$ are constrained to be positive using

* The work is supported by the U.S. DOE Office of Science-Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.
† chenran.xu@anl.gov

a softplus transform,

$$(L_{\boldsymbol{\theta}})_{jj} = \log\left(1 + \exp\left(\theta_{jj,\mathrm{raw}}\right)\right), \tag{2}$$

while the off-diagonal elements are used directly as optimization variables. This parameterization ensures that the resulting covariance matrix is positive definite without requiring additional constraints in the gradient optimizer.

For each optimizer step, a fixed reference macroparticle distribution is transformed to match the parametrized covariance matrix. Let $\boldsymbol{X}_0$ be the reference distribution with centroid $\boldsymbol{\mu}_0$ and Cholesky factor L_0. The incoming distribution for the current trial matrix is generated as

$$\boldsymbol{X}_{\mathrm{in}}(\boldsymbol{\theta}) = (\boldsymbol{X}_0 - \boldsymbol{\mu}_0)(L_{\boldsymbol{\theta}} L_0^{-1})^{\top} + \boldsymbol{\mu}_0. \tag{3}$$

Equivalently, this produces macroparticles distributed according to

$$\boldsymbol{X}_{\mathrm{in}}(\boldsymbol{\theta}) \sim \Sigma_{\mathrm{in}}(\boldsymbol{\theta}). \tag{4}$$

The transformed particles are tracked through the transport-line model for each quadrupole setting. For quadrupole setting i and screen s, the predicted screen images are

$$\boldsymbol{Y}^{(i,s)}(\boldsymbol{\theta}) = f_s\left(\boldsymbol{X}_{\mathrm{in}}(\boldsymbol{\theta}) \mid \boldsymbol{k}_{Q,i}\right), \tag{5}$$

where f_s denotes tracking from the entrance of the BTS line to screen s. As the reconstruction in this work assumes a nearly Gaussian beam distribution, each predicted transverse profile is reduced to its second moments,

$$\boldsymbol{Y}^{(i,s)}(\boldsymbol{\theta}) = \left(\sigma_{xx}, \sigma_{yy}, \sigma_{xy}\right)^{(i,s)}. \tag{6}$$

This scalarized objective is more computationally efficient than direct image matching while still capturing the beam sizes and transverse coupling relevant to sigma-matrix reconstruction.

The target data consist of measured moments $\boldsymbol{Y}^{(i,s)}_{\mathrm{target}}$ for a set of quadrupole settings and diagnostic screens. The loss function is defined as the mean squared difference between the measured and predicted moments,

$$l(\boldsymbol{\theta}) = \frac{1}{3MN} \sum_{s=1}^{M} \sum_{i=1}^{N} \sum_{j=1}^{3} \left(\boldsymbol{Y}^{(i,s)}_{\mathrm{target},j} - \boldsymbol{Y}^{(i,s)}_{\mathrm{pred},j}(\boldsymbol{\theta})\right)^2, \tag{7}$$

where N is the number of quadrupole settings and M is the number of screens. The multi-screen formulation adds independent constraints from different lattice locations and helps reduce degeneracy among beam sigma matrices that can produce similar moments at a single screen.

Because each operation in the workflow is differentiable, the gradient $\nabla_{\boldsymbol{\theta}} l$ is obtained directly through AD. The Cholesky parameters can then be updated with gradient descent,

$$\boldsymbol{\theta}_{k+1} \leftarrow \boldsymbol{\theta}_k - \eta \nabla_{\boldsymbol{\theta}} l(\boldsymbol{\theta}_k), \tag{8}$$

or with adaptive gradient-based optimizers. In this study, we use Adam [10], which adjusts the effective learning rate for each parameter and improves convergence for parameters with different numerical scales.

MULTI-SCREEN MEASUREMENT AT BTS TRANSPORT LINE

The multi-screen reconstruction was tested on the APS BTS transport line. The scan varied the AQ5, BQ2, and BQ3 quadrupoles, while the beam moments were measured on the downstream FS3 and FS4 diagnostic screens. Two dipoles are located between AQ5 and BQ2, and the region between FS3 and FS4 includes skew quadrupoles used for emittance exchange. The two screens therefore sample the beam after different dispersive and coupled transport sections, providing complementary constraints on the incoming sigma matrix. For each run, 30 quadrupole settings were selected from simulation. The quadrupole strengths were randomly sampled within their allowed ranges, and only settings with full beam transmission were retained. The scan sets were also filtered by the predicted beam size on the screens: the maximum accepted beam size was 1.5 mm for Run1, 2.0 mm for Run2, and 2.5 mm for Run3.

In the reconstruction process, 1000 macroparticles were tracked, and the optimization was performed with Adam for 2000 steps, corresponding to around 4 minutes of computation time on a laptop. The Run2 result is shown in Fig. 2, where the predicted moments from the reconstructed sigma matrix are compared with the measured moments on FS3 and FS4, with $\sigma_{x,y}$ being the horizontal and vertical beam sizes, and $\rho_{xy} = \sigma_{xy}/(\sigma_x \sigma_y)$ being the correlation coefficients. The agreement across both screens indicates that the fitted incoming distribution reproduces the measured beam response over the scanned optics.

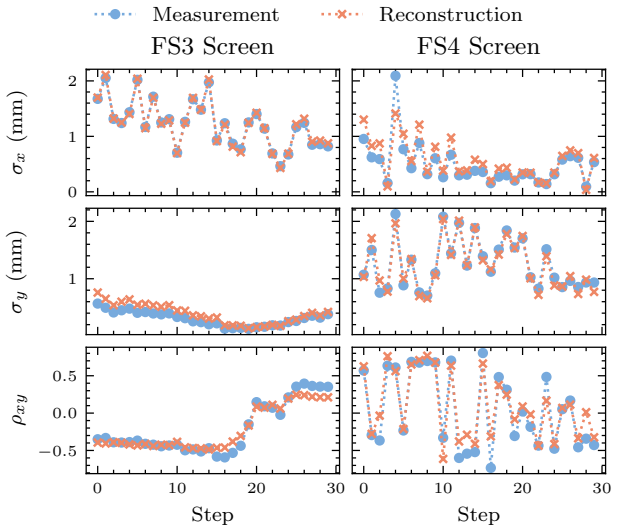

Figure 2: Transverse beam moments measured on the FS3 and FS4 diagnostic screens compared with the predicted moments from the reconstruction.

The reconstructed beam parameters from the three runs are listed in Table 1. Run2 and Run3 were measured on the same day, while each run used a different scan set (see above). The reconstructed parameters are broadly consistent across the three measurements: the horizontal emittance is near 60 nm, and the vertical emittance is near 1.2 nm. All

Table 1: Results of the beam matrix estimated from three measurements. Each scan used a different set of quadrupole settings. Run2 and Run3 were measured on the same day.

Parameter	Run1	Run2	Run3
ϵ_x (nm)	60.6851	61.9216	59.9041
β_x (m)	12.5346	12.4223	11.9061
α_x	2.2871	2.3125	2.2326
η_x (m)	−0.3867	−0.3597	−0.3987
$\eta_{x'}$	0.0818	0.0740	0.0781
ϵ_y (nm)	1.1734	1.2196	1.1038
β_y (m)	3.3092	2.9092	3.2008
α_y	−0.9478	−0.5663	−0.8293
η_y (m)	0.0193	0.0341	0.0256
$\eta_{y'}$	0.0057	0.0051	0.0087
θ_{xy}	−0.0314	−0.0478	−0.0354
σ_p	0.0012	0.0013	0.0013

three runs also predict a horizontal dispersion of approximately −0.4 m and a transverse tilt angle of order −0.03 rad to −0.05 rad, which may indicate an optics mismatch in the booster.

As the inverse problem is nonlinear, the gradient-based optimization may converge to local minima. However, the consistency obtained from three independently selected scan sets suggests that the relevant minima are close in the physical beam-parameter space. This repeatability is an important indication that the multi-screen formulation improves the robustness of the reconstruction.

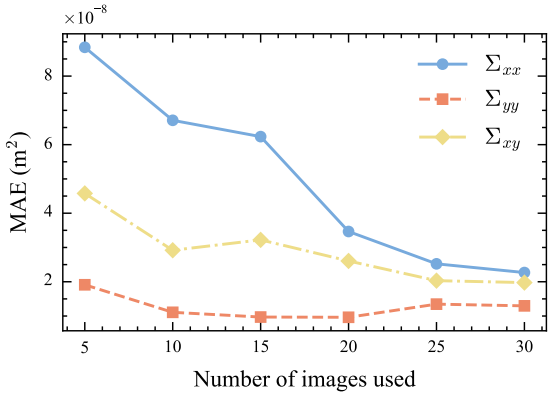

Figure 3: mean absolute error (MAE) of the beam moments on FS3 screen when using different numbers of images for the reconstruction. The images are down-selected from the set of 30 measured images.

Incorporating multi-screen measurements could further reduce the number of required images, and therefore the beam time needed for a measurement. Figure 3 shows the reconstruction result using only a subset of the measured images. As expected, the reconstruction error in the transverse beam moments decreases with increasing number of images. This is consistent with the fact that, although the problem is formally overdetermined, it remains partially degenerate in practice. The five-dimensional sigma matrix has 15 degrees of freedom, while each image provides three observables, so in principle only five images are required. In practice, however, additional measurements help better constrain correlations and reduce ambiguities in the reconstructed initial beam distribution. For this study, using around 20 images appears to preserve most of the reconstruction accuracy. This should be regarded as an initial estimate, since the current result is based only on down-selecting from an existing dataset. If fewer images are to be acquired in practice, the quadrupole settings could instead be re-optimized to maximize the information content of the measurement set.

The present reconstruction does not yet account for several real-world effects. In particular, beam orbit variation during the quadrupole scan can lead to variable contribution to the beam sizes. Small but not precisely known quadrupole misalignments also influence the result by introducing spurious dispersion to the measured beam moments. Extending the reconstruction model to include these effects is expected to further improve both the accuracy and robustness of the inferred beam parameters.

CONCLUSION

We have developed and implemented a novel method for reconstructing the five-dimensional beam sigma matrix in a transport line from quadrupole-scan measurements using the differentiable simulation code Cheetah. The approach is computationally efficient, robust, and yields physically consistent solutions. Its formulation is also flexible, making it readily extensible to more general measurement configurations, such as multi-screen setups, and broadly applicable to transport lines with nonzero dispersion. Experimental results demonstrate good consistency of the reconstructed sigma matrices across measurements taken on different days, indicating that the method is sufficiently reproducible for practical operation. Incorporation of real-world effects like beam and magnet offsets is expected to further improve the reconstruction accuracy. These promising results motivates its future online deployment in automated beam characterization workflows [11] and provide a natural foundation for further extensions, such as integrated system identification and adaptive measurement strategies.

ACKNOWLEDGMENT

We gratefully acknowledge the computing resources provided on Swing, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory.

REFERENCES

[1] M. Borland, V. Sajaev, and K. P. Wootton, “Promise and Challenges of a Method for 5x5 Sigma Matrix Measurement in a Transport Line”, in *Proc. NAPAC’22*, Albuquerque, NM, USA, Aug. 2022, pp. 382–385. doi:10.18429/JACoW-NAPAC2022-TUPA18

[2] M. Borland, "ELEGANT: A flexible SDDS-compliant code for accelerator simulation", Argonne National Laboratory, Lemont, IL, USA, Rep. LS-287, Aug. 2000. doi:10.2172/761286

[3] J. Kaiser, C. Xu, A. Eichler, and A. Santamaria Garcia, "Bridging the gap between machine learning and particle accelerator physics with high-speed, differentiable simulations", *Phys. Rev. Accel. Beams*, vol. 27, no. 5, p. 054601, May 2024. doi:10.1103/PhysRevAccelBeams.27.054601

[4] R. Roussel *et al.*, "Advancements in backwards differentiable beam dynamics simulations for accelerator design, model calibration, and machine learning", in *Proc. LINAC2024*, Chicago, IL, USA, Aug. 2024, pp. 768–771. doi:10.18429/JACoW-LINAC2024-THPB068

[5] D. Sagan, G. Hoffstaetter, M. Signorelli, and A. Coxe, "Bmad-Julia: a Julia environment for accelerator simulations including machine learning", in *Proc. IPAC'24*, Nashville, TN, USA, May 2024, pp. 2612–2615. doi:10.18429/JACoW-IPAC2024-WEPR52

[6] J. Wan, H. Alamprese, C. Ratcliff, J. Qiang, and Y. Hao, "JuTrack: A Julia package for auto-differentiable accelerator modeling and particle tracking", *Comput. Phys. Commun.*, vol. 309, p. 109497, Apr. 2025. doi:10.1016/j.cpc.2024.109497

[7] R. Roussel *et al.*, "Phase Space Reconstruction from Accelerator Beam Measurements Using Neural Networks and Differentiable Simulations", *Phys. Rev. Lett.*, vol. 130, no. 14, p. 145001, Apr. 2023. doi:10.1103/PhysRevLett.130.145001

[8] R. Roussel *et al.*, "Efficient six-dimensional phase space reconstructions from experimental measurements using generative machine learning", *Phys. Rev. Accel. Beams*, vol. 27, no. 9, p. 094601, Sep. 2024. doi:10.1103/PhysRevAccelBeams.27.094601

[9] C. Xu, L. Emery, M. Borland, O. Mohsen, Y. Sun, and R. Roussel, "Towards accurate beam sigma matrix determination in a transport line using differentiable simulation", in *Proc. NAPAC2025*, Sacramento, CA, USA, Aug. 2025, pp. 264–267. doi:10.18429/JACoW-NAPAC2025-MOP093

[10] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", in *Proc. ICLR 2015*, San Diego, CA, USA, May 2015. doi:10.48550/arXiv.1412.6980

[11] R. Roussel *et al.*, "Autonomous operation of the DIAG0 diagnostic line for 6D phase-space monitoring at LCLS-II", Apr. 2026. doi:10.48550/arXiv.2604.20125