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

R. Roussel, A. Edelen, C. Mayes, D. Ratner, J. P. Gonzalez-Aguilera, S. Kim, E. Wisniewski, and J. Power

Phys. Rev. Lett. **130**, 145001 — Published 5 April 2023
DOI: 10.1103/PhysRevLett.130.145001

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

R. Roussel,* A. Edelen, C. Mayes, and D. Ratner
SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

J.P. Gonzalez-Aguilera
Department of Physics, University of Chicago, Chicago, Illinois 60637, USA

S. Kim, E. Wisniewski, J. Power
Argonne National Laboratory, Argonne, Illinois 60439, USA
(Dated: January 24, 2023)

Characterizing the phase space distribution of particle beams in accelerators is a central part of accelerator understanding and performance optimization. However, conventional reconstruction-based techniques either use simplifying assumptions or require specialized diagnostics to infer high-dimensional (> 2D) beam properties. In this Letter, we introduce a general-purpose algorithm that combines neural networks with differentiable particle tracking to efficiently reconstruct high-dimensional phase space distributions without using specialized beam diagnostics or beam manipulations. We demonstrate that our algorithm accurately reconstructs detailed 4D phase space distributions with corresponding confidence intervals in both simulation and experiment using a single focusing quadrupole and diagnostic screen. This technique allows for the measurement of multiple correlated phase spaces simultaneously, which will enable simplified 6D phase space distribution reconstructions in the future.

Increasingly precise control of the distribution of particles in position-momentum phase space is needed for emerging applications of accelerators [1]. This includes, for example, new experiments at free electron lasers [2–6] and novel acceleration schemes that promise higher-energy beams in compact spaces [7]. Numerous techniques have been developed for precision shaping of beam distributions [8]; however, the effectiveness of these techniques relies on accurate measurements of the 6D phase space distribution, which is a challenging task unto itself.

Tomographic measurement techniques are used in accelerators to determine the density distribution of beam particles in phase space $\rho(x, p_x, y, p_y, z, p_z)$ from limited measurements [9–14]. The simplest form of this uses scalar metrics, such as second-order moments, to describe observations of the transverse beam distribution when projected onto a scintillating screen. [15–17]. This process however discards significant amounts of information about the beam distribution captured by high-resolution diagnostic screens and only predicts scalar quantities of the beam distribution. In contrast, methods using projections of the beam image, including filtered back-projection [12, 18], algebraic reconstruction [19–21], and maximum entropy tomography (MENT) [13, 22] produce more accurate reconstructions.

The MENT algorithm is particularly well-suited to reconstructing beams from limited and/or partial information sources about the beam distribution, as is the case in most experimental accelerator measurements. MENT solves for a phase space distribution that maximizes entropy (and, as a result, likelihood), subject to the constraint that the distribution accurately reproduces experimental measurements. While these techniques have been shown to effectively reconstruct 2D phase spaces from image projections using algebraic methods, application to higher-dimensional spaces requires independence assumptions between the phase spaces of principal coordinate axes (x, y, z), complicated phase space rotation procedures [20, 23], or simultaneous measurement of multiple 2D sub-spaces with specialized diagnostic hardware [24].

Numerical optimization methods can also be used to infer beam distributions from experimental data. For example, arbitrary beam distributions can be parameterized by a set of principal components [25] whose relative weights can be optimized to produce a beam distribution that, when tracked through a simulation, reproduces experimental measurements. Alternatively, heuristics can be used to delete or generate particles in a distribution until particle tracking results match experiments [26, 27]. Unfortunately, these methods suffer from increasing computational cost when extending them to reconstructing high-dimensional phase space distributions, primarily due to the cost associated with optimizing the large number of free parameters needed to represent detailed beam characteristics in high-dimensional phase spaces.

In this Letter we describe a new method that provides detailed reconstructions of the beam phase space using simple and widely-available accelerator elements and diagnostics. To achieve this, we take advantage of recent developments in machine learning to introduce two new concepts (shown in Fig. 1): a method for parameterizing arbitrary beam distributions in 6D phase space, and a differentiable particle tracking simulation that allows

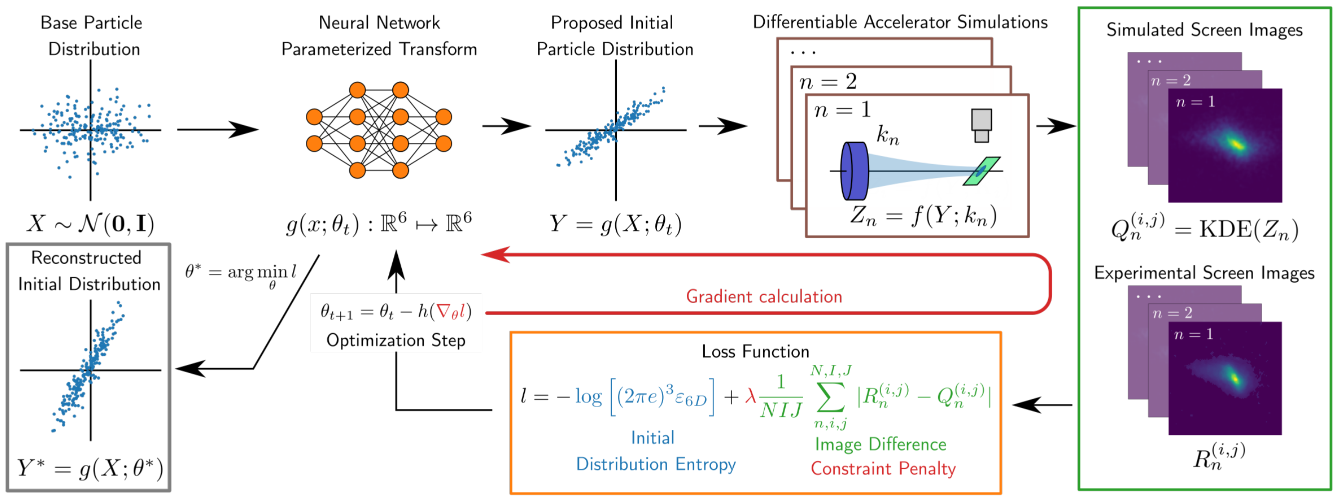

FIG. 1. Description of our approach for reconstructing phase space beam distributions. First, a 6D base distribution is transformed via neural network, parameterized by θ_t, into a proposed initial distribution. This distribution is then transported through a differentiable accelerator simulation of the tomographic beamline. The quadrupole is scanned to produce a series of images on the screen, both in simulation and on the operating accelerator. The images produced both from the simulation $Q_n^{(i,j)}$ and the accelerator $R_n^{(i,j)}$ are then compared with a custom loss function, which attempts to maximize the entropy of the proposal distribution, constrained on accurately reproducing experimental measurements. This loss function is then used to update the neural network parameters $\theta_t \to \theta_{t+1}$ via gradient descent. The neural network transformation that minimizes the loss function generates the beam distribution that has the highest likelihood of matching the real initial beam distribution.

us to learn the beam distribution from arbitrary downstream accelerator measurements. We examine how this method extracts detailed 4-dimensional phase space distributions from measurements in simulation and experiment, using a simple diagnostic beamline, containing a single quadrupole, drift and diagnostic screen to image the transverse (x, y) beam distribution. Finally, we discuss current limitations of this method as well as future directions for the design of novel accelerator diagnostics using this technique.

We first demonstrate our algorithm using a synthetic example, where we attempt to determine the distribution of a 10-MeV beam given a predefined structure in 6D phase space. The propagation of a synthetic beam distribution through a simple diagnostic beamline containing a 10 cm long quadrupole followed by a 1.0 m drift is simulated using a custom implementation of Bmad [28] referred to here as Bmad-X. To illustrate the capabilities of our technique, the synthetic beam contains multiple higher order moments between each phase space coordinates (see Supplemental Materials for details). To simulate an experimental measurement, we simulate particles traveling through the diagnostic beamline while the quadrupole strength k is scanned over N points. The final transverse distribution of the beam is measured at each quadrupole strength using a simulated 200×200 pixel screen, with a pixel resolution of 300 µm (image data can be viewed in the Supplemental Materials). The set of images, where the intensity of pixel (i, j) on the n'th image is represented by $R_n^{(i,j)}$, is then collected with the corresponding quadrupole strengths to create the data set, which is then split into training and testing subsets by selecting every other sample as a test sample, resulting in 10 samples for each data subset.

The reconstruction algorithm begins with the generation of arbitrary initial beam distributions (referred to here as proposal distributions) through the use of a neural network transformation. A neural network, consisting of only 2 fully-connected layers of 20 neurons each, is used to transform samples drawn from a 6D normal distribution centered at the origin to macro-particle coordinates in real 6D phase space (where positional coordinates are given in meters and momentum coordinates are in radians for transverse momenta). As a result, the coordinates of particles in the proposal distribution are fully parameterized by the neural network parameter set θ_t.

Fitting neural network parameters to experimental measurements is done by minimizing a loss function to determine the most likely initial beam distribution, subject to the constraint that it reproduces experimental measurements; this is similar to the MENT algorithm [22]. The likelihood of an initial beam distribution in phase space is maximized by maximizing the distribution entropy, which is proportional to the log of the 6D beam emittance ε_{6D} [29]. Thus, we specify a loss function that minimizes the negative entropy of the proposal beam distribution, penalized by the degree to which the proposal distribution reproduces measurements of the transverse beam distribution at the screen location. To evaluate the penalty for a given proposal distribution, we track

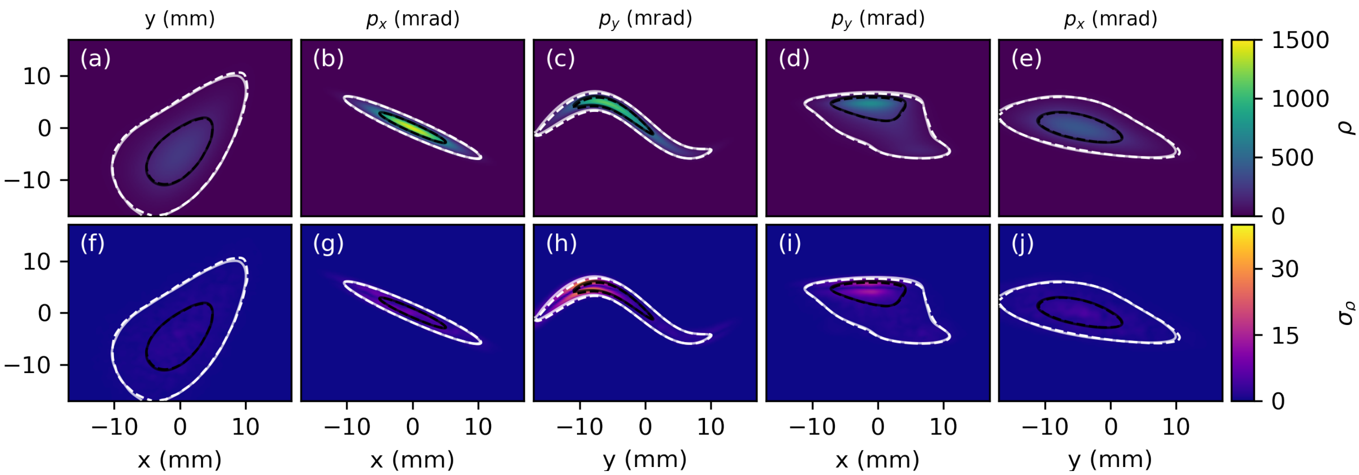

FIG. 2. Comparisons between the synthetic and reconstructed beam probability distributions using our method. (a-e) Plots of the mean predicted phase space density projections in 4D transverse phase space. Contours that denote the 50^{th} (black) and 95^{th} (white) percentiles of the synthetic ground truth (dashed) and reconstructed (solid) distributions. (f-j) Plots of the predicted phase space density uncertainty.

the proposal distribution through a batch of accelerator simulations that mimic experimental conditions to generate a set of simulated images $Q_n^{(i,j)}$ to compare with experimental measurements. The total loss function is given by

$$l = -\log\left[(2\pi e)^3 \varepsilon_{6D}\right] + \lambda \frac{1}{NIJ} \sum_{n,i,j}^{N,I,J} |R_n^{(i,j)} - Q_n^{(i,j)}| \quad (1)$$

where λ scales the distribution loss penalty function relative to the entropy term and is chosen empirically based on the resolution of the images.

However, the large ($> 10^3$) number of free parameters contained in the neural network transformation used to generate proposal distributions necessitates the use of gradient-based optimization algorithms such as Adam [30] to minimize the loss function. Thus, we need to implement computation of the loss function such that it supports backward differentiation [31] (referred to here as *differentiable* computations), allowing us to cheaply compute loss function derivatives with respect to every neural network parameter. This requires that every step involved in calculating the loss function is also differentiable, including computing the beam emittance and tracking particles through the accelerator. Unfortunately, to the best of our knowledge, no particle tracking codes currently support backwards differentiation. To satisfy this requirement, we implement particle tracking in Bmad-X using the machine learning library *PyTorch* [32]. We estimate screen pixel intensities from a discrete particle distribution with a differentiable implementation of kernel density estimation [33].

Results from our reconstruction of the initial beam phase space using synthetic images are shown in Fig. 2. We characterize the uncertainty of our reconstruction using snapshot ensembling [34]. During model training, we cycle the learning rate of gradient descent in a periodic fashion which encourages the optimizer to explore multiple possible solutions (if they exist). After several of these cycles (known as a "burn-in" period), we save model parameters at each minima of the learning rate cycle, as shown in Fig. 3(a). We then weight predictions from each model equally, using them to predict a mean initial beam density distribution Fig. 2(a-e) with associated confidence intervals Fig. 2(f-j). Performing this analysis by tracking 10^5 particles for each image took less than 30 seconds per ensemble sample using a professional grade GPU ($<$ 60 ms per iteration, 500 steps per ensemble sample).

TABLE I. Predicted Emittances Compared to True Values

Parameter	Ground truth	RMS Prediction	Reconstruction	Unit
ε_x	2.00	2.47	2.00 ± 0.01	mm-mrad
ε_y	11.45	14.10	10.84 ± 0.04	mm-mrad
ε_{4D}	18.51	34.83*	17.34 ± 0.08	mm^2-mrad^2

* Assumes x-y phase space independence

We see excellent agreement between the average reconstructed and synthetic projections in both transverse correlated and uncorrelated phase spaces. Furthermore, the prediction uncertainty from ensembling is on the order of a few percent relative to the predicted mean, providing confidence that the overall solution found during optimization is unique. As shown in Table I, reconstructions of the beam distribution from image data predicts transverse phase space emittances that are closer to ground truth values than those predicted from second-order moment measurements of the transverse beam distribution.

This results from non-linearities and cross-correlations present in the 4-D transverse phase space distribution.

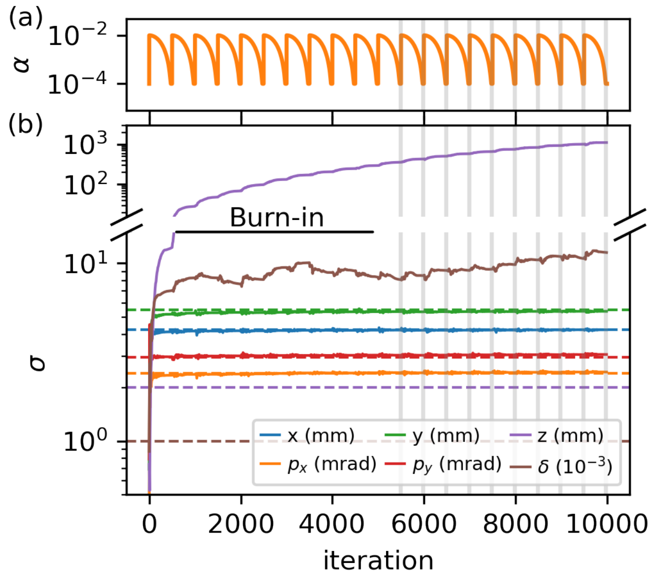

FIG. 3. Evolution of the proposal distribution during training on synthetic data. (a) Learning rate schedule for snapshot ensembling. (b) Second order moments of beam reconstruction during training for each phase space coordinate. Dashed lines denote ground truth values. Vertical lines denote snapshot locations after burn-in period.

It is instructive to examine the evolution of the proposal distribution during model training. In Fig. 3(b) we examine second order scalar metrics of the proposal distribution after each training iteration for each phase space coordinate. The entropy term in Eq. 1 causes the distribution to expand in 6D phase space until constrained by experimental evidence. Phase space components that have the strongest impact on beam transport through the beamline as a function of quadrupole strength converge quickly to the true values, whereas the ones that have little-to-no impact (e.g. the longitudinal distribution characteristics) continue to grow. In other cases, there is weak coupling between the experimental measurements and beam properties; for example, chromatic focusing effects due to the energy spread σ_δ of the beam weakly affect the measured images. Here, the reconstruction can only provide an upper-bound estimate of the energy spread, since small changes in transverse beam propagation due to chromatic aberrations are overshadowed by statistically dominated particle motion. Convergence of the proposal distribution thus provides a useful indicator of which phase space components can be reliably reconstructed from arbitrary sets of measurements.

We now describe a demonstration of our method on an experimental example at the Argonne Wakefield Accelerator (AWA) [35] facility at Argonne National Laboratory. Our objective is to identify the phase space distribution of 65-MeV electron beams at the end of the primary accelerator beamline. The focusing strength of a quadrupole, with an effective length of 12 cm, is scanned while imaging the beam at a transverse scintillating screen located 3.38 m downstream. Charge windowing, image filtering, thresholding and downsampling were used to generate a set of 3 images for each quadrupole setting (see the Supplemental Materials for additional details).

TABLE II. Predicted Emittances from Experimental Data

Parameter	RMS Prediction	Reconstruction	Unit
$\varepsilon_{x,n}$	4.18 ± 0.71	4.23 ± 0.02	mm-mrad
$\varepsilon_{y,n}$	3.65 ± 0.36	3.42 ± 0.02	mm-mrad

We developed a differentiable simulation in Bmad-X of the experimental beamline, including details of the diagnostics used, such as the location and properties of beamline elements and the per-pixel resolution of the imaging screen. With this simulation, we used our method to reconstruct the beam distribution from experimentally-measured transverse beam images. The results, as shown in Figure 4 and Table II, demonstrate good agreement between experimental measurements of the beam distribution and predictions from our reconstruction. Scalar predictions of the beam emittances from the image-based reconstruction are consistent with those calculated from RMS measurements. Additionally, our reconstruction method accurately reproduces fine features of the transverse beam distribution that were not present in the training data set.

In this work, we have demonstrated how differentiable particle tracking simulations, combined with neural network based representations of beam distributions, can be used to interpret common image-based diagnostic measurements. Our method produces detailed reconstructions of 4-dimensional transverse phase space distributions from limited data sets, without the use of complex phase space manipulations or specialized diagnostics. Additionally, our reconstruction identifies limitations in resolving certain aspects of the beam distribution based on available measurements. This analysis is enabled by inexpensive gradient calculations provided by backwards differentiable physics simulations. As a result, we are able to determine thousands of free parameters used to describe complex beam distributions on a time scale similar to the time it takes to perform the physical tomographic measurements themselves. Thus, our reconstruction technique is suitable for inferring detailed beam distributions in an online fashion, i.e. during accelerator operations.

As with any new algorithmic technique, there are areas for future improvement. Uncertainty estimates provided by the reconstruction algorithm only capture systematic

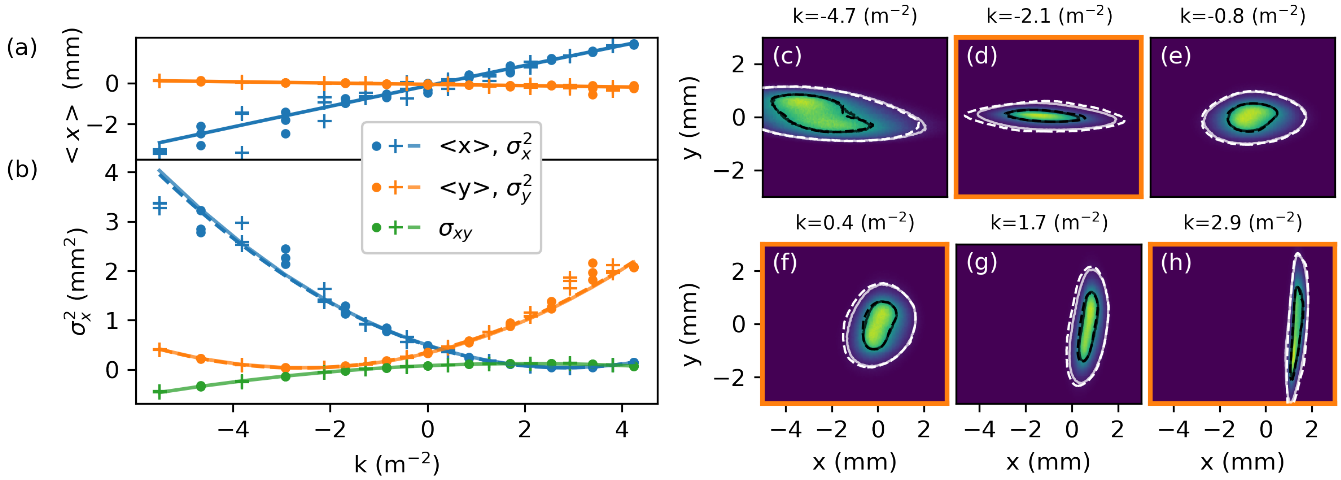

FIG. 4. Reconstruction results from experimental measurements at AWA. Comparison between measured and predicted beam centroids (a) and second-order beam moments (b) on the diagnostic screen as a function of geometric quadrupole focusing strength (k). Points denote training samples and crosses denote test samples. Dashed line shows second order polynomial fit of training data and solid line shows predictions from image-based phase space reconstruction. We also compare (c-h) screen images and reconstructed predictions for a subset of quadrupole strengths. Contours denote the 50^{th} (black) and 95^{th} (white) percentiles of the measured (dashed) and predicted (solid) screen distributions. Orange borders denote test samples.

uncertainties from optimizing the loss function, Eq. 1; thus it ignores systematic uncertainties of the physical measurement and stochastic noise inherent in real accelerators. Future work will incorporate Bayesian analysis techniques into the reconstruction to provide calibrated uncertainty estimates to experimental measurements. Also, while our method significantly increases the speed of high-dimensional phase space reconstructions, achieving this requires substantial amounts of memory to store the derivative information of each macro-particle at every tracking step ($\sim$ 4 GB for each snapshot in the analysis performed here). Peak memory consumption can be reduced through the use of checkpointing [36] or pre-computing derivatives associated with tracking particles through the entire beamline. Finally, this method is limited by the availability of accurate, computationally efficient, backwards differentiable particle tracking simulations. In order to expand the range of diagnostic measurements that can be analyzed by this technique, further investment in differentiable implementations of particle tracking simulations is needed.

This new reconstruction approach opens the door to efficient, detailed characterization of 6-dimensional phase space distributions and new types of compound diagnostic measurements. By adding longitudinal beam manipulations, such as transverse deflecting cavities paired with dipole spectrometers, to the beamline used here, full phase space distributions can be characterized through a series of quadrupole strength and deflecting cavity phase scans.

The authors would like to thank Lukas Heinrich and Michael Kagan for useful discussion during the early conceptual development of this work. This work was supported by the U.S. Department of Energy, under DOE Contract No. DE-AC02-76SF00515, the Office of Science, Office of Basic Energy Sciences and the Center for Bright Beams, NSF award PHY-1549132. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award ERCAP0020725.

Author Contributions: R.R. and A.E. conceived of the idea to combine differentiable simulations with machine learning for phase space tomography. R.R. led the studies and performed the work for phase space reconstruction. A.E. and D.R. provided technical guidance and feedback. J.P.G. developed the differentiable simulation with guidance from R.R. and C.M.. S.K., E.W. and J.P. assisted with experimental studies at AWA. R.R. and A.E. wrote the manuscript. J.P.G., C.M. provided substantial edits to the manuscript. All authors provided feedback on the manuscript.

* rroussel@slac.stanford.edu

[1] S. Nagaitsev, Z. Huang, J. Power, J. L. Vay, P. Piot, L. Spentzouris, J. Rosenzweig, Y. Cai, S. Cousineau, M. Conde, M. Hogan, A. Valishev, M. Minty, T. Zolkin, X. Huang, V. Shiltsev, J. Seeman, J. Byrd, Y. Hao, B. Dunham, B. Carlsten, A. Seryi, and R. Patterson, Accelerator and beam physics research goals and oppor-

tunities (2021).

[2] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F. J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, P. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H. D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda, First lasing and operation of an aangstrom-wavelength free-electron laser, Nature Photonics **4**, 641 (2010), publisher: Nature Publishing Group.

[3] H. Li, Y. Sun, J. Vila-Comamala, T. Sato, S. Song, P. Sun, M. H. Seaberg, N. Wang, J. B. Hastings, M. Dunne, P. Fuoss, C. David, M. Sutton, and D. Zhu, Generation of highly mutually coherent hard-x-ray pulse pairs with an amplitude-splitting delay line, Phys. Rev. Research **3**, 043050 (2021).

[4] Y. Sun, M. Dunne, P. Fuoss, A. Robert, D. Zhu, T. Osaka, M. Yabashi, and M. Sutton, Realizing split-pulse x-ray photon correlation spectroscopy to measure ultrafast dynamics in complex matter, Phys. Rev. Research **2**, 023099 (2020).

[5] A. Marinelli, D. Ratner, A. A. Lutman, J. Turner, J. Welch, F. J. Decker, H. Loos, C. Behrens, S. Gilevich, A. A. Miahnahri, S. Vetter, T. J. Maxwell, Y. Ding, R. Coffee, S. Wakatsuki, and Z. Huang, High-intensity double-pulse x-ray free-electron laser, Nature Communications **6**, 10.1038/ncomms7369 (2015).

[6] F.-J. Decker, K. L. Bane, W. Colocho, S. Gilevich, A. Marinelli, J. C. Sheppard, J. L. Turner, J. J. Turner, S. L. Vetter, A. Halavanau, C. Pellegrini, and A. A. Lutman, Tunable x-ray free electron laser multipulses with nanosecond separation, Scientific Reports **12**, 10.1038/s41598-022-06754-y (2022).

[7] *Advanced Accelerator Development Strategy Report: DOE Advanced Accelerator Concepts Research Roadmap Workshop*, Tech. Rep. (USDOE Office of Science, Washington, DC (United States), 2016).

[8] G. Ha, K.-J. Kim, J. Power, Y. Sun, P. Piot, *et al.*, Bunch shaping in electron linear accelerators, Reviews of Modern Physics **94**, 025006 (2022).

[9] C. McKee, P. O'Shea, and J. Madey, Phase space tomography of relativistic electron beams, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **358**, 264 (1995).

[10] S. Hancock, M. Lindroos, E. McIntosh, and M. Metcalf, Tomographic measurements of longitudinal phase space density, Computer Physics Communications **118**, 61 (1999).

[11] D. Stratakis, R. A. Kishek, I. Haber, M. Walter, R. B. Fiorito, S. Bernal, J. Thangaraj, K. Tian, C. Papadopoulos, M. Reiser, and P. G. O'Shea, Phase space tomography of beams with extreme space charge, in *2007 IEEE Particle Accelerator Conference (PAC)* (2007) pp. 2025–2029.

[12] V. Yakimenko, M. Babzien, I. Ben-Zvi, R. Malone, and X.-J. Wang, Electron beam phase-space measurement using a high-precision tomography technique, Physical Review Special Topics - Accelerators and Beams **6**, 122801 (2003).

[13] M. Röhrs, C. Gerth, H. Schlarb, B. Schmidt, and P. Schmüser, Time-resolved electron beam phase space tomography at a soft x-ray free-electron laser, Physical Review Special Topics - Accelerators and Beams **12**, 050704 (2009).

[14] M. Gordon, W. Li, M. Andorf, A. Bartnik, C. Duncan, M. Kaemingk, C. Pennington, I. Bazarov, Y.-K. Kim, and J. Maxson, Four-dimensional emittance measurements of ultrafast electron diffraction optics corrected up to sextupole order, Physical Review Accelerators and Beams **25**, 084001 (2022).

[15] A. Green and Y.-M. Shin, Implementation of Quadrupole-scan Emittance Measurement at Fermilab's Advanced Superconducting Test Accelerator (ASTA), in *6th International Particle Accelerator Conference* (2015) p. MOPMA052.

[16] E. Prat and M. Aiba, Four-dimensional transverse beam matrix measurement using the multiple-quadrupole scan technique, Physical Review Special Topics - Accelerators and Beams **17**, 052801 (2014).

[17] A. Mostacci, M. Bellaveglia, E. Chiadroni, A. Cianchi, M. Ferrario, D. Filippetto, G. Gatti, and C. Ronsivalle, Chromatic effects in quadrupole scan emittance measurements, Phys. Rev. ST Accel. Beams **15**, 082802 (2012).

[18] S. Webb, *The Physics of Medical Imaging* (CRC Press, Boca Raton, 1987).

[19] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging* (SIAM, 2001).

[20] A. Wolski, D. Christie, B. Militsyn, D. Scott, and H. Kockelbergh, Transverse phase space characterization in an accelerator test facility, Physical Review Accelerators and Beams **23**, 032804 (2020).

[21] A. Wolski, M. A. Johnson, M. King, B. L. Militsyn, and P. H. Williams, Transverse phase space tomography in the clara accelerator test facility using image compression and machine learning, arXiv preprint arXiv:2209.00814 (2022).

[22] K. M. Hock and M. G. Ibison, A study of the maximum entropy technique for phase space tomography, Journal of Instrumentation **8** (02), P02003.

[23] K. Hock and A. Wolski, Tomographic reconstruction of the full 4d transverse phase space, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **726**, 8 (2013).

[24] J. C. Wong, A. Shishlo, A. Aleksandrov, Y. Liu, and C. Long, 4D transverse phase space tomography of an operational hydrogen ion beam via noninvasive 2D measurements using laser wires, Physical Review Accelerators and Beams **25**, 042801 (2022).

[25] A. Scheinker, Adaptive Machine Learning for Robust Diagnostics and Control of Time-Varying Particle Accelerator Components and Beams, Information **12**, 161 (2021), number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[26] M. Wang, Z. Wang, D. Wang, W. Liu, B. Wang, M. Wang, M. Qiu, X. Guan, X. Wang, W. Huang, and S. Zheng, Four-dimensional phase space measurement using multiple two-dimensional profiles, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **943**, 162438 (2019).

[27] B. Hermann, V. A. Guzenko, O. R. Hürzeler, A. Kirchner, G. L. Orlandi, E. Prat, and R. Ischebeck, Electron beam transverse phase space tomography using

nanofabricated wire scanners with submicrometer resolution, Physical Review Accelerators and Beams **24**, 022802 (2021).

[28] D. Sagan, Bmad: A relativistic charged particle simulation library, *Computational accelerator physics. Proceedings, 8th International Conference, ICAP 2004, St. Petersburg, Russia, June 29-July 2, 2004*, Nucl. Instrum. Meth. **A558**, 356 (2006), proceedings of the 8th International Computational Accelerator Physics Conference.

[29] J. Lawson, R. Gluckstern, and P. M. Lapostolle, Emittance, entropy and information, Part. Accel. **5**, 61 (1973).

[30] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, arXiv:1412.6980 [cs] (2017), arXiv: 1412.6980.

[31] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, Efficient BackProp, in *Neural Networks: Tricks of the Trade: Second Edition*, Lecture Notes in Computer Science, edited by G. Montavon, G. B. Orr, and K.-R. Müller (Springer, Berlin, Heidelberg, 2012) pp. 9–48.

[32] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, PyTorch: An Imperative Style, High-Performance Deep Learning Library, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d. Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019) pp. 8024–8035.

[33] M. Rosenblatt, Remarks on some nonparametric estimates of a density function, The annals of mathematical statistics , 832 (1956).

[34] G. Huang, Y. Li, G. Pleiss, Z. Liu, J. E. Hopcroft, and K. Q. Weinberger, Snapshot ensembles: Train 1, get m for free, arXiv preprint arXiv:1704.00109 (2017).

[35] M. Conde, S. Antipov, D. Doran, W. Gai, Q. Gao, G. Ha, C. Jing, W. Liu, N. Neveu, J. Power, *et al.*, Research program and recent results at the argonne wakefield accelerator facility (awa), Proc. IPAC'17 , 2885 (2017).

[36] B. Dauvergne and L. Hascoët, The data-flow equations of checkpointing in reverse automatic differentiation, in *International Conference on Computational Science* (Springer, 2006) pp. 566–573.