

Online optimization of nonlinear lattice using a data-driven chaos indicator

Minghao Song[✉] and Yongjun Li^{*}*Brookhaven National Laboratory, Upton, New York 11973, USA*

(Received 26 April 2026; accepted 20 July 2026; published 12 August 2026)

We report the experimental implementation of a data-driven chaos indicator (DDCI) [Y. Li *et al.*, Data-driven chaos indicator for nonlinear dynamics and applications on storage ring lattice design, Nucl. Instrum. Methods Phys. Res. A **1024**, 166060 (2022)] for online optimization of the National Synchrotron Light Source II storage ring. The DDCI quantifies the predictability of electron beam dynamics using turn-by-turn beam position monitor data. A surrogate model of the one-turn map is first trained, and its out-of-sample predictive uncertainty is then employed as a measurable indicator of chaos. By tuning sextupole magnets to mitigate nonlinear effects, a clear enlargement of the dynamic aperture is achieved, accompanied by a corresponding improvement in injection efficiency.

DOI: [10.1103/PhysRevAccelBeams.29.L081601](https://doi.org/10.1103/PhysRevAccelBeams.29.L081601)

The nonlinear beam dynamics of storage rings determine key operational performance metrics, including injection efficiency and beam lifetime, which are closely correlated with the dynamic aperture (DA) and momentum acceptance (MA). During the design stage, storage ring lattices are typically optimized to enhance DA and MA using physics-based accelerator models and/or simulation-driven objectives. However, the “as-built” machine inevitably contains unmodeled imperfections, such as magnet field errors, alignment and calibration residuals, hysteresis effects, and stray fields. These discrepancies can significantly degrade nonlinear performance relative to design expectations. Consequently, online optimization based directly on experimental data is strongly motivated.

A key challenge in the online optimization of nonlinear lattices lies in defining a robust objective function that (1) is experimentally measurable within limited beam time, (2) is sensitive to the nonlinear dynamics governing DA and MA, and (3) is sufficiently smooth and repeatable to effectively guide an optimizer under operational constraints. Traditional chaos indicators, such as Lyapunov exponents [1–3], forward-reverse integration error [4,5], and tangent map methods [6,7], are powerful in simulation but are not readily quantifiable in real-world accelerators.

This paper presents the experimental validation of a data-driven chaos indicator (DDCI), originally introduced and demonstrated offline using simulated data in Ref. [8]. The method exploits the fundamental property that chaos

limits predictability. When beam motion is represented by turn-by-turn trajectories, predictability can be quantified by training a surrogate model to learn the one-turn map from beam position monitor (BPM) data and evaluating its out-of-sample prediction performance. For a fixed surrogate architecture and data quality, reduced predictability reflects increased chaos in the underlying system. This behavior can be leveraged as an objective for optimization by tuning the available control knobs.

Consider the transverse one-turn map of the storage ring,

$$X_{n+1} = M(X_n), \quad X = (x, x', y, y'),$$

where M denotes the effective one-turn transformation at the observation point and X_n represents the transverse phase-space coordinate at turn n with the closed-orbit components subtracted. In an ideal symplectic description, the map M is deterministic and measure-preserving for single-particle motion. Chaos is then characterized by exponential sensitivity to initial conditions, meaning that small differences in initial conditions can grow exponentially over time. In the linear regime, two close initial conditions, separated by ΔX_n , propagate linearly:

$$\Delta X_{n+1} \approx J_M(X_n) \Delta X_n,$$

where J_M is the Jacobian of M . However, for chaotic trajectories, the deviation evolves approximately as

$$\|\Delta X_{n+1}\| \sim e^{\lambda \tau} \|\Delta X_n\|,$$

with λ denoting a finite-time Lyapunov exponent (with τ the revolution time of electrons in storage rings). Even without explicitly computing λ , this exponential growth

^{*}Contact author: yli@bnl.gov

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

implies a limited prediction window beyond which accurate prediction of the trajectory becomes unreliable.

Under real machine conditions, the observed map also includes measurement and dynamical effects. When we attempt to predict X_{n+1} from X_n using a surrogate $\hat{M}$ trained from data, the one-turn prediction error can be approximately decomposed into three contributions:

$$\begin{aligned} e_{n+1} &\equiv X_{n+1} - \hat{M}(X_n) \\ &= \underbrace{[M(X_n) - \hat{M}(X_n)]}_{\text{model bias}} + \underbrace{r_{n+1}}_{\text{random effects}} + \underbrace{d_{n+1}}_{\text{slow drifts}}, \quad (1) \end{aligned}$$

where r_{n+1} represents measurement noise, for example, fluctuations in BPM readings, radiation damping, and decoherence. d_{n+1} describes slow drifts, such as changes in orbit or optics over the evaluation window. Under the present experimental conditions, measurement noise and slow drift are mitigated by high-precision beam diagnostics and a confined acquisition window, while the surrogate structure is fixed across evaluations. The National Synchrotron Light Source II (NSLS-II) BPM electronics are thermally regulated ($\pm 0.1^\circ\text{C}$) and incorporate advanced electronic stabilization, providing a $1.0\ \mu\text{m}$ beam position resolution [9,10]. In addition, each DDCI evaluation is completed within approximately 60 min, minimizing the impact of slow orbit and optics drifts. For optimization, variations in prediction error are therefore dominated by the underlying nonlinear dynamics, allowing predictability to serve as a practical proxy for chaos.

When the oscillation amplitude is sufficiently small, nearby points in phase space evolve smoothly, so the map M can be well approximated by a surrogate of fixed complexity, leading to a small bias. However, as the amplitude increases, the map exhibits stronger chaos, and a progressively more complex surrogate is needed. If we insist on using a fixed-complexity surrogate, chaotic motion can be characterized as irreducible error.

The use of predictability as a practical chaos indicator can be further justified by the exponential amplification of uncertainty and finite data resolution. In chaotic systems, small uncertainties grow exponentially over time, quickly limiting predictive accuracy. At the same time, finite data make learning more difficult, especially as nearby points in phase space diverge and follow different dynamics. As a result, reduced predictability reflects stronger chaotic behavior. Under fixed surrogate class, data acquisition, and noise conditions, variations in prediction error primarily reflect changes in the underlying dynamical complexity rather than modeling artifacts.

Although predictability is not equivalent to a Lyapunov exponent, the DDCI is not intended to provide a rigorous measure of chaos. Rather, it serves as an empirical indicator of the predictability of the underlying dynamics under a fixed surrogate complexity. In storage rings, local

nonlinear distortions captured by the one-turn map are strongly correlated with global stability properties such as the dynamic aperture. Consequently, reduced predictability provides a practical objective for online nonlinear optimization.

The above considerations establish predictability as a practical and measurable chaos indicator. We now describe its experimental implementation. The experiment is conducted on the NSLS-II storage ring, a 3 GeV electron storage ring comprising 30 cells with a double-bend achromat lattice [11]. While its baseline nonlinear lattice was optimized during the design stage, the operational machine exhibits residual nonlinear performance limitations. In this study, six families of harmonic sextupoles are used as tuning knobs to suppress chaos without altering the linear chromaticity.

For each sextupole configuration evaluated by the optimizer, the beam is excited with a pinger magnet to four progressively increasing amplitudes in the transverse (x - y) plane to probe nonlinearities. At each amplitude, three repeated excitations are performed to assess repeatability and provide multiple datasets for training and testing statistics. Then a dedicated BPM pair separated by a pure drift is used to acquire turn-by-turn beam positions [12], which are then used to construct the four-dimensional phase-space coordinate after nonlinear calibration [13]. Instead of generating many independent trajectories, successive turns from a single excitation are used to form input-output pairs. The main reason is that the measurement can be completed within a finite operational time window. In the meantime, the slow drift in Eq. (1) can be neglected.

The surrogate is obtained by training on a randomly selected 90% subset of input-output pairs, while the remaining pairs are used for out-of-sample prediction testing. The prediction performance is quantified as a chaos indicator to drive the optimizer to adjust the sextupole knobs. To assess robustness against sampling variance, each dataset is randomly shuffled for cross validation. The complete closed-loop optimization procedure is illustrated in Fig. 1.

As discussed in Ref. [8], the DDCI measures the predictability of the underlying dynamics rather than the performance of a particular surrogate model, provided that the model complexity is kept fixed. As illustrated in Fig. 2, three surrogate models—a Gaussian process (GP) [14], a seventh-order polynomial, and a neural network (NN)—all exhibit the same qualitative behavior: prediction errors remain small at low amplitudes but increase as the motion becomes more nonlinear. This growing irreducible prediction error forms the basis of the DDCI, while the consistent trends across different surrogate models demonstrate its model independence.

In our beam experiment, we employ a GP surrogate with an RBF kernel, whose hyperparameters, i.e., the amplitude

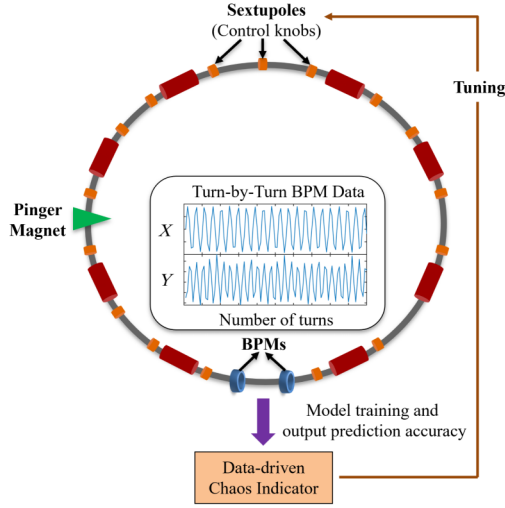


FIG. 1. Schematic of the online optimization loop using DDCI. For each sextupole setting, the stored beam is excited by the pinger magnet and turn-by-turn data are recorded by a BPM pair. The reconstructed phase-space samples form one-turn input-output pairs to train a fixed-structure surrogate. The surrogate's out-of-sample predictive accuracy defines the DDCI objective. Bayesian optimization proposes new sextupole settings to minimize the DDCI, and the loop repeats until convergence.

variance, length scale, and parameter bounds, are held constant throughout the optimization. The DDCI objective, for a given excitation amplitude, is quantified by the standard deviation of the prediction errors on out-of-sample test sets,

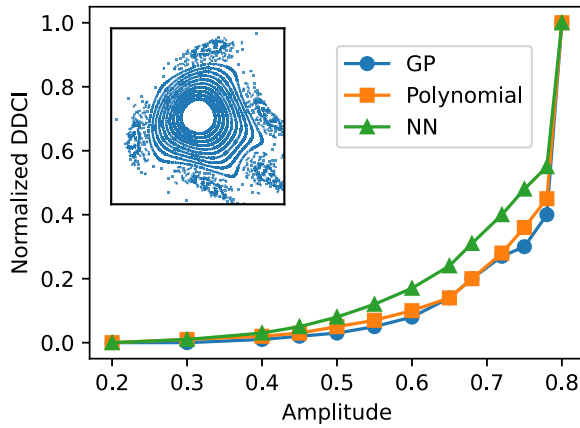


FIG. 2. Illustration of the DDCI principle using three fixed-complexity surrogate models. For the sextupole-driven Hénon map (inset), the surrogate accurately reproduces the turn-by-turn dynamics at small amplitudes, resulting in a small out-of-sample prediction error. As the amplitude increases, the same surrogate becomes progressively unable to capture the increasingly non-linear dynamics, leading to larger prediction errors and, consequently, a larger DDCI. For comparison, the DDCI obtained with each surrogate model is independently normalized to the interval $[0, 1]$, highlighting its amplitude dependence across the three models.

$$E_q = \sqrt{\frac{1}{N} \sum_{j=1}^N \Delta q_j^2}, \quad \Delta q = q_{\text{meas.}} - q_{\text{pred.}},$$

where N denotes the total number of test sets, and $q \in X$. To suppress chaos across amplitudes and mitigate scale differences, for each sextupole configuration, its DDCI is normalized by a baseline value $E_{q,0}$, which is measured with the initial sextupole settings. The total objective minimized by the optimizer is

$$E = \sum_{i=1}^4 \sum_{q \in X} \left(\frac{E_q}{E_{q,0}} \right)_i,$$

where $i = 1, \dots, 4$ denotes four progressively increasing pinger excitation amplitudes, uniformly spanning the ranges $[0.5, 1.2]$ kV in the horizontal plane and $[0.3, 0.6]$ kV in the vertical plane, as illustrated in Fig. 3. Lower E indicates higher predictability, i.e., more regular beam motion. Equal weighting is adopted to avoid bias toward any single phase-space coordinate and to provide a balanced measure of the overall predictability. It is worth noting that the BPM position resolution ($< 1 \mu\text{m}$) is more than an order of magnitude smaller than the measured prediction errors (10–50 μm), indicating that the DDCI is dominated by the underlying beam dynamics rather than BPM measurement noise.

A Bayesian optimizer was adopted due to its suitability for optimizing noisy, expensive-to-evaluate objective functions under a limited evaluation budget. The optimizer iteratively updates a probabilistic objective function using

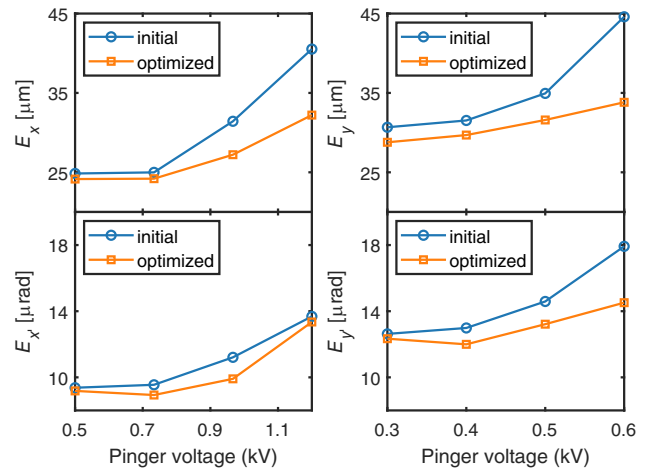


FIG. 3. Mean standard deviation of the prediction errors for (x, x', y, y') at four excitation amplitudes (pinger voltages) for the initial sextupole configuration and the optimized one. Prediction errors are computed on out-of-sample test sets with a 90%/10% split. The optimized setting reduces the DDCI across amplitudes, with relatively larger reduction at higher excitation where nonlinear effects are stronger.

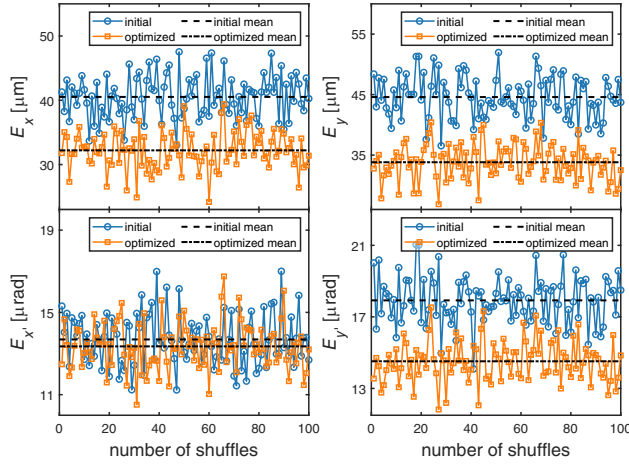


FIG. 4. Cross validation with 100 random shuffles for the largest excitation amplitude (1.2 kV horizontal, 0.6 kV vertical). The optimized configuration consistently exhibits reduced standard deviation of the prediction errors across phase-space coordinates. Variations among individual coordinates reflect differences in sensitivity to nonlinear dynamics and measurement noise.

previous measurements and proposes new sextupole settings by balancing exploration of uncertain regions and exploitation of promising solutions [15]. Each optimization epoch typically required approximately 40 evaluations, corresponding to about 1 h of machine time. This duration aligns well with routine user-facility operations, during which slow system drift is also negligible. After optimization, the prediction performance is improved across the four pinger excitations, as shown in Fig. 3, which is consistent with stronger sensitivity to nonlinearities at larger oscillation amplitudes.

To confirm that the observed reduction in the DDCI is not an artifact of a specific data partition, but instead reflects systematic chaos suppression, cross validation, as shown in Fig. 4, was adopted. While most phase-space coordinates exhibit a clear reduction in predictive uncertainty, one component x' shows weaker improvement and larger statistical fluctuations. This is acceptable, as the optimization minimizes an aggregate objective rather than individual coordinates, and different phase-space directions can have varying sensitivity to sextupole tuning and measurement noise. Despite this, the overall reduction in the aggregated DDCI and the independently observed improvement in DA, as demonstrated in the following measurements, confirm the effectiveness of the optimization.

With the optimized configuration, an enlarged DA is confirmed by a direct measurement. In this measurement, the pinger kick strength is progressively increased until the stored beam is completely lost from the DA. Figure 5 compares the beam-loss patterns for the initial and optimized configurations, where I/I_0 represents the fraction of beam current remaining relative to the initial current after being kicked. Under the same measurement

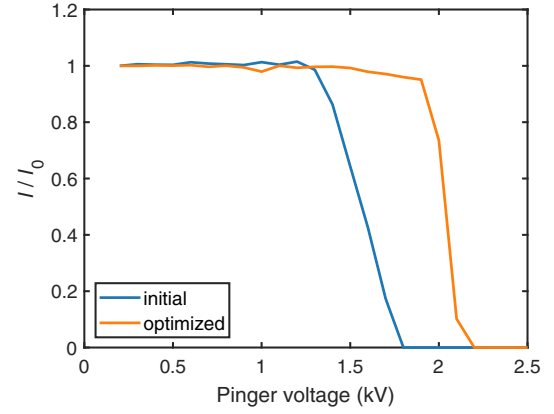


FIG. 5. Measurement of the horizontal DA with the initial/optimized sextupole configurations. The stored beam is kicked with progressively increasing pinger voltage until beam loss occurs. The fraction of remaining beam current after being kicked, I/I_0 , is plotted as a function of pinger voltage. The optimized configuration can tolerate larger kicks, indicating an enlarged DA.

protocol, the optimized setting sustains larger kick amplitudes before beam loss occurs, demonstrating an obvious enlargement of the horizontal DA. In line with the improvement in DA, the off-axis injection efficiency—formerly adopted as the optimization objective [16,17]—improves from about 70% to above 90%. The beam lifetime remains similar under the same beam filling conditions, suggesting that the optimization improves transverse DA without degrading MA.

The online optimization was repeated under multiple operational configurations, including various insertion device (ID) gap settings, both with IDs open and with all IDs engaged. The latter represents a particularly challenging operating condition, as the IDs not only distort the linear optics and break the machine symmetry but also reduce the DA through residual multipole field components. Since user operations are routinely performed with IDs engaged, effective and efficient online nonlinear optimization under these conditions is essential [16]. Injection efficiency improved in all tested machine configurations (including all IDs open and all IDs engaged), demonstrating the robustness and operational reliability of this accelerator-model-independent approach.

We note the differences between using DDCI in simulation [8] and measurement. In offline optimization with simulation data, the one-turn map is obtained from a high-fidelity accelerator model, and a large ensemble of independent trajectories can be generated in parallel across phase space. The surrogate is trained on abundant synthetic data with controlled initial conditions, and chaos arises purely from the dynamical system itself. In contrast, online optimization uses measurement data that include random errors, a limited data pool, and a finite time

window to complete each optimization loop. Therefore, although the theoretical foundation of DDCI is shared between offline and online contexts, the data acquisition strategy and robustness analysis differ substantially.

In summary, we demonstrated online optimization of nonlinear lattice performance at NSLS-II using a DDCI derived from the predictability of turn-by-turn beam motion. Unlike the previously proposed Shannon entropy indicator [18], which requires phase-space transformations, DDCI relies directly on turn-by-turn data. The timescale of each optimization epoch is compatible with operational constraints. While demonstrated on NSLS-II, the method is expected to be transferable to other facilities with turn-by-turn diagnostics owing to its model independence. Further validation on other facilities would be valuable.

Acknowledgments. This work is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, under Contract No. DE-SC0012704 and Field Work Proposal No. 2025-BNL-PS040.

Data availability. The data that support the findings of this study are available upon reasonable request and subject to standard U.S. National Laboratory data-sharing policies.

-
- [1] S. Habib and R.D. Ryne, Symplectic calculation of Lyapunov exponents, *Phys. Rev. Lett* **74**, 70 (1995).
 - [2] F. Schmidt, F. Willeke, and F. Zimmermann, Comparison of methods to determine long-term stability in proton storage rings, *Part. Accel.* **35**, 249 (1991).
 - [3] W. Fischer, An Experimental Study on the Long-Term Stability of Particle Motion in Hadron Storage Rings, Ph.D. thesis, DESY, 1995.
 - [4] Y. Li, Y. Hao, K. Hwang, R. Rainer, A. He, and A. Liu, Fast dynamic aperture optimization with forward-reversal integration, *Nucl. Instrum. Methods Phys. Res., Sect. A* **988**, 164936 (2021).
 - [5] F. Panichi, K. Goździewski, and G. Turchetti, The reversibility error method (REM): A new, dynamical fast indicator for planetary dynamics, *Mon. Not. R. Astron. Soc.* **468**, 469 (2017).
 - [6] C. Skokos, Alignment indices: A new, simple method for chaos detection, *J. Phys. A:Math. Gen.* **34**, 10029 (2001).
 - [7] J. Qiang, J. Wan, A. Qiang, and Y. Hao, Fast chaos indicator from auto-differentiation for dynamic aperture optimization, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1087**, 171427 (2026).
 - [8] Y. Li, J. Wan, A. Liu, Y. Jiao, and R. Rainer, Data-driven chaos indicator for nonlinear dynamics and applications on storage ring lattice design, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1024**, 166060 (2022).
 - [9] S. Sharma, L. Doom, A. Jain, P. Joshi, F. Lincoln, and V. Ravindranath, Optimization of magnet stability and alignment for NSLS-II, in *Proceedings of the 2011 Particle Accelerator Conference (PAC'11)* (JACoW Publishing, Geneva, Switzerland, 2011), pp. 2082–2086.
 - [10] W. X. Cheng, K. Ha, J. Mead, B. Podobedov, O. Singh, Y. Tian, and L. Yu, Characterization of NSLS-II storage ring beam orbit stability, in *Proceedings of the 4th International Beam Instrumentation Conference (IBIC2015)* (JACoW, Melbourne, Australia, 2015), pp. 625–629.
 - [11] S. Dierker, *NSLS-II Preliminary Design Report* (Upton, New York, USA, 2007).
 - [12] Y. Li, K. Ha, D. Padrazo, B. Kosciuk, B. Bacha, M. Seegitz, R. Rainer, J. Mead, X. Yang, Y. Tian, R. Todd, V. Smaluk, and W. Cheng, Dedicated beam position monitor pair for model-independent lattice characterization at NSLS-II, *Nucl. Instrum. Methods Phys. Res., Sect. A* **1065**, 169557 (2024).
 - [13] W. Cheng, B. Bacha, and O. Singh, NSLS2 beam position monitor calibration, in *Proceedings of the 15-th Beam Instrumentation Workshop* (JACoW Publishing, Geneva, Switzerland, 2012).
 - [14] C. Williams and C. Rasmussen, *Gaussian Processes for Machine Learning* (MIT Press, Cambridge, MA, 2006), Vol. 2.
 - [15] R. Roussel *et al.*, Bayesian optimization algorithms for accelerator physics, *Phys. Rev. Accel. Beams* **27**, 084801 (2024).
 - [16] X. Huang, *Beam-Based Correction and Optimization for Accelerators* (CRC Press, Boca Raton, Florida, USA, 2019).
 - [17] Z. Zhang, M. Song, and X. Huang, Online accelerator optimization with a machine learning-based stochastic algorithm, *Mach. Learn.: Sci. Technol.* **2**, 015014 (2020).
 - [18] Y. Li, K. Anderson, D. Xu, Y. Hao, K. Ha, Y. Hidaka, M. Song, R. Rainer, V. Smaluk, and T. Shaftan, Online regularization of Poincaré map of storage rings with Shannon entropy, *Phys. Rev. Accel. Beams* **28**, 034001 (2025).