

Single-waveform time-series modeling via kernel-based dynamic mode decomposition for real-time dynamics reconstruction

Faya Wang^{✉*}

SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA



(Received 6 November 2025; accepted 21 May 2026; published 31 July 2026)

Real-time monitoring and predictive modeling of complex accelerator subsystems demand data-efficient approaches that can capture essential dynamics from minimal measurements. This study introduces a framework based on kernel-based dynamic mode decomposition that constructs surrogate dynamical models from a single reference waveform, enabling accurate time-series reconstruction with substantially reduced data requirements compared to conventional methods. The approach formulates the system dynamics as a continuous ordinary differential equation (ODE) in a delay-embedded phase space, where a sparse dictionary and polynomial kernel provide both computational efficiency and physics-aligned interpretability through the underlying ODE structure. The method is validated on three beam kicker systems and two pulse-forming network (PFN) modulator systems at the Stanford Synchrotron Radiation Lightsource (SSRL) at SLAC, demonstrating accurate waveform reconstruction using only one training waveform per system. A comparative study against exponential smoothing and linear trend models confirms the method's superior predictive capability, particularly in extrapolative tail regions. These results establish a lightweight, scalable foundation for real-time waveform monitoring, anomaly detection, and potential integration into broader facility-level digital twin frameworks in accelerator science and beyond.

DOI: [10.1103/vbnq-lj9j](https://doi.org/10.1103/vbnq-lj9j)

Introduction. Real-time monitoring of complex engineering systems requires models that accurately capture system dynamics while remaining computationally tractable. For particle accelerators, subsystems such as beam kickers and pulse-forming network (PFN) modulators exhibit intricate pulsed dynamics that are critical for beam control and must be monitored continuously to ensure operational reliability. Traditional physics-based modeling of these subsystems requires detailed knowledge of component parameters and is often computationally intensive, making it impractical for real-time applications.

Data-driven approaches have seen significant development within the accelerator community, including surrogate models, predictive models, and optimization algorithms [1–8]. These methods leverage neural networks, Gaussian processes, and Bayesian optimization for efficient modeling but typically require substantial training datasets. A complementary, data-efficient alternative is desirable for scenarios where only limited waveform data are available.

Dynamic mode decomposition (DMD), introduced by Schmid [9], provides a powerful data-driven framework for extracting dynamic modes from time-series data without prior system assumptions. Kernel-based extensions of DMD [11] map the data into higher-dimensional feature spaces to accommodate nonlinear dynamics, while sparse dictionary methods [12,13] enhance computational efficiency. The LANDO linear and nonlinear disambiguation optimization (LANDO) framework [13] combines these ideas into a unified algorithm for learning nonlinear dynamics from data.

This study applies the kernel-based DMD framework (LANDO) to construct surrogate dynamical models for accelerator subsystems using only a single reference waveform per system. By formulating the problem as a continuous ordinary differential equation (ODE) in a delay-embedded phase space [14], the approach captures the essential dynamics from minimal data while maintaining physical interpretability through the ODE structure. The key contributions are as follows: (i) demonstration that a single pulsed waveform provides sufficient information for accurate dynamics reconstruction in accelerator kicker and modulator systems; (ii) systematic validation across five different subsystems at SLAC spanning three years of operational data; and (iii) a comparative study against alternative time-series methods confirming the approach's effectiveness. While the present work focuses on unidirectional monitoring (physical system to model), the

*Contact author: fywang@slac.stanford.edu

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.

lightweight surrogate models produced here could serve as building blocks for broader, facility-level digital twin efforts in the future [15,16].

Methodology.

Signal preprocessing: The raw time-series data from kicker and modulator systems consist of pulsed waveforms that require preprocessing to ensure uniformity before DMD analysis. Let $x_{\text{raw}}(t)$ denote a raw waveform of length N . The preprocessing pipeline consists of three steps applied sequentially.

Step 1: dc offset removal. The baseline offset is estimated from the first N_{off} samples (typically $N_{\text{off}} = 10$), where no pulse activity is present:

$$x_{\text{cen}}(t) = x_{\text{raw}}(t) - \frac{1}{N_{\text{off}}} \sum_{i=1}^{N_{\text{off}}} x_{\text{raw}}(t_i). \quad (1)$$

Step 2: Amplitude normalization. Signals are normalized using the mean amplitude within the full width half maximum (FWHM) region. Let $A_{\text{max}} = \max_t |x_{\text{cen}}(t)|$. The FWHM region is defined as $\mathcal{S} = \{t: |x_{\text{cen}}(t)| > A_{\text{max}}/2\}$, and the normalization factor is

$$\bar{A} = \frac{1}{|\mathcal{S}|} \sum_{t \in \mathcal{S}} x_{\text{cen}}(t), \quad x_{\text{norm}}(t) = \frac{x_{\text{cen}}(t)}{|\bar{A}|}. \quad (2)$$

Step 3: Temporal alignment. Waveforms are aligned to a reference signal x_{ref} via cross-correlation. The optimal lag τ^* is determined by

$$\tau^* = \arg \max_{\tau} \sum_t x_{\text{ref}}(t) x_{\text{norm}}(t - \tau), \quad (3)$$

and the aligned waveform is obtained by shifting x_{norm} by τ^* samples, with zero padding applied to vacated regions.

For the results presented here, the most recent waveform in each dataset serves as the reference. These preprocessing parameters (offset region, FWHM normalization constant, and alignment lag) are computed independently for each new waveform using the fixed reference, which is important for real-time deployment: only one stored reference waveform is needed, and the per-waveform preprocessing involves only subtraction, division, and a cross-correlation, all of which are readily implementable on field-programmable gate array (FPGA) hardware.

Dynamic mode decomposition framework: The system dynamics are modeled as a continuous ODE:

$$\frac{d\mathbf{X}}{dt} = F(\mathbf{X}), \quad (4)$$

where $\mathbf{X}(t) \in \mathbb{R}^D$ is the system state in an embedded phase space, and F is the unknown dynamics function. The continuous ODE formulation reflects the inherently smooth evolution of physics-based systems such as kickers (driven

by electromagnetic fields and current flow) and modulators (governed by RLC circuit dynamics) [10], avoiding discretization artifacts while supporting flexible numerical integration.

The dynamics function F is approximated using a kernel-based surrogate:

$$\frac{d\mathbf{X}}{dt} \approx \mathbf{Y} = \mathbf{W} k(\mathbf{X}_d, \mathbf{X}), \quad (5)$$

where $\mathbf{Y} = d\mathbf{X}/dt$ is the time derivative (computed via numerical gradient of the training waveform), $\mathbf{W}$ is the trained weight matrix, $k(.,.)$ is a kernel function, and $\mathbf{X}_d$ is a sparse dictionary of representative states. Equation (5) is an approximation; the “ $\approx$ ” reflects that $\mathbf{Y}$ is a data-driven surrogate for the true dynamics $F(\mathbf{X})$.

Phase-space construction via delay embedding: A single scalar time series $V(t)$ is insufficient to represent the full system state. Phase-space reconstruction via Takens’ delay embedding theorem [14] constructs a higher-dimensional state from delayed copies of the signal:

$$\mathbf{X}(t) = \begin{bmatrix} V(t) \\ V(t + t_g) \\ \vdots \\ V(t + (D-1) \cdot t_g) \end{bmatrix} \in \mathbb{R}^D, \quad (6)$$

where t_g is the embedding lag (in samples), and D is the embedding dimension. Takens’ theorem guarantees that, for a sufficiently large D (specifically $D > 2d$, where d is the dimension of the underlying attractor), this embedding is a diffeomorphism of the original attractor, preserving its topological properties. For the accelerator waveforms studied here, the pulsed nature of the signals provides inherent richness: the rising edge, peak, and decaying tail traverse a substantial portion of the system’s state space within a single pulse, providing the diversity of states needed for accurate model identification. This is analogous to the persistence of excitation condition in system identification [17]: the training signal must explore a sufficiently rich region of state space to constrain the model parameters.

For kicker systems, only a current signal is available. For PFN modulator systems, both voltage and current signals are available; the current signal was used for the primary analysis (*Modulator Systems* section), and a comparison of voltage-versus-current modeling is presented in the *Voltage Versus Current Signal Comparison* section.

The embedding parameters t_g and D are system-specific and optimized via grid search to minimize the L_1 loss $\sum_t |V_{\text{true}}(t) - V_{\text{pred}}(t)|$ evaluated on a held-out test waveform. The constraint $t_g \times D \leq N_{\text{max}}$ (with $N_{\text{max}} = 350$ samples for kickers) ensures that the number of initial-condition points required for prediction remains practical.

The L_1 loss is evaluated on the tail region (samples beyond $N_{\max}$), which represents the purely extrapolative prediction where no input data are available.

Sparse dictionary construction and kernel training: The dictionary $\mathbf{X}_d$ is a sparse subset of the phase-space states from the training waveform, selected automatically by the LANDO algorithm [13] during training. Starting from the first state vector, the algorithm sequentially evaluates each new state: if it is approximately linearly dependent in the kernel-induced feature space on the existing dictionary members, it is discarded; otherwise, it is added to the dictionary. Formally, a state $\mathbf{x}_t$ is added to the dictionary if:

$$\delta_t = k(\mathbf{x}_t, \mathbf{x}_t) - \mathbf{k}_t^T \boldsymbol{\pi}_t > \nu, \quad (7)$$

where $\mathbf{k}_t = k(\mathbf{X}_d, \mathbf{x}_t)$ is the vector of kernel evaluations between the dictionary and the candidate, $\boldsymbol{\pi}_t$ solves $\mathbf{C}^T \mathbf{C} \boldsymbol{\pi}_t = \mathbf{k}_t$ (with $\mathbf{C}$ the Cholesky factor of the kernel

matrix), and ν is the sparsification threshold. This almost linearly dependent (ALD) test ensures that the dictionary captures the essential structure of the data while remaining compact. In practice, the dictionary typically retains $\sim 5\%$ – 10% of the original phase-space points (e.g., 26 dictionary elements from 375 phase-space states for K2 modulator current, a sparsity ratio of 6.9%), providing substantial computational savings for prediction.

A polynomial kernel is employed:

$$k(\mathbf{x}, \mathbf{y}) = (1 + \alpha \cdot \mathbf{x}^T \mathbf{y})^p, \quad (8)$$

with $p = 1$ and $\alpha = 0.1$ throughout this study, yielding $k(\mathbf{x}, \mathbf{y}) = 1 + 0.1 \mathbf{x}^T \mathbf{y}$. While $p = 1$ produces a kernel that is linear in the inner product, it is not equivalent to standard linear DMD. The constant offset term (the “1” in the kernel) provides an important bias that standard linear DMD lacks, enabling the model to handle dc components and baseline variations in the signal naturally. Furthermore, the kernel

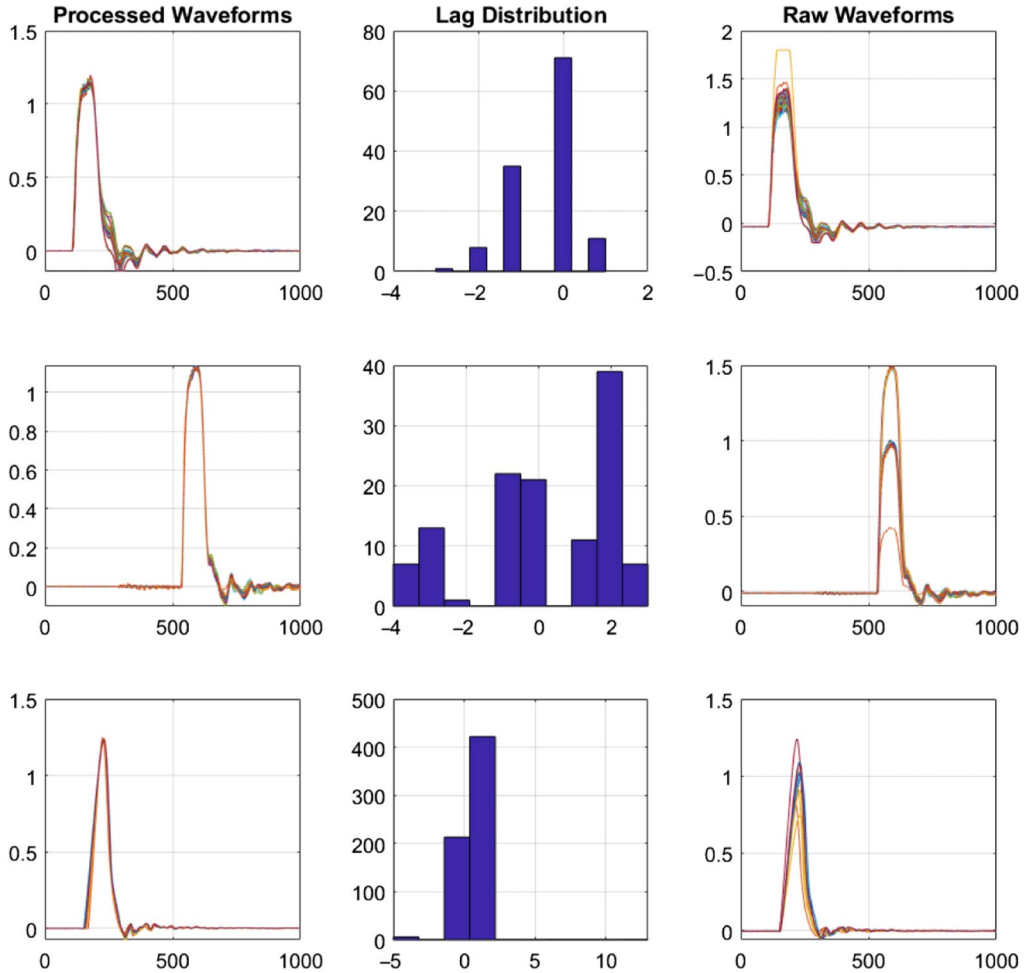


FIG. 1. SSRL kicker system waveforms. Raw waveforms are displayed on the right (top: booster extraction kicker, middle: booster injection kicker, bottom: SPEAR3 injection kicker). These waveforms are sampled at 100 MHz, each spanning 10 μs . The left side presents the processed waveforms after alignment and normalization. The middle column shows the lag distribution for each waveform set, obtained through the alignment process using the most recent waveform as reference.

acts in the delay-embedded phase space (dimension $D = 25\text{--}55$), so the resulting model captures dynamics that are nonlinear in the original scalar signal space. This choice balances model expressiveness with computational efficiency and avoids overfitting when training on a single waveform. Higher-order kernels ($p > 1$) or Gaussian kernels could be employed for systems with stronger nonlinearities, though the polynomial $p = 1$ kernel proved sufficient for all systems studied here.

The weight matrix $\mathbf{W}$ is obtained by solving the regularized least-squares problem:

$$\mathbf{W} = \arg \min_{\mathbf{W}} \|\mathbf{Y} - \mathbf{W} k(\mathbf{X}_d, \mathbf{X})\|_2, \quad (9)$$

via the pseudoinverse of $k(\mathbf{X}_d, \mathbf{X})$ computed using SVD with a regularization cutoff of 10^{-7} for numerical stability [10]. An ensemble of $N_{\text{ens}} = 5$ models is trained, each on a different random permutation of the training data columns, and predictions are averaged to improve robustness.

Prediction via ODE integration: With the trained weight matrix $\mathbf{W}$ and dictionary $\mathbf{X}_d$, predictions for a new waveform are generated as follows: (i) The new waveform is preprocessed (*Signal Preprocessing* section). (ii) The first D samples at intervals of t_g provide the initial condition $\mathbf{X}(0)$ via Eq. (6). (iii) The ODE $d\mathbf{X}/dt = \mathbf{W} k(\mathbf{X}_d, \mathbf{X})$ is integrated forward in time. (iv) The multirow prediction matrix is assembled into a single time series by taking the first row as the primary signal and overwriting subsequent segments with appropriately shifted later rows.

This procedure requires only $N = D$ input points (the initial condition) to reconstruct the full waveform, enabling

TABLE I. Optimized embedding parameters for kicker systems.

Kicker system	Lag t_g	Dim D	Dict. size
Booster extraction	10	35	22
Booster injection	12	29	18
SPEAR injection	14	25	20

applications such as: (i) Real-time monitoring: Continuous tracking of waveform shape for drift or degradation. (ii) Anomaly detection: Elevated L_1 loss between prediction and measurement flags abnormal behavior. (iii) Predictive simulation: Forecasting future states from partial initial measurements.

Results.

Kicker systems: Three kicker systems at the SSRL are analyzed: the booster extraction kicker, the booster injection kicker, and the SPEAR3 Injection kicker. Each system provides current waveforms sampled at high rate (~ 400 samples per pulse), as shown in Fig. 1. The dataset spans approximately three years of operation (2021–2024), with several hundred waveforms per system.

For each kicker, the most recent waveform serves as the training (reference) signal, and the second-most-recent waveform is used for hyperparameter optimization. The embedding parameters (t_g, D) are optimized via grid search over $t_g \in [8, 20]$ and $D \in [20, 55]$, subject to the constraint $t_g \times D \leq 350$. Table I summarizes the optimized parameters.

Figure 2 shows the prediction results for all three kicker systems. In each panel, the black dots indicate

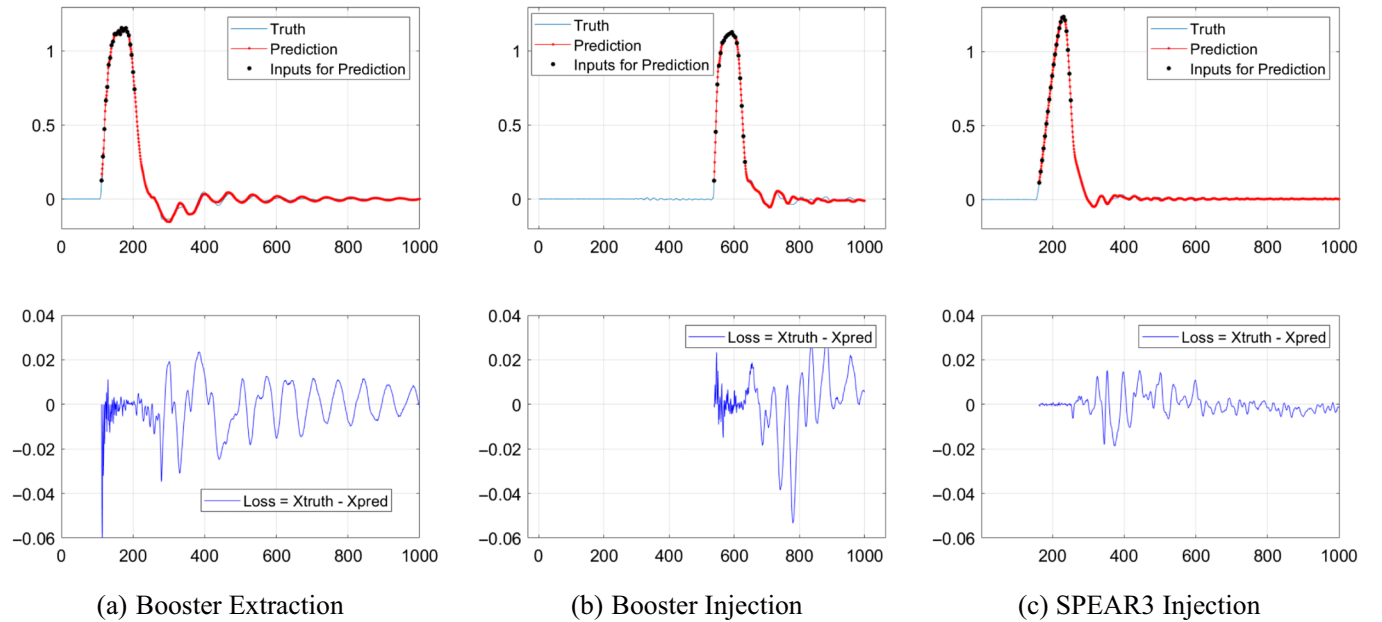


FIG. 2. Kicker system predictions obtained by applying the initial conditions (black dots) to the learned ODE [Eq. (4)]. Red: prediction; Blue: measured truth. All three systems show excellent agreement between the model and measurement.

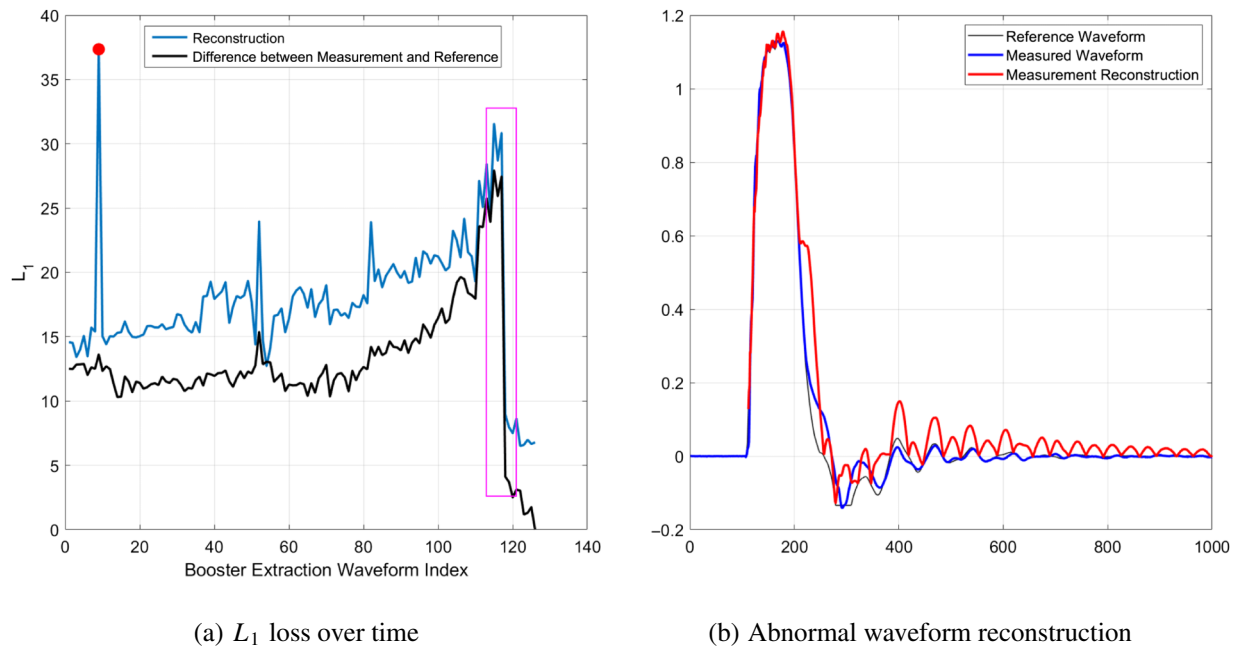


FIG. 3. (a) Booster extraction kicker L_1 loss over three years. The loss jump within the pink box corresponds to a hardware modification; the red dot marks the largest anomaly. (b) Reconstruction of the anomalous waveform, showing the shape distortion detected by the model.

the D initial-condition points extracted from the test waveform, and the red curve shows the model prediction obtained by integrating Eq. (4) forward from these initial conditions. The blue curve is the measured truth. The model accurately captures the pulse shape, including the rising edge, peak amplitude, and decaying tail, across all three systems.

Long-term monitoring and anomaly detection. To assess the model's ability to track system changes over time, the trained model is applied to all waveforms in the booster extraction dataset spanning three years. Figure 3(a) shows the L_1 loss for each waveform as a function of time. The

loss remains low and stable for normal operation, with a clear jump within the pink-shaded region corresponding to a known hardware modification. The red dot marks the waveform with the largest loss spike. Figure 3(b) shows the reconstruction of this abnormal waveform, revealing a shape distortion that the model correctly identifies as anomalous. This demonstrates the framework's utility for long-term monitoring: systematic drift or abrupt hardware changes are reflected in elevated L_1 loss, while normal operational variations produce consistently low loss values. The baseline L_1 loss of approximately 15 (in normalized

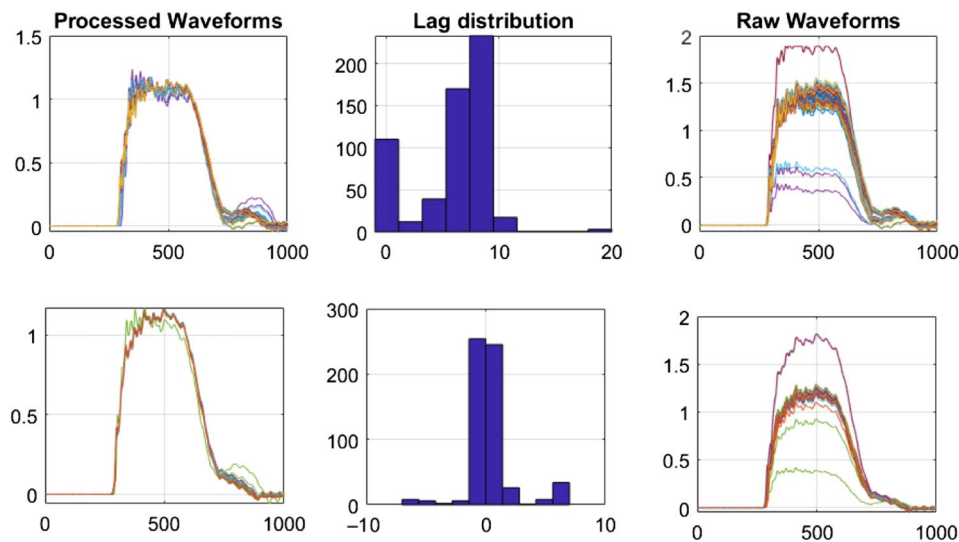


FIG. 4. SSRL Linac PFN modulator current waveforms for K2 (top) and K3 (bottom), sampled at 100 MHz. Left: processed (aligned and normalized) waveforms. Middle: lag distribution from the alignment process. Right: raw waveforms.

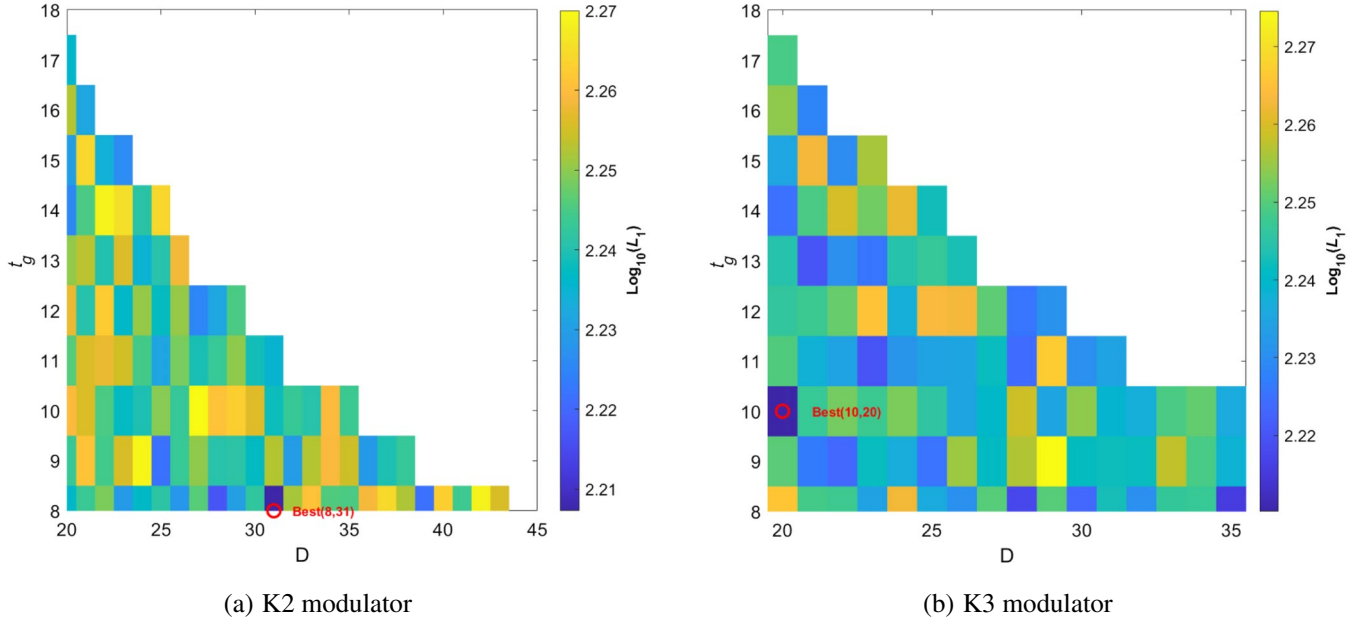


FIG. 5. 2D heatmaps of $\log_{10}(L_1)$ loss as a function of embedding lag t_g and dimension D for the (a) K2 and (b) K3 modulator current signals. Red circles indicate the optimal (t_g, D) pairs.

units) represents the combined effect of measurement noise, minor pulse-to-pulse jitter, and residual alignment imperfections, model uncertainty, and serves as a practical threshold for anomaly detection.

Modulator systems: Two PFN modulator systems (K2 and K3) on the SSRL Linac are analyzed. These systems provide both current and voltage waveforms sampled at 100 MHz. Figure 4 shows the raw and processed current waveforms for both modulators, along with the lag distributions obtained during alignment.

The embedding parameters are optimized as described for the kickers, with the results shown in Figs. 5(a) and 5(b) as 2D heatmaps of $\log_{10}(L_1)$ over the (t_g, D) grid. The red circles mark the optimal parameter sets. These heatmaps are computed using a single training waveform and a single test waveform; the smooth landscape indicates that the optimization is not sensitive to the specific training waveform chosen, as verified by repeating the scan with different training waveforms, which produce qualitatively similar landscapes.

Figures 6(a) and 6(b) show the prediction results for the K2 and K3 modulator current waveforms, respectively. As with the kickers, the model achieves excellent agreement with the measured waveforms.

Voltage versus current signal comparison For the PFN modulator systems, both voltage and current signals are available. To assess whether the choice of signal affects reconstruction quality, the same modeling procedure was applied to the voltage waveforms of both K2 and K3 modulators. Table II summarizes the results. The voltage

signal yields significantly lower reconstruction error for both modulators (approximately $40\times$ lower MSE), with correspondingly higher correlation. This is likely because the voltage signal has a smoother waveform shape that is better suited to the polynomial kernel and delay embedding, while the current signal contains sharper transients that are more challenging to capture. For systems where both signals are available, the voltage signal is therefore recommended for surrogate model development. However, the current signal results remain highly accurate (correlation > 0.999) and are sufficient for monitoring applications.

Comparison with alternative time-series methods: To validate the claim that a single waveform provides sufficient training data, the kernel-based DMD method is compared against two alternative time-series approaches, all trained on the same single K2 modulator current waveform: (i) Exponential smoothing: A first-order exponential smoothing filter with $\alpha = 0.3$ is applied to the test waveform. This is a causal filter that uses only past values. (ii) Linear trend model: A linear regression is fit to the training waveform and used to predict the test waveform shape.

Table III summarizes the results, with metrics computed on the tail region (samples > 350 , the purely extrapolative region beyond the initial-condition window).

The kernel-based DMD model learns the underlying dynamics and generates predictions by integrating an ODE from initial conditions, making it robust to pulse-to-pulse variations and capable of detecting anomalies (as demonstrated in the *Kicker Systems* section). Among the methods

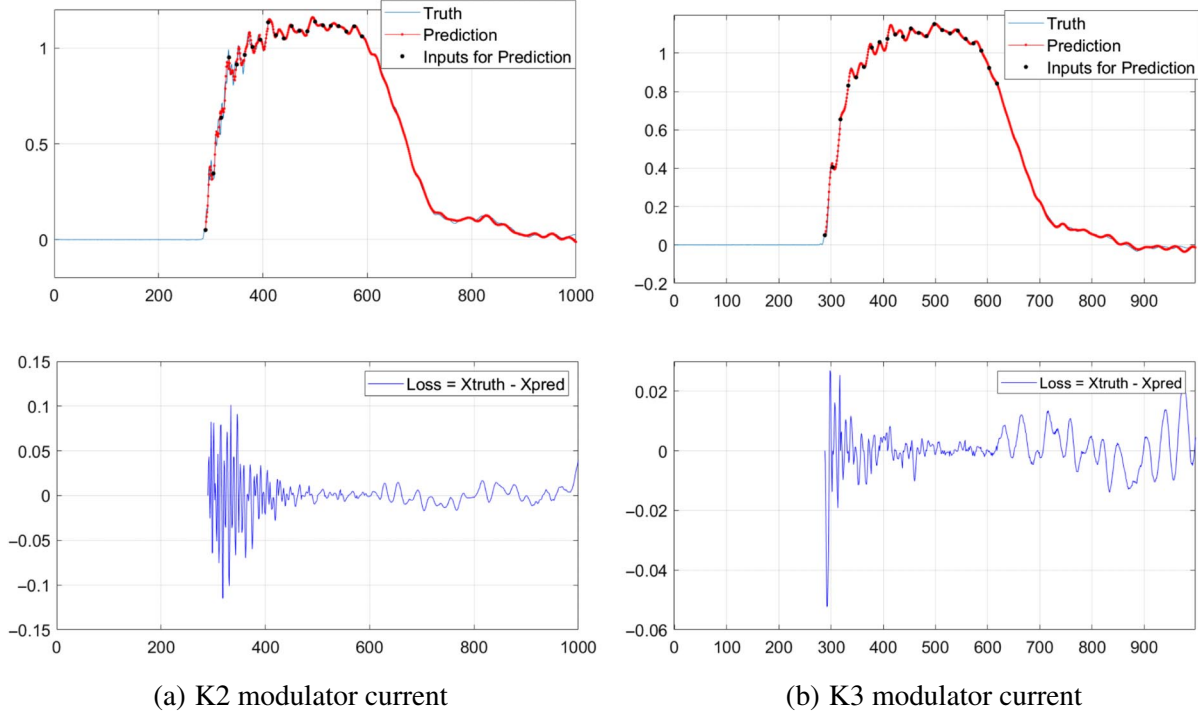


FIG. 6. Reconstruction of modulator current waveforms for (a) K2 and (b) K3. Black dots: initial condition points; Red: prediction from ODE integration; Blue: measured truth.

that genuinely learn from data, the kernel-based DMD outperforms both exponential smoothing ($1.7\times$ lower tail MSE) and the linear trend model ($> 100\times$ lower tail MSE), confirming its effectiveness for single-waveform training.

Discussion. The results demonstrate that the kernel-based DMD framework can accurately reconstruct waveform dynamics across five distinct accelerator subsystems using only a single training waveform per system. Several aspects of these results merit discussion.

Why does single-waveform training work? The success of single-waveform training can be understood through two complementary perspectives. First, Takens' embedding theorem [14] guarantees that a sufficiently rich time series can reconstruct the system's attractor, and the pulsed accelerator waveforms studied here satisfy this requirement: the pulse traverses the full dynamic range of the system (rising edge, peak, decay) within a single shot, effectively exploring the relevant region of state space. This

is analogous to the persistence of excitation condition in classical system identification [17]. Second, the high pulse-to-pulse repeatability of these systems means that a single representative waveform captures the essential dynamics with minimal information loss.

Interpretability of the kernel-based DMD model. The framework provides interpretability at multiple levels: (i) the ODE formulation [Eq. (4)] directly represents the system dynamics as a differential equation, providing a physics-aligned structure; (ii) the dictionary $\mathbf{X}_d$ constitutes a sparse, identifiable set of representative states in phase space; (iii) the weight matrix $\mathbf{W}$ encodes the linear relationship between kernel similarities and the time derivatives, which can be inspected or linearized to extract effective eigenvalues and modes; and (iv) the polynomial kernel $k(\mathbf{x}, \mathbf{y}) = 1 + 0.1 \mathbf{x}^T \mathbf{y}$ is a simple, transparent function. Together, these elements provide a level of interpretability that is absent from black-box approaches such as deep neural networks.

TABLE II. Voltage vs current signal reconstruction comparison for modulator systems.

Modulator	Signal	MSE	Correlation	Rel. RMSE (%)
K2	Current	4.6×10^{-4}	0.9990	4.7
	Voltage	1.1×10^{-5}	0.9999	0.8
K3	Current	8.2×10^{-5}	0.9996	1.9
	Voltage	2.0×10^{-6}	0.9999	0.3

TABLE III. Comparison of time-series methods on K2 modulator current (tail-region metrics, samples > 350).

Method	MSE (Tail)	L_1 (Tail)	Corr. (Tail)	Rank
Kernel-based DMD	4.6×10^{-4}	5.4	0.9980	1
Exp. smoothing	7.8×10^{-4}	7.1	0.9960	2
Linear trend	7.3×10^{-2}	68.5	0.6800	3

Role of the dictionary and kernel choice. The sparse dictionary $\mathbf{X}_d$ serves as a compact basis for the model, retaining only $\sim 5\text{--}10\%$ of the original phase-space states while preserving the essential dynamics. The polynomial kernel with $p = 1$ was found to be sufficient for all systems studied here. While this kernel is linear in the inner product, it operates in the delay-embedded phase space (dimension 25–55), so the resulting model captures dynamics that are nonlinear in the original scalar signal. The constant offset term in the kernel provides an important bias not present in standard linear DMD, improving numerical stability and baseline handling.

Sensitivity of hyperparameter optimization. The 2D heatmaps (Fig. 5) show that the L_1 loss landscape is smooth and contains a well-defined optimum, indicating that the hyperparameter selection is robust. The optimization was performed using a single training and test waveform pair; repeating the scan with different waveform pairs produces consistent optima, confirming that the results are not artifacts of a particular waveform choice.

Preprocessing considerations for real-time deployment. The three preprocessing steps (offset removal, FWHM normalization, temporal alignment) are computationally lightweight and deterministic. For FPGA deployment, the offset removal requires a running average over N_{off} samples; the normalization requires computing the half-maximum threshold and a conditional mean; and the alignment requires a cross-correlation, which can be efficiently implemented using dedicated DSP blocks. The reference waveform and model parameters ($\mathbf{W}$, $\mathbf{X}_d$, kernel coefficients) are stored as constants, requiring minimal on-chip memory.

Limitations. Several limitations of the current framework should be noted: (i) System stationarity assumption: The model assumes that the system dynamics are stationary—i.e., that a waveform captured at one time is representative of the dynamics at all times. The long-term monitoring results (Fig. 3) show that hardware changes produce detectable L_1 loss increases, but the model does not adapt automatically; retraining with a new reference waveform would be required after significant system modifications. (ii) Linear kernel regime: The polynomial kernel with $p = 1$ is effective for the systems studied here, where the dynamics in the embedded phase space are approximately linear. Systems with strongly nonlinear dynamics (e.g., chaotic or multimodal behavior) may require higher-order kernels or Gaussian kernels, at the cost of increased model complexity and potential overfitting with limited training data. (iii) Hyperparameter sensitivity: While the L_1 loss landscape is smooth for the systems studied (Fig. 5), the grid search over (t_g, D) requires evaluation of multiple models. For deployment on new systems, this optimization step must be performed off-line. (iv) Unidirectional implementation: The current implementation is unidirectional, providing monitoring and prediction from physical

measurements to the model. Full bidirectional interaction (model-to-physical-system feedback), as required for a complete digital twin [15,16], is not yet implemented but could be realized by using the L_1 loss signal to trigger control actions. (v) Scope of validation: The method has been validated on pulsed waveforms from kicker and modulator systems. Generalization to other signal types (e.g., continuous rf signals, nonrepetitive transients) has not been tested and may require modifications to the preprocessing or embedding strategy.

Conclusion and future work. This study presents a time-series modeling framework based on kernel-based DMD that constructs surrogate dynamical models from a single reference waveform. Validated on three kicker systems and two PFN modulator systems at the SSRL at SLAC, the approach demonstrates accurate waveform reconstruction with minimal data requirements. The continuous ODE formulation, sparse dictionary representation, and polynomial kernel provide a physically interpretable model that is suitable for real-time monitoring and anomaly detection.

A comparative study confirms that the kernel-based DMD outperforms exponential smoothing and linear trend models when trained on a single waveform, and the long-term monitoring analysis demonstrates the method's ability to detect hardware changes over multiyear timescales. The voltage-versus-current comparison for modulator systems provides practical guidance for signal selection in future deployments.

Future work will focus on (i) integrating the single-waveform surrogate models into broader facility-level monitoring frameworks, where they could serve as components of a comprehensive digital twin architecture; (ii) implementing bidirectional feedback by coupling the L_1 loss signal to automated control actions; (iii) hardware-accelerated deployment on FPGAs for sub-millisecond prediction; and (iv) extending the methodology to non-pulsed signal types and multisignal systems using stacked delay embeddings.

Acknowledgments. This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-76SF00515. The author acknowledges the SSRL Accelerator team for assistance with data access.

Data availability. The data that support the findings of this article are openly available in the GitHub repository SingleWaveReconstructionData, hosted by SLAC National Accelerator Laboratory [18].

-
- [1] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes, and A. Adelman, Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems, *Phys. Rev. Accel. Beams* **23**, 044601 (2020).

- [2] A. A. Mishra, A. Edelen, A. Hanuka, and C. Mayes, Uncertainty quantification for deep learning in particle accelerator applications, *Phys. Rev. Accel. Beams* **24**, 114601 (2021).
- [3] J. Duris, D. Kennedy, A. Hanuka, J. Shtalenkova, A. Edelen, P. Baxevanis, A. Egger, T. Cope, M. McIntire, S. Ermon, and D. Ratner, Bayesian optimization of a free-electron laser, *Phys. Rev. Lett.* **124**, 124801 (2020).
- [4] A. Hanuka, X. Huang, J. Shtalenkova, D. Kennedy, A. Edelen, Z. Zhang, V. R. Lalchand, D. Ratner, and J. Duris, Online tuning and light source control using a physics-informed Gaussian process, *Phys. Rev. Accel. Beams* **24**, 072802 (2021).
- [5] 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* **27**, 054601 (2024).
- [6] R. Roussel *et al.*, Bayesian optimization algorithms for accelerator physics, *Phys. Rev. Accel. Beams* **27**, 084801 (2024).
- [7] A. Scheinker, A. Edelen, D. Bohler, C. Emma, and A. Lutman, Demonstration of model-independent control of the longitudinal phase space of electron beams in the Linac-Coherent Light Source with femtosecond resolution, *Phys. Rev. Lett.* **121**, 044801 (2018).
- [8] C. Emma, A. Edelen, M. J. Hogan, B. O'Shea, G. White, and V. Yakimenko, Machine learning-based longitudinal phase space prediction of particle accelerators, *Phys. Rev. Accel. Beams* **21**, 112802 (2018).
- [9] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, *J. Fluid Mech.* **656**, 5 (2010).
- [10] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems* (SIAM, Philadelphia, 2016), .
- [11] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition, *J. Nonlinear Sci.* **25**, 1307 (2015).
- [12] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3932 (2016).
- [13] P. J. Baddoo, B. Herrmann, B. J. McKeon, and S. L. Brunton, Kernel learning for robust dynamic mode decomposition: Linear and nonlinear disambiguation optimization, *Proc. R. Soc. A* **478**, 20210830 (2022).
- [14] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Warwick 1980*, edited by D. Rand and L. S. Young, Lecture Notes in Mathematics Vol. 898 (Springer, Berlin, 1981), pp. 366–381.
- [15] K. Willcox and B. Segundo, The role of computational science in digital twins, *Nat. Comput. Sci.* **4**, 147 (2024).
- [16] *National Academies of Sciences, Engineering, and Medicine*, Foundational Research Gaps and Future Directions for Digital Twins (National Academies Press, Washington, DC, 2024), 10.17226/26894.
- [17] L. Ljung, *System Identification: Theory for the User*, 2nd ed. (Prentice Hall, 1999).
- [18] F. Wang, SingleWaveReconstructionData, SLAC National Accelerator Laboratory, <https://github.com/slaclab/SingleWaveReconstructionData> (2026).