

Beam-based rf station fault identification at the SLAC Linac Coherent Light Source

Ryan Humble^{1,*}, Finn H. O'Shea,² William Colucho², Matt Gibbs², Helen Chaffee,² Eric Darve,^{1,†} and Daniel Ratner²

¹*Institute for Computational and Mathematical Engineering,
Stanford University, Stanford, California 94305, USA*

²*SLAC National Laboratory, Menlo Park, California 94025, USA*



(Received 10 June 2022; accepted 17 November 2022; published 16 December 2022)

Accelerators produce too many signals for a small operations team to monitor in real time. In addition, many of these signals are only interpretable by subject matter experts with years of experience. As a result, changes in accelerator performance can require time-intensive consultations with experts to identify the underlying problem. Herein, we focus on a particular anomaly detection task for radio-frequency (rf) stations at the SLAC Linac Coherent Light Source (LCLS). The existing rf station diagnostics are bandwidth limited, resulting in slow, unreliable signals. As a result, anomaly detection is currently a manual process. We propose a beam-based method, identifying changes in the accelerator status using shot-to-shot data from the beam position monitoring system; by comparing the beam-based anomalies to data from rf stations, we identify the source of the change. We find that our proposed method can be fully automated while identifying more events with fewer false positives than the rf station diagnostics alone. Our automated fault identification system has been used to create a new dataset for investigating the interaction between the rf stations and accelerator performance.

DOI: [10.1103/PhysRevAccelBeams.25.122804](https://doi.org/10.1103/PhysRevAccelBeams.25.122804)

I. INTRODUCTION

The increasing complexity of particle accelerators is creating new challenges in operating user facilities with high-uptime. The tasks of predicting, identifying, and recovering from anomalies have already grown beyond what a human operator can monitor unassisted [1]; for example, the control system for SLAC's Linac Coherent Light Source (LCLS) includes 200,000 different process variables (PV) stream while the next-generation LCLS-II will record 2 million. This growth in data rates opens new opportunities for data science techniques to improve the automation of accelerator operations.

The problem of anomaly detection, the task of finding abnormal events or data, is common beyond particle accelerators. Nearly all modern data-heavy systems, such as complex internet, industrial, and engineering systems, produce thousands of real-time signals, overwhelming the human operators tasked with monitoring performance.

Moreover, research has shown that humans, even trained operators, are severely limited when detecting rare events [2–4]. Nonetheless, human operators have the necessary causal reasoning and deep institutional knowledge to investigate and understand the root cause of anomalies. To better detect anomalies in these complex systems, we seek to delegate the constant monitoring of a large number of complex data streams to an automated artificial intelligence (AI) agent and refocus human attention, cognition, and causal reasoning to anomaly investigation and recovery.

A. Anomalies at LCLS

LCLS, like many complex engineering systems, experiences unplanned downtime and unexpected behavior that is difficult to pinpoint and correct. LCLS loses at least 3% of availability [5]—more than 180 h per year—to unplanned downtime, equivalent to more than three full user experiments. We also focus on more subtle changes to performance which degrade the free-electron laser (FEL) performance, even when the beam is still delivered. Furthermore, fluctuations in performance can restrict the operational range of the accelerator (e.g., if spare rf stations must be held in reserve, the range of x-ray energy available to users is limited) or contribute noise to experiments through changes in x-ray photon energy. Therefore, our goal is to minimize the number of interruptions to stable

*ryhumble@stanford.edu

†Also at Department of Mechanical Engineering, Stanford University, Stanford, California, USA.

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

beam rather than integrated downtime. Due to limited bandwidth across the multikilometer accelerator, subsystem diagnostics are typically low fidelity, and existing methods for identifying faults purely on subsystem diagnostics have both too many false positives and false negatives to be deployed in operation.

The motivation for this work is threefold: (i) develop an algorithm that will alert an operator that an anomaly has occurred and propose a candidate source; (ii) automate logging of anomalous events; and (iii) create a new labeled anomaly dataset that can be used for future studies.

First, by alerting operators that an anomaly has occurred and localizing a candidate source, we aim to increase uptime and reliability, decrease time to restore the beam to user operation, lower maintenance costs, and, indirectly, expand the effective operational limits of the machine. Second, automating the anomaly detection process enables automated logging of events, which can in turn identify subsystems in need of repair prior to failures or subsystems that are recurrently problematic. These goals align with ongoing work at other organizations, such as the Smart LINAC project at CERN [6], which is utilizing AI-based anomaly detection to maximize uptime and decrease maintenance costs for LINACs.

Finally, the creation of labeled, voluminous, datasets is often the first step in tackling any machine learning (ML) problem [7]. Thus, in addition to the operational benefits just mentioned, an algorithm that can automatically detect anomalous behavior can also enable the creation of a labeled dataset. The problem of assembling datasets has been mentioned in a number of recent works on applying ML at accelerators. Specifically, they point to the burden created by the amount of effort required to produce a suitable dataset [8–10] or they expressed a desire for more data to improve the performance of their algorithms [11,12]. In the rf station example, a dataset labeled with slow rf station data could enable a future algorithm based entirely on fast, beam-based signals.

B. Contribution

In this work, we introduce a fully automated, beam-based diagnostic method for rf station fault detection at LCLS. Our key contribution is the combination of high fidelity beam position monitoring data with bandwidth limited (and noisy) rf station diagnostic data to achieve highly accurate, nearly real-time fault identification. This approach, of using beam-based data as a way of understanding subsystem performance, is applicable beyond rf stations at LCLS to anomaly problems at LCLS-II and other accelerators. We primarily present an interpretable unsupervised anomaly detection algorithm that exploits knowledge of LCLS’s design and seeks to mimic a human operator’s fault analysis process.

Our paper makes several contributions. First, we describe the current problem of identifying rf station faults

at LCLS, notably the bandwidth limitation and inaccuracy of the existing rf station diagnostics system, and explain how pulse-by-pulse beam position data can be used to supplement these diagnostics (see Sec. III). Second, we present how to identify possible rf station faults, which we refer to as anomaly candidates, from two primary existing rf station diagnostics. See Sec. IV A. Third, we define a statistics-based unsupervised time-series anomaly detection algorithm on the beam position data. The algorithm assigns an anomaly score to each anomaly candidate to determine whether an rf station fault occurred. See Sec. IV B. Fourth, we compare our fully automated system to an existing detection system, over a labeled study period, and find our system reduces the number of false positives by a factor of 20x while correctly identifying 96% of the faults. Over a longer unlabeled period, we analyze the faults identified by our approach and correlate our findings with faults that were manually logged by LCLS operators. Finally, we describe the creation of a public anomaly detection dataset from the rf station diagnostics, beam position data, and hand labels and demonstrate that a proof of concept DNN trained on this data achieves comparable results to our original (unsupervised) system. See Sec. V.

II. RELATED WORK

Fault identification and recovery reduce downtime and improve performance metrics. Within the last 5 years, various accelerators have investigated machine learning algorithms to identify unusual (i.e., anomalous) accelerator states to increase the performance of the accelerator: isolation forests to improve the quality of optics measurements by detecting faulty BPMs [13]; autoencoders to show sensitivity to parameter drifts in drift-tube LINACs [14]; Gaussian mixture models and isolation forests to detect anomalous behavior in injection kicker magnets [15]; transfer learning to detect anomalous vacuum pressure changes at a synchrotron [16]; principal component analysis and local outlier factor to detect collimation system settings that need improvement [17]; Siamese neural network model to predict anomalous pulses at a proton LINAC and prevent damage to the accelerator [18]; and recurrent autoencoders to detect failing rf power modulators [19]. Reference [20] applies a variety of different model-based and nonmodel-based techniques for detecting anomalies in a cryogenic system, including an online update to the definition of anomaly. References [8,21] also apply a number of methods to detect anomalies of different types in the rf power of a proton source. Several algorithms have been used to monitor the cryogenic system at the European XFEL, such as a residual of the eigenvalues of a state matrix to detect quenches in superconducting cavities at a reduced repetition rate [22], and anomaly detection using a model-based parity space algorithm and classification using support

vector machines [23,24]. In the latter case, anomalies are defined on an *ad hoc* basis by separating clusters.

III. rf STATION ANOMALIES

The health of LCLS is judged by the quality of the x-ray beam delivered to users, typically through metrics such as energy per pulse, stability of photon energy, and x-ray bandwidth. For the purposes of this work, we define an anomaly to have occurred when there is an unexpected change in the state of the accelerator that either degrades the performance of the x-ray beam or causes a machine fault that halts operation entirely. After observing the onset of degraded performance, operators retune based on previous operational experience, a collection of physical models, and intuition. With strong anomalies, the machine-protection system will turn off beam delivery, and the same techniques are used to recover accelerator operation. Searching for the root cause of an anomaly can be time-consuming, both due to a large number of process variables and due to the complex interaction between the accelerator and its various automated control systems, which may blur cause and effect.

Out of LCLS' various failure modes, we focus on the challenge of identifying anomalous behavior in the 82 rf stations that power the x-ray laser's accelerator. Changes in the power delivered by an rf station can range from nearly imperceptible impact to the electron beam to completely stopping operation (i.e., a machine fault). Fault identification and recovery currently require an experienced operator to search through hundreds of data streams in real time after the fault occurs. Anomalies in the rf stations are an interesting source of study for multiple reasons: (i) they are a major failure mode, with several drops in rf station performance per hour during beam operation; (ii) they directly affect the beam energy, causing degradation of free-electron laser (FEL) performance; and (iii) they have an existing but inaccurate detection system.

A. Current method

There are two existing methods for detecting rf station faults at LCLS. A drop in rf station performance may be observed either through (i) system diagnostics that directly monitor the rf station or (ii) through the subsequent impact on the accelerator itself.

Unfortunately, the rf station diagnostics have two shortcomings: the data are reported no faster than once every 5 s ($600 \times$ slower than the beam rate) due to limited bandwidth in the control system; and, more fundamentally, the diagnostic info is often incorrect, falsely reporting normal performance as anomalous and missing real dips in performance (as shown in Sec. VA).

Consequently, LCLS primarily uses a human-centered approach following a drop in accelerator performance, with operators checking by hand a set of rf station diagnostics

for a possible cause of the anomalous behavior. This human-driven fault detection and recovery have drawbacks, requiring the attention of a human operator for routine activity. Moreover, because the search for an anomaly is prompted by an operator, the human-centered approach only addresses significant and prolonged changes in performance that catch the operator's attention and misses less noticeable—but possibly frequent—drops in performance that may still impact user experiments or may predict future failures.

B. Combining LCLS data sources

As an alternative to the human-centered approach, we combine two sources of data available at LCLS: rf station diagnostic data, which are sensitive to the source of the fault, and beam-based diagnostics, which are sensitive to pulse-by-pulse accelerator performance. Combining the two data streams makes it possible to, accurately and with full autonomy, both detect rf station anomalies and identify the responsible rf station, thereby avoiding the shortcomings of the inaccurate and manual current method.

The rf station diagnostic data consist of roughly a dozen status bits (0—healthy, 1—unhealthy) as well as several real-valued signals. Most notably, there is a klystron AMM (amplitude mean out of tolerance) status bit that reports if the station klystron's accelerating rf amplitude is too high or low and a downsampled version of the real-valued AMPL (amplitude) signal. Klystron amplitude is related to the energy gain of the electron bunch so a change in amplitude will change the energy of the bunch and degrade the ability to transport the bunch through the magnetic lattice. The amplitude is derived from directional couplers at the rectangular waveguides that feed the accelerating structures. When searching for a fault by hand, operators typically use the AMM status bit. However, since both the AMM and AMPL are inaccurate, we cannot rely on this source of data alone. We continue to explore using the unused rf station diagnostic information, including phase information PHAS, for improved anomaly detection.

The second source is beam-based diagnostic data. The human-centered evaluation of FEL performance leans on x-ray diagnostics, which measure the quantities of interest for the FEL users. However, x-ray performance is sensitive to nearly every potential type of failure. Instead, we focus on stripline beam position monitor (BPM) [25] data in dispersive regions, which are sensitive primarily to fluctuations in beam energy. The BPM data are included in the beam-synchronous acquisition (BSA) service [26], which aligns a subset of the machine signals to the individual beam pulses; this enables BPM data to be read and recorded at the full LCLS operational rate of 120 Hz. As the “ground truth” for the beam status, the BPM data allow anomalies to be discovered at the full beam rate; however, as the BSA service does not record any rf station diagnostic

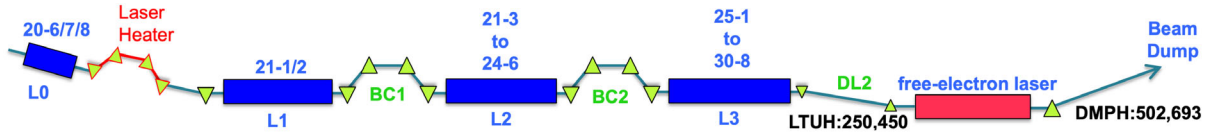


FIG. 1. Layout of the LCLS showing the location of the klystrons included in this work above the accelerating sections (blue rectangles: L0, L1, L2, and L3) and the names of the BPMs used to detect changes in beam energy: LTUH:250, LTUH:450, DMPH:502, and DMPH:693. The green triangles show dipoles and the red rectangles represent the free-electron laser undulators.

information, attributing the anomaly to a particular component of the accelerator is challenging.

Each of the 175 BPMs along the accelerator produces three signals: X, Y, and TMIT. X and Y measure transverse position; TMIT (transmitted intensity) measures the passing beam charge. Of the 175 BPMs, there are 4 BPMs of interest to our work: BPMS:LTUH:250, BPMS:LTUH:450, BPMS:DMPH:502, and BPMS:DMPH:693. As shown in Fig. 1, all of these BPMs are located in regions of the accelerator with nonzero dispersion in either X or Y, making them sensitive to changes in beam energy, and are downstream of all rf stations. We take the four beam positions in the dispersive plane and the four TMIT signals as our input BPM data.

We then seek to use the BPM data to corroborate and enrich the rf station status that is reported by the diagnostic data. By cross-referencing the faster, more sensitive BPM data with the slower, easier-to-interpret diagnostic data, the algorithm adopts the strengths of both: changes in the rf station diagnostic that are not corroborated by the BPMs are rejected, and subtle changes to the rf station that were previously rejected can be confirmed by the BPMs. We evaluate the performance of this strategy on archival data at LCLS.

IV. METHOD

Our method can be broken down into two steps. The first step is anomaly candidate identification, where we use the

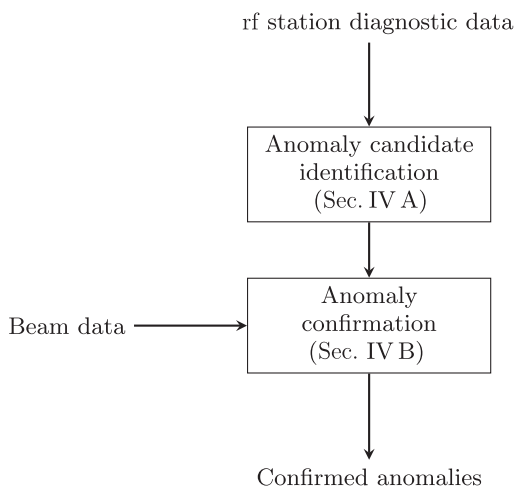


FIG. 2. Two-stage rf station anomaly identification method.

slowly updating rf station diagnostic data to identify possible anomalies. The second is to confirm the anomaly using the full-rate BPM data. A flowchart for this process is shown in Fig. 2.

A. Anomaly candidate identification

The first step generates candidates independently for each rf station based on each station's diagnostic data. Formally, given diagnostic data $h_{k,t}$ from rf station k at time t , we want to identify anomaly candidates $C = \{C_i\}$. A candidate is $C_i = (s_i, e_i, k_i)$ for start time s_i , end time e_i , and station k_i . Candidates are merged across all rf stations with overlapping anomaly windows. We then drop candidates associated with more than one rf station for the sake of the interpretability of the performance metrics. This yields a set of anomaly candidates with a unique rf station association: $C = \{(s_i, e_i, k_i)\}$ where $s_1 < e_1 < s_2 < \dots < e_n$. (We note that multiple stations can fail simultaneously, for example, after a fault in the substation. The uniqueness requirement is only to simplify the analysis metric and can be discarded during deployment.)

Generating the set of candidates for each station requires defining a function that takes the diagnostic data $\{h_{k,t}\}$ and produces a set of candidates. We require this function to be computationally simple enough to run in real time and produce no overlapping candidates. For the AMM status bit (0—healthy, 1—unhealthy), the function returns a candidate for every time range where the bit is continuously 1. (See Sec. III for details.) Creating anomaly candidates from the real-valued AMPL signal is slightly more involved. LCLS has several automatic and operator-driven controls that constantly but slowly adjust the rf amplitudes. Therefore, we subtract the 3.5-min rolling median value from the current AMPL value and threshold the absolute difference to create a new AMPL-based status bit. Using this computed status bit, we can recover anomaly candidates as with the AMM bit. The 3.5-min rolling window duration was chosen to be substantially longer than short-duration dips, to prevent cases where the machine recovers on its own from strongly disturbing the median, but shorter than the time scale of machine drift. The threshold (0.5%) was chosen to be deliberately tight (i.e., it will be prone to false positives) to flag subtle anomaly candidates, improving the sensitivity over the AMM-derived candidates. Additional details on how these diagnostic signals are converted into candidates can be found in Appendix A.

B. Anomaly confirmation

In the second stage, we use the BPM data to corroborate or reject anomaly candidates $C_i = (s_i, e_i, k_i)$. We require this step to be: unsupervised (there is a high volume of data, and manually labeling anomaly candidates is very time-intensive); drift tolerant (the BPM data distribution is constantly changing due to operator input and automated feedback mechanisms); jump sensitive (sensitive to sudden and possibly short-lived changes in BPM readings); interpretable (the anomaly confirmation should be easily interpreted by an operator and closely match their intuition); and real-time capable (anomalies should be confirmed in “human time”, i.e., on the order of several seconds).

We use a simple strategy based on univariate median absolute deviation (MAD) [27] and a geometric mean aggregation rule. This strategy requires no training, uses only two hyperparameters, and depends on only the last ~ 10 s of data (making it drift tolerant, easily computed, and available in near real time). Specifically, given a multivariate time series $x_{s,t}$ for signal s at time t , we define the anomaly score as follows. First, we run MAD on each BPM signal individually to get per-signal anomaly scores $a_{s,t}$, using a 5-s rolling window. Note that we use MAD as a robust z score, where the mean and standard deviation estimates are replaced by robust median-based counterparts. The standard z -score, which measures the number of standard deviations above or below the mean, is quite sensitive to outlier behavior (e.g., the standard deviation will increase quickly and aggressively due to outlier data, lowering the z -scores of subsequent outlier data and thereby underweighting the outlier severity). Next, we aggregate across all signals to suppress beam noise and boost detection of beam energy changes. We use a geometric mean to more heavily penalize disagreement as compared to the arithmetic mean. This yields a single anomaly time series a_t . Then, we aggregate across several consecutive pulses. This boosts the score for sustained beam anomalies. As before, we use a geometric mean. Lastly, we threshold the maximum score in the anomaly window to determine whether an anomaly has occurred.

The full definition is given in Appendix B. A key assumption here is that the beam is in a stable, non-anomalous condition prior to the rf station diagnostics identifying a potential problem. Therefore, the rolling windows used to define the anomaly score should capture the expected beam position and provide a meaningful baseline.

V. RESULTS

We assess the performance of our method relative to the existing AMM-based detection system for the period stretching from November 2, 2020, to February 10, 2022. We only analyze periods of operation where the beam is being delivered to experiments with certain

operational conditions to simplify data processing. We also remove time periods in which the BSA data-recording service itself had anomalous performance, which we identify using the BOCPD (Bayesian online changepoint detection) [28] algorithm. See Appendix C for full details on the data filtering.

As mentioned in Sec. IV B, labeling anomaly candidates is time-intensive. However, for the purposes of evaluating our approach, we hand-labeled the anomaly candidates generated during the start of our study window, specifically from November 2, 2020, to December 10, 2020. Labels follow the operators’ process for identifying anomalous rf station behavior, selecting as anomalous any rf station diagnostic warnings that are nearly simultaneous with unusual changes in beam energy as measured by the BPMs.

A. Improvement over existing system

During the labeled period, the AMM status bit identifies 1237 anomaly candidates. However, we found that only 385 (31%) coincided with visible drops in accelerator performance, meaning the vast majority of predicted anomalies are false positives. Additionally, we will see later that the AMM bit misses a significant number of true anomalies as well. In sum, the AMM bit alone is an unreliable anomaly detector.

The premise of this paper is that corroborating the AMM status bit (or other rf station diagnostics) with the BPM data using the procedure in Sec. IV improves the precision of anomaly detection. With corroboration, we reduce the number of false positives to 37 (3%), a factor of 20 lower than using the AMM bit alone (69%). At the same time, we still correctly identify 368 of the 385 anomalies present in the candidate set, meaning the corroboration only reduces the number of anomalies found (i.e., recall) by 4%. The full confusion matrix is given in Table I.

Since we suspected that the AMM bit was missing anomalies, we also investigated using the underlying AMPL signal to improve the recall of the rf station anomaly detection. With a permissive threshold of just 0.5%, the AMPL signal identifies 2845 anomaly candidates during the labeled period, more than twice as many identified by

TABLE I. Confusion matrix for BPM corroboration of AMM-based candidates. Of the 385 anomalies, 368 are correctly confirmed; only 37 nonanomalies are mistakenly confirmed. Uses a BPM threshold that maximizes the F_1 score. Note that the 17 false negatives consider only candidates selected by the AMM bit; true anomalies missed by the AMM bit itself are not considered here.

		True class	
		Anomaly	No anomaly
Predicted class	Anomaly	368	37
	No anomaly	17	815

TABLE II. Confusion matrix for BPM corroboration of AMPL-based candidates. Of the 553 anomalies, 504 are correctly confirmed; only 68 nonanomalies are mistakenly confirmed. Uses a BPM threshold that maximizes the F_1 score. Note that the 49 false negatives consider only candidates selected by the AMPL signal; true anomalies missed by the AMPL signal itself are not considered here.

		True class	
		Anomaly	No anomaly
Predicted class	Anomaly	504	68
	No anomaly	49	2224

the AMM bit. Of these, we find 553 (19.4%) coincide with drops in accelerator performance according to the BPM data. Using the underlying AMPL signal, therefore, identifies at least 168 anomalies that are missed by AMM, increasing recall by 44%. Note that both AMM and AMPL likely fail to identify true anomalies as candidates due to the bandwidth and noise issues mentioned in Sec. III, leading to additional false negatives. Future work will aim to predict anomalies solely from full beam-rate BPM data to capture additional anomalies (see Sec. V C). The full confusion matrix and interpolated precision-recall (PR) curve for the AMPL signal are shown in Table II and Fig. 3, respectively. To obtain the PR curve and confusion matrix, we vary only the threshold on the BPM-based anomaly score from Sec. IV B, keeping the 0.5% permissive AMPL deviation threshold fixed.

The precision and the number of anomalies found with all three methods are given in Table III. (Note that the raw number of anomalies is given rather than recall because there are likely additional anomalies that are undiscovered

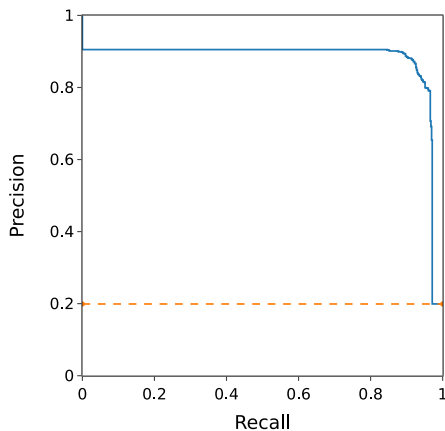


FIG. 3. Interpolated precision-recall curve for the BPM threshold on AMPL-based candidates. The best F_1 score is 0.897, at a BPM anomaly score threshold of 2.848, corresponding to an accuracy of 96%. The dashed line is the precision-recall curve for a random classifier. Note that recall here does not include anomalies missed by the AMPL threshold and is only an upper limit on the recall for the full algorithm.

TABLE III. Precision and number of anomalies found using the AMM bit alone, AMM bit corroborated by BPMs, and AMPL corroborated by BPMs. The corroborated data use the BPM thresholds that maximize the F_1 score.

	Precision	Anomalies found
AMM-only	0.31	385
AMM + BPM	0.91	368
AMPL + BPM	0.88	504

by any of the existing methods.) The addition of the BPM corroboration to the AMM bit results in only a small drop in the number of anomalies found while greatly improving precision. Switching the trigger to the AMPL signal significantly improves the number of anomalies found at the expense of a small drop in precision.

Both the AMM and AMPL signals are evidently noisy and inaccurate to differing degrees. By combining these signals with the anomaly detection algorithm on the full-rate BPM data, we are able to significantly improve detection accuracy by rejecting numerous false positives. Furthermore, using the AMPL signal in addition to the AMM status bit reduces the number of missed anomalies (false negatives). In particular, of the 385 true anomalies detected by the AMM signal, the AMPL signal flagged 305 of them while only missing 13 anomalies. The remaining 67 anomalies were given a different or multiple attributions according to the AMPL signal. Of the 553 true anomalies detected by the AMPL signal, the AMM signal flagged 307 of them but missed 245; there was 1 anomaly that AMPL attributed to KLYS:LI29:11 but AMM attributed it to KLYS:LI27:41.

As a by-product of our analysis, during hand labeling of the anomaly candidates, we discovered a natural grouping of the candidates into two types: (i) *Faults*: The beam is lost (TMIT goes to 0). Figure 4 shows an example. (ii) *Sustained anomalies*: initial large deviation followed by a recovery period lasting several hundred pulses. Figure 5 shows an example, in which AMM misses the anomaly.

Of the 553 anomalies detected by combining the AMPL signal and BPM data, 217 were *Faults* and 336 were *Sustained anomalies*. The majority of the anomalies, including both examples (Figs. 4 and 5), are short-lived, lasting less than ~ 10 s. Human operators are likely to miss these anomalies, as by the time they notice a degradation in beam performance, the problem is likely resolved and the klystron health signal reverted to normal. Particularly in the case of *Sustained anomalies*, human operators might not even notice the moderate performance degradation, despite the impact on FEL energy and therefore user experiments.

B. Anomalies at LCLS

We now apply the same algorithm in an unsupervised manner to the entire dataset to search for anomalies. After

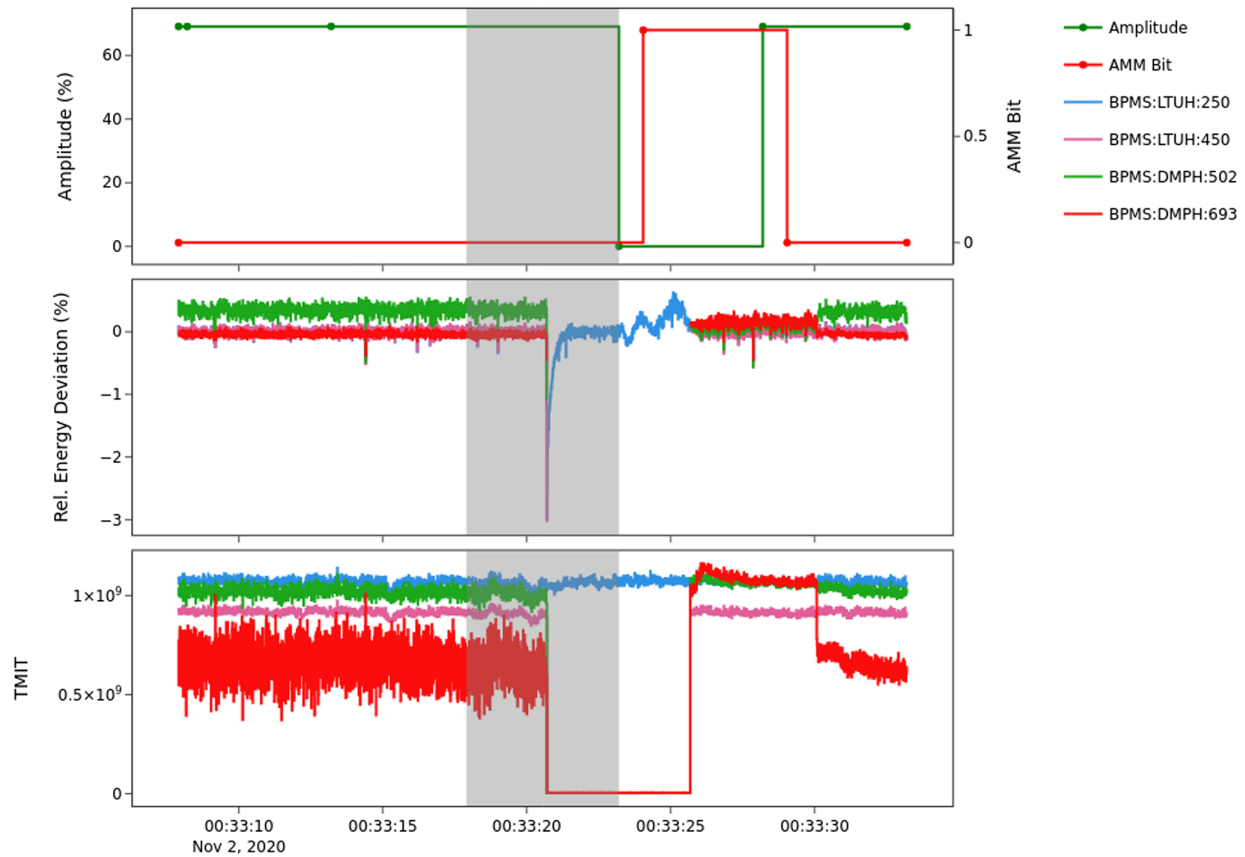


FIG. 4. Example of an rf station fault at KLYS:LI29:11. The top plot shows the klystron diagnostic data (both the AMM status bit and the AMPL signal). The middle plot shows the relative energy deviation (calculated from the beam position in the dispersive plane). The bottom plot shows the TMIT readings (which is the approximate number of electrons, subject to some calibration). The gray region defines the implied anomaly window, due to the delayed reporting of the klystron health information. This extreme example shows the AMPL drop nearly to 0, which causes the beam to be lost. The missing pulses in the energy deviation plot are the result of there being no beam to measure.

applying the criteria for data selection at the beginning of Sec. V, we have approximately 80 days worth of data to survey. Over this time period, we find 29,739 anomaly candidates in the rf station diagnostic data: 3615 from the AMM bit alone, 22,570 from the AMPL signal alone, and 1777 found by both. Of these 6529 (21.9%) exceed the BPM anomaly threshold found in the previous section and are labeled as anomalous. Of the more restrictive AMM candidates, 47.5% are labeled as anomalous, while only 16.3% of the more permissive AMPL signal candidates are labeled as anomalous. Figure 6 shows that anomalies are not distributed equally among the rf stations; in particular, the AMPL signal usually identifies many more anomalies in the rf stations that experienced the most anomalies. The top 5 rf stations (of 82) account for $\sim 36\%$ of the identified anomalies.

It is possible that the BPM data contain anomalylike behavior even in the absence of an rf station anomaly, which would lead to corroboration of false positives. To investigate the false positive rate, we take 6453 random samples (approximately half of a day) of the BPM data

independent of the rf station diagnostics. All samples of the BPM data are of the same length as the candidate samples. We find 176 ± 13 ($2.7 \pm 0.2\%$) samples exceed the anomaly threshold and are labeled anomalies despite no apparent anomaly in the rf station data (the 1-sigma confidence interval assumes the random samples are independent identically distributed Bernoulli random variables). Taking the conservative assumption that all events found in the random stream are false positives sets an upper limit on the false-positive rate.

A secondary benefit of an automated detection scheme is the ability to log large numbers of events to track performance, predict future failures, or enable later analysis. At present, rf station anomalies are logged through the manual CATER system [29]. The CATER reports are limited by the need for operator attention; operators may not notice a transient anomaly, or may not have the time to manually log an event, and there is no way for one or two operators to watch thousands of potential signals. Moreover, the CATER system does not try to quantify the anomaly rate for any particular subsystem.

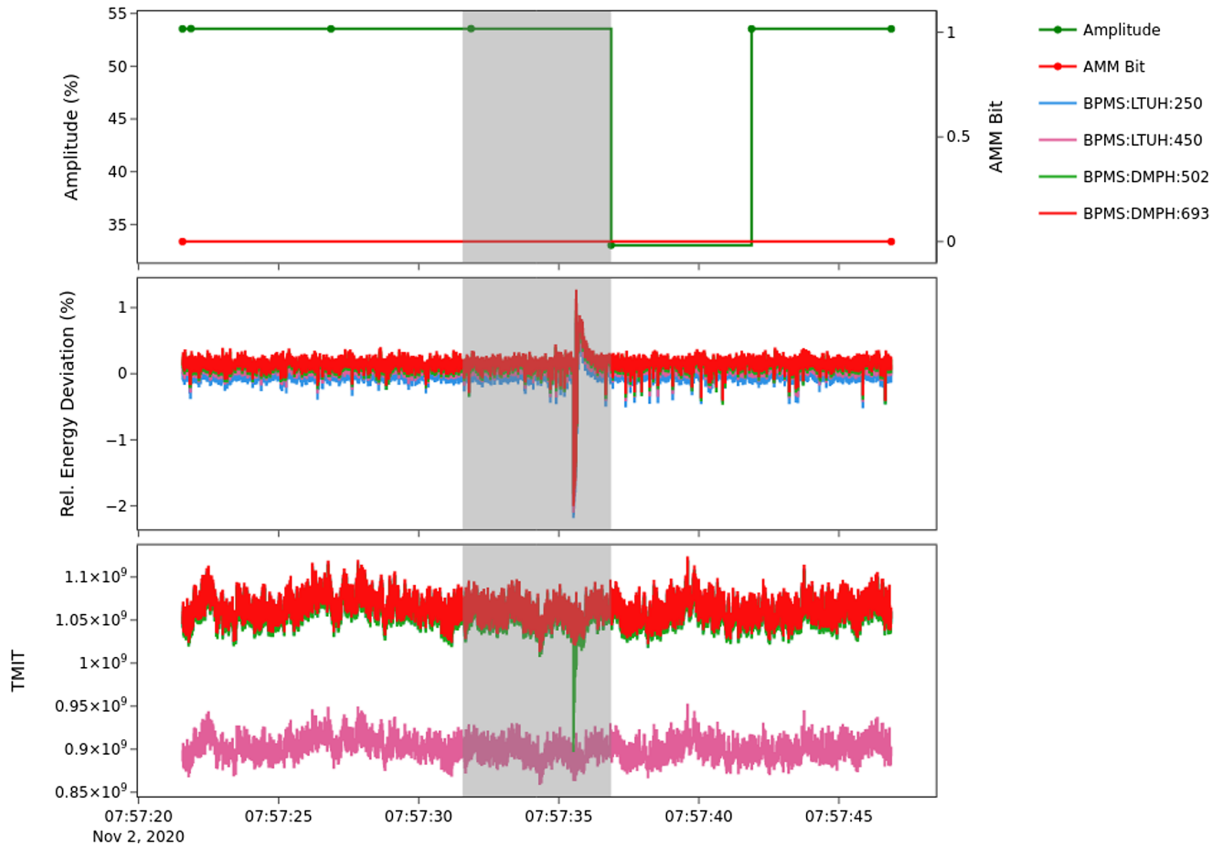


FIG. 5. Example of a sustained anomaly at KLYS:LI28:11. The top plot shows the klystron diagnostic data (both the AMM status bit and the AMPL signal). The middle plot shows the relative energy deviation (calculated from the beam position in the dispersive plane), with a brief but large deviation followed by a recovery period. The bottom plot shows the TMIT readings (which is the approximate number of electrons, subject to some calibration). The gray region defines the implied anomaly window, due to the delayed reporting of the klystron health information. Note this anomaly is missed by AMM.

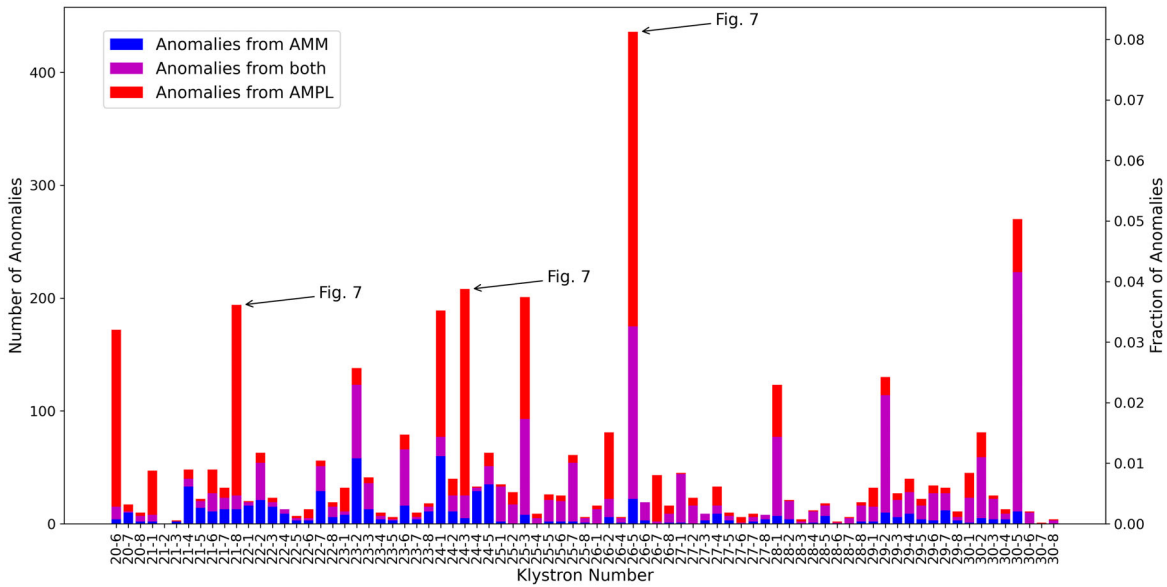


FIG. 6. Count of anomaly candidates identified by either the AMM bit, the AMPL signal, or both that were labeled anomalies using our anomaly detection routine on the BPM data. The red part of the bar is anomaly candidates detected only by the AMPL signal, the blue part of the bar is anomaly candidates detected only by the AMM bit, and the magenta part of the bar is anomaly candidates that were detected by both. The klystrons are enumerated along the LCLS accelerator from low energy (left) to high energy (right).

To assess our automated approach, we compare the automated logging to CATER events during the studied time period. In total, there were 597 requests for work made to the groups charged with maintaining the rf stations. 444 of these reference a specific rf station, compared to 6529 for the automated approach despite the BSA service being active for only a subset of the time period. All 82 stations studied in this work are referenced at least once. Figure 7 shows predictions for three of the stations showing the most anomalies: 26-5, 24-3, and 21-8.

Station 26-5 had the most anomalies of any station. Figure 7 (top) shows a histogram of the anomalies (binned by week) and the times at which problems with the rf station were reported through CATER. Both the AMM bit and the AMPL signal capture events. The performance on the data from stations 25-3 and 30-5 (not shown) was similar. For stations 21-8 and 24-3, the importance of the AMPL signal specifically is more apparent (see Fig. 7 middle and bottom). For 24-3, there are two periods, in April and November 2021, where our algorithm detects anomalies that were totally missed by the AMM bit. These periods of high anomaly counts were followed by several CATER events. Station 21-8 shows more overlap (in time) of the anomalies identified using the AMM bit and the AMPL signal, but there are still periods of anomalies missed entirely by the AMM bit.

C. rf station anomaly dataset and classifier

For the labeled study period, we have compiled a dataset of the rf station diagnostic data and the BPM signals, alongside the hand labels, as a public anomaly detection dataset. The dataset consists of two HDF5 files (one for AMM and one for AMPL) containing the raw data, two CSV files containing information about the candidates, and two CSV files containing the hand labels. We also include in the HDF5 files random samples of beam operation to serve as negative or background examples. A full description of the files can be found in Appendix D. The files can be accessed at [30].

While our study focused on unsupervised anomaly detection, the existence of a labeled dataset opens the possibility to train more powerful supervised models, including deep neural networks (DNNs). For example, we have only looked at the most direct rf station and BPM signals in the amplitude and beam energy, respectively. Future work could extend to more subtle diagnostics, such as phase information for the rf station or nondispersive BPMs. rf phase information is expected to reveal additional anomalies not observed solely through amplitude variation, but the signal was too complex to be analyzed with an unsupervised method. A DNN may be more successful. Similarly, nondispersive BPMs located throughout the accelerator are both full beam rate and have rf station specificity and may be able to sidestep the bandwidth limitations of the rf station diagnostics entirely. While the

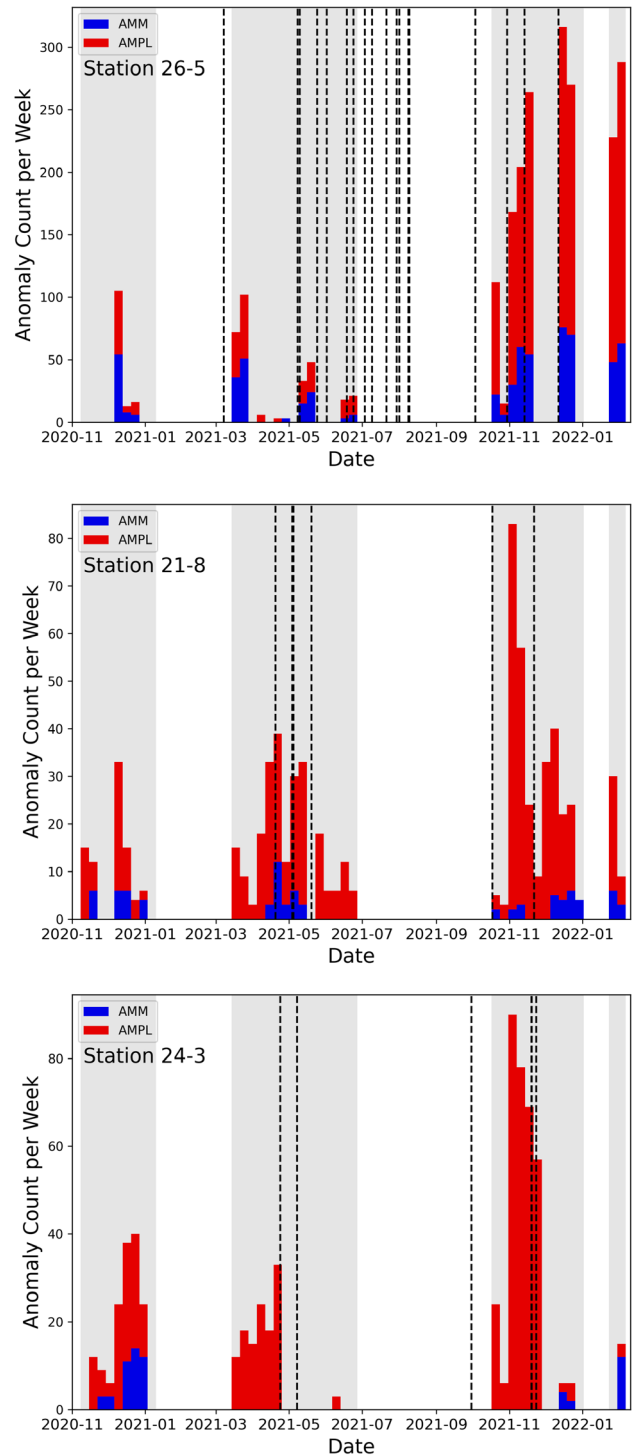


FIG. 7. Count of the number of anomalies detected by AMPL (red bars), the AMM bit (blue bars), and problems reported in CATER (black dashed lines) for various stations. The anomaly counts have been binned by week to produce the histograms. The shaded regions show weeks where the BPM data were available through the BSA service. Top: station 26-5. For this station, both the AMPL and AMM bit signals detect anomalies at approximately the same time. Middle: station 21-8. AMPL detects anomalous activity more often than the AMM bit does. Bottom: station 24-3. AMPL detects anomalous activity in April and November 2021 that the AMM bit misses entirely.

nondispersive BPMs are sensitive to many other types of anomalies that would confuse our unsupervised approach, a DNN trained on a labeled dataset may be capable of distinguishing anomaly categories and identifying the faulty rf station. As a brief demonstration, we trained a deep binary classifier using the raw BPM signals as input. The experimental setup is described in Appendix E. On the 20% held-out test set, which contains 569 candidates, our trained model achieved an F1 score of 0.88, correctly identifying 95%, showing the DNN is at least as accurate as the algorithm presented in Sec. IV B. We present the full precision-recall curve in Appendix E. Further improvement will require an extension of the dataset to include rf phase diagnostics or nondispersive BPMs. We leave the problem of identifying rf station performance changes solely from BPM data to future work. We emphasize that this proof-of-concept DNN is not targeted toward deployment in the operation room but is mostly a suggestion for potential future work.

VI. CONCLUSION AND DISCUSSION

This work describes a beam-based approach to automating the detection of anomalous behavior at LCLS. We focus on a particularly common failure mode, rf station anomalies, and improve upon the existing diagnostic system. We have developed a fully automated system that combines slow-updating, inaccurate rf station diagnostic data with full-rate BPM data. By corroborating the diagnostic data with measurements of the beam, we are able to filter out more than 95% of the false positives signaled by the rf station diagnostic data. Moreover, by using the underlying AMPL signal, in addition to the current AMM status bit, we detect significantly more rf station drops and faults, reducing the number of anomalies missed (false negatives). On a hand-labeled dataset, our approach improved precision and the number of events found from 0.31 and 385 for the raw AMM bit to 0.88 and 504 for AMPL corroborated by the beam. On an extended unlabeled dataset, our algorithm found a total of 6529 events, compared to 597 CATER events logged during the same time period. The proposed algorithm will give operators a higher degree of confidence that rf station faults are being detected and the responsible station identified and provides rf technicians with a detailed history of faults to guide repairs.

The need to combine expert knowledge with anomaly detection systems will only grow as accelerator systems become more complex. Moreover, as the data volume grows, data collection cannot always be at the full system rate, forcing some signals to be recorded at slower rates. The beam-based method presented, leveraging high-fidelity diagnostics on the beam to corroborate low-fidelity subsystem diagnostics, may find common use in other accelerator anomaly prediction tasks. Future work will expand the existing dataset and can serve as a test bed for the development of new beam-based anomaly detection

algorithms for accelerators. We also plan to detect and classify rf station anomalies using data from both dispersive and nondispersive BPMs, as different rf stations will affect the beam to different degrees and in different manners.

ACKNOWLEDGMENTS

Use of the LINAC Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

APPENDIX A: CREATING CANDIDATES

In practice, the AMM status bit is a noisy and unwieldy signal, so creating candidates from it is not as trivial as taking consecutive regions where the status bit is 1. The AMM status bit is itself the result of applying an operator-configurable threshold to internal klystron diagnostic information. Unfortunately, this threshold can often be misconfigured, leading to $h_{k,AMM,t} = 1$ for implausibly long periods of time. This affects the candidate identification step by technically overlapping with many other possible candidates, which for the sake of interpretability, we would then drop. To counteract this, and attempt a fairer comparison with AMPL, we will ignore $h_{k,AMM,t}$ if it exceeds 10% unhealthy over a period of otherwise apparently normal beam operation. This trouble in generating anomaly candidates from AMM further supports the conclusion that AMPL is a better diagnostic for rf station health.

Also, each klystron is not constantly in use, with at least a couple held in reserve for operators to swap out malfunctioning klystrons for healthy ones. Therefore, before assessing the diagnostic signal (AMM or AMPL), we need to first check if the klystron is in use. This information is stored in a different PV in the EPICS Archiver, which we fetch alongside the diagnostic signal itself.

Finally, the klystron signals are only recorded in the EPICS Archiver infrequently (~ 0.2 Hz for AMM and AMPL) and only when the value changes. This has the effect of time stamping the AMM and AMPL data in the Archiver with a delay possibly as large as the inverse of the logging rate. Therefore, we say AMM or AMPL implies a candidate window in the ~ 5 s before the timestamp actually stored in the Archiver. Also, when we take rolling windows in creating the AMPL-based candidates, we need to carefully weigh the data due to the nonuniform sampling in time.

APPENDIX B: ANOMALY CONFIRMATION ALGORITHM

Given the multivariate time series BPM data $x_{s,t}$ for signal s at time t , we define the per-signal anomaly score as $a_{s,t} = \text{MAD}(x_{s,t-2l}, \dots, x_{s,t})$ where l is the lagging rolling window used to compute the median and median-absolute deviation of the signal $x_{s,t}$. Specifically,

we estimate $\hat{\mu}_{s,t} = \text{median}(x_{s,t-l}, \dots, x_{s,t-1})$ and $\hat{\sigma}_{s,t} = k \text{median}(|x_{s,t-l} - \hat{\mu}_{s,t-l}|, \dots, |x_{s,t-1} - \hat{\mu}_{s,t-1}|)$. The extra factor k ensures that $\hat{\sigma}_{s,t}$ is a consistent estimator of the true standard deviation, where the underlying data are normally distributed. Rigorously, $k = 1/\Phi^{-1}(3/4) \approx 1.4286$ (where $\Phi(x)$ is the cumulative distribution function of a normal distribution) [27]. Thus, $\text{MAD}(x_{s,t-l}, \dots, x_{s,t}) = \hat{\sigma}_{s,t}^{-1}(x_{s,t} - \hat{\mu}_{s,t})$.

Now, we want to aggregate into a single anomaly score a_t . Recall that we are using both the dispersive X/Y signals and the TMIT signals from the four dispersive BPMs at the end of the beamline. Unfortunately, if the anomaly is severe enough to cause loss of beam at a BPM (TMIT lower than 1×10^8), the X/Y reading for that pulse is simply repeated from the last known value. This can make the dispersive reading appear perfectly normal. Therefore, we define the anomaly score for each BPM as the anomaly score associated with the TMIT signal if the beam is lost (as the TMIT anomaly score will be large); otherwise, we define it to be the anomaly score associated with the dispersive signal.

Given an anomaly score for each BPM then, we finally create a single anomaly score using a geometric mean rule: $a_t = (\prod_s a_{s,t})^{1/n}$. We further aggregate across ten consecutive pulses to get the final score

$$a_{\text{AGG},t} = \left(\prod_{i=0}^9 a_{s,t-i} \right)^{1/10}$$

to boost the score for sustained beam anomalies. Finally, each candidate (as identified in **IVA**) suggests an anomaly might have occurred during some time period from start time s to end time e . We say the BPM data confirms the anomaly if

$$\max_{s \leq t \leq e} a_{\text{AGG},t} \geq \tau$$

for a configurable threshold τ .

APPENDIX C: BEAM and BSA STATUS FILTERS

During our study period from November 2, 2020, to February 10, 2022 (457 days), we applied a series of data filters, ultimately yielding ~ 80 days of data. The aim of these filters is to ensure the beam is in a steady operation mode within normal operational bounds and that the recorded BPM data is quality. We also restrict this analysis to hard x-ray operation, because the hard and soft lines use different BPMs; this restriction can be lifted in the future with additional data processing work.

We start by imposing a number of checks on the beam state: (i) Does the beam have a standard charge at the injector? We require $\text{BPMS:IN20:221:TMITCUH} > 0.5 \times 10^9$ and be logged every second in the EPICS

Archiver. A lower charge or the charge not being logged can indicate the beam is not being operated in a standard operational mode. (ii) Is the beam stopper being used? We require $\text{STPR:BSYH:2:STD2INA} = 0$, indicating the beam stopper is out. (iii) Is the beam rate 120 Hz? We require $\text{IOC:BSY0:MP01:PCRATE} = 8$. (iv) Is the entire beam being delivered to the hard x-ray line? We require $\text{IOC:IN20:EV01:RG02ACTRATE} = 10$. We do allow temporary (< 90 s) violations of these conditions to not disallow short periods of “nonstandard” beam operation caused by automatic feedback or protection-based control mechanisms. This reduces the 457-day study period to ~ 153 days of beam operation (~ 103 days are removed due to charge violations, ~ 76 days are removed due to beam delivery violations, and ~ 125 days are removed due to both).

Separately, we check if the data from the BSA service [26] exists during these periods of beam operation and was recorded correctly. Due to a number of engineering issues, significant portions of our study window do not have good quality data. The BSA service records over 1000 PVs synchronously at the beam operational rate into hour-long HDF5 data files. To verify the quality of the data, we check the following: (i) Do the data files exist? Many data files were simply never written due to issues with the BSA service. (ii) Can we read the data from the file? Some data files are corrupted and cannot be read. (iii) Was real data recorded? We drop any data where NaNs are recorded. (iv) Was the data recorded at the correct rate? We expect data to be recorded at 120 Hz. This check was somewhat challenging, so we discuss it at more length below. Of the 457-day study period, 231 days (just over 50%) must be dropped due to a violation of one of these conditions. The primary reason was missing data files: roughly 154 days of data were not recorded. The majority of the remaining 72 invalidated days were due to an invalid record rate.

Unfortunately, the rate recording issues varied widely and required additional work to discover. Let t_i be the timestamp for the i th row of data. Since the standard operational rate of LCLS is 120 Hz, we expect $\Delta_i = t_i - t_{i-1} \approx 1/120$ s. We first define a timestamp mask $m_i = \mathbf{1}(|\Delta_k - 1/120| \leq 2.5 \times 10^{-4})$; this defines whether the current row is recorded at the correct rate with respect to the previous row, up to a tolerance. To identify time periods with bad data recording, we employed the BOCPD (Bayesian online changepoint detection) [28] algorithm. Using a constant hazard $\lambda = 1 \times 10^5$ and a binomial likelihood with prior $p = 0.9$ with strength (pseudo-observations) 2, we ran the BOCPD algorithm on the mask m signal. We define a change point at time i if the probability of having a run of length 100 at time $i + 100$ exceeds 10%. This yields a list of time periods whose mask behavior is statistically distinct from their neighbors. Finally then, we drop time periods where the average mask value was below

85%. All of these hyperparameters were chosen on a best effort/empirical basis to remove known periods of invalid rate recording. We set the seemingly permissive 85% threshold because the rate recording issue can in fact often be repaired. The most common observed rate failure mode was a double logging issue, where $\Delta_i \approx 1/120$, $\Delta_{i+1} \approx 0$, and the $i+1$ data row was an exact duplicate of the previous row i . Therefore, we simply repaired these periods by dropping the offending duplicate rows.

To further refine our search period to periods of constant beam delivery, we only analyze periods of operation where all of the above are continuously satisfied for at least 5 min. Combined together, these conditions reduced our study period from 457 days to only ~ 80 days of operation that could be analyzed. Finally, as the intricacy of LCLS, there are certain operation modes that forcibly kick pulses at a configurable cadence out of the beamline. This has the effect of making the TMIT reading at our dispersive BPMs crash toward 0. Without additional handling, these kicks appear to an observer as single pulse anomalies, despite this phenomenon being a deliberate, operator-controlled behavior. Therefore, we apply a preprocessing step to remove these kicked pulses from our data.

APPENDIX D: DATASET DETAILS

The rf station anomaly dataset is composed of several files. For both AMPL and AMM, there are three files: “klys_anom_dset_{type}.h5,” “candidates_{type}.csv,” and “labels_{type}.csv.”

The HDF5 file has two top level HDF5 groups, one for *candidates* and one for *samples*, corresponding to the anomaly candidates and random samples of the beam operation, respectively. Each top-level group is composed of a number of subgroups corresponding to the individual examples. The subgroups are named according to the end time of the anomaly candidate (or randomly sampled window); the time is stored as nanoseconds since the epoch. Each subgroup contains two datasets called *health* and *bpm*. Each dataset is written as a 2D array of data, where the rows correspond to individual measurements and the columns correspond to the signal name. The timestamps of the rows are stored as the attribute *index* on the dataset; the values denote the number of nanoseconds past the epoch. The column names are stored as the attribute *columns* on the dataset. The subgroups within *candidates* additionally have a group attribute called *klys*, which lists the rf station associated with the candidate.

The *health* dataset stores the rf diagnostic data of the given type over a period of 3.5 min (which corresponds to the look-back duration of the algorithm defined in IV A for AMPL). It will have a shape $(n, 82)$ where n is the number of measurements and 82 is the number of rf stations. Because the rf station diagnostic data are not beam synchronous, the reporting timestamps are not aligned with one another nor the beam, meaning n can vary from

dataset to dataset. In particular, it can be assumed that the AMM and AMPL diagnostic data can be reported with a delay of up to roughly 5 s. All of the NaN values correspond to no new data for that rf station at that timestamp, and the actual value can be assumed to be equal to the last known value. If there is no good last value, then the rf station was not in use during that period of time.

The *bpm* dataset stores the BPM data over a period of roughly 17 s (which corresponds to the look-back duration of the algorithm defined in IV B). It will have a shape $(n, 8)$ where n is the number of pulses and 8 is the number of signals (two signals for each of the four dispersive BPMs of interest). The number of pulses n can vary due to the rate recording and pulse-kicking behaviors mentioned in C. These BPM data were additionally modified from the original raw data by several transforms. First, if the BPM’s TMIT value was below 1×10^8 , there was not enough beam to record a reliable value of the dispersion, so the raw data merely repeat the last known value. This often erroneously repeats a “normal” value, so we artificially set the dispersive reading to 100 (well outside its feasible range) when the minimum TMIT condition is violated. Second, we apply a dispersive scaling to each BPM to place it on the order of beam energy loss. The scalings are as follows: (BPMS:LTUH:250, 0.125), (BPMS:LTUH:450, -0.125), (BPMS:DMPH:502, 0.964), and (BPMS:DMPH:693, 0.469).

The candidates CSV file lists all of the candidates and has six columns: (i) start (earliest time anomaly could have occurred); (ii) end (latest time anomaly could have occurred); (iii) klys: (associated rf station); (iv) source (reporting diagnostic signal - AMM or AMPL); (v) corroborate (detailed below); and (vi) corr anomaly list (lists the associated rf stations and the reporting diagnostic source). The corroborate column can take several different values, depending on what the other diagnostic signal (AMM if considering AMPL candidates and vice versa) reports: (i) missing (other diagnostic signals did not flag this candidate); (ii) reinforced (other diagnostic signal flagged the candidate and supports the attribution); or (iii) disagree (other diagnostic signal flagged the candidate but disagrees on the rf station attribution).

The labels CSV file gives the hand labels for the candidates. Like the candidates’ file, it has start and end columns. It has two additional columns: (i) is_anom: Boolean indicating whether the candidate is an anomaly (ii) anom_type: Either “f” for *Fault*, “s” for *Sustained anomaly*, or blank for no anomaly

APPENDIX E: CLASSIFIER EXPERIMENTAL SETUP

We trained a deep neural network on the BPM data and labels described in Appendix D. We use the last 720 pulses before the candidate’s end time to create a constant input size of (8, 720) to our network, which we defined using

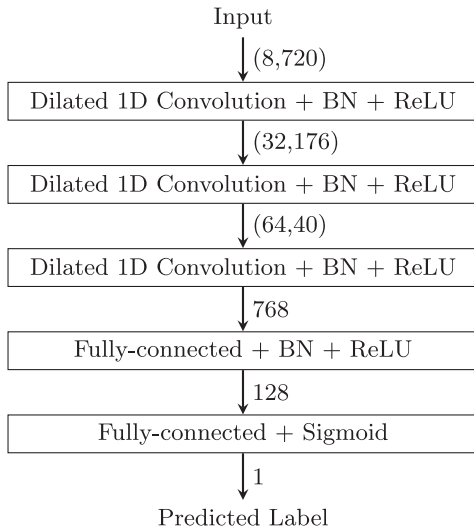


FIG. 8. The model architecture of the demonstration DNN, showing the input size to each layer.

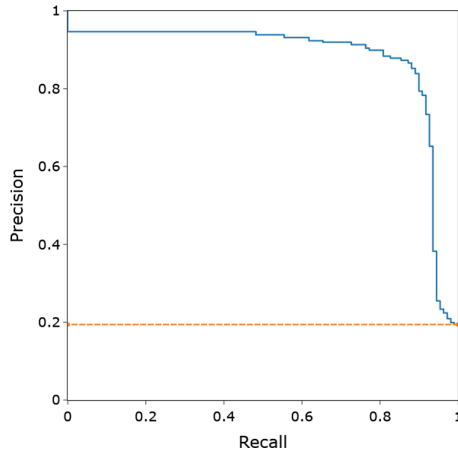


FIG. 9. Interpolated precision-recall curve for the DNN classifier on our new rf station anomaly dataset. The dashed line is the precision-recall curve for a random classifier.

PyTorch [31] and initialized using Kaiming’s uniform distribution [32]. As shown in Fig. 8, our network consists of three 1D convolutions with kernel size 8, padding 1, dilation 3, stride 4, and out channels of 32, 64, 128, respectively, followed by two FC layers of output size 128 and 1. All layers but the final layer are followed by a 1D batch normalization and ReLU activation. We split the dataset into a train and test set with a 80%/20% split and normalize the data using the median and MAD of the train data. We use a binary cross entropy loss with positive weighting equal to 4.15 since the candidate set is only 19.4% anomalous. We use the Adam optimizer [33] with a learning rate of 10^{-4} and weight decay of 10^{-6} and train for 240 epochs. The resulting precision-recall curve is shown in Fig. 9. We also trained the classifier with several different random splits and with a split done in time (i.e., the most

recent examples become the test set). We found no significant performance difference due to different seeds; however, splitting in time does cause the F1 score to drop to 0.78 from 0.88. We suspect that this is due to a notable difference in the anomaly percentage between the train and test sets when splitting in time: the train set is 21.2% anomalous while the test set is only 12.5% anomalous. Random splits of the dataset do not exhibit this same discrepancy. The practical problem of dealing with data drift in the rf station anomaly dataset, such as anomaly fraction, is deferred to future work.

- [1] M. Fazio, G. Laramore, S. Pillai, A. Badruzzaman, H. Martz, J. Buchsbaum, D. Jaffray, M.-K. Boss, A. Schmidt, J. Calame *et al.*, Basic research needs Workshop on Compact Accelerators for Security and Medicine: Tools for the 21st century, U.S. DOE Office of Science (SC), Technical Report, 2019.
- [2] T. Drew, M. L.-H. Vö, and J. M. Wolfe, The invisible gorilla strikes again, *Psychol. Sci.* **24**, 1848 (2013).
- [3] D. J. Simons and C. F. Chabris, Gorillas in our midst: Sustained inattention blindness for dynamic events, *Perception* **28**, 1059 (1999).
- [4] J. M. Wolfe, T. S. Horowitz, and N. M. Kenner, Rare items often missed in visual searches, *Nature (London)* **435**, 439 (2005).
- [5] Only downtime events lasting longer than 6 min are tracked. The downtime statistics are therefore underestimates.
- [6] Smart linac—smart anomaly detection and maintenance planning platform for linear accelerators, <https://kt.cern/kt-fund/projects/smart-linac-smart-anomaly-detection-and-maintenance-planning-platform-linear> (2019), accessed June 3, 2021.
- [7] P. Domingos, A few useful things to know about machine learning, *Commun. ACM* **55**, 78 (2012).
- [8] Y. Donon, A. Kupriyanov, D. Kirsh, A. D. Meglio, R. Paringer, P. Serafimovich, S. Syomic, Y. Donon, A. Kupriyanov, D. Kirsh, A. D. Meglio, R. Paringer, P. Serafimovich, and S. Syomic, Anomaly detection and breakdown prediction in RF power source output: A review of approaches, 27th International Symposium Nuclear Electronics and Computing, Budva, Montenegro (2019), pp. 99–104.
- [9] C. Tennant, A. Carpenter, T. Powers, A. S. Solopova, L. Vidyaratne, and K. Iftekharuddin, Superconducting radio-frequency cavity fault classification using machine learning at Jefferson Laboratory, *Phys. Rev. Accel. Beams* **23**, 114601 (2020).
- [10] M. Wielgosz, A. Skoczeń, and M. Mertik, Using LSTM recurrent neural networks for monitoring the LHC superconducting magnets, *Nucl. Instrum. Methods Phys. Res., Sect. A* **867**, 40 (2017).
- [11] S. Li, M. Zacharias, J. Snuverink, J. C. de Portugal, F. Perez-Cruz, D. Reggiani, and A. Adelmann, A novel approach for classification and forecasting of time series in particle accelerators, *Information* **12**, 121 (2021).

- [12] M. Rescic, R. Seviour, and W. Bloklund, Predicting particle accelerator failures using binary classifiers, *Nucl. Instrum. Methods Phys. Res., Sect. A* **955**, 163240 (2020).
- [13] E. Fol, R. Tomás, J. C. D. Portugal, and G. Franchetti, Detection of faulty beam position monitors using unsupervised learning, *Phys. Rev. Accel. Beams* **23**, 102805 (2020).
- [14] J. P. Edelen and C. C. Hall, Autoencoder based analysis of RF parameters in the fermilab low energy Linac, *Information (Switzerland)* **12**, 238 (2021).
- [15] T. Dewitte, W. Meert, E. V. Wolputte, K. Leuven, and B. P. V. Trappen, Anomaly detection for CERN beam transfer installations using machine learning, *17th International Conference on Accelerator and Large Experimental Physics Control Systems* (JACoW Publishing, Geneva, Switzerland, 2019), pp. 1066–1070, [10.18429/JACoW-ICALEPCS2019-WEMPR010](https://doi.org/10.18429/JACoW-ICALEPCS2019-WEMPR010).
- [16] M. Piekarski, J. Jaworek-Korjakowska, A. I. Wawrzyniak, and M. Gorgon, Convolutional neural network architecture for beam instabilities identification in Synchrotron Radiation Systems as an anomaly detection problem, *Measurement* **165**, 108116 (2020).
- [17] G. Valentino, R. Bruce, S. Redaelli, R. Rossi, P. Theodoropoulos, and S. Jaster-Merz, Anomaly detection for beam loss maps in the large hadron collider, *J. Phys. Conf. Ser.* **874**, 012002 (2017).
- [18] W. Bloklund, P. Ramuhalli, C. Peters, Y. Yucesan, A. Zhukov, M. Schram, K. Rajput, and T. Jeske, Uncertainty aware anomaly detection to predict errant beam pulses in the SNS accelerator, [arXiv:2110.12006](https://arxiv.org/abs/2110.12006).
- [19] M. I. Radaideh, C. Pappas, J. Walden, D. Lu, L. Vidyaratne, T. Britton, K. Rajput, M. Schram, and S. Cousineau, Time series anomaly detection in power electronics signals with recurrent and convlstm autoencoders, *SSRN Electron. J.*, [10.2139/ssrn.4069225](https://doi.org/10.2139/ssrn.4069225) (2021).
- [20] F. Tilaro, B. Bradu, M. Gonzalez-Berges, M. Roshchin, F. Varela *et al.*, Model learning algorithms for anomaly detection in cern control systems, in *Proceedings of the 16th International Conference on Accelerator and Large Experimental Control Systems, ICALEPCS'17, Barcelona, Spain* (JACoW, Geneva, 2017), pp. 8–13, [10.18429/JACoW-ICALEPCS2017-TUCPA04](https://doi.org/10.18429/JACoW-ICALEPCS2017-TUCPA04).
- [21] Y. Donon, A. Kupriyanov, D. Kirsh, A. D. Meglio, R. Paringer, I. Rytsarev, P. Serafimovich, and S. Syomic, Extended anomaly detection and breakdown prediction in LINAC 4's RF power source output, *Proceedings of ITNT 2020- 6th IEEE International Conference on Information Technology and Nanotechnology* (Institute of Electrical and Electronics Engineers Inc., 2020), [10.1109/ITNT49337.2020.9253296](https://doi.org/10.1109/ITNT49337.2020.9253296).
- [22] A. S. Nawaz, S. Pfeiffer, G. Lichtenberg, and H. Schlarb, Self-organized critical control for the European XFEL using black box parameter identification for the quench detection system, *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)* (IEEE, New York, 2016), [10.1109/systol.2016.7739750](https://doi.org/10.1109/systol.2016.7739750).
- [23] A. Nawaz, S. Pfeiffer, G. Lichtenberg, and P. Rostalski, Anomaly detection for the european XFEL using a non-linear parity space method, *IFAC-Papers On Line* (Elsevier B. V., New York, 2018), pp. 1379–1386, [10.1016/j.ifacol.2018.09.554](https://doi.org/10.1016/j.ifacol.2018.09.554).
- [24] A. Nawaz, S. Pfeiffer, G. G. Lichtenberg, and P. Rostalski, Anomaly detection for cavity signals-results from the european XFEL, in *Proceedings of the 9th International Particle Accelerator Conference, Vancouver, BC, Canada* (JACoW, Geneva, Switzerland, 2018), Vol. 29.
- [25] T. Straumann, R. Akre, R. G. Johnson, D. Kotturi, P. Krejcik, E. Medvedko, J. Olsen, and S. Smith, LCLS beam-position monitor data acquisition system, in *Proceedings of International Conference on Accelerator and Large Experimental Physics Control Systems, ICALEPCS 2007, Knoxville, TN* (2007), WPPB40.
- [26] K. Kim, S. Allison, T. Straumann, and E. Williams, Real-time performance improvements and consideration of parallel processing for beam synchronous acquisition (BSA), in *Proceedings of the 15th International Conference on Accelerator and Large Experimental Physics Control Systems, ICALEPCS'15, Melbourne, Australia* (JACoW, Geneva, Switzerland, 2015), pp. 992–994, [10.18429/JACoW-ICALEPCS2015-WEPGF122](https://doi.org/10.18429/JACoW-ICALEPCS2015-WEPGF122).
- [27] F. R. Hampel, The influence curve and its role in robust estimation, *J. Am. Stat. Assoc.* **69**, 383 (1974).
- [28] R. P. Adams and D. J. MacKay, Bayesian online changepoint detection, [arXiv:0710.3742](https://arxiv.org/abs/0710.3742).
- [29] C. Allen, S. Anderson, R. Erickson, W. Linebarger, J. C. Sheppard, and M. Stanek, Opportunistic or event-driven maintenance at the Stanford Linear Accelerator Center, Stanford Linear Accelerator Center, Technical Report No. SLAC-PUB-7424, 1997.
- [30] <https://www.slac.stanford.edu/grp/ad/ard/rfanom/rfanom.html>.
- [31] A. Paszke *et al.*, PyTorch: An imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems*, edited by H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, and R. Garnett (Curran Associates, Inc., 2019), pp. 8024–8035.
- [32] K. He, X. Zhang, S. Ren, and J. Sun, in *Proceedings of 2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, 2015* (IEEE Computer Society, New York, 2015), pp. 1026–1034.
- [33] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, 2015*, edited by Y. Bengio and Y. LeCun (2015), [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).