Improving absorption spectroscopy measurements in attosecond x-ray free-electron laser experiments using machine learning

K. Hu,* O. G. Alexander, F. Egun, and J. P. Marangos†
Department of Physics, Imperial College London, South Kensington, London SW7 2AZ, United Kingdom

(Received 16 March 2026; accepted 17 August 2026; published 18 September 2026)

X-ray free-electron lasers (XFELs) provide ultrashort, high-brightness pulses for probing matter on femtosecond and attosecond timescales, but their stochastic shot-to-shot fluctuations make x-ray absorption spectroscopy (XAS) challenging, as both incident and transmitted spectra are required to be measured on the same shot. Conventional approaches that rely on averaged reference spectra have limited accuracy, and full-shot characterizations are impractical. Here, we demonstrate a machine learning approach to bypass these limitations. An artificial neural network predicts the incident spectral shape from recorded diagnostics, effectively providing the equivalent of simultaneous incident spectrum measurements based on recorded machine diagnostics. Applied to XAS of Mylar at the oxygen K-edge, the method improves the accuracy in the pre-edge region (<530.5 eV), reducing the root-mean-squared error from 0.0415 to 0.0082 OD (optical density). This approach remains highly robust with small target datasets and reduces experimental acquisition time, offering a strategy for more accurate XAS measurements in XFEL experiments where noninvasive, single-shot monitoring of the incident spectrum is not feasible.

DOI: 10.1103/8vc5-pt1f

I. INTRODUCTION

X-ray free-electron lasers (XFELs) [1–5] are among the most advanced light sources available today. Major XFEL facilities such as SLAC [2], SACLA [4], PAL-XFEL [6], SwissFEL [7], and the European XFEL [8] have enabled a wide range of experiments that were previously inaccessible with conventional synchrotron light sources [9]. With recent technological advances [10–12], the XFEL can now generate bright, coherent, and tunable attosecond x-ray pulses, enabling breakthroughs across a wide range of fields including attosecond science [13–16]. A major challenge is that XFEL pulses exhibit strong shot-to-shot stochastic fluctuations in the temporal and spectral domains [17]. These arise from the self-amplified spontaneous emission (SASE) process and variations in the electron beam properties delivered by the accelerator. As a result, consecutive XFEL shots can differ substantially, which complicates experiments that require precise spectral knowledge.

To overcome this challenge, one approach is to perform full x-ray characterization of every single XFEL shot, retaining only the pulses with desired properties. However, this is prohibitively expensive in terms of beamtime resources, and in practice selected diagnostics must be prioritized. In addition, some diagnostics, such as x-ray spectrometers, typically intercept the entire beam. Splitting the x-ray pulse into reference and experimental beams is not feasible for many experiments, as it increases the complexity of the experimental configuration and reduces photon flux to the sample. One has reported an approach that used two identical dispersive spectrometers to capture the reference and transmitted spectra simultaneously [18]. It requires precise alignment to correct for the intrinsic spatial chirp, and the XFEL operates in the overcompressed mode [19,20] to reduce the noise. Moreover, the analysis of these massive datasets is constrained by data transfer bandwidth and computational infrastructure, thereby impeding high-repetition-rate beam characterization [21,22].

A common application of XFELs is time-resolved x-ray absorption spectroscopy (XAS), which measures the absorption cross section of a material as a function of photon energy. In XAS measurements of condensed matter targets, absorption is determined directly by measuring both the incoming and transmitted spectra. This is not always practical because of the physical constraints of the experimental setups. It may be simpler to measure the products of the x-ray transitions such as total or partial fluorescence yields. However, because these yields are proportional to the incident intensity, stochastic SASE fluctuations are directly imprinted onto the signal. Without precise shot-to-shot normalization, this leads to statistical noise and spectral broadening.

In low-density matter, because the optical density of the target is very low, the total photoelectron or ion yield is typically measured. In this case, high-resolution XAS using techniques such as spectral domain ghost imaging requires characterization of the incoming spectrum [23], although this is simpler than in the condensed phase because it can be measured after transmission. In condensed phase targets,

*Contact author: kaiyu.hu19@imperial.ac.uk
†Contact author: j.marangos@imperial.ac.uk

when the incoming and transmitted spectra cannot be measured simultaneously, the incoming spectrum is typically estimated to be the mean of spectra that are recorded with the sample removed. The absorption is then typically approximated by the ratio of the mean transmitted spectrum to the mean incoming spectrum. This approximation is equivalent to assuming laboratory-time independence of the incoming spectrum. However, this generally fails because of large shot-to-shot fluctuations and long-term drifts at XFELs.

Machine learning (ML) methods, and, in particular, artificial neural networks (ANNs) [24], offer a promising way to infer spectral information from auxiliary diagnostics. Importantly, ANNs can exploit dedicated low-cost measurements and incidentally recorded machine parameters, reducing the need for complete characterization of every shot. A few studies have demonstrated the feasibility of this approach, for example, in predicting the central pulse energy [25], temporal power profiles [26], and spectral profiles [27,28] of pulses. In addition, real-time or quasi-real-time ML diagnostics have been integrated into XFEL control systems to automate facility operations [29,30]. However, the systematic application of ANNs to directly improve absorption spectroscopy remains unaddressed.

Here, we introduce an ANN-based framework for predicting incident XFEL spectra from machine parameters and auxiliary diagnostics. In practice, this approach provides the functional equivalent to measuring the incident and transmitted spectra simultaneously: The ANN prediction supplies the missing incident spectrum for each transmitted shot. We demonstrate that this approach significantly improves absorption measurements in the attosecond XFEL experiment, reducing noise and acquisition time compared to the traditional method.

II. METHODS

A. XFEL experiment setup

Using the same experimental configuration as described in Ref. [13], the experiment was performed at the ChemRIXS beamline of the Linac Coherent Light Source (LCLS) XFEL. The XFEL was operated in the x-ray laser-enhanced attosecond pulse (XLEAP) mode [10], which utilizes enhanced self-amplified spontaneous emission to generate high-power attosecond pulses. X-ray pulses tuned to just below the oxygen K-edge were delivered to the beamline at a repetition rate of 120 Hz. Based on a measured spectral bandwidth of 7 eV and a time-bandwidth product of approximately 1.5 times the transform limit—noting that for standard SASE this product is typically many times larger—the pulse duration is estimated to be 400 as full width at half maximum, consistent with prior angular streaking measurements [10].

The x-ray pulses were focused into a vacuum chamber using a pair of Kirkpatrick-Baez (KB) mirrors. The horizontal and vertical components of the wave front could be adjusted independently by translating the corresponding KB mirrors along the beam axis, allowing control over the focal position. The pulse spectrum was characterized using a Hettrick-Underwood x-ray spectrometer comprising an elliptical mirror and a variable-line-spacing reflection grating. The dispersed x-rays were recorded with a charge-coupled device camera operated in full vertical binning mode, enabling data acquisition at the full 120 Hz repetition rate.

The sample consisted of a freestanding Mylar (polyethylene terephthalate, PET) film with a nominal thickness of 0.9 μm and a density of $1.39\,\mathrm{g\,cm^{-3}}$. The Mylar film was nonmetallized and mounted normal to the incident x-ray beam for transmission measurements.

B. Machine learning framework

Figure 1 shows how the machine learning method is used to improve absorption measurements. The absorbance A is calculated from the transmitted spectrum $S_{\mathrm{meas,in}}$ and the corresponding incident (reference) spectrum. In the machine learning approach [Fig. 1(b)], the incident spectrum $S_{\mathrm{pred,in}}$ is predicted on a shot-by-shot basis using the same set of experimental parameters recorded when measuring the transmitted spectrum. It essentially provides a functional equivalent to measuring both spectra simultaneously. In contrast, the traditional method [Fig. 1(c)] approximates the incident spectrum $S_{\mathrm{meas,out}}$ by measuring the spectra with the sample removed at a different laboratory time.

We first trained a base model on a large reference dataset and then fine-tuned it using target experiment reference shots. The training dataset for our base model contains more than 300 000 "sample-out" shots at a nominal central pulse energy of 529 eV. For the absorption experiment, where only a limited number of reference shots were collected, we retrained the base model on this dataset. The retrained model is used to predict the incident spectrum for each transmitted shot, which is then combined with the measured transmitted spectrum to calculate the absorbance. This transfer learning approach reduces both training and acquisition times, ensuring that reference and transmitted spectra are recorded closer in laboratory time. Therefore, the data are less affected by the drifts at XFEL facilities.

C. Feature construction and data processing

The inputs consist of 28 parameters selected manually based on experimental availability and their *a priori* relationship to undulator lasing or XFEL radiation output. These parameters encompass pulse energy measurements from gas monitor detectors (GMDs) and fluorescence intensity monitors (FIMs), alongside electron beam properties. We calculated the permutation importance [31] of each parameter to assess feature relevance (see Appendix B). GMD readings exhibit the largest influence on the model, and electron beam properties also rank highly. FIM signals show smaller contributions than GMD and electron beam features. Since GMDs provide a global pulse energy reference and FIMs monitor the pulse energy at the sample position, both sets of measurements are retained in the model.

The input pipeline extracts 20 beam-diagnostic parameters and the raw FIM data from eight detector channels. The FIM data are integrated via an absolute sum for each channel, providing the remaining eight features of the 28-parameter input matrix (number of shots × 28 features). Shot-to-shot filtering uses an experimental light flag to drop shots with

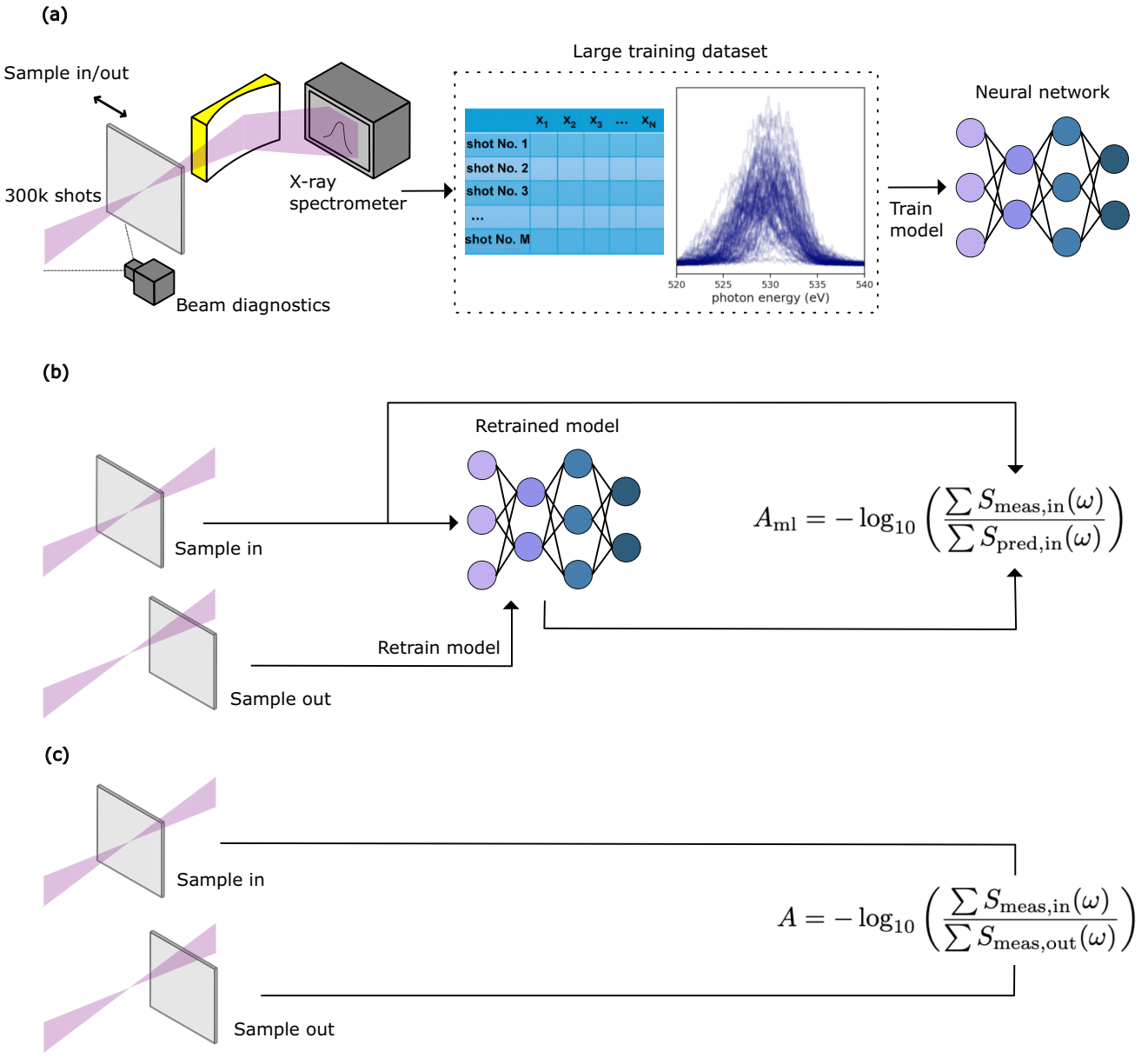

FIG. 1. Machine learning method framework. (a) Experimental data collection. Approximately 300 000 XFEL shots acquired at the Linac Coherent Light Source with a nominal photon energy of 529 eV are used to train the base model. All spectra are overlaid to illustrate intrinsic shot-to-shot fluctuations. (b) Machine learning method. A small dataset of the experiment (∼5000 shots, sample removed) is used to retrain the base model. For measurements with the sample in place, the incidentally recorded machine parameters are passed through the retrained model to predict the incident spectrum $S_{\mathrm{pred,in}}$, and the transmitted spectrum $S_{\mathrm{meas,in}}$ is measured by the x-ray spectrometer. The absorbance A_{ml} is calculated from the predicted incident and measured transmitted spectra. (c) Traditional method. The transmitted spectra and reference spectra $S_{\mathrm{meas,out}}$ are measured at different times.

no x-ray light, along with the immediate subsequent shot to eliminate transient effects. Features with zero variance or entirely nonfinite values are discarded. Any rows containing missing values (NaN) are purged, and low-power shots with a GMD pulse energy below 0.02 mJ are removed. A detector calibration offset is corrected by applying a −24 pixel shift for all experimental runs past index 150. Finally, input parameters are scaled using a standard scaler to ensure equal weighting across features.

For the target outputs, the x-ray spectrometer spectra are cropped to an energy window of 520–540 eV, yielding 198 discrete data points. The spectral data are subjected to the identical shot-level filtering applied to the input features. Principal component analysis (PCA) [32] is then employed to compress these spectral shapes into 20 principal components for efficient network training. While 198 points are computationally manageable, this compression maps the correlated bins into an orthogonal space to prevent the network from overfitting to high-frequency noise. It also ensures the pipeline scales well if future diagnostics use much higher spectral resolutions. This dimensionality reduction study (detailed in Appendix C) demonstrates that the first 20 principal

components account for 98.4% of the total variance in the spectral data. Because the reconstruction error plateaus beyond this point and the network prediction accuracy has no significant gains with higher dimensions, the target output is fixed at 20 components. During training, the model learns to predict these 20 component amplitudes, and the full 198-point spectrum is reconstructed via an inverse PCA transformation.

D. Neural network architecture

We use a multilayer perceptron (MLP) [33], which has been applied to XFEL pulse prediction tasks [25,27,28,34]. The model learns a mapping from the experimental parameters to the incident spectrum through a series of nonlinear transformations. While convolutional neural networks (CNNs) have proven effective for image-based diagnostics, such as electron beam longitudinal phase-space images [35] and angular streaking spectrograms [36], they offer no clear advantage for the present one-dimensional (1D) spectral prediction task. We benchmarked the MLP against both a CNN and a long short-term memory (LSTM) network. The MLP achieved slightly better predictive performance than either alternative architecture (Appendix A) and was adopted for all subsequent analyses.

The neural network was implemented using TensorFlow/Keras [37]. It consists of an input layer spanning 28 features, three hidden layers containing 32, 64, and 32 neurons, respectively, and a linear output layer with 20 neurons, corresponding to 20 principal components of the spectrum of the XFEL pulses. Nonlinearity in the hidden layers is mediated via rectified linear unit (ReLU) activation functions [38]. The dataset is partitioned into training, validation, and testing subsets using a 70:15:15 ratio, and the Adam optimizer [39] is employed to minimize a mean-squared error (MSE) loss function. Base model training utilizes a learning rate of 0.001, a batch size of 64, and 128 epochs. For the retrained model, the learning rate is adjusted to 0.0001 with a batch size of 64 and 64 epochs. These hyperparameter configurations are optimized via a grid search.

E. Model evaluation

Prediction quality was quantified by the cosine similarity, defined as the normalized dot product between predicted and measured spectra [Eq. (1)]. Unlike MSE, which is highly sensitive to small peak shifts, this metric captures the overall spectral overlap. This approach represents standard practice. Li *et al.* [27] employed an identical metric, and Sanchez-Gonzalez *et al.* [28] used a vector-normalized agreement metric with a conceptually equivalent form:

$$\text{Similarity}(\mathbf{A},\mathbf{B}) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\,\|\mathbf{B}\|}. \tag{1}$$

For XAS, the peak photon energy is very important, so we also calculate the gap between the predicted and measured peak energies. The peak position is defined as the intensity-weighted centroid of the spectrum over the full spectral range (520–540 eV):

$$\bar{E} = \frac{\sum_i E_i \times I_i}{\sum_i I_i}, \tag{2}$$

where E is the photon energy and I is the intensity. Negative intensity values (detector noise artifacts) are clipped to zero before computing the centroid.

III. RESULTS

A. Validation of the retrained model for incident-spectrum prediction

To demonstrate the applicability of our approach, we performed XAS on Mylar near the oxygen K-edge. The finite bandwidth of the attosecond pulses [10] presents a significant challenge for accurate absorption measurements, particularly in the spectral wings where the signal-to-noise ratio is lowest. This provides a stringent test case for our ML-based reconstruction, as it requires the model to accurately recover the incident spectral shape even in regions of low photon flux.

We measured only ∼5000 shots of reference spectra at 120 Hz, taking less than 1 min. Approximately 2 min later, we measured ∼5000 spectra in transmission. In both cases, the XFEL was tuned so that the center of the x-ray spectrum was nominally at 532 eV, i.e., at the oxygen K-edge. These experimental data were used to retrain the ML model.

To evaluate the performance of the retraining, we split the reference spectra into train, validation, and test data with a ratio of 70:15:15 and used the test data to evaluate the similarity [Eq. (1)] of the predicted spectra to the measured spectra. Figure 2 illustrates the performance of the retrained model. The mean similarity is 0.9081 which is lower than that of the base model (0.9427) due to the smaller training set. 71.6% shots have a similarity greater than 0.9. In Fig. 2(b), the average peak mismatch [peak defined by Eq. (2)] is close to zero, while a few shots have a large peak mismatch due to the spike at ∼524 eV [Fig. 2(d)], which is far away from the nominal energy. 88.1% of shots have peak energy mismatch within 1 eV and 1.4% shots are outliers beyond ±2 eV. The mean residual across all energy points is approximately zero [Fig. 2(e)], and the residual distribution is roughly symmetrical around x axis, indicating that the model produces unbiased intensity predictions on average. The residual standard deviation is highest at the absorption edge (∼530–532 eV).

B. Improved x-ray absorption spectroscopy of Mylar at the oxygen K-edge

We predict the incident spectra when measuring in transmission and used the predictions to calculate the absorbance of Mylar (Fig. 3). At a central photon energy of 532 eV, the transmitted spectra have significant spectral intensity between approximately 522 and 535 eV, with a skew in average spectrum having a longer low-energy tail than higher-energy tail. This allows us to measure the first two peaks of the oxygen K-edge and an ∼10 eV region below the first peak. Above approximately 535 eV, the XFEL photon flux becomes insufficient for accurate measurement of the transmitted spectrum. In transient absorption spectra, the pre-edge region is typically the most sensitive to changes in the electronic configuration, so accurate measurement of this region is necessary. Before the first peak, the only absorption channel is valence ionization, so we expect the energy dependence to be approximately flat. This behavior is consistent with previously reported

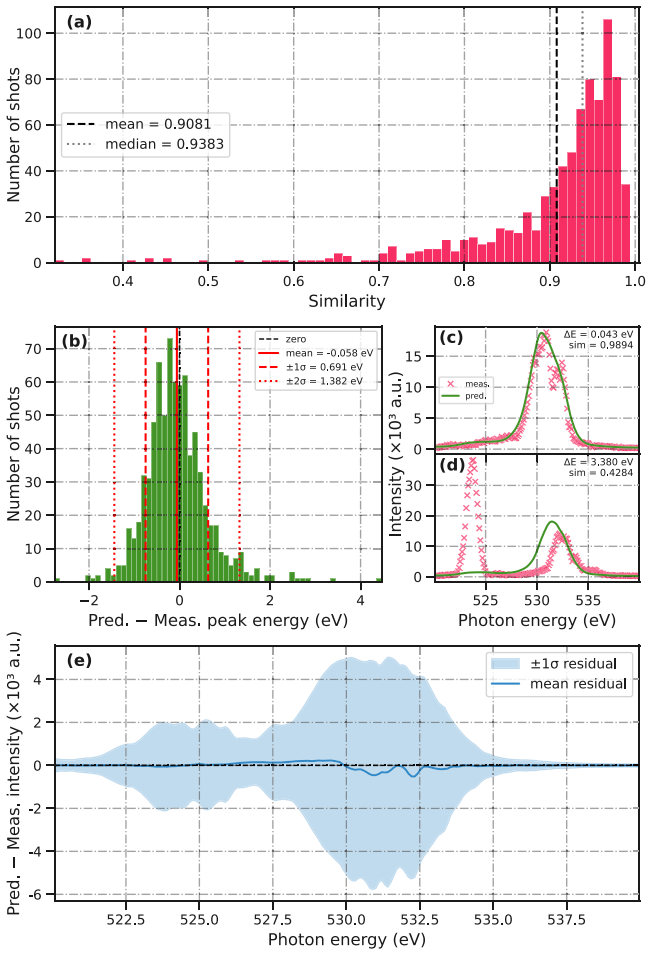

FIG. 2. Evaluation of the retrained model on the test data. (a) The distribution of the similarity [Eq. (1)] between predictions and measurements. (b) The distribution of the peak photon energy mismatch ΔE. (c) An example of "good" shot prediction with low ΔE and high similarity. (d) An example of "bad" shot prediction with high ΔE and low similarity. (e) Residuals between predictions and measurements.

x-ray absorption spectra of Mylar measured using synchrotron radiation, which we use as a benchmark [40]. Similar spectral features have been widely observed in earlier studies [41–43].

Our ML method improves the measurement of the spectrum in the pre-edge region while preserving the positions of the three absorption peaks. It reduces the root-mean-squared error (RMSE) in the pre-edge region (photon energy <530.5 eV) from 0.0415 OD (traditional) to 0.0082 OD (ML). The RMSE in the energy range 527–535 eV decreased from 0.0415 to 0.0230 OD.

C. Robust absorbance recovery with limited sample size

We also tested the robustness of the method with smaller sample sizes. Random subsets of transmission spectra were selected, and the absorbance was calculated with and without ML correction. As shown in Fig. 4(a), the ML method consistently outperformed the traditional approach across all sample sizes. Notably, even with as few as 100 shots, the ANN still yielded a relatively accurate absorbance spectrum.

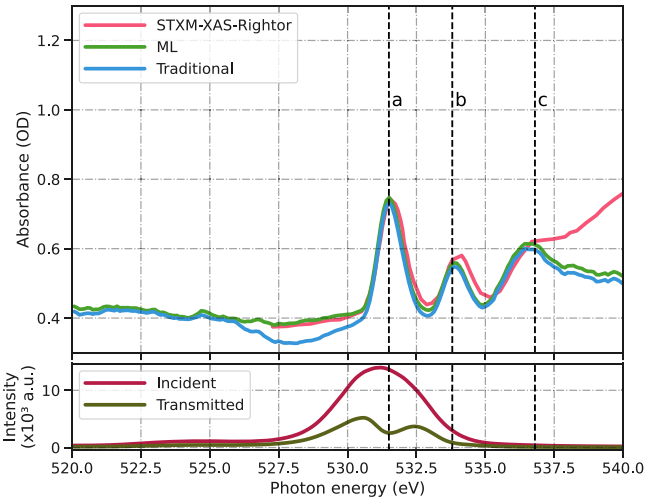

FIG. 3. Implementation of machine learning method on the calculation of x-ray absorbance of Mylar at the oxygen K-edge. Upper panel: The red curve shows a reference x-ray absorption spectrum measured by scanning transmission x-ray microscopy (STXM-XAS) [40], used as a benchmark. The green and blue curves represent the Mylar absorbance calculated using the machine learning method and the traditional method, respectively. The STXM-XAS data were originally reported as oscillator strength per oxygen atom after background subtraction; for comparison, the spectrum is converted to optical density (OD) and the same background offset as in the XFEL measurements is applied. Lower panel: Mean incident XFEL spectra and corresponding transmitted spectra used to calculate the absorbance.

IV. DISCUSSIONS

The results presented here demonstrate that ANN prediction of incident spectra can improve absorption measurements at XFELs. In the Mylar experiment, the framework enables reliable recovery of the incident spectrum on a shot-to-shot basis, leading to an improvement in the reconstructed absorbance, particularly in the pre-edge region. This improved reconstruction of the pre-edge region is critical because this spectral range is highly sensitive to changes in electronic structure, and has previously been difficult to access accurately at XFELs due to shot-to-shot spectral fluctuations.

Our workflow can also be understood as a form of transfer learning: A base model trained on a large dataset (300k sample-out shots at 529 eV) is successfully adapted to the Mylar experiment with a slightly different nominal energy at 532 eV. A key advantage is the reduction in acquisition time: In our demonstration, reliable retraining was achieved with only about 5000 reference shots, corresponding to less than 1 min of data collection at 120 Hz. This is an important benefit for attosecond experiments, where beamtime is limited and machine stability windows are short. As shown in Fig. 4, the ML method yields good absorbance spectra already at a sample size of only 100 shots, with minor performance gains when scaling to the full dataset. This greatly reduces the required exposure time, which is particularly important for radiation-sensitive samples such as biological specimens and thin films, where prolonged illumination can cause irreversible damage. This ML approach mitigates radiation damage constraints, enabling the characterization of highly sensitive specimens.

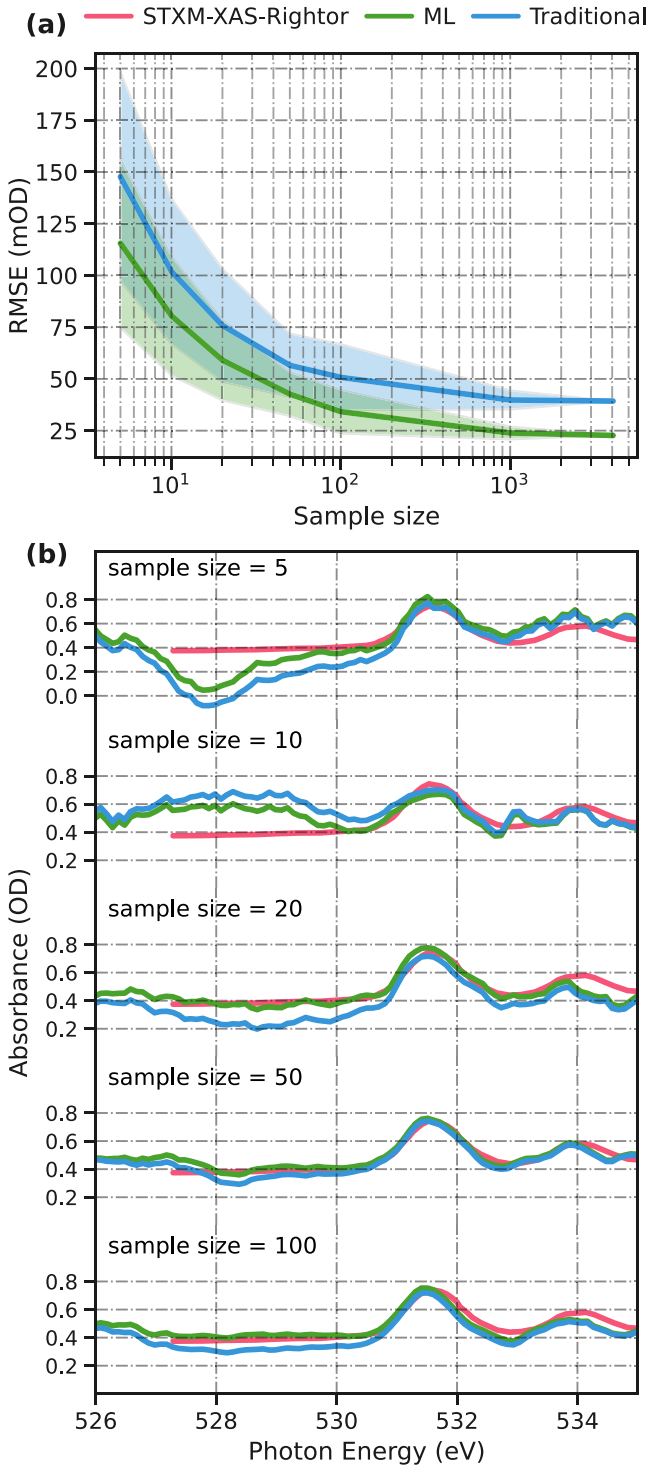

FIG. 4. Performance of the machine learning approach at different sample sizes. (a) The average root-mean-squared error of Mylar absorbance between 527 and 535 eV against the sample size for the machine learning method (green) and the traditional method (red). The average RMSE is calculated over 100 iterations and the shadow area represents the range within one standard deviation. Shots are randomly chosen in each sample. (b) Examples of the absorbances obtained by machine learning method (green), experimental XFEL data (blue), and experimental synchrotron data (red) [40].

Our method relies only on diagnostics that are routinely available at XFELs and does not require beam splitting or invasive spectral monitoring. It can be integrated into existing experimental setups with minimal hardware modification. In pump-probe spectroscopy, where both incident and transmitted spectra must be known for each delay, the approach could enable high-throughput measurements without compromising accuracy. Similarly, in resonant inelastic x-ray scattering (RIXS) and other techniques that depend on precise spectral characterization, ANN could serve as a practical surrogate for direct monitoring. More generally, integration of this framework into beamline diagnostics could provide real-time correction for spectral fluctuations, enhancing stability across a wide range of experiments.

Despite these advantages, several limitations remain. The base model was trained exclusively under specific machine conditions at a nominal photon energy of 529 eV (within a 520–540 eV spectral window for the oxygen K-edge) and in XLEAP mode. Its generalizability across disparate photon energies, pulse durations, or accelerator configurations remains to be verified. If the XFEL is tuned for an entirely different absorption edge (e.g., the carbon or nitrogen K-edge), new datasets must be acquired to train the base model. Although we have demonstrated that the retrained model successfully reconstructs absorbance for an experiment with a slightly shifted central photon energy, datasets with vastly different photon energies or bandwidths would lie outside the network's current interpolation regime. In such scenarios, the base model must be trained on the new domain before the transfer learning framework can be applied. In addition, while the ANN achieves excellent predictive performance, it acts as a statistical mapping tool rather than an explicit physical model of SASE fluctuations. However, the present results demonstrate that ANN-based virtual diagnostics can serve as an effective surrogate for simultaneous incident and transmitted spectral measurements. Future work will focus on extending this framework to broader energy ranges and exploring online adaptation strategies to compensate for slow drift in machine pointing or accelerator gradients during extended experimental runs.

V. CONCLUSIONS

We have demonstrated a machine learning framework for predicting incident x-ray spectra to enhance the accuracy of absorption spectroscopy in attosecond XFEL experiments. By utilizing an ANN trained on routinely recorded electron and photon diagnostics, and adapting it with a modest set of *in situ* reference shots, we effectively reconstruct the incident spectrum for every transmitted shot. This approach provides a functional equivalent to simultaneous reference measurements, bypassing the need for invasive beam-splitting optics or dedicated reference beamlines.

Using Mylar at the oxygen K-edge as a benchmark, the method significantly improves spectral reconstruction in the sensitive pre-edge region, outperforming conventional mean-reference approaches and accurately reproducing

synchrotron-validated baselines. The RMSE in the pre-edge region of the ML method is 0.0082 OD, five times lower than that of the traditional method. Notably, the transfer learning framework requires less than 1 min of data collection at 120 Hz for retraining. Furthermore, the model demonstrates remarkable robustness, maintaining high predictive accuracy with a sample size of as few as 100 shots. Reduction of both total acquisition time and the cumulative radiation dose is particularly advantageous for sensitive specimens and high-throughput pump-probe studies, where machine stability windows are narrow.

Future work will target extending the model's spectral range, implementing online adaptation to track long-term accelerator drifts, and improving the physical interpretability of the diagnostic correlations. These results establish machine learning as a practical and powerful tool for the next generation of high-repetition-rate XFEL experiments.

ACKNOWLEDGMENTS

We would like to acknowledge the support of ChemRIXS beamline scientists Georgi Dakovski, Douglas Garret, Stefan Moeller, and Ming-Fu Lin and the LX52 collaboration. Funding by United Kingdom Engineering and Physical Sciences Research Council under Research Grants No. EP/X026094/1 and No. EP/V026690/1.

DATA AVAILABILITY

The data and code supporting the findings of this study, including the experimental dataset, machine learning pipeline, and data analysis scripts, are publicly available via Zenodo [44].

APPENDIX A: BENCHMARKING ALTERNATIVE NEURAL NETWORK ARCHITECTURES

In addition to the MLP adopted in the main text, we benchmarked two alternative neural architectures (CNN and LSTM) on the same dataset and data split (70:15:15 ratio with seed 42). Inputs were standardized per feature. Predictive quality is quantified by the similarity [Eq. (1)] between the predicted and measured spectra. Figure 5 shows probability-density histograms with shared bin edges. For each model, hyperparameters were selected by grid searches.

A convolutional neural network was included because each spectrum is a one-dimensional signal, effectively a 1D image along the photon-energy axis. Local convolutions provide a natural inductive bias for features such as peaks and shoulders. Our CNN takes each per-shot feature vector $\mathbf{x} \in \mathbb{R}^d$, reshapes it to a single 1D channel, applies two 1D convolutional layers ($1 \to 16$ and $16 \to 32$ channels, kernel size 3, padding 1) with ReLU activations, flattens the result, and maps through a fully connected layer to a linear output that produces the spectrum. A grid search over learning rate and hidden-layer width selected the best configuration with hidden size 256, which achieved a mean similarity of 0.937 on the test set.

We also evaluated an LSTM network to capture potential time-dependent drifts or correlations in machine parameters (e.g., electron-beam properties). Time-ordered windows were formed within stable machine segments; each training sample used the preceding window of parameters to predict the spectrum at the final time step. Concretely, windows of length 50 were standardized across time/features and passed to a single-layer LSTM with 64 hidden units, followed by a linear head. The optimized configuration achieved a mean similarity of 0.919.

The CNN and LSTM distributions lie slightly to the left of the MLPs, indicating modestly lower predictive fidelity. Combined with their higher training cost and lower interpretability for our application, these results motivated our use of the MLP.

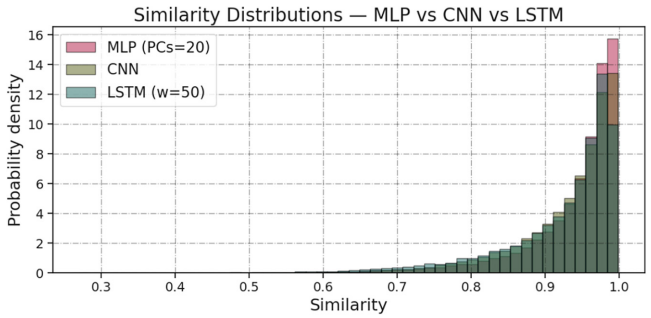

FIG. 5. Normalized (probability-density) histograms of the similarities. The MLP (see main text) exhibits the most right-skewed distribution. The best CNN model has a mean similarity of 0.937 and the optimized LSTM model has 0.919.

APPENDIX B: FEATURE IMPORTANCE ANALYSIS

Table I lists the 28 experimental parameters used to train the ANN.

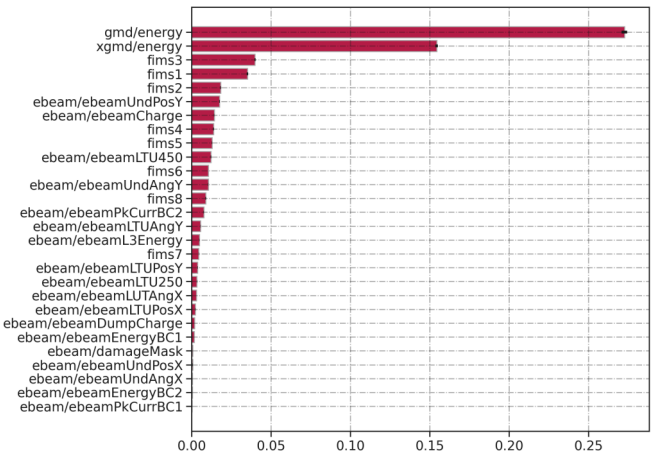

FIG. 6. Permutation importance of experimental parameters for the machine learning model, evaluated on the test set using the similarity metric. The baseline model performance without feature shuffling was 0.94. Bars indicate the average decrease in similarity score when the corresponding feature was randomly permuted; error bars denote the standard deviation across ten shuffling repetitions. Larger drops signify greater feature relevance to the model's predictive accuracy.

TABLE I. Key and meaning of parameters used to train the ANN.

Key	Meaning
ebeam/damageMask	Electron beam damage or validity mask flag for the shot
ebeam/ebeamCharge	Total electron bunch charge
ebeam/ebeamDumpCharge	Electron charge measured at the beam dump
ebeam/ebeamEnergyBC1	Electron beam energy measured at bunch compressor 1 (BC1)
ebeam/ebeamEnergyBC2	Electron beam energy measured at bunch compressor 2 (BC2)
ebeam/ebeamL3Energy	Electron beam energy after linac 3 (final beam energy)
ebeam/ebeamLTU250	Electron beam energy or setting at the linac-to-undulator (LTU) location 250 m
ebeam/ebeamLTU450	Electron beam energy or setting at the LTU location 450 m
ebeam/ebeamLTUAngY	Vertical electron beam angle in the LTU section
ebeam/ebeamLTUPosX	Horizontal electron beam position in the LTU section
ebeam/ebeamLTUPosY	Vertical electron beam position in the LTU section
ebeam/ebeamLUTAngX	Horizontal electron beam angle in the LTU transport
ebeam/ebeamPkCurrBC1	Peak electron beam current near bunch compressor 1 (BC1)
ebeam/ebeamPkCurrBC2	Peak electron beam current near bunch compressor 2 (BC2)
ebeam/ebeamUndAngX	Horizontal electron beam angle in the undulator section
ebeam/ebeamUndAngY	Vertical electron beam angle in the undulator section
ebeam/ebeamUndPosX	Horizontal electron beam position in the undulator section
ebeam/ebeamUndPosY	Vertical electron beam position in the undulator section
gmd/energy	X-ray pulse energy measured by the gas monitor detector upstream
xgmd/energy	X-ray pulse energy measured by the gas monitor detector downstream in beam transport hall
rix_fim0/integral_0	Integrated fluorescence signal from channel 0 of FIM detector 0
fims1	Integrated fluorescence signal from channel 1 of FIM detector 1
fims2	Integrated fluorescence signal from channel 2 of FIM detector 1
fims3	Integrated fluorescence signal from channel 3 of FIM detector 1
fims4	Integrated fluorescence signal from channel 4 of FIM detector 1
fims5	Integrated fluorescence signal from channel 5 of FIM detector 1
fims6	Integrated fluorescence signal from channel 6 of FIM detector 1
fims7	Integrated fluorescence signal from channel 7 of FIM detector 1
fims8	Integrated fluorescence signal from channel 8 of FIM detector 1

We use permutation importance to rank how much each experimental parameter contributes to the model's predictions. The method shuffles the values of a single feature across the test set to disrupt its relationship with the target and then measures the drop in the similarity metric. A larger drop in performance means the model relies more heavily on that feature.

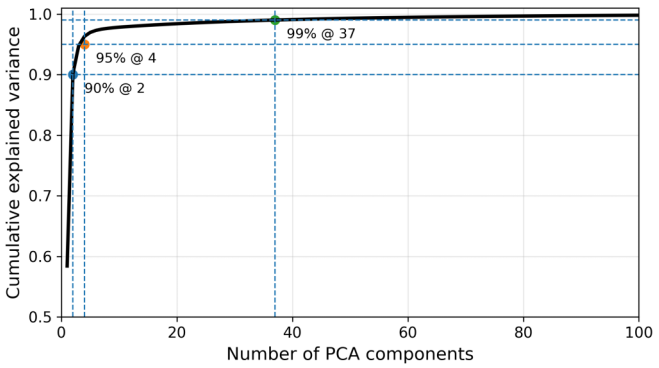

FIG. 7. Cumulative explained variance ratio (EVR) of the spectra. The first few principal components capture the vast majority of the variance, with 2 components explaining 90%, 4 components explaining 95%, and 37 components explaining 99%. Using 20 components already accounts for 98.4% of the total variance.

As shown in Fig. 6, the most critical parameters are electron beam energy, trajectory, and peak current. This matches physical expectations, as these variables directly dictate the XFEL resonance condition and gain process. X-ray intensity monitors, specifically the upstream and downstream gas monitor detectors (GMD and XGMD), also rank highly because they act as direct proxies for pulse energy. Fluorescence imaging monitor channels show a smaller contribution, suggesting that these secondary diagnostics offer complementary

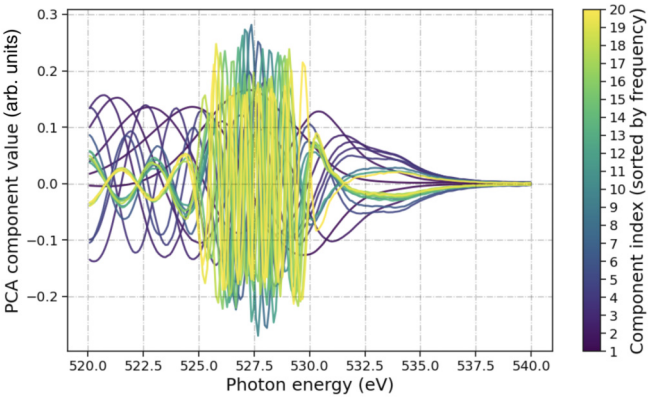

FIG. 8. Twenty PCA components sorted by dominant frequencies.

data on beam stability. This ranking confirms that the model relies on sound physical correlations rather than numerical artifacts.

APPENDIX C: PRINCIPAL COMPONENT ANALYSIS

We used principal component analysis to determine the effective dimensionality of the measured spectra. By calculating the explained variance ratio (EVR) and cumulative EVR for each principal component, we can track how much total spectral variance is captured as more components are added (Fig. 7). This analysis justifies truncating the target space to a reduced number of components (e.g., the subset explaining >95% of the variance). This cutoff preserves core spectral features while filtering out random measurement noise and redundant channels.

Because the cumulative explained variance levels off quickly, the spectral fluctuations clearly occupy a low-dimensional subspace. This means that a few underlying physical modes—like shifts in central photon energy, bandwidth changes, and intensity variations—drive most of the changes in these seemingly complex single-shot spectra. Truncating the spectra to a smaller PCA basis captures these primary drivers while filtering out shot noise. Figure 8 presents the 20 PCA components retained in the analysis, ordered according to their dominant frequencies.

[1] W. Ackermann *et al.*, Operation of a free-electron laser from the extreme ultraviolet to the water window, Nat. Photon. **1**, 336 (2007).

[2] P. Emma *et al.*, First lasing and operation of an ångstrom-wavelength free-electron laser, Nat. Photon. **4**, 641 (2010).

[3] E. Allaria *et al.*, Highly coherent and stable pulses from the FERMI seeded free-electron laser in the extreme ultraviolet, Nat. Photon. **6**, 699 (2012).

[4] T. Ishikawa *et al.*, A compact X-ray free-electron laser emitting in the sub-ångstrom region, Nat. Photon. **6**, 540 (2012).

[5] E. Allaria *et al.*, Two-stage seeded soft-X-ray free-electron laser, Nat. Photon. **7**, 913 (2013).

[6] H.-S. Kang *et al.*, Hard X-ray free-electron laser with femtosecond-scale timing jitter, Nat. Photon. **11**, 708 (2017).

[7] C. J. Milne *et al.*, SwissFEL: The Swiss X-ray free-electron laser, Appl. Sci. **7**, 720 (2017).

[8] W. Decking *et al.*, A MHz-repetition-rate hard X-ray free-electron laser driven by a superconducting linear accelerator, Nat. Photon. **14**, 391 (2020).

[9] S. E. Posen, Science opportunities and capabilities enabled by LCLS-II and LCLS-II HE, Technical Report FERMILAB-POSTER–20-005-TD, US Government Office of Science and Technology, 2020.

[10] J. Duris *et al.*, Tunable isolated attosecond X-ray pulses with gigawatt peak power from a free-electron laser, Nat. Photon. **14**, 30 (2020).

[11] S. Huang, Y. Ding, Y. Feng, E. Hemsing, Z. Huang, J. Krzywinski, A. A. Lutman, A. Marinelli, T. J. Maxwell, and D. Zhu, Terawatt-attosecond hard X-ray free-electron laser at high repetition rate, Nat. Photon. **18**, 1293 (2024).

[12] S. Huang, Y. Ding, Y. Feng, E. Hemsing, Z. Huang, J. Krzywinski, A. A. Lutman, A. Marinelli, T. J. Maxwell, and D. Zhu, Generating single-spike hard X-ray pulses with nonlinear bunch compression in free-electron lasers, Phys. Rev. Lett. **119**, 154801 (2017).

[13] O. Alexander *et al.*, Attosecond impulsive stimulated X-ray Raman scattering in liquid water, Sci. Adv. **10**, eadp0841 (2024).

[14] S. Li *et al.*, Attosecond-pump attosecond-probe X-ray spectroscopy of liquid water, Science **383**, 1118 (2024).

[15] T. Driver *et al.*, Attosecond delays in X-ray molecular ionization, Nature (London) **632**, 762 (2024).

[16] Z. Guo *et al.*, Experimental demonstration of attosecond pump–probe spectroscopy with an X-ray free-electron laser, Nat. Photon. **18**, 691 (2024).

[17] R. Bonifacio, L. De Salvo, P. Pierini, N. Piovella, and C. Pellegrini, Spectrum, temporal structure, and fluctuations in a high-gain free-electron laser starting from noise, Phys. Rev. Lett. **73**, 70 (1994).

[18] M. Harmand, M. Cammarata, M. Chollet, A. G. Krygier, H. T. Lemke, and D. Zhu, Single-shot X-ray absorption spectroscopy at X-ray free electron lasers, Sci. Rep. **13**, 18203 (2023).

[19] Y. Ding, K. L. F. Bane, W. Colocho, F. J. Decker, P. Emma, J. Frisch, M. W. Guetg, Z. Huang, R. Iverson, J. Krzywinski, H. Loos, A. Lutman, T. J. Maxwell, H. D. Nuhn, D. Ratner, J. Turner, J. Welch, and F. Zhou, Beam shaping to improve the free-electron laser performance at the Linac Coherent Light Source, Phys. Rev. Accel. Beams **19**, 100703 (2016).

[20] J. L. Turner, F.-J. Decker, Y. Ding, Z. Huang, R. Iverson, J. Krzywinski, H. Loos, A. Marinelli, T. J. Maxwell, H.-D. Nuhn, D. Ratner, T. Smith, J. Welch, and F. Zhou, FEL overcompression in the LCLS, in *Proceedings of the 36th International Free Electron Laser Conference (FEL2014)* (JACoW Publishing, Basel, Switzerland, 2014), p. TUB03.

[21] M. Harmand, R. Coffee, M. R. Bionta, M. Chollet, D. French, J. M. Glownia, S. Lee, H. T. Lemke, D. Zhu, and D. M. Fritz, Achieving few-femtosecond time-sorting at hard X-ray free-electron lasers, Nat. Photon. **7**, 215 (2013).

[22] V. Kimberg, N. Rohringer, K. Bennett, M. Kowalewski, J.-E. Rubensson, B. W. J. McNeil, I. S. Averbukh, R. Santra, and S. Mukamel, Stimulated X-ray Raman scattering: A critical assessment of the building block of nonlinear X-ray spectroscopy, Faraday Discuss. **194**, 305 (2016).

[23] T. Driver *et al.*, Attosecond transient absorption spooktroscopy: A ghost imaging approach to ultrafast absorption spectroscopy, Phys. Chem. Chem. Phys. **22**, 2704 (2020).

[24] B. Cheng and D. M. Titterington, Neural networks: A review from a statistical perspective, Stat. Sci. **9**, 2 (1994).

[25] K. K. Alaa El-Din, O. G. Alexander, L. J. Frasinski, F. Mintert, Z. Guo, J. Duris, Z. Zhang, D. B. Cesar, P. Franz, T. Driver, P. Walter, J. P. Cryan, A. Marinelli, J. P. Marangos, and R. Mukherjee, Efficient prediction of attosecond two-colour pulses from an X-ray free-electron laser with machine learning, Sci. Rep. **14**, 7267 (2024).

[26] X. Ren, Y. Li, X. Yang, Y. Ding, Z. Huang, G. Marcus, T. Maxwell, A. Brachmann, G. Hays, R. Coffee, J. P. Cryan, J. M. Glownia, M. P. Minitti, W. F. Schlotter, D. Zhu, F. Wang, and T. Zhang, Temporal power reconstruction for an X-ray free-electron laser using convolutional neural networks, Phys. Rev. Accel. Beams **23**, 040701 (2020).

[27] K. Li, G. Zhou, Y. Liu, J. Wu, M.-F. Lin, X. Cheng, A. A. Lutman, M. Seaberg, H. Smith, P. A. Kakhandiki, and A. Sakdinawat, Prediction on X-ray output of free electron laser based on artificial neural networks, Nat. Commun. **14**, 7183 (2023).

[28] A. Sanchez-Gonzalez *et al.*, Accurate prediction of X-ray pulse properties from a free-electron laser using machine learning, Nat. Commun. **8**, 15461 (2017).

[29] D. E. F. de Lima, T. Michelat, E. Sobolev, A. Davtyan, O. Turkot, F. Dall'Antonia, M. Sikorski, A. Round, R. de Wijn, J. Möller, S. Birnsteinova, R. Bean, and L. Gelisio, Core principles for machine learning at European XFEL, J. Phys.: Conf. Ser. **3010**, 012138 (2025).

[30] D. E. Ferreira de Lima, A. Davtyan, J. Laksman, N. Gerasimova, T. Maltezopoulos, J. Liu, P. Schmidt, T. Michelat, T. Mazza, M. Meyer, J. Grünert, and L. Gelisio, Machine-learning-enhanced automatic spectral characterization of X-ray pulses from a free-electron laser, Commun. Phys. **7**, 400 (2024).

[31] L. Breiman, Random forests, Mach. Learn. **45**, 5 (2001).

[32] I. T. Jolliffe and J. Cadima, Principal component analysis: A review and recent developments, Philos. Trans. R. Soc. A **374**, 20150202 (2016).

[33] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, Nature (London) **323**, 533 (1986).

[34] T. Korten, V. Rybnikov, P. Steinbach, and N. Mirian, Virtual pulse reconstruction diagnostic for single-shot measurement of free electron laser radiation power, Phys. Rev. Accel. Beams **28**, 030703 (2025).

[35] G. Goetzke, R. Plumley, G. Hartmann, T. Maxwell, F.-J. Decker, A. Lutman, M. Dunne, D. Ratner, and J. J. Turner, Femto-PIXAR: A self-supervised neural network method for reconstructing femtosecond X-ray free electron laser pulses, Opt. Express **33**, 31235 (2025).

[36] K. Dingel, T. Otto, L. Marder, L. Funke, A. Held, S. Savio, A. Hans, G. Hartmann, D. Meier, J. Viefhaus, B. Sick, A. Ehresmann, M. Ilchen, and W. Helml, Artificial intelligence for online characterization of ultrashort X-ray free-electron laser pulses, Sci. Rep. **12**, 17809 (2022).

[37] F. Chollet *et al.*, Keras, 2015, https://keras.io.

[38] V. Nair and G. E. Hinton, Rectified linear units improve restricted Boltzmann machines, in *Proceedings of the 27th International Conference on Machine Learning* (Omnipress, Madison, 2010), pp. 807–814.

[39] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)* (OpenReview.net, Alameda, CA, 2015).

[40] E. G. Rightor, A. P. Hitchcock, H. Ade, R. D. Leapman, S. G. Urquhart, A. P. Smith, G. Mitchell, D. Fischer, H. J. Shin, and T. Warwick, Spectromicroscopy of poly(ethylene terephthalate): Comparison of spectra and radiation damage rates in X-ray absorption and electron energy loss, J. Phys. Chem. B **101**, 1950 (1997).

[41] I. Ouchi *et al.*, Structure and core electron absorption spectra of polyester films, Jpn. J. Appl. Phys. **38**, 183 (1999).

[42] T. Okajima, K. Teramoto, R. Mitsumoto, H. Oji, Y. Yamamoto, I. Mori, H. Ishii, Y. Ouchi, and K. Seki, Polarized NEXAFS spectroscopic studies of poly(butylene terephthalate), poly(ethylene terephthalate), and their model compounds, J. Phys. Chem. A **102**, 7093 (1998).

[43] J. Wang, G. A. Botton, M. M. West, and A. P. Hitchcock, Quantitative evaluation of radiation damage to polyethylene terephthalate by soft X-rays and high-energy electrons, J. Phys. Chem. B **113**, 1869 (2009).

[44] K. Hu, O. Alexander, F. Egun, and J. Marangos, Improving absorption spectroscopy measurements in attosecond X-ray free electron laser experiments using machine learning [Dataset], Zenodo, 2026, https://doi.org/10.5281/zenodo.20816497