



High-sensitivity extreme-ultraviolet transient absorption spectroscopy enabled by machine learning

TOBIAS GUTBERLET,¹  HUNG-TZU CHANG,¹ 
SERGEY ZAYKO,^{1,2}  MURAT SIVIS,^{1,2}  AND CLAUS ROPERS^{1,2,*}

¹Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany

²4th Physical Institute – Solids and Nanostructures, University of Göttingen, 37077 Göttingen, Germany

*claus.ropers@mpinat.mpg.de

Abstract: We present a novel denoising scheme for spectroscopy experiments employing broadband light sources and demonstrate its capabilities using transient absorption measurements with a high-harmonic source. Our scheme relies on measuring the probe spectra before and after interacting with the sample while capturing correlations between spectral components through machine learning approaches. With the present setup we achieve up to a tenfold improvement in noise suppression in XUV transient absorption spectra compared to the conventional pump on/pump off referencing method. By utilizing strong spectral correlations in source fluctuations, the use of an artificial neural network facilitates pixel-wise noise reduction without requiring wavelength calibration of the reference spectrum. Our method can be adapted to a wide range of experiments and may be particularly advantageous for low repetition-rate systems, such as free electron lasers as well as laser-driven plasma and HHG sources. The enhanced sensitivity enables the investigation of subtle electron and lattice dynamics in the weak excitation regime, which is relevant for studying photovoltaics and photo-induced phase transitions in strongly correlated materials.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

Table-top optical-extreme-ultraviolet (XUV) transient absorption spectroscopy is a powerful tool to investigate photoinduced electronic and structural dynamics in atoms, molecules, and solids [1–4]. In this approach, a sample material is excited by an optical pulse and subsequently probed in reflection or transmission with a time-delayed XUV pulse produced by high-harmonic generation (HHG). Photons in the XUV spectrum can excite core electrons, and thus enable element-specific probing of the optically induced dynamics with intrinsic attosecond timing precision due to the phase-locked generation [5]. In addition, the HHG process allows for single isolated attosecond pulses. These properties enable unprecedented time resolution for the observation of sub-cycle electron dynamics [1,6,7].

Despite its wide range of applications, table-top optical-XUV transient absorption spectroscopy has been mainly utilized in the high-pump-fluence regime ($>1 \text{ mJ/cm}^2$) that induce large transient signals. This is largely due to the strong nonlinearities in HHG that greatly amplify fluctuations of the driving laser pulse, leading to a typical noise floor of 0.1 - 1 mOD in state-of-the-art experiments [8–12], which presents a formidable challenge for the observation of small pump-induced XUV absorbance changes. This far-from-shot-noise-limited sensitivity hinders studies of dynamics in the low-excitation regime, as involved in carrier dynamics in photovoltaics and electronic and structural phase transitions in solids [13–15]. For example, in a solar cell, the excited carrier density by sunlight is typically on the order of 10^{14} - 10^{17} cm^{-3} [16,17], whereas the carrier density in a 200 nm thick silicon membrane excited by 800 nm laser irradiation with fluence of 1 mJ/cm^2 exceeds 10^{18} cm^{-3} [18]. Substantial efforts have already been made to enhance the signal-to-noise ratio of transient absorption spectroscopy in the context of HHG sources. The most common method to trace long-term fluctuations involves a successive recording of pumped

and un-pumped spectra [9,10]. Additionally, noise reduction is possible by referencing, using the laser power [19], the spectral regions unaffected by the pump-induced dynamics [8,20], or a second spectrum of the XUV pulse before the sample [21,22]. It is also possible to use an almost common path referencing by splitting up the XUV beam [23–25]. However, these approaches typically impose requirements on the sample and experimental geometry and can not be easily used in a large variety of beam lines. Although approaches with a second spectrum may also overcome noise due to both long and short term fluctuations, imperfectly matched spectra can lead to far-from-ideal noise suppression.

In this work, we utilize machine learning to find the optimal referencing given an imperfect reference spectrum. Linear and polynomial regression and a neural network model are evaluated for signal enhancement of experimental XUV transient absorption spectra. In experiments with a 60 nm thick 1T-TiSe₂ sample at low fluences (0.66 mJ/cm²), we find that a three-layer neural network surpasses all other methods and enables studies of coherent lattice vibrations and a phase transition in this material [26]. Moreover, we exploit the capability of performing a pixel-wise referencing with sub harmonic spectral resolution using the neural network.

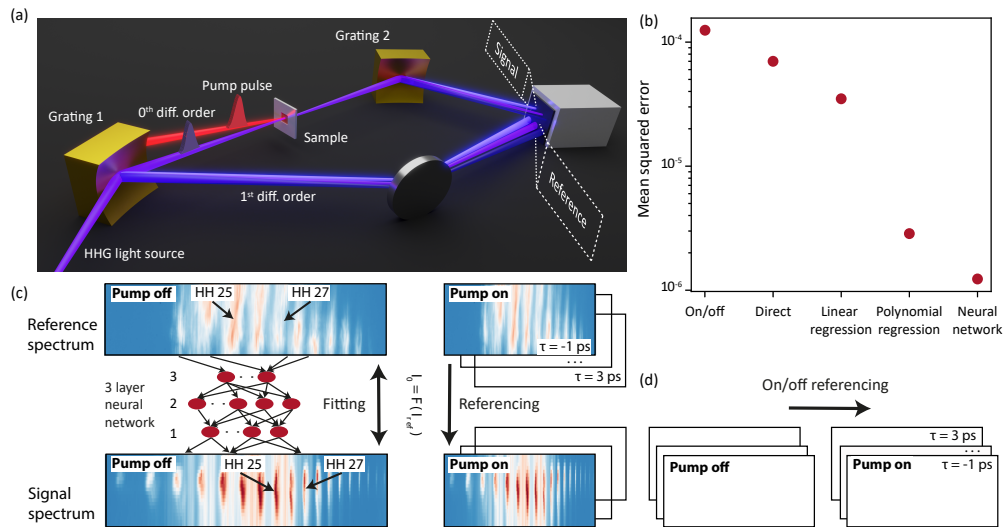


Fig. 1. (a) Experimental setup for extreme ultraviolet transient absorption spectroscopy with a reference spectrum. (b) Comparison of different referencing algorithms. (c) Scheme for referencing with a neural network, trained on the un-pumped data. The model is applied at every pump-probe time delay τ . (d) Noise control scheme by comparing pumped and un-pumped signal (on/off) spectrum without additional reference spectrum.

The experimental setup is illustrated in Fig. 1(a). We investigate a 1T-TiSe₂ sample with XUV pulses produced by HHG with a 2 mJ, 35-fs-duration, 800-nm-wavelength driving pulse and its second harmonic from a Ti:sapphire laser operating at 1 kHz repetition rate with 0.28 % RMS shot-to-shot stability. A second femtosecond optical pump pulse at 2 μm wavelength produced by optical parametric amplification from the same laser system is used to excite the sample. We employ discrete harmonics that are even and odd multiples of the 1.55 eV fundamental photon energy and fluctuate by 4 % RMS. The total flux exceeds 6×10^9 photons/s/harmonic for 38 nm wavelength at the source. In contrast to conventional XUV absorption experiments [9], where the XUV beam is focused onto the sample with a toroidal mirror, here the focusing mirror before the sample is replaced by a toroidal grating (Grating 1) and the first-order diffracted XUV beam from the grating is taken as the *reference* spectrum. The zeroth-order beam probes the sample

and is dispersed by a second grating producing the *signal* spectrum. The signal and the reference XUV beam illuminate the upper and lower half of a charged-coupled device (CCD, 1024×255 pixels), respectively. By recording the reference beam intensity I_{ref} , the noise of each acquisition of the signal spectrum I_{sig} may be directly compensated. However, noise originating from the beam path after the reference grating such as vibrations of the sample cannot be accounted for. Transient absorption experiments measure the change of absorbance $\Delta A = -\log_{10}(I_{sig}/I_0)$, defined as the logarithm of the pumped signal spectrum normalized by I_0 which corresponds to the transmitted probe spectrum without the pump. The time-dependent information of I_{sig} is obtained by varying the delay τ between pump and probe pulse.

Most experiments use successive recording of pumped and unpumped (on/off) signal spectra (I_0 corresponds to a subsequently acquired signal spectrum without the pump) to trace fluctuations on timescales exceeding the camera acquisition time (see Fig. 1(d)). However, noise originating from fluctuations on the acquisition timescale can only be fully suppressed if I_0 and I_{sig} are known simultaneously. In a two-spectrometer configuration, the normalization spectrum I_0 can be calculated by the simultaneously acquired reference spectrum $I_0 = F(I_{ref})$. Here, F is a function that relates the reference spectrum with the unpumped sample absorption.

In its simplest form, F merely contains the averaged ratio between the intensity of the unpumped signal and reference spectrum (*direct* referencing):

$$F(\lambda, I_{ref}) = I_{ref}(\lambda) \times \langle I_{sig}(\lambda)/I_{ref}(\lambda) \rangle, \quad (1)$$

with λ denoting the wavelength. In practice, however, the signal and reference beams may be subjected to different noise levels which prevents successful signal extraction. For example, the reference and signal beams may vary drastically in intensity and spectral purity induced by the differences in optics and path lengths between the two beams. Furthermore, the two spectra may cover different spectral regions due to their respective beam paths and it can be very challenging to precisely calibrate CCD pixels and corresponding wavelength for both spectra at the same time. To address these issues, machine learning, which is capable of *learning* intensity relations between different wavelengths [8,20], is utilized to establish the pixel-wise relationship between the signal and reference spectrum. The model function F is trained by 58240 sets (80%) of data to learn the relation between I_0 and I_{ref} . The training data is acquired without optical pump. The efficacy of the models are tested by another 14560 sets (20%) of unpumped data. In the following, we introduce three different machine learning models: linear and polynomial regression, and a three-layer neural network. The performance of the three approaches are evaluated by comparing their mean squared error to *direct* referencing and on/off referencing using the same dataset (Fig. 1(b)). Due to the discrete nature of the harmonics in our experiment and the described problems for *direct* and on/off approaches, we use the individual harmonics (7 in the reference and 15 in the signal spectrum) as pixels in the comparison.

In the nonlinear polynomial regression model of degree N , the intensity of I_0 at pixel s is expressed as

$$I_{0,s} = F(s, I_{ref}) = a_0^s + \sum_{k=1}^N \sum_{r_1, \dots, r_k} a_{r_1, \dots, r_k}^s I_{ref, r_1} \times \dots \times I_{ref, r_k}, \quad (2)$$

where indices r_i run through all pixels containing the reference beam and a_0^s and a_{r_1, \dots, r_k}^s 's are fitting coefficients. The linear regression model uses $N = 1$ with $(7 + 1) \times 15 = 120$ fitting parameters. As the number of fitting parameters increases exponentially with N , nonlinear polynomial regression models with different N 's are tested by 5-fold cross validation to avoid over-fitting [27]. Here, 20% of the randomly ordered training data are permutatively taken out to evaluate the degree of over-fitting. In this data set, over-fitting sets in at polynomial orders larger than 5 which correspond to 30030 fitting parameters.

The neural network model can be viewed as a further generalization, capable of fitting an arbitrary, non-polynomial function $F(I_{ref})$ to the data [28]. In the present work, we employ a

three-layer feed-forward neural network (see Fig. 1(c)) constructed with libraries *Keras* and *TensorFlow* [29]. The code is available at [30]. In the model, the number of neurons from hidden layers 1-3 are 250, 100, 70, and 200, 100, 20 for correlating intensity data of each pixel and each harmonic, respectively. These numbers have been empirically tested on the presented data and equivalent data sets recorded with the same setup and slightly changed HHG parameters. Larger models were found to yield no further accuracy improvements. A rectified-linear-unit activation function [31,32] was used. The networks of 24435 and 193170 trainable parameters, corresponding to individual harmonic and pixel wise models, are optimized with stochastic gradient descent on the mean-squared error with the Adam optimizer [33] and batches of 50 data points [34]. Training of the larger model by 58240 data sets is achieved in less than three hours on a personal computer without parallelization, which presents a very moderate computational cost. In order to avoid over-fitting, the fitting is stopped when the error minimization of the testing data saturates. The presented architecture of a simple 3 layer neural network poses a good foundation for a model to describe the connection of signal and reference spectrum without any prior knowledge. In the future, more sophisticated designs may be employed to further improve the already great performance and accuracy.

Among the different models, the mean squared error (Fig. 1(b)) of the pumped spectrum before time zero and a subsequently acquired unpumped spectrum (on/off) is the largest. Marginal improvements are obtained by *direct* referencing (Eq. (1)) and linear regression. The polynomial regression method can reach better approximation with much smaller error, but is still surpassed by the neural network model. This indicates the importance to account for non-polynomial intensity noise imprinted by the highly nonlinear HHG process. To further analyze the performance of the different approaches, the models are applied to experimental XUV transient absorption data before time zero. Here, the pump pulse arrives *after* the XUV probe and we expect the transient absorption in this material system to be zero. In Fig. 2(a) the absorbance change of the 25th and 27th harmonic are compared between the on/off method, *direct* referencing and the neural network. The on/off approach shows the largest noise which is given by the standard deviation of the histograms. *Direct* referencing improves the noise level, but is highly dependent on the harmonic order. This may result from various effects including clipping of some intensity and the unequal imaging condition such as optical aberrations, imperfect simultaneous alignment of both focal planes, and spectral impurity caused by the different beam paths and optical elements. Significantly better results are achieved with the neural network yielding similar noise levels at both harmonics.

A comprehensive analysis of the noise level derived from the standard deviations as a function of harmonic order is presented in Fig. 2(b). While the *direct* referencing improves the on/off approach at almost all wavelengths, the resulting noise strongly varies between harmonics. Please note that due to geometric constraints, only harmonic 24 to 30 can be recorded in the reference spectrum, limiting the spectral applicability of the *direct* referencing. The machine learning approaches which includes linear and polynomial regression and the neural network show significantly better performance with less deviation over the whole spectral region. We find that the neural network outperforms the other techniques at almost all orders of harmonics for the presented data. While the linear regression shows already good results and could be applied as a simpler model, the difference to the neural network approach is still significant. The results are compared with the photon shot noise that defines the lower noise limit. At the utilized 200 msec integration time and 3 MHz readout rate, the dark current of 0.02 counts and the readout noise of 30 counts are negligible in comparison to the measured signal of $>10^5$ counts per pixel. The machine learning referencing reduces the noise level significantly and roughly matches the photon shot noise spectral dependence. The remaining discrepancy is likely to result from imperfect matching of signal and reference spectrum due to the different optical elements and

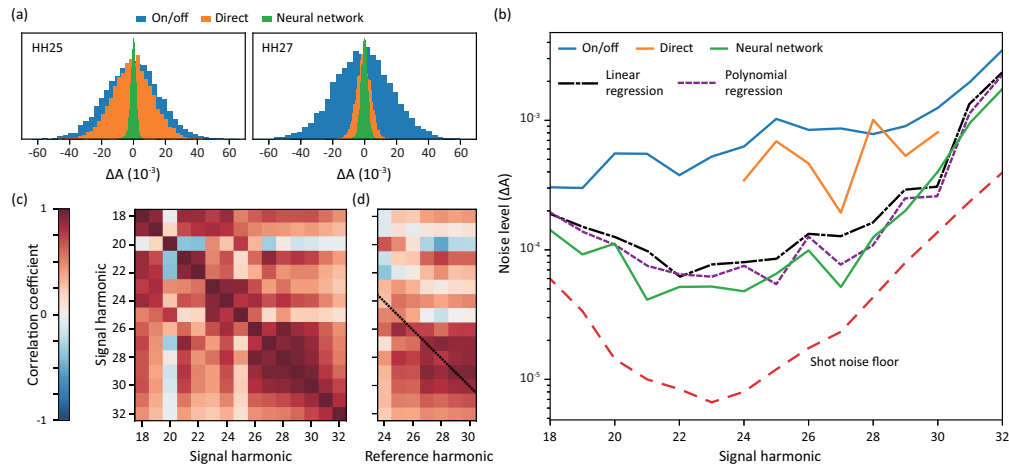


Fig. 2. (a) Histograms of optical density difference before time zero for harmonics 25 and 27. (b) Noise levels of the discussed referencing techniques derived by the histogram standard deviation as a function of the harmonic order. (c) Autocorrelation of the signal spectrum. (d) Cross-correlation between reference and signal spectrum. Diagonal elements (black line) are used for *direct* referencing.

beam paths. Furthermore, the spectral region which is not covered by the reference spectrum is prone to additional error.

To validate the predictive power when the reference and signal spectra do not share the exact same wavelength ranges, we study the intensity correlation of different harmonic orders between the unpumped signal and reference spectra. Figures 2(c) and 2(d) show the correlation coefficients among harmonics of the signal spectrum and between signal and reference spectrum, respectively. Most off-diagonal elements in both correlation maps have >0.4 magnitude, indicating that the intensities of different harmonics are strongly correlated by the HHG process. The black line in Fig. 2(d) shows components that are used in *direct* referencing. As the intensities between different regions of the harmonic spectrum are highly correlated, it is feasible to predict the signal spectrum using a reference spectrum with partially overlapped spectral region. In addition, the off-diagonal components provide additional information, which are used by the linear regression model. Nonlinear correlations fitted by the polynomial regression and the neural network model are not captured by the correlation coefficients. In principle, any other correlated quantity, like laser power, HHG source variables or total XUV intensity could be used for de-noising with machine learning.

Finally, we demonstrate the signal enhancement capability by showing a pump-probe trace of the 25th harmonic for on/off, *direct* and neural network referencing in Fig. 3(a). The 9.4×10^{-4} OD standard deviation of the spectrum with on/off referencing aligns with the typical noise level ≈ 1 mOD found in most transient absorption studies [8] and can be improved by a factor of two with *direct* referencing. An improvement by more than one order of magnitude in sensitivity to 9×10^{-5} OD is achieved with the neural network. In addition, the neural network model can be applied to obtain a pixel-wise correlation between the signal and reference spectra. As most attosecond sources offer a broadband supercontinuum covering large portions of the XUV region, pixel-wise correlation offers noise-reduction without the loss of spectral resolution due to binning. Transient absorption spectra as a function of pump-probe delays and XUV photon energies are compared in Fig. 3(b) and Fig. 3(c) between the established on/off approach and the neural network referencing, respectively. For clarity, the spectral regions with low XUV intensities (owing to the discrete nature of harmonics) are omitted. While the pump induced

response remains at a similar level, the neural network referencing drastically improves the visibility of the pump probe data. Again one order of magnitude improvement down to roughly 10^{-4} OD is visible. Note that while polynomial regression can also handle nonlinearities in the spectral correlation, it is computationally unfeasible to fit single pixel spectra due to the large number of fit parameters. For example, in the case of a 1024-pixel detector, 10^{14} fit parameters would be needed to model a polynomial of 5th degree in polynomial regression.

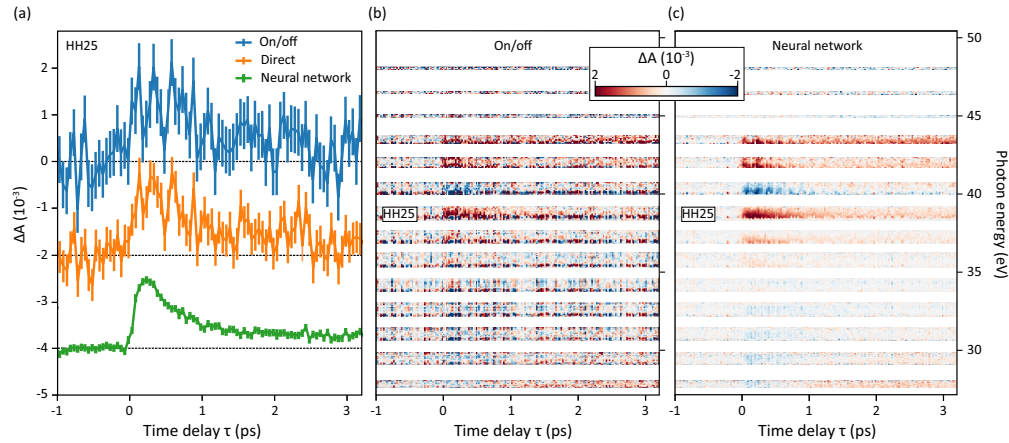


Fig. 3. (a) Transient absorption of the 25th harmonic evaluated with different referencing techniques. The error bars are identical for every delay and defined by the standard deviation at the particular harmonic. (b) Pixel-wise spectrum obtained with the conventional on/off approach. (c) Same as (b) but with the neural network referencing.

The feasibility of single pixel referencing renders the neural network algorithm ideally suited for normalizing the broadband spectra from a wide range of attosecond beam lines. The reference spectrum can be acquired without wavelength calibration or complete coverage of the signal spectrum. Furthermore, instead of implementing a second spectrum, any measured quantity which is correlated to the harmonic intensity such as laser power, total XUV intensity, pulse duration, pulse shape, gas pressure or beam position may be used as input in the machine learning algorithm. As a side note, the neural network should always be trained during the data acquisition on the unpumped data such that the model can accurately capture fluctuations of the HHG source during the measurement. Comparing with the widely used on/off referencing, the neural network method does not require additional experiment time since pumped data are recorded by default for the on/off referencing in quick succession to un-pumped spectra. Note that when using a pre-trained model for noise-reduction of datasets in a different experiment, the performance of the model may degrade due to differences in HHG parameters.

In summary, we developed an algorithm for optimal noise correction in transient absorption spectroscopy with a reference spectrum. Thereby, we extend the applications of data driven algorithms in optics and photonics [35] which already showed great success in X-ray diffraction and spectroscopy for noise suppression and signal extraction [36]. Among the tested algorithms, the three-layer neural network approach shows the best result and improves the noise level by an order of magnitude compared to the established on/off referencing method. This sensitivity increase opens the field of femtosecond to attosecond XUV transient absorption spectroscopy in the weak excitation regime such as photoinduced phase transitions in photovoltaics *in operando* or strongly correlated materials like 1T-TiSe₂. As the neural network approach utilizes the nonlinear intensity correlation between different pixels, it does not require wavelength-calibrated reference spectra, nor does it rely on signal and reference beams covering the same spectral

region. This largely relaxes the conditions for experimental instrumentation and highlights the wide application of our method to both narrowband and broadband sources in the field of transient spectroscopy. On a broader scale, the presented combination of referencing with machine learning may be applied to a large variety of experiments with fluctuating probes to enhance the sensitivity.

Funding. Alexander von Humboldt-Stiftung; Deutsche Forschungsgesellschaft; Gottfried Wilhelm Leibniz Prize; Max-Planck-Gesellschaft.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. F. Krausz and M. Ivanov, "Attosecond physics," *Rev. Mod. Phys.* **81**(1), 163–234 (2009).
2. R. Geneaux, H. J. B. Marroux, A. Guggenmos, D. M. Neumark, and S. R. Leone, "Transient absorption spectroscopy using high harmonic generation: A review of ultrafast X-ray dynamics in molecules and solids," *Phil. Trans. R. Soc. A.* **377**(2145), 20170463 (2019).
3. H. Liu, I. M. Klein, J. M. Michelsen, and S. K. Cushing, "Element-specific electronic and structural dynamics using transient XUV and soft X-ray spectroscopy," *Chem* **7**(10), 2569–2584 (2021).
4. S. Biswas and L. R. Baker, "Extreme ultraviolet reflection-absorption spectroscopy: Probing dynamics at surfaces from a molecular perspective," *Acc. Chem. Res.* **55**(6), 893–903 (2022).
5. M. Huppert, I. Jordan, and H. J. Wörner, "Attosecond beamline with actively stabilized and spatially separated beam paths," *Rev. Sci. Instrum.* **86**(12), 123106 (2015).
6. F. Krausz and M. I. Stockman, "Attosecond metrology: From electron capture to future signal processing," *Nat. Photonics* **8**(3), 205–213 (2014).
7. K. Zhao, Q. Zhang, M. Chini, Y. Wu, X. Wang, and Z. Chang, "Tailoring a 67 attosecond pulse through advantageous phase-mismatch," *Opt. Lett.* **37**(18), 3891 (2012).
8. R. G eneaux, H.-T. Chang, A. M. Schwartzberg, and H. J. B. Marroux, "Source noise suppression in attosecond transient absorption spectroscopy by edge-pixel referencing," *Opt. Express* **29**(2), 951 (2021).
9. M. Z urch, H.-T. Chang, L. J. Borja, P. M. Kraus, S. K. Cushing, A. Gandman, C. J. Kaplan, M. H. Oh, J. S. Prell, D. Prendergast, C. D. Pemmaraju, D. M. Neumark, and S. R. Leone, "Direct and simultaneous observation of ultrafast electron and hole dynamics in germanium," *Nat. Commun.* **8**(1), 15734 (2017).
10. V. Stoof, M. Hartmann, P. Birk, G. D. Borisova, T. Ding, A. Bl attermann, C. Ott, and T. Pfeifer, "XUV-beamline for attosecond transient absorption measurements featuring a broadband common beam-path time-delay unit and in situ reference spectrometer for high stability and sensitivity," *Rev. Sci. Instrum.* **90**(5), 053108 (2019).
11. J. Weisshaupt, A. Rouz e, M. Woerner, M. J. J. Vrakking, T. Elsaesser, E. L. Shirley, and A. Borgschulte, "Ultrafast modulation of electronic structure by coherent phonon excitations," *Phys. Rev. B* **95**(8), 081101 (2017).
12. K. Kato, H. Mashiko, Y. Kunihashi, H. Omi, H. Gotoh, and K. Oguri, "Highly sensitive transient reflection measurement in extreme ultraviolet region for tracking carrier and coherent phonon dynamics," *Opt. Express* **28**(2), 1595–1602 (2020).
13. D. N. Basov, R. D. Averitt, D. van der Marel, M. Dressel, and K. Haule, "Electrodynamics of correlated electron materials," *Rev. Mod. Phys.* **83**(2), 471–541 (2011).
14. M. Imada, A. Fujimori, and Y. Tokura, "Metal-insulator transitions," *Rev. Mod. Phys.* **70**(4), 1039–1263 (1998).
15. E. M ohr-Vorobeva, S. L. Johnson, P. Beaud, U. Staub, R. De Souza, C. Milne, G. Ingold, J. Demsar, H. Schaefer, and A. Titov, "Nonthermal Melting of a Charge Density Wave in TiSe 2," *Phys. Rev. Lett.* **107**(3), 036403 (2011).
16. J. Wu, H. Cha, T. Du, Y. Dong, W. Xu, C.-T. Lin, and J. R. Durrant, "A Comparison of Charge Carrier Dynamics in Organic and Perovskite Solar Cells," *Adv. Mater.* **34**(2), 2101833 (2022).
17. P. Uprety, I. Subedi, M. M. Junda, R. W. Collins, and N. J. Podraza, "Photogenerated Carrier Transport Properties in Silicon Photovoltaics," *Sci. Rep.* **9**(1), 19015 (2019).
18. S. K. Cushing, A. Lee, I. J. Porter, L. M. Carneiro, H.-T. Chang, M. Z urch, and S. R. Leone, "Differentiating Photoexcited Carrier and Phonon Dynamics in the Δ , L , and Γ Valleys of Si(100) with Transient Extreme Ultraviolet Spectroscopy," *J. Phys. Chem. C* **123**(6), 3343–3352 (2019).
19. M. Volkov, J. Pupekis, C. R. Phillips, F. Schlaepfer, L. Gallmann, and U. Keller, "Reduction of laser-intensity-correlated noise in high-harmonic generation," *Opt. Express* **27**(6), 7886–7895 (2019).
20. D. Faccial a, B. W. Toulson, and O. Gessner, "Removal of correlated background in a high-order harmonic transient absorption spectra with principal component regression," *Opt. Express* **29**(22), 35135–35148 (2021).
21. F. Willems, C. von Korff Schmising, C. Str uber, D. Schick, D. W. Engel, J. K. Dewhurst, P. Elliott, S. Sharma, and S. Eisebitt, "Optical inter-site spin transfer probed by energy and spin-resolved transient absorption spectroscopy," *Nat. Commun.* **11**(1), 871 (2020).
22. P. C. Johnsen, S. A. Ryan, C. Gentry, A. Grafov, H. Kapteyn, and M. Murnane, "A beamline for ultrafast extreme ultraviolet magneto-optical spectroscopy in reflection near the shot noise limit," *Rev. Sci. Instrum.* **94**(3), 033001 (2023).

23. T. Heldt, J. Dubois, P. Birk, G. D. Borisova, G. M. Lando, C. Ott, and T. Pfeifer, "Attosecond Real-Time Observation of Recolliding Electron Trajectories in Helium at Low Laser Intensities," *Phys. Rev. Lett.* **130**(18), 183201 (2023).
24. A. Camper, H. Park, S. J. Hageman, G. Smith, T. Auguste, P. Agostini, and L. F. DiMauro, "High relative-phase precision beam duplicator for mid-infrared femtosecond pulses," *Opt. Lett.* **44**(22), 5465 (2019).
25. J. J. Abel, F. Wiesner, J. Nathanael, J. Reinhard, M. Wünsche, G. Schmidl, A. Gawlik, U. Hübner, J. Plentz, C. Rödel, G. G. Paulus, and S. Fuchs, "Absolute EUV reflectivity measurements using a broadband high-harmonic source and an in situ single exposure reference scheme," *Opt. Express* **30**(20), 35671 (2022).
26. T. Heinrich, H.-T. Chang, S. Zayko, K. Rossnagel, M. Sivis, and C. Ropers, "Electronic and structural fingerprints of charge-density-wave excitations in extreme ultraviolet transient absorption spectroscopy," *Phys. Rev. X* **13**(2), 021033 (2023).
27. G. C. Cawley and N. L. C. Talbot, "On over-fitting in model selection and subsequent selection bias in performance evaluation," *J. Mach. Learn. Res.* **11**, 2079–2107 (2010).
28. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks* **2**(5), 359–366 (1989).
29. F. Chollet, "Keras," <https://keras.io> (2015).
30. T. Heinrich, "Referencing-neural-network," <https://github.com/tobheinrich/referencing-neural-network>.
31. V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ICML 2010*, (2010), pp. 807–814.
32. X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, G. Gordon, D. Dunson, and M. Dudík, eds. (PMLR, 2011), pp. 315–323.
33. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv*, arXiv:1412.6980 (2014).
34. N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," *arXiv*, arXiv:1609.04836 (2016).
35. G. Genty, L. Salmela, J. M. Dudley, D. Brunner, A. Kokhanovskiy, S. Koltsev, and S. K. Turitsyn, "Machine learning and applications in ultrafast photonics," *Nat. Photonics* **15**(2), 91–101 (2021).
36. Z. Chen, N. Andrejevic, N. C. Drucker, T. Nguyen, R. P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D. A. Tennant, M. Chan, and M. Li, "Machine learning on neutron and x-ray scattering and spectroscopies," *Chem. Phys. Rev.* **2**(3), 031301 (2021).