Multivariate analysis of x-ray scattering using a stochastic source

Sioan Zohar[1,*] **and Joshua J. Turner**[1,2]

[1]Linac Coherent Light Source, SLAC National Accelerator Laboratory, 2575 Sand Hill Rd., Menlo Park, California 94025, USA
[2]Stanford Institute for Materials and Energy Sciences, Stanford University, Stanford, California 94305, USA
*Corresponding author: sioan@slac.stanford.edu

The normalization of scattered intensity by incident flux is a crucial step in analyzing data from stochastic x-ray free electron laser sources and is complicated by non-linearities traditionally attributed to detector saturation. Here we show that such non-linearities can also arise when the sample spectra are non-uniform within the monochromator bandwidth. A method for modeling and removing this non-linearity using multivariate regression with shot-by-shot x-ray photon energy as an independent variable is presented. This approach demonstrates the benefit of event building and will allow for a reconsideration of data which has proven challenging to normalize.

X-ray free electron lasers (X-FELs) [1] are extraordinary instruments capable of producing high-brilliance X-ray pulses less than 50 fs in duration with up to 10^{13} photons per pulse. Over the past decade, X-FELs have made a tremendous impact on a number of scientific fields [2] ranging from astrophysics [3] and time-resolved femtochemistry [4] to structural biology [5] and condensed matter physics [6]. More recently, X-FEL coherence properties [7], together with two-color pump-probe techniques [8,9], have been used to probe topological Skyrmion lattices [10] and offer the potential for studying quantum fluctuations.

Despite the tremendous progress in probing femtosecond-resolved phenomena, the chaotic nature of X-FELs makes high-precision measurements such as x-ray absorption spectroscopy challenging to perform. Compared to synchrotrons, normalizing scattered x-ray intensity by incident power is of paramount importance at X-FELs, especially when searching for subtle effects. Efforts aimed at addressing these challenges have focused on developing diagnostics such as gas monitor detectors (GMDs) [11–13], power meters [14], and two-dimensional (2D) detectors [15,16]. Even with state-of-the-art diagnostics, non-linearities that contaminate data still appear. Traditionally, these issues are dealt with by cross-calibrating multiple devices with known linear detectors, or by aggressively filtering events as a function of x-ray pulse fluence to within a narrow linear regime.

In this Letter, we demonstrate that non-linearities traditionally attributed to the onset of detector saturation can also arise when the sample spectra are non-uniform within the monochromator bandwidth. A method for modeling this behavior using the photon energy estimated from electron beam diagnostics to recover a linear response is presented.

Experiments were performed at the Linac Coherent Light Source (LCLS) [1] using the soft x-ray research (SXR) instrument [17]. The beamline and undulator are shown in Fig. 1. X-rays with 70 fs pulse widths and 120 Hz repetition rates are generated by electrons accelerated into the X-FEL undulator via a self amplified spontaneous seeding (SASE) process [18]. The electron position is measured using a beam position monitor (BPM) in the accelerator, and a correction to the average x-ray photon energy, E, is estimated from calibrated electron position-energy dispersion curves [19]. The x-ray intensity, I_1, exiting the undulator is measured using a gas detector located in the front end enclosure (FEE) [13]. The x-ray beam energy is vertically dispersed using a variable line-spacing monochromator [20]. The exit slit was set to 320 μm, or an X-ray bandwidth of 2.25 at 914 eV [12]. The x-ray intensity downstream from these slits, I_0, is measured using a GMD [11]. X-ray scattering intensity, I, is recorded using an avalanche photodiode (APD) housed inside a soft x-ray scattering chamber [16], upgraded for terahertz experiments [21]. More details of the SXR instrument can be found in Schlotter *et al.* [22]. A total of $\sim 2.6 \times 10^5$ events containing I, I_0, I_1, and E, were recorded on a per pulse basis and event built using the LCLS data acquisition system and analyzed using the photon science analysis (PSANA) framework [23–25]. Correlations and patterns in the data were visually explored using glue visualization [26].

Simulations demonstrating how distinct spectra for I and I_0 give rise to non-linear I versus I_0 curves are shown for three different scenarios in Figs. 2(a), 2(c), and 2(e). The scattered intensity, I, versus normalization diagnostic, I_0, for the different scenarios displayed in panels (a), (c), and (e) are shown in panels (b), (d), and (f), respectively. In general, the intensity scattered from a sample is proportional to the x-ray intensity measured using upstream diagnostics [Fig. 2(a)], and I versus

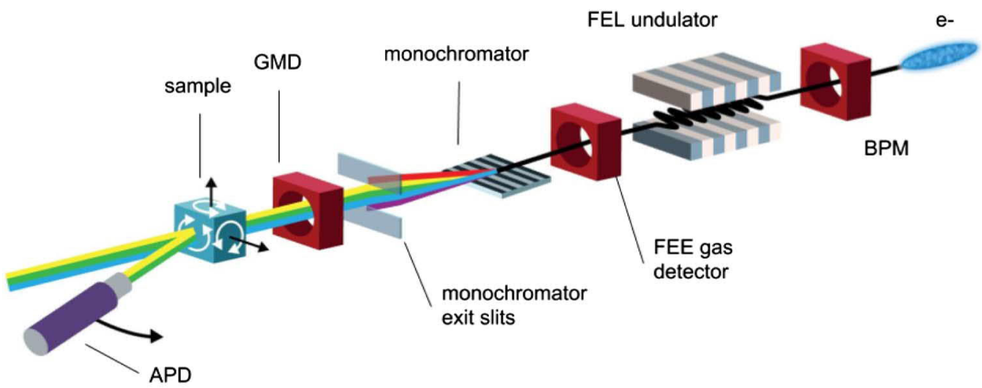

Fig. 1. X-rays with 70 fs pulse widths and 120 Hz repetition rates are generated by electrons accelerated into the FEL undulator via a SASE process. The x-ray photon energy, E, is estimated using a combination of calibrated electron BPMs and position-energy dispersion curves [19]. The x-ray intensity, I_1, exiting the undulator is measured using the FEE gas detector. The variable line-spacing monochromator [20], together with the slits downstream from the monochromator, reduces the x-ray bandwidth from ~10 to 2.25 eV centered at E_0. The x-ray intensity downstream from these slits, I_0, is measured using the GMD [11,12]. X-ray scattering is measured in the end-station using an APD. Steering, focusing, and other optics and diagnostics are omitted for clarity.

I_0 is linear [Fig. 2(b)]. When the sample spectrum has a gradient over the monochromator bandwidth, then the center of the scattered intensity distribution appears to be shifted with respect to the I_0 spectrum [Fig. 2(c)], and I versus I_0 traces out a loop [Fig. 2(d)]. When the spectral shapes are different for I and I_0 [Fig. 2(e)], then I versus I_0 is non-linear [Fig. 2(f)].

In Fig. 3(a), I versus I_0 is shown on an event-resolved basis with each point representing a single x-ray pulse. To compare with the simulations in Fig. 2, the APD intensity represents the scattered intensity I, and the GMD intensity represents I_0. The red trend line is estimated by taking the mean value of I for binned sections of I_0, and its non-linearity is typically interpreted as the onset of detector saturation. The main result of this Letter is captured in Fig. 3(b) which shows I versus I_0 after correcting for the sample spectral gradient within the monochromator transmission spectrum using Eqs. (5) and (6), and will be discussed below.

The photon energy dependence of I and I_0, together with the histogram of E, is shown in Fig. 4 panels (a), (b), and (c), respectively. A histogram of the x-ray intensity measured at the FEE gas detector I_1 is shown (d). The three curves in Figs. 4(a) and 4(b) are color coded to correspond to the three colored slices of I_1 shown in Fig. 4(d) from which they were filtered. The shift between the peak values of I and I_0 shown in Figs. 4(a) and 4(b) can arise from a spectral gradient within the bandwidth of the measurement and is described below.

The dependence of I_0 upon E can be approximated by a Gaussian whose amplitude depends on I_1. That is, I can be expressed as

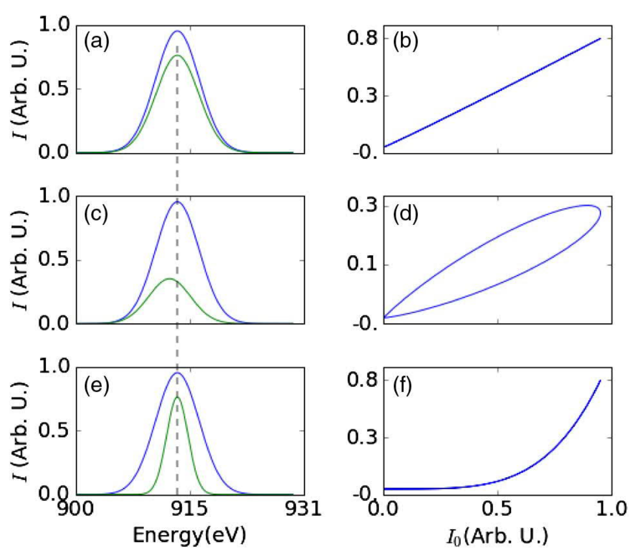

Fig. 2. Simulations demonstrating the origin of a non-linear I versus I_0 relationship. The Gaussian curves in panels (a), (c), and (e) correspond to simulated scattered intensity scattered from the sample I (green) and normalizing upstream diagnostic, I_0 (blue). The scattered and normalization intensities shown in panels (a), (c), and (e) are graphed with respect to each other in panels (b), (d), and (f), respectively.

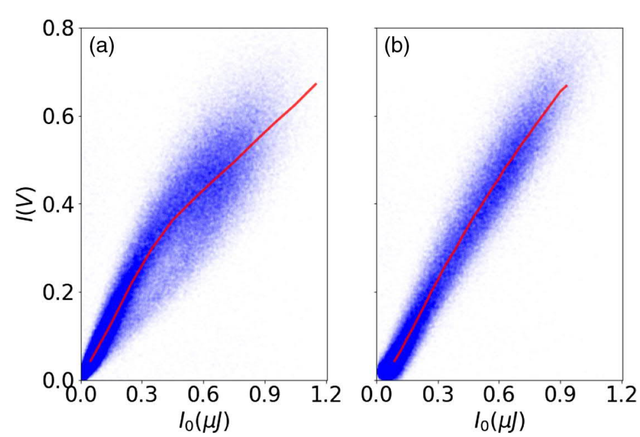

Fig. 3. (a) Measured I versus I_0 where each point corresponds to measurements from a single x-ray pulse. The red line is calculated from the blue scatter points by taking the mean of I in binned intervals of I_0 and shows non-linear behavior. (b) The corrected I versus I_0 curve after accounting for the gradient in the sample spectrum observable within the monochromator bandwidth.

$$I(E) = \alpha(E) I_0(E) = \alpha(E) I_1 \exp\left[\frac{-(E_0 - E)^2}{2\sigma^2}\right], \quad (1)$$

where σ is the standard deviation of the monochromatized spectrum, and $\alpha(E)$ is a general scattering parameter that can be used to represent the fraction of incident x-rays that are either transmitted, reflected, or diffracted from the sample, and can be approximated using a Taylor expansion by

$$\alpha(E) = \alpha_0 + (E_0 - E)\alpha', \quad (2)$$

where α_0 and α' are the zeroth- and first-order coefficients of the sample spectrum expanded about the monochromator peak transmission, E_0. Equations (1) and (2), together with the fundamental theorem of calculus, are combined to produce

$$I(E) = \alpha_0 \; I_1 \; \exp\left[\frac{-(E_0' - E)^2}{2\sigma^2}\right], \quad (3)$$

where $E_0' = E_0 - \Delta E$ and ΔE is defined by

$$\Delta E = -\frac{\sigma^2 \alpha'}{\alpha_0}, \quad (4)$$

and represents the apparent shift of the spectrum scattered from the sample, $I(E)$, with respect to the monochromator transmission spectrum, $I_0(E)$.

The simulations from Figs. 2(c) and 2(d) are compared to the experimental 2D histogram of I versus I_0 in Fig. 5. The blue points were generated by taking I versus E and I_0 versus E curves [similar to those in Figs. 4(a) and 4(b)] and graphing them with respect to one another. The resultant loop agrees with the simulated behavior shown in Figs. 2(c) and 2(d). The heat asymmetry is due to increased sampling of the higher energy portion of the loop in Fig. 5. The linear response (green trend line) can be recovered by graphing the I versus E maxima against the I_0 versus E maxima [Figs. 4(a) and 4(b)] for narrow bins of I_1.

This behavior is modeled by including the product of E and I_0 as an extra regressor in addition to I_0. That is,

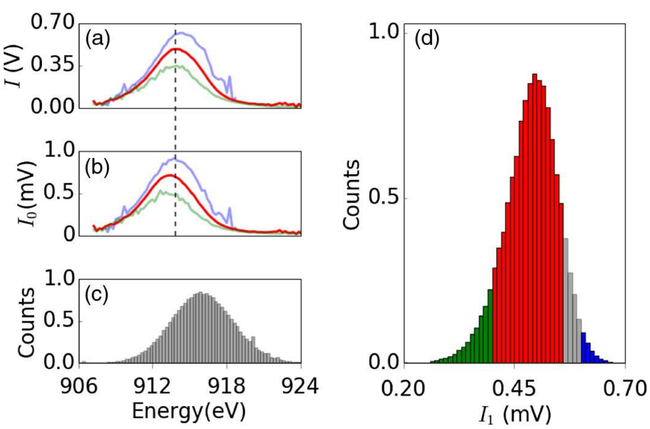

Fig. 4. (a) Scattered intensity, I, and (b) the incident intensity, I_0, binned along energy E in steps of 0.2 eV, and binned in interval I_0 (FEE) values of 0.2 to 0.4 (green), 0.4 to 0.6 (red), and 0.6 to 1.0 (blue) volts. Panels (a) and (b) highlight the peak position shift between I and I_0 due to the spectral distribution for the same I_0 slice. (d) Histogram of the x-ray beam intensity I_1 measured upstream of the monochromator in the FEE. The y-axes in panels (c) and (d) have been normalized to 1 for clarity.

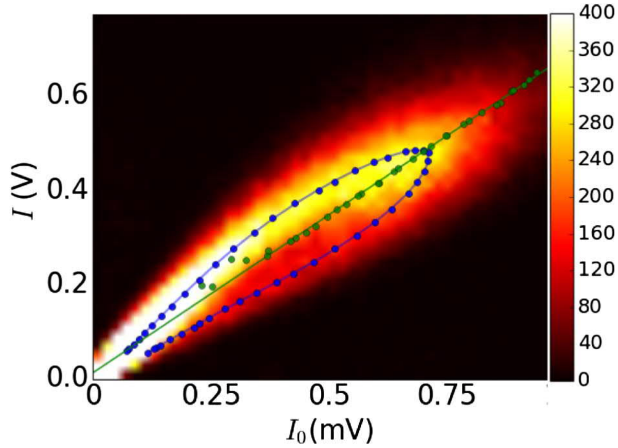

Fig. 5. Heat map for a 2D histogram of experimentally measured I versus I_0. This relationship is not linear due to the energy shift between I and I_0 shown in Fig. 2. The scale bar axis indicates the number of events. The blue dots and trend line are $I(E)$ versus $I_0(E)$ shown in Figs. 4(a) and 4(b).

$$\mathbf{X}\beta = \mathbf{I}, \quad (5)$$

where $\mathbf{I}$ is a $(n \times 1)$ vector containing the scattered intensity, I, for each event; n is the number of events; β is a (4×1) vector whose elements are the regression coefficients; and $\mathbf{X}$ is the matrix of independent variables given by

$$\mathbf{X} = [\vec{\mathbf{1}}, \mathbf{I_0} - \langle \mathbf{I_0} \rangle, \mathbf{E} - \langle \mathbf{E} \rangle, (\mathbf{E} - \langle \mathbf{E} \rangle)(\mathbf{I_0} - \langle \mathbf{I_0} \rangle)], \quad (6)$$

where the boldface variables represent $(n \times 1)$ vectors whose elements are the experimentally measured values for the non-boldface counterparts. The $\langle\rangle$ denotes the arithmetic mean over all elements of the vector. Carrying out this analysis recovers linearity and can be observed by subtracting the contribution from the product of E and I_0 from I, as shown in Fig. 3(b).

This Letter demonstrates how correlations between x-ray science data and additional electron beam diagnostics can be used to improve statistics of X-FEL measurements by modeling and removing systematic errors. To date, only two studies have exploited correlations between event-built photon science and electron beam data [27,28]. This work, however, demonstrates that event building at X-FELs enables data to be normalized by correcting for the sample wavelength dependence within the monochromator bandwidth. These results are important for experiments where the monochromator exit slit is opened to increase fluence. While the trade-off between photon fluence and energy resolution for wide monochromator exit slits is well understood, the increased non-linearity has not been previously identified. We predict that the spectral gradient estimation approach demonstrated here could be used to correct for energy smearing and chromatic aberrations present in x-ray scattering and imaging experiments, respectively, that are introduced by larger x-ray bandwidth experiments.

In conclusion, we have demonstrated that non-linear scattered intensity versus normalization curves typically attributed to detector saturation can also be identified as originating from sub-monochromator bandwidth-resolved photon energy dispersion. A multivariate model using regressors constructed from electron beam diagnostics to account for the photon energy distribution recovers linearity.

Funding. U.S. Department of Energy (DOE), Office of Science (SC), Office of Basic Energy Sciences, through the Materials Sciences and Engineering Division (contract DE-AC02-76SF00515).

Acknowledgment. This work and the use of the LCLS, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, through the Materials Sciences and Engineering Division. J. Turner acknowledges support from the U.S. Department of Energy, Office of Science, Basic Energy Sciences through the Early Career Research Program. The authors thank M. Seaberg and A. Reid for helpful comments with this Letter.

REFERENCES

1. P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F. J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, P. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H. D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky, and J. Galayda, Nat. Photonics **4**, 641 (2010).
2. C. Bostedt, S. Boutet, D. M. Fritz, Z. Huang, H. J. Lee, H. T. Lemke, A. Robert, W. F. Schlotter, J. J. Turner, and G. J. Williams, Rev. Mod. Phys. **88**, 015007 (2016).
3. O. Ciricosta, S. M. Vinko, B. Barbrel, D. S. Rackstraw, T. R. Preston, T. Burian, J. Chalupský, B. I. Cho, H. K. Chung, G. L. Dakovski, K. Engelhorn, V. Hájková, P. Heimann, M. Holmes, L. Juha, J. Krzywinski, R. W. Lee, S. Toleikis, J. J. Turner, U. Zastrau, and J. S. Wark, Nat. Commun. **7**, 11713 (2016).
4. M. Dell'Angela, T. Anniyev, M. Beye, R. Coffee, A. Föhlisch, J. Gladh, T. Katayama, S. Kaya, O. Krupin, J. LaRue, A. Møgelhøj, D. Nordlund, J. K. Nørskov, H. Öberg, H. Ogasawara, H. Öström, L. G. M. Pettersson, W. F. Schlotter, J. A. Sellberg, F. Sorgenfrei, J. J. Turner, M. Wolf, W. Wurth, and A. Nilsson, Science **339**, 1302 (2013).
5. J.-P. Colletier, M. R. Sawaya, M. Gingery, J. A. Rodriguez, D. Cascio, A. S. Brewster, T. Michels-Clark, R. H. Hice, N. Coquelle, S. Boutet, G. J. Williams, M. Messerschmidt, D. P. DePonte, R. G. Sierra, H. Laksmono, J. E. Koglin, M. S. Hunter, H.-W. Park, M. Uervirojnangkoorn, D. K. Bideshi, A. T. Brunger, B. A. Federici, N. K. Sauter, and D. S. Eisenberg, Nature **539**, 43 (2016).
6. R. Mankowsky, A. Subedi, M. Först, S. O. Mariager, M. Chollet, H. T. Lemke, J. S. Robinson, J. M. Glownia, M. P. Minitti, A. Frano, M. Fechner, N. A. Spaldin, T. Loew, B. Keimer, A. Georges, and A. Cavalleri, Nature **516**, 71 (2014).
7. I. A. Vartanyants, et al., Phys. Rev. Lett. **107**, 144801 (2014).
8. F. Decker, S. Gilevich, Z. Huang, H. Loos, A. Marinelli, C. A. Stan, J. L. Turner, Z. V. Hoover, and S. Vetter, *Free Electron Laser Conference (FEL)* (2015), pp. 634–638.
9. A. A. Lutman, T. J. Maxwell, J. P. MacArthur, M. W. Guetg, N. Berrah, R. N. Coffee, Y. Ding, Z. Huang, A. Marinelli, S. Moeller, and J. C. U. Zemella, Nat. Photonics **10**, 745 (2016).
10. M. H. Seaberg, B. Holladay, J. C. T. Lee, M. Sikorski, A. H. Reid, S. A. Montoya, G. L. Dakovski, J. D. Koralek, G. Coslovich, S. Moeller, W. F. Schlotter, R. Streubel, S. D. Kevan, P. Fischer, E. E. Fullerton, J. L. Turner, F.-J. Decker, S. K. Sinha, S. Roy, and J. J. Turner, Phys. Rev. Lett. **119**, 067403 (2017).
11. S. Moeller, G. Brown, G. Dakovski, B. Hill, M. Holmes, J. Loos, R. Maida, E. Paiser, W. Schlotter, J. J. Turner, A. Wallace, U. Jastrow, S. Kreis, A. A. Sorokin, and K. Tiedtke, J. Synchrotron Radiat. **22**, 606 (2015).
12. K. Tiedtke, A. A. Sorokin, U. Jastrow, P. Juranić, S. Kreis, N. Gerken, M. Richter, U. Arp, Y. Feng, D. Nordlund, R. Soufli, M. Fernández-Perea, L. Juha, P. Heimann, B. Nagler, H. J. Lee, S. Mack, M. Cammarata, O. Krupin, M. Messerschmidt, M. Holmes, M. Rowen, W. Schlotter, S. Moeller, and J. J. Turner, Opt. Express **22**, 21214 (2014).
13. S. Moeller, J. Arthur, A. Brachmann, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, P. Emma, Y. Feng, A. Fisher, J. Frisch, J. Galayda, S. Gilevich, J. Hastings, G. Hays, P. Hering, Z. Huang, R. Iverson, J. Krzywinski, S. Lewis, H. Loos, M. Messerschmidt, A. Miahnahri, H.-D. Nuhn, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, B. White, J. Wu, G. Yocky, R. Bionta, E. Ables, B. Abraham, C. Gardener, K. Fong, S. Friedrich, S. Hau-Riege, K. Kishiyama, T. McCarville, D. McMahon, M. McKernan, L. Ott, M. Pivovaroff, J. Robinson, D. Ryutov, S. Shen, R. Soufli, and G. Pile, Nucl. Instrum. Methods Phys. Res. Sect. A **635**, S6 (2011).
14. P. Heimann, S. Moeller, S. Carbajo, S. Song, G. Dakovski, D. Nordlund, and D. Fritz, J. Synchrotron Radiat. **25**, 72 (2018).
15. D. J. Higley, K. Hirsch, G. L. Dakovski, E. Jal, E. Yuan, T. Liu, A. A. Lutman, J. P. MacArthur, E. Arenholz, Z. Chen, G. Coslovich, P. Denes, P. W. Granitzka, P. Hart, M. C. Hoffmann, J. Joseph, L. Le Guyader, A. Mitra, S. Moeller, H. Ohldag, M. Seaberg, P. Shafer, J. Stohr, A. Tsukamoto, H.-D. Nuhn, A. H. Reid, H. A. Durr, and W. F. Schlotter, Rev. Sci. Instrum. **87**, 033110 (2016).
16. D. Doering, Y.-D. Chuang, N. Andresen, K. Chow, D. Contarato, C. Cummings, E. Domning, J. Joseph, J. S. Pepper, B. Smith, G. Zizka, C. Ford, W. S. Lee, M. Weaver, L. Patthey, J. Weizeorick, Z. Hussain, and P. Denes, Rev. Sci. Instrum. **82**, 073303 (2011).
17. G. L. Dakovski, P. Heimann, M. Holmes, O. Krupin, M. P. Minitti, A. Mitra, S. Moeller, M. Rowen, W. F. Schlotter, and J. J. Turner, J. Synchrotron Radiat. **22**, 498 (2015).
18. R. Bonifacio, C. Pellegrini, and L. Narducci, Opt. Commun. **50**, 373 (1984).
19. J. Welch, private communication, 2018.
20. P. Heimann, O. Krupin, W. F. Schlotter, J. Turner, J. Krzywinski, F. Sorgenfrei, M. Messerschmidt, D. Bernstein, J. Chalupský, V. Hájková, S. Hau-Riege, M. Holmes, L. Juha, N. Kelez, J. Lüning, D. Nordlund, M. Fernandez Perea, A. Scherz, R. Soufli, W. Wurth, and M. Rowen, Rev. Sci. Instrum. **82**, 093104 (2011).
21. J. J. Turner, G. L. Dakovski, M. C. Hoffmann, H. Y. Hwang, A. Zarem, W. F. Schlotter, S. Moeller, M. P. Minitti, U. Staub, S. Johnson, A. Mitra, M. Swiggers, P. Noonan, G. I. Curiel, and M. Holmes, J. Synchrotron Radiat. **22**, 621 (2015).
22. W. F. Schlotter, J. J. Turner, M. Rowen, P. Heimann, M. Holmes, O. Krupin, M. Messerschmidt, S. Moeller, J. Krzywinski, R. Soufli, M. Fernández-Perea, N. Kelez, S. Lee, R. Coffee, G. Hays, M. Beye, N. Gerken, F. Sorgenfrei, S. Hau-Riege, L. Juha, J. Chalupsky, V. Hajkova, A. P. Mancuso, A. Singer, O. Yefanov, I. A. Vartanyants, G. Cadenazzi, B. Abbey, K. A. Nugent, H. Sinn, J. Lüning, S. Schaffert, S. Eisebitt, W.-S. Lee, A. Scherz, A. R. Nilsson, and W. Wurth, Rev. Sci. Instrum. **83**, 043107 (2012).
23. J. Thayer, D. Damiani, C. Ford, I. Gaponenko, W. Kroeger, C. O'Grady, J. Pines, T. Tookey, M. Weaver, and A. Perazzo, J. Appl. Crystallogr. **49**, 1363 (2016).
24. J. Thayer, D. Damiani, C. Ford, M. Dubrovin, I. Gaponenko, C. P. O'Grady, W. Kroeger, J. Pines, T. J. Lane, A. Salnikov, D. Schneider, T. Tookey, M. Weaver, C. H. Yoon, and A. Perazzo, Adv. Struct. Chem. Imaging **3**, 3 (2017).
25. D. Damiani, M. Dubrovin, I. Gaponenko, W. Kroeger, T. J. Lane, A. Mitra, C. P. O'Grady, A. Salnikov, A. Sanchez-Gonzalez, D. Schneider, and C. H. Yoon, J. Appl. Crystallogr. **49**, 672 (2016).
26. C. Beaumont, T. Robitaille, and M. Borkin, "Glue: Linked data visualizations across multiple files," Astrophysics Source Code Library (2014).
27. T. B. van Driel, K. S. Kjaer, E. Biasin, K. Haldrup, H. T. Lemke, and M. M. Nielsen, Faraday Discuss. **177**, 443 (2015).
28. T. B. van Driel, S. Herrmann, G. Carini, M. M. Nielsen, and H. T. Lemke, J. Synchrotron Radiat. **22**, 584 (2015).