

Correction of complex nonlinear signal response from a pixel array detector

Tim Brandt van Driel,^{a,b} Sven Herrmann,^a Gabriella Carini,^a
Martin Meedom Nielsen^b and Henrik Till Lemke^{a,*}

^aLinac Coherent Light Source, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA, and ^bDepartment of Physics, Technical University of Denmark, 2800 Kgs Lyngby, Denmark.
*Correspondence e-mail: lemke@slac.stanford.edu

Received 22 November 2014

Accepted 17 March 2015

Edited by I. Schlichting, Max Planck Institute for Medical Research, Germany

Keywords: FEL; detector nonlinearity; diffuse scattering.

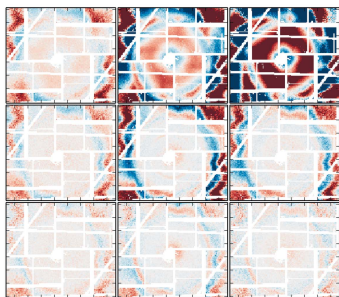
Supporting information: this article has supporting information at journals.iucr.org/s

The pulsed free-electron laser light sources represent a new challenge to photon area detectors due to the intrinsic spontaneous X-ray photon generation process that makes single-pulse detection necessary. Intensity fluctuations up to 100% between individual pulses lead to high linearity requirements in order to distinguish small signal changes. In real detectors, signal distortions as a function of the intensity distribution on the entire detector can occur. Here a robust method to correct this nonlinear response in an area detector is presented for the case of exposures to similar signals. The method is tested for the case of diffuse scattering from liquids where relevant sub-1% signal changes appear on the same order as artifacts induced by the detector electronics.

1. Introduction

Free-electron lasers (FELs) represent a new type of X-ray light source with different properties and requirements than synchrotron sources. These pose a challenge to experimenters, predominantly from the synchrotron community where the beam parameters are significantly more stable. The self-amplified spontaneous emission (SASE) process creates spontaneously emitted, very intense and transverse coherent X-ray pulses (~ 1 mJ per pulse), which fluctuate in properties such as pointing, intensity (up to 100%), photon energy and arrival time. In order to sort and select pulses for these properties, single-pulse detection is essential (≤ 120 Hz pulse rate for presently operating facilities). For area detectors, this has led to special developments of direct detection and fast-readout integrating detectors (Henrich *et al.*, 2011; Weidenpointner *et al.*, 2011; Koch *et al.*, 2013; Mozzanica *et al.*, 2014). At the Linac Coherent Light Source (LCLS), the Cornell–SLAC Pixel Array Detector (CSPAD) has been developed and deployed. This 14-bit hybrid pixel detector, based on ASICs with 194×185 $110 \mu\text{m} \times 110 \mu\text{m}$ pixels, reads out at 120 Hz (Philipp *et al.*, 2010, 2011; Herrmann *et al.*, 2013; Hart *et al.*, 2012; Carini *et al.*, 2014; Blaj *et al.*, 2015). The detector has down to single photon sensitivity (at 8 keV and at the few photons per pixel limit) and can be combined to a 170 mm \times 170 mm 2.3 megapixel camera for applications in crystallography or diffuse scattering (Boutet *et al.*, 2012; Trigo *et al.*, 2013; Sellberg *et al.*, 2014; Arnlund *et al.*, 2014).

While the first prototypes of the CSPAD were characterized by a significant non-uniformity of the ASIC, further iterations led to a much more uniform response and improved noise performance of the detector (Herrmann *et al.*, 2014). Another artifact exacerbated by the fast readout is pixel crosstalk, which depends highly on the exposure of the detector (Herr-



mann *et al.*, 2013). In effect, the pixel sensitivity depends both on the intensity and spatial distribution of the exposure. For strong diffuse signals, *i.e.* from liquid scattering, correction for such effects is very challenging, because the entire detector module is flooded with signal and normalization to local reference data in the detector pattern is not possible. Especially in the case of ultrafast pump/probe experiments, one of the major applications of LCLS, the signal changes after excitation with an optical light pulse are often below 1%. Nonlinear effects, for example due to crosstalk, here often outweigh the physical signal significantly and make it impossible to measure any sensible signal without special data treatment. Efficient correction methods based on separation of dominant components in measured signals using singular value decomposition (SVD) have been presented for binned rotationally symmetric data (Haldrup, 2014) and for a general intensity field of a megapixel detector (van Driel *et al.*, 2015).

In the following we describe a simple formalism for a general complex spatially dependent nonlinear detector response. We show that under the boundary of a similar spatial signal distribution the problem becomes solvable and the nonlinear behavior can be approximated by that of individual pixels for which calibration curves can be experimentally determined. After deriving a nonlinear correction formula for each individual pixel in a most general form, we show a specific way of parametrization of the nonlinearity and application of the correction to data obtained using the CSPAD in the v1.2 generation as example. We show that the method is particularly suited for the challenging case of small changes within an intense diffuse signal and can reduce systematic measurement errors by around a factor of ten.

In our approach, the systematic behavior is isolated from random detector-induced effects and parametrized model-independently in relation to a calibration of known intensities in order to determine the detector gain field that may depend on the exposed pattern. As an example of application, pump/probe scattering data are treated by this method, where optically excited signal changes are tracked as a function of delay between the exciting optical laser pulse and the probe X-ray pulse.

A description of variables used during the mathematical derivation of a correction equation is provided in Table 1.

2. Signal composition

The raw digital signal of one single image measured by a pixel array detector (PAD), an array $\mathbf{D} = [D_1, D_2, D_3, \dots, D_N]$, can be decomposed into a constant dark-current component \mathbf{d}_{dark} , random photon exposure independent components which average to zero over many exposures $\mathbf{d}_{\text{rand},o}$ ($\langle \mathbf{d}_{\text{rand},o} \rangle = 0$), as well as a signal that is due to the exposure to photons \mathbf{d} :

$$\mathbf{D} = \mathbf{d} + \underbrace{\mathbf{d}_{\text{dark}} + \sum_o \mathbf{d}_{\text{rand},o}}_{\text{photon independent}}. \quad (1)$$

The random component can consist of different sources (accounted for by index o), such as individual pixel noise and

Table 1

Description of important variable names used in the article; when applicable, the element names of array variables which describe parameters for each pixel of an area detector are listed in the ‘Elements’ column.

Symbol	Type	Elements	Description
\mathbf{D}	Array	D_n	Raw digital area detector data
N	Integer	–	Total number of pixels in detector
n	Index	–	Pixel index (1, 2, 3, ..., N)
\mathbf{d}	Array	d_n	Photon-dependent digital data contribution
\mathbf{s}	Array	s_n	Photon signal on each pixel
\mathbf{I}	Array	I_n	Given constant intensity distribution, normalized to 1
i	Scalar	–	Total detector intensity
\mathbf{c}	Array	c_n	i dependence of \mathbf{d} at constant intensity distribution, used as calibration dataset
i_c	Scalar	–	i reference point for calibration across pixels
\mathbf{d}_c	Array	d_{cn}	\mathbf{d} at i_c
\mathbf{s}_c	Array	s_{cn}	The expected signal expected at i_c , used to calibrate the absolute units.
\mathbf{q}	Array	q_n	Gainmap
$\Delta \mathbf{d}$	Array	δd_n	Deviation of the measured intensity from the calibration $\mathbf{d} - \mathbf{c}(i)$
G	Integer	–	Maximum Taylor expansion order used for used approximation of \mathbf{c}
\mathbf{a}	Matrix	a_{ng}	Polynomial parameters used to approximate \mathbf{c} ; polynomial order g ranges from 1 to G
p	Scalar	–	General physical example parameter
\mathbf{d}_p	Array	d_{pn}	Detector signal reduced for dependency to parameter p
\mathbf{c}_p	Array	c_{pn}	p dependence of \mathbf{d} at constant intensity distribution

components common to an entire detector segment. In the latter case, correction algorithms can be applied which analyse the intensity statistics of an entire image \mathbf{D} or a subset of that (*e.g.* physical detector segments), or use the signals from purposely unexposed pixels as noise reference. For an average sample of images ($\langle \mathbf{D} \rangle$), the randomly fluctuating signals average to zero by definition, *i.e.* the random detector behavior, if not accounted for, can be reduced by averaging over measurements of an identical signal. For this reason, the constant dark offset can be measured with high accuracy by averaging a large number of unexposed detector image acquisitions.

In this article, we focus on the photon-signal-dependent components summarized in \mathbf{d} , in which both the ideally linear response to physical signals but also potential nonlinear artifacts from, for example, pixel crosstalk are combined.

3. Example CSPAD dataset

As an exemplary test case, we use a reference diffuse X-ray scattering dataset from a free-flowing 100 μm -thick liquid sheet jet of acetonitrile which was recorded for a range of different FEL X-ray intensities using the CSPAD. The data were recorded at the LCLS XPP instrument (Chollet *et al.*, 2015), where the FEL intensity can be controlled by inserting single-crystalline silicon attenuator blades of different thicknesses into the beam path before the sample. The data represent a high exposure of the CSPAD in low-gain mode with 5 ADU (analog-to-digital units) per 9.5 keV photon, *i.e.*

the maximum shown in Fig. 1 represents approximately 2000 photons per pixel in the liquid peak and about 4×10^8 photons on the entire CSPAD.

The X-ray pulse intensity delivered to the sample was measured by a diode-based transmissive intensity monitor (Feng *et al.*, 2011). As the hereby measured intensities do not

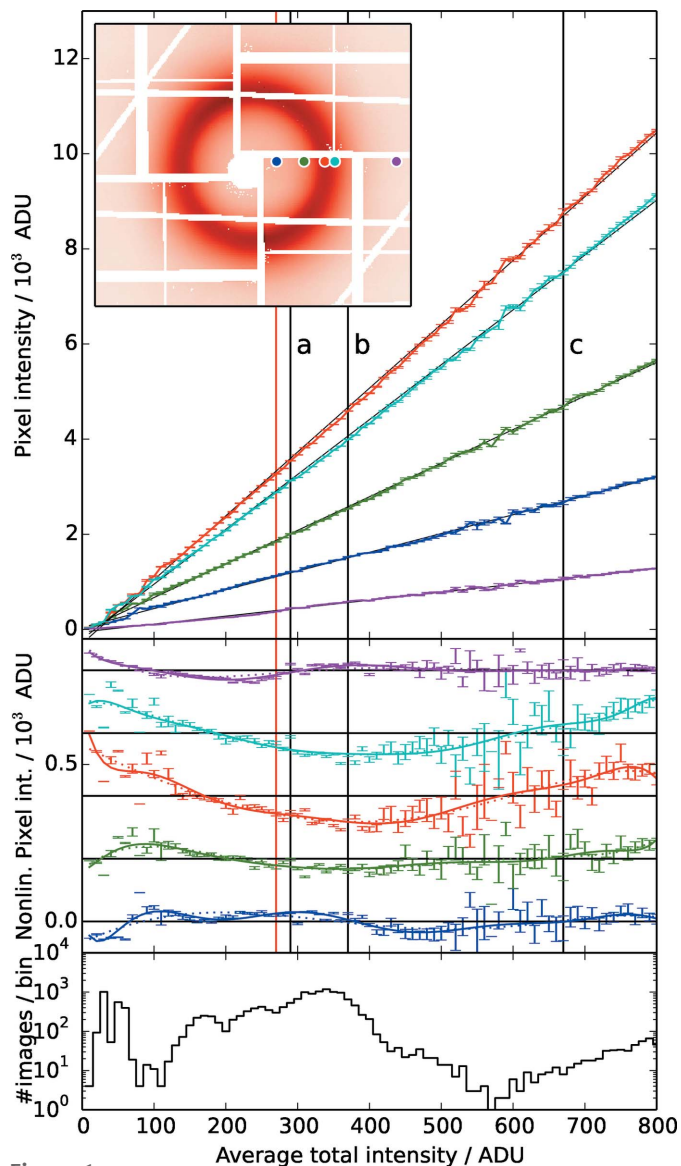


Figure 1
 Top panel: total intensity dependence of five CSPAD pixels at characteristic positions in a diffuse scattering intensity distribution (indicated by round color symbols in the inset). The intensity units (ordinate), in analog digital units (ADU), represent the pixel-dependent parameter d_n used in the mathematical descriptions. The data from multiple acquired images were averaged into bins of the total intensity (abscissa) given in units of the mean ADU intensity over the entire CSPAD pattern represented by i in the text [see also discussion after equation (10)]. The thin solid black lines represent calibration fits as described in §5. The middle panel shows the nonlinear residuals after subtracting this first-order polynomial fit. The residuals from each pixel have been offset for visibility and error bars corresponding to the standard deviation in each bin have been added. The varying statistics in each intensity interval effect the point-to-point variation; the bottom panel shows the number of images in each bin. The red (i_c) and black vertical lines (a , b , c) are example intensities used in Fig. 2.

reflect changes of sample volume due to fluctuations in the liquid jet thickness, the total exposed intensity (after subtraction of the dark component) was used as the total intensity estimate (see remarks about this procedure in §7). After subtraction of the dark component, the X-ray patterns were sorted into equally spaced intensity intervals and averaged. The data in each interval were averaged in order to reduce the magnitude of random components in the digital data or the dependency to sample jet fluctuations.

A spatially dependent nonlinear detector response becomes visible when comparing this dataset of identical signal distribution measured at different total X-ray intensities (Fig. 1). While the signals from pixels at exemplary locations in the scattering pattern show predominantly a dependence on the total intensity, additional nonlinear deviations become visible after subtracting a fitted linear component. The nonlinear behavior clearly varies qualitatively and quantitatively at the different locations in the scattering pattern. Neighboring pixels which measure a similar signal of the smooth scattering pattern, however, behave qualitatively alike: when comparing average patterns measured at different intensities that were normalized to their total intensity, residual regions with amplitudes up to 10% of the actual signals are observed where pixel groups form elevated and reduced intensity regions [Figs. 2(a)–2(c)]. These residuals, due to spatially dependent nonlinearity, increase with the difference of total X-ray intensity at which the compared patterns are measured. The intensity-feature shapes appear to correlate with the shape of the exposed signal pattern, which is a sign of pixel crosstalk. Additionally, features that are given from the individual

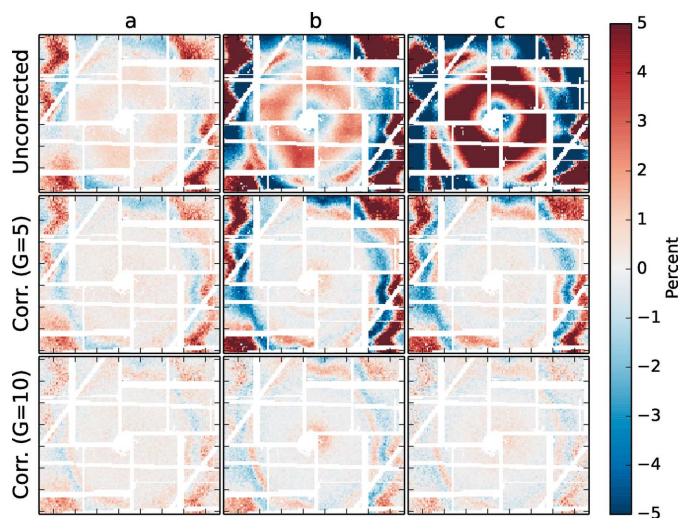


Figure 2
 Comparison of the shape of acetonitrile diffuse scattering patterns recorded at different incoming X-ray intensities i (indicated a , b and c in Fig. 1) with respect to a reference intensity (*cf.* i_c in the text, indicated by a red vertical line in Fig. 1). After the raw intensity patterns were normalized to their total intensity, their relative changes to the normalized pattern at the reference intensity are shown in color contrast as a percentage [a , b , c as columns, first row, *cf.* $(d_x/i_x)/(d_c/i_c - 1)$ where x represents a , b and c]. The same comparison was repeated after correcting the patterns with equation (10) and parametrization with equation (11) up to to orders five and ten (maximum order G ; middle and bottom row).

detector tile appear overlaid with the intensity-dependent part of nonlinear components.

4. Decomposition and correction of the photon-dependent signal at similar exposures

In order to contemplate the cross dependencies of the pixels we decompose the pixel array \mathbf{d} into its elements, the intensities per pixel d_n for the N pixels of the detector,

$$\mathbf{d} = [d_1, d_2, d_3, \dots, d_N]. \quad (2)$$

In a functional dependency view, the measured intensity on an individual pixel is a function of the physical photon exposure s_n of that pixel but also of the measured intensities in all other pixels, when assuming a most generalized pixel crosstalk,

$$d_n = d_n(s_n, d_1, d_2, d_3, \dots, d_{n-1}, d_{n+1}, \dots, d_N). \quad (3)$$

This is eventually equivalent to a dependence on the individual photon exposure experienced by each pixel or the total intensity field $\mathbf{s} = [s_1, s_2, s_3, \dots, s_N]$:

$$d_n = d_n(s_1, s_2, s_3, \dots, s_N) = d_n(\mathbf{s}). \quad (4)$$

A full correction for the detector behavior would be possible with knowledge of the intensity dependence of each pixel to all possible intensity distributions on the detector. An experimental measurement of such a general calibration function, however, is highly unpractical. For a parametrization into M intensity intervals for each pixel, M^N calibration measurements have to be taken which makes a general signal calibration for ≥ 14 -bit megapixel detectors practically impossible. However, special assumptions about the detector and the photon signal \mathbf{s} described by the total intensity field allow to some extent a quantitative comparison of the measurements.

For a specific constant photon density distribution \mathbf{I} on the detector, the intensity dependence equation (4) can be reduced to a single non-constant scalar variable:

$$d_n = d_n(i\mathbf{I}) \Big|_{\mathbf{I}=\text{const.}} = d_n(i)|_{\mathbf{I}}; \quad \sum_{n=1}^N I_n = 1, \quad (5)$$

where i is a value that scales with the total photon intensity and \mathbf{I} is normalized to unity. In this particular case the single-pixel response becomes a one-dimensional function dependent only on the scalar photon intensity i and a detector calibration dataset requires only $M \times N$ measured calibration values for the constant photon distribution under study. These can be acquired by measuring an identical signal on the entire detector as a function of the single external variable i which is proportional to the total intensity (*i.e.* the dataset described above can serve as a calibration dataset). We describe this general detector response for a given intensity distribution \mathbf{I} as a set of calibration functions c_n for each pixel of the detector:

$$d_n|_{\mathbf{I}} = c_n(i). \quad (6)$$

The functions $c_n(i)$ describe both the expected digital reading for each pixel given \mathbf{I} and the sensitivity to a change in

intensity $c'_n(i) = dc/di$. In order to quantitatively compare these N functions relative to each other, a calibration across all pixels is required, *i.e.* one point of exposure with a known signal on the calibration curves for each pixel. Here, we assume a calibration measurement at total intensity i_c with a measured intensity $\mathbf{d}_c = \mathbf{c}(i_c)$ and the known calibration signal \mathbf{s}_c , so that we can calculate a gain field $\mathbf{q} = [q_1, q_2, q_3, \dots, q_N]$ of ‘signal units per digital reading’ (*cf.* ‘gain map’ or ‘flatfield’ for a single intensity¹) at that i_c whose elements are

$$q_n = s_{cn}/d_{cn}. \quad (7)$$

Note that we have chosen our correction around a point of reference at a general intensity value i_c and not at the zero limit typical for nonlinearity correction of diodes. A reference at finite intensity values can often be the better choice because PAD detector artifacts like crosstalk may affect especially patterns of low exposure (Herrmann *et al.*, 2013). From that single reference point and assuming a constant signal distribution \mathbf{I} we can now trivially construct the linear corrected signal proportional to the total intensity scalar i as

$$s_n(i)|_{\mathbf{I}} = \frac{s_{cn}}{i_c} i. \quad (8)$$

After we have defined calibration functions for each pixel under the assumption of a constant intensity distribution, we now consider a small intensity variation in a subset of pixels, *i.e.* the measured set of total intensity and individual pixel $[i, d_n = c_n(i) + \Delta d_n]$ deviates from the ideally calibrated tuple $[i, d_n = c_n(i)]$. For such small changes from the given \mathbf{I} , to first-order approximation, effects due to changes in pixel crosstalk would be very small. The intensity distribution, and therefore the distribution of charge load on the detector, is still very similar, and the calibration set given by c_n can be approximated as invariant. The total intensity parameter i can be regarded as defining a working point on the calibration curves c_n , selecting a set of correction parameters. As the nonlinear behavior of c_n around a measured value i depends on the total load on the detector characterized by i , it does not determine the photon sensitivity of the individual pixels at that working point. The gradient $c'_n(i)$, locally defined at i , represents the best estimate of the readout response Δd_n due to a change of external photon signal Δs_n . In general, $c'_n(i)$ differs from the sensitivity at the calibration point $c'_n(i_c)$ for which the gain q_n was determined by a reference dataset. For approximating the change in corrected signal Δs_n that is due to Δd_n , the gain therefore needs to be corrected by the ratio between the sensitivities at both points before being applied to the deviation Δd_n (see also schematic visualization in Fig. 3):

$$\Delta s_n(\Delta d_n)|_{\mathbf{I},i} = \frac{c'_n(i_c)}{c'_n(i)} q_n \Delta d_n. \quad (9)$$

Under the assumption that the calibration data \mathbf{s}_c and c stay constant for small deviations from an intensity distribution \mathbf{I} ,

¹ Note that the definition of gain used here for flatfield correction is the reciprocal of the definition often used when describing the electrical signal gain of hardware.

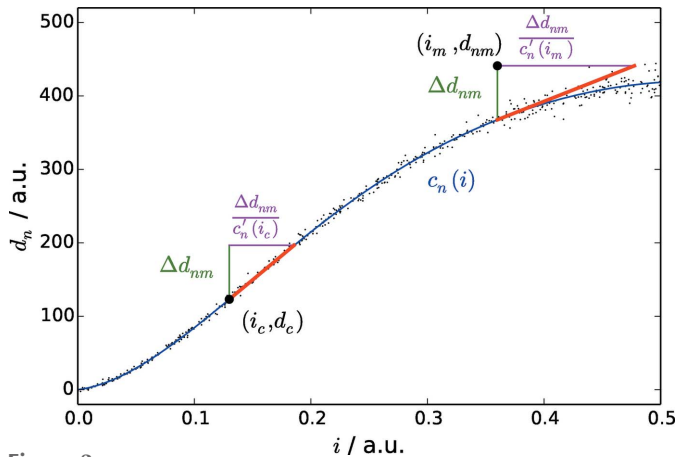


Figure 3 Schematic example of the correction of gain for a single pixel. Artificial example data of exaggerated nonlinearity show the dependence of a single pixel reading d_n as a function of the total intensity i at a constant intensity distribution on the entire detector (small black dots). The data have been parametrized to form a calibration curve $c_n(i)$ (blue line). An example data tuple (i_m, d_{nm}) shows a deviation Δd_{nm} from the expected calibrated value $c(i_m)$. This deviation is subject to a different intensity dependence (gain) than at the intensity i_c at which a quantitative calibration dataset was taken. This deviation in detector response can be accounted for by the ratio $c'_n(i_c)/c'_n(i_m)$.

the corrected intensity can be calculated from the measured tuple (i, d_n) by

$$s_n(i, d_n)|_{\mathbf{I}} = \underbrace{\frac{s_{cn}}{i_c} i}_{s_n(i)|_{\mathbf{I}}} + \underbrace{\frac{c'_n(i_c)}{c'_n(i)} q_n [d_n - c_n(i)]}_{\Delta s_n(\Delta d_n)|_{\mathbf{I},i}}. \quad (10)$$

The result of this correction function has the same unit as the calibration dataset s_c . As the unit of the total intensity value i cancels out in this final correction function, any parameter that is proportional to the total intensity can be used in the correction process.

5. Parametrization of calibration and numerical data correction

For practical application of the correction given by equation (10), an external calibration signal s_c (see also discussion in §7) and a calibration dataset of the nonlinear detector defining $\mathbf{c}(i)$ are required. An analytical differentiable parametrization of the N function elements in \mathbf{c} is preferable for efficient numerical evaluation. Additionally, an approximation by a fitted analytical function can help reduce undesired effects from statistical noise in the reference dataset. The best choice of such a function depends on the specific nonlinear effects and what physical models can be used to describe the behavior.

In the examples shown here, the calibration data are approximated by a Taylor expansion around the cross-calibration point i_c ,

$$c_n(i) = \sum_{g=0}^G a_{ng} (i - i_c)^g, \quad (11)$$

for which efficient least-squares fitting algorithms can find the parameters a_g that match the calibration data best. The polynomial order can be truncated at degree G as required to characterize the features of the calibration data while minimizing the computational effort.

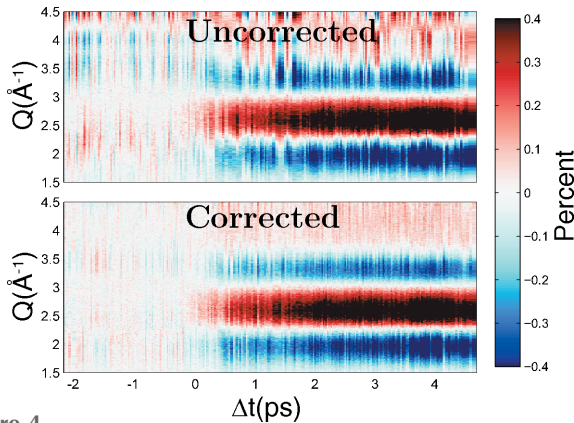
The dataset used in Fig. 1 was parametrized by this Taylor approximation for \mathbf{c} up to order $G = 5$ and $G = 10$ and the comparison of identical signals at different total X-ray intensities was repeated after the data had been corrected by equation (10) using \mathbf{c} given by equation (11) (Fig. 2, second and third row). The systematic distortions could be significantly reduced, and the overall remaining residuals are reduced by a factor ≥ 10 . The degree G of correction clearly influences the quality of correction. For our example dataset, a five-order correction approximation of \mathbf{c} does not provide enough flexibility to represent the calibration dataset with sufficient detail.

For the polynomial approximation chosen here, it is straightforward to form the derivative c' , required in equation (10), and the gain at the calibration intensity i_c is given by the first-order parameters $c'(i_c) = a_1$. For practical numerical correction, the parameter a_{ng} , an $N \times G$ -element dataset, can be kept in computer memory, and the correction reduced to a series of simple array operations which are ‘embarrassingly parallel’ to be run efficiently in numerical environments of high-level programming languages and/or on multiprocessor systems.²

As an example of realistic application, the described nonlinear intensity corrections were applied to data from a pump/probe X-ray diffuse scattering experiment with signal changes below 1%. The diffuse scattering from a 100 μm -thick jet of water was measured after exciting bending and stretching modes of the water molecules directly through a ~ 75 fs short pulse of 1950 nm light. The non-corrected and corrected images were azimuthally averaged and binned into radial profiles of units of the absolute wavevector transfer \mathbf{Q} . Difference profiles between excited and unexcited water were calculated and sorted for laser-to-X-ray time delay Δt as measured by a timing diagnostics (Harmand *et al.*, 2013; Fig. 4). Each time delay bin contains an average of 50 difference scattering images resulting in ~ 30 fs bin sizes for the presented dataset. After the pump/probe delay $\Delta t = 0$, when the two pulses overlap in time, a difference signal characteristic for heated water (Cammarata *et al.*, 2006; Kjær *et al.*, 2013) grows in.

In the uncorrected case, strong fluctuations of similar magnitude as the physical light-induced signal changes dominate the extracted signal, even though the shown data represent averages of hundreds of pixels and multiple images. The average is therefore insufficient to accurately determine the 0.1% signal changes of interest. After correction, the fluctuations in direction of the pump/probe time delay Δt are reduced to below 0.1% of the total signal and the experimental information can more readily be extracted. In this

² An example high-level code written in the languages Python and MATLAB is available as supporting information to this article.


Figure 4

Example application of the nonlinear intensity correction presented in the article. The correction was applied to a typical X-ray diffuse scattering pump/probe experiment measuring the scattering response of prompt heating of water pumped with a ~ 75 fs, 1950 nm laser pulse. The individual scattering images have been corrected, subtracted from unexcited reference images and azimuthally integrated to form difference scattering curves. The color contrast shows the relative change in scattered intensity for different scattering vectors as a function of time-delay between pump and probe pulse. After correction (lower panel) the signal fluctuations between time-bins are significantly reduced.

example the signal shape fluctuations were dominated by intensity-dependent effects.

6. General application

The method described in §4 and §5, in essence, finds the correlations of each pixel in a pixel detector with respect to a single parameter describing a physical property of one X-ray pulse, by averaging all acquired data in fine intervals of that parameter and thereby reducing the influence of all other fluctuation effects. Subsequently, a linear signal is calculated from the parameter, and deviations from the calibrated dependence are weighted according to the corresponding selected working point.

Instead of regarding deviations from an expected dependence, the parameter calibration can also be used to remove the correlation, again under the limit of similar exposure patterns. This can be useful for all those parameters which fluctuate, can be measured by a scalar signal and induce distortions image of the area detector. Removal of distortions caused by parameter p can be described by

$$d_{pn}(p)|_{\mathbf{I}} = d_n - [c_{pn}(p) - c_{pn}(p_c)], \quad (12)$$

where $d_{pn}(p)|_{\mathbf{I}}$ represents a ‘corrected’ digital reading without dependence on the parameter p , $c_{pn}(p)$ denotes calibration functions of the read signal as a function of parameter p , and p_c is a chosen calibration center.

FEL-related examples for parasitic signal distortions due to an external parameter p are the fluctuating photon energy and beam pointing parameters which can be measured by FEL diagnostics. In the example case of liquid scattering the fluctuating photon energy would change the wavevector transfer \mathbf{Q} probed by an area detector which leads to small radial

motions of the scattering features (van Driel *et al.*, 2015). As the photon energy can be measured for each pulse, either directly by a transmissive spectrometer (Zhu *et al.*, 2012) or indirectly by the energy of the FEL electron bunch, a calibration dataset can be sorted and averaged for that photon energy parameter and a calibration \mathbf{c}_p can be formed where p in equation (12) represents the photon energy. As a calibration parameter value p_c the average photon energy would be a well suited choice for the nominal value, and the images would be corrected to match the signal at this p_c . Also for other FEL-specific fluctuations related to beam pointing or temporal properties of the pulses, a parametrized correction as described in this article can be envisioned.

7. Practical considerations

The procedure described here corrects complex nonlinear signal distortions model-independently through signal-dependent calibration. The nonlinearity correction relies on a representative calibration dataset of a constant signal measured at different intensities that span over the intensity distribution in the real measurement [defining $\mathbf{c}(i)$]. In the case of typical FEL intensity fluctuations, a measurement over a large number of X-ray pulses along with the data would typically serve that purpose, as the intensity fluctuations between this dataset and the actual experiment would be comparable. A larger intensity range can be covered by purposely attenuating the incoming X-ray pulse intensity as described in §3. Such a dataset can directly be used to determine the calibration field $\mathbf{c}(i)|_{\mathbf{I}}$ for the measured i range, *e.g.* by polynomial approximation as given in equation (11). In the case of additional fluctuations like common mode detector noise, sample volume fluctuations or similar, the effect of those parameters on the nonlinearity calibration dataset can be reduced by averaging the data of similar intensity into bins (performed in the example dataset shown in Figs. 1 and 2).

The quality of the nonlinearity corrections depends on the quality of the calibration dataset. But even data at lower statistics can lead to a significant reduction of nonlinearity-induced fluctuations. In case a dedicated calibration dataset has not been recorded, suitable reference measurements can often be identified within the recorded data. In the case of pump/probe measurements, a set of unexcited time-delays or reference images are usually acquired during the experiment. Such data are well suited for corrections as long as the intensity distribution and effective fluctuations are representative of those data to be corrected.

In the examples presented here, the CSPAD in the v1.2 generation was corrected at high intensity levels. The systematic nonlinear effects had been significantly improved compared with the v1.0 generation. As the crosstalk effects have been shown to affect also very low measured intensities with few photons per pixel (Herrmann *et al.*, 2014), the correction procedures can improve all intensity ranges at which the CSPAD is used. The newer v1.5 and the presently active v1.6 CSPAD generation represent another substantial step of reduction of crosstalk effects. Nevertheless, we have

gained significant improvements of pump/probe data fluctuations by correcting for intensity-dependent nonlinearity for those newest generations. The large fluctuations of pulse intensities typical for FELs cause a high sensitivity to nonlinear effects. It should be noted that very small crosstalk effects can be expected at other types of integrating area detectors whose effects turn out to be less obvious when using very stable light sources like synchrotrons.

Other experimental techniques, not tested in this article, often regard single-image differences from a well established average of all images, and a correction of nonlinear detector response described here might reduce systematic detector-induced errors. In such cases the binning technique described in this article can help to generate a calibration dataset with reduced dependence on the individual signal changes.

In macromolecular crystallography, a major field of application of FEL radiation, artifacts of structure factor determination through detector-induced spatial and intensity-dependent signal contributions can be typically reduced by relating the Bragg reflection intensities to the background signal measured in their vicinity. Strong diffuse background signals due to amorphous liquid, however, can represent the major signal measured by the area detector, and thereby cause nonlinear artifacts in the measurement of Bragg reflections through pixel crosstalk. The correction concepts described here could help to reduce such effects because the small Bragg reflections represent a comparably small deviation from a constant intensity distribution, if the background is strong. Quantification of the background signal as well as measurement of a calibration dataset based on the background signal, independently of the strong Bragg reflections, can require advanced numerical methods like spatial frequency filtering. Correction of such data has not been tested. It should be noted that the systematic errors due to nonlinearity we present here in the CSPAD examples do not exceed a few percent of the signal and might be outweighed by other systematic uncertainties in a crystallography experiment.

As derived from equation (10), requirements for the total intensity parameter i are relaxed to a parameter which is proportional to the intensity. Nevertheless, a good measure of the total intensity can be difficult to obtain in practice, for example due to additional nonlinearity of intensity monitors or fluctuations in the sample as in the case for a liquid jet. In the presented example, the total intensity measured by the CSPAD detector was used as a reference intensity monitor. This procedure appeared justified by cross-correlation with an intensity monitor which suggested that the nonlinear effects of the individual pixels appear reduced when averaging over a large ensemble of pixels at different intensities. Depending on the level of nonlinearity and the exposure level of the detector, such a procedure might be inappropriate and return a nonlinear signal itself, and other means for an intensity monitor have to be identified.

The practical relative or even absolute gain calibration across all pixels using a single reference image s_c can represent a challenge, depending on the detector behavior. In the case of negligible pixel crosstalk at the low exposure limit, the zero

approximation ($i_c = 0$) can serve as a good and most general reference point, which would be independent of the regarded intensity distribution. In this case a low-exposure flatfield measurement of a large number of pulses can serve as calibration and the important gain calibration $\mathbf{c}'(i = 0)$ can be determined by extrapolation of \mathbf{c} to $i = 0$. In case the calibration curves cannot be easily parametrized at the low exposure limit, for example due to discontinuous behavior through crosstalk [like in the case of the CSPAD, cf. Herrmann *et al.* (2014)], the actual intensity distribution might be required to serve as intensity calibration. In many cases the measured signal distribution (e.g. a liquid scattering pattern) can be or has been measured on a highly linear detector at a synchrotron source and can be used to calibrate intensities. For the direct comparison of such a reference, parameters dependent on the individual setup geometry (e.g. detector efficiency, scattering angles, X-ray polarization) have to be taken into account. Even without an intensity calibration dataset, the nonlinear corrections can be applied and the gain properties are determined by $\mathbf{c}(i = i_c)$.

The method presented here independently corrects any array of individual elements with respect to an intensity variable valid for all those elements. Its field of application is therefore more general compared with the approach on independent components which gains sensitivity with the dimensionality of the detector array (van Driel *et al.*, 2015). As the variable of fluctuation is required to be known and the dependence of the calibration to other variables can be reduced by averaging into intensity bins, correction of lower dimension data does not reduce its efficiency and validity. It can therefore also be applied to smaller subsets of pixels with applied masks or to reduced data where one element of the array includes information from multiple pixels. In those cases the numerical effort can therefore be reduced. The methods can be combined in cases where the origin of signal fluctuations is not unknown. Hence, the SVD-based approach would help to identify parameters that dominate the fluctuations, like the total intensity or other physical X-ray beam parameters (see §6). Thereafter, the calibration steps presented here would be applied in order to determine the nonlinear dependence of each of those identified parameters and reduce either dependencies or a nonlinear detector response.

Acknowledgements

The authors would like to thank Kelly Gaffney for providing access to the acetonitrile data used in this article, which were taken during his user experiment. HTL thanks Marco Cammarata for valuable input and fruitful discussions. Authors TBvD and MMN acknowledge support from DANSCATT and support from the ReLIable project (Project No. 11-116792) funded by the Danish Council for Strategic Research Programme Commission on Sustainable Energy and Environment. Portions of this research were carried out at the Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory. LCLS is an Office of Science User

Facility operated for the US Department of Energy Office of Science by Stanford University.

References

- Arnlund, D. *et al.* (2014). *Nat. Methods*, **11**, 923–926.
- Blaj, G., Caragiulo, P., Carini, G., Carron, S., Dragone, A., Freytag, D., Haller, G., Hart, P., Hasi, J., Herbst, R., Herrmann, S., Kenney, C., Markovic, B., Nishimura, K., Osier, S., Pines, J., Reese, B., Segal, J., Tomada, A. & Weaver, M. (2015). *J. Synchrotron Rad.* **22**, 577–583.
- Boutet, S. *et al.* (2012). *Science*, **337**, 362–364.
- Cammarata, M., Lorenc, M., Kim, T. K., Lee, J. H., Kong, Q. Y., Pontecorvo, E., Lo Russo, M., Schiró, G., Cupane, A., Wulff, M. & Ihee, H. (2006). *J. Chem. Phys.* **124**, 124504.
- Carini, G. A., Boutet, S., Chollet, M., Dragone, A., Haller, G., Hart, P. A., Herrmann, S. C., Kenney, C. J., Koglin, J., Messerschmidt, M., Nelson, S., Pines, J., Robert, A., Song, S., Thayer, J. B., Williams, G. J. & Zhu, D. (2014). *J. Phys. Conf. Ser.* **493**, 012011.
- Chollet, M., Alonso-Mori, R., Cammarata, M., Damiani, D., Defever, J., Delor, J. T., Feng, Y., Glowina, J. M., Langton, J. B., Nelson, S., Ramsey, K., Robert, A., Sikorski, M., Song, S., Stefanescu, D., Srinivasan, V., Zhu, D., Lemke, H. T. & Fritz, D. M. (2015). *J. Synchrotron Rad.* **22**, 503–507.
- Driel, T. B. van, Kjaer, K. S., Biasin, E., Haldrup, K., Lemke, H. T. & Nielsen, M. M. (2015). *Faraday Discuss.* **177**, 443–465.
- Feng, Y., Feldkamp, J. M., Fritz, D. M., Cammarata, M., Aymeric, R., Caronna, C., Lemke, H. T., Zhu, D., Lee, S., Boutet, S., Williams, G., Tono, K., Yabashi, M. & Hastings, J. B. (2011). *Proc. SPIE*, **8140**, 81400Q.
- Haldrup, K. (2014). *Philos. Trans. R. Soc. B*, **369**, 20130336.
- Harmand, M., Coffee, R., Bionta, M. R., Chollet, M., French, D., Zhu, D., Fritz, D. M., Lemke, H. T., Medvedev, N., Ziaja, B., Toleikis, S. & Cammarata, M. (2013). *Nat. Photon.* **7**, 215–218.
- Hart, P., Boutet, S., Carini, G., Dubrovin, M., Duda, B., Fritz, D., Haller, G., Herbst, R., Herrmann, S., Kenney, C., Others, Kurita, N., Lemke, H., Messerschmidt, M., Nordby, M., Pines, J., Schafer, D., Swift, M., Weaver, M., Williams, G., Zhu, D., Van Bakel, N. & Morse, J. (2012). *Proc. SPIE*, **8504**, 85040C.
- Henrich, B., Becker, J., Dinapoli, R., Goettlicher, P., Graafsma, H., Hirsemann, H., Klanner, R., Krueger, H., Mazzocco, R., Mozzanica, A., Perrey, H., Potdevin, G., Schmitt, B., Shi, X., Srivastava, A., Trunk, U. & Youngman, C. (2011). *Nucl. Instrum. Methods Phys. Res. A*, **633**(Suppl. 1), S11–S14.
- Herrmann, S., Boutet, S., Duda, B., Fritz, D., Haller, G., Hart, P., Herbst, R., Kenney, C., Lemke, H., Messerschmidt, M., Pines, J., Robert, A., Sikorski, M. & Williams, G. (2013). *Nucl. Instrum. Methods Phys. Res. A*, **718**, 550–553.
- Herrmann, S., Hart, P., Dragone, A., Freytag, D., Herbst, R., Pines, J., Weaver, M., Carini, G. A., Thayer, J. B., Shawn, O., Kenney, C. J. & Haller, G. (2014). *J. Phys. Conf. Ser.* **493**, 012013.
- Kjaer, K. S., van Driel, T. B., Kehres, J., Haldrup, K., Khakhulin, D., Bechgaard, K., Cammarata, M., Wulff, M., Sørensen, T. J. & Nielsen, M. M. (2013). *Phys. Chem. Chem. Phys.* **15**, 15003–15016.
- Koch, A., Hart, M., Nicholls, T., Angelsen, C., Coughlan, J., French, M., Hauf, S., Kuster, M., Sztuk-Dambietz, J., Turcato, M., Carini, G. A., Chollet, M., Herrmann, S. C., Lemke, H. T., Nelson, S., Song, S., Weaver, M., Zhu, D., Meents, A. & Fischer, P. (2013). *JINST*, **8**, C11001.
- Mozzanica, A., Bergamaschi, A., Cartier, S., Dinapoli, R., Greiffenberg, D., Johnson, I., Jungmann, J., Maliakal, D., Mezza, D., Ruder, C., Schaedler, L., Schmitt, B., Shi, X. & Tinti, G. (2014). *JINST*, **9**, C05010.
- Philipp, H. T., Hromalik, M., Tate, M., Koerner, L. & Gruner, S. M. (2011). *Nucl. Instrum. Methods Phys. Res. A*, **649**, 67–69.
- Philipp, H. T., Koerner, L. J., Hromalik, M. S., Tate, M. W. & Gruner, S. M. (2010). *IEEE Trans. Nucl. Sci.* **57**, 3795–3799.
- Sellberg, J. A. *et al.* (2014). *Nature (London)*, **510**, 381–384.
- Trigo, M., Fuchs, M., Chen, J., Jiang, M. P., Cammarata, M., Fahy, S., Fritz, D. M., Gaffney, K., Ghimire, S., Higginbotham, A., Johnson, S. L., Kozina, M. E., Larsson, J., Lemke, H., Lindenberg, A. M., Ndabashimiye, G., Quirin, F., Sokolowski-Tinten, K., Uher, C., Wang, G., Wark, J. S., Zhu, D. & Reis, D. A. (2013). *Nat. Phys.* **9**, 790–794.
- Weidenspointner, G., Andritschke, R., Aschauer, S., Fischer, P., Granato, S., Hansen, K., Lechner, P., Lutz, G., Moch, D., Porro, M., Sandow, C., Sant, T. & Struder, L. (2011). *In 2011 IEEE Nuclear Science Symposium Conference Record*, pp. 468–473. IEEE.
- Zhu, D., Cammarata, M., Feldkamp, J. M., Fritz, D. M., Hastings, J. B., Lee, S., Lemke, H. T., Robert, A., Turner, J. L. & Feng, Y. (2012). *Appl. Phys. Lett.* **101**, 034103.