

Absolute contrast estimation for soft X-ray photon fluctuation spectroscopy using a variational droplet model

Nicolas Georges Burdet (✉ roibabar@slac.stanford.edu)

SLAC National Accelerator Laboratory

Vincent Esposito

SLAC National Accelerator Laboratory

Matthew Seaberg

SLAC National Accelerator Laboratory

Chun Hong Yoon

SLAC National Accelerator Laboratory

Joshua Turner

SLAC National Accelerator Laboratory

Research Article

Keywords: XPFS, Linac Coherent Light, low-energy physics, X-ray, X-FELs

Posted Date: May 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-474179/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Absolute contrast estimation for soft X-ray photon fluctuation spectroscopy using a variational droplet model

N.G Burdet^{1,2,*}, V. Esposito^{1,2}, M.H Seaberg², C.H. Yoon², and J.J. Turner^{1,2}

¹Stanford Institute for Materials and Energy Science, Stanford University and SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA

²Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA

*roibabar@slac.stanford.edu

ABSTRACT

X-ray photon fluctuation spectroscopy (XPFS) using a two-pulse mode at the Linac Coherent Light Source has great potential for the study of quantum fluctuations in materials as it allows for exploration of low-energy physics. However, the complexity of the data analysis and interpretation still prevent recovering real-time results during an experiment, and can even complicate post-analysis processes. This is particularly true for high-spatial resolution applications using CCDs with small pixels, which can decrease the photon mapping accuracy resulting from the large electron cloud generation at the detector. Droplet algorithms endeavor to restore accurate photon maps, but the results can be altered by their hyper-parameters. We present numerical modeling tools through extensive simulations that mimic previous fluctuation spectroscopy experiments. By a modification of a fast droplet algorithm, our results demonstrate how to optimize the precise parameters that lift the intrinsic counting degeneracy impeding accuracy in extracting the speckle contrast. These results allow for an absolute determination of the summed contrast from multi-pulse x-ray speckle diffraction, the *modus operandi* by which the correlation time for spontaneous fluctuations can be measured.

Introduction

Spontaneous nanoscale fluctuations manifest over a wide range of material systems and X-FELs offers a new range of time and energy scales over which their affect on material properties can be studied. In x-ray photon correlation spectroscopy (XPCS), the sample dynamics are revealed by correlating a time sequence of images emanating from the coherent scattering of a pulse train^{1,2}. More specifically, the minute changes with time delay τ , are captured by the function $g^{(2)}(q, \tau)$, which measures the temporal auto-correlation of the scattered intensity $I(q, \tau)$. The sensitivity comes from the coherent nature of the scattered photons, which interfere to forge a speckle pattern that uniquely corresponds to the fingerprint of the sample complex domain distribution.

New versions of XPCS using a split-pulse or a double-bunch pulse pair, what we refer to here as x-ray photon fluctuation spectroscopy (XPFS), are methodologies which are being developed in which the photons are not correlated, but counted. The coherent photon distribution follows a well-subscribed recipe which can be related back to what is known as the contrast function, and hence $g^{(2)}(q, \tau)$ through the speckle visibility relation. In other words, the photon fluctuations finally yield the photon correlation function. This has been recently demonstrated³ to remove the detector read-out limitations that limit the shortest measurement times in XPCS. It relies on measuring the change in contrast $C(q, \tau)$ from the recorded sum of two diffracted patterns delayed by a variable time τ , which has been proven equivalent to measuring $g^{(2)}(q, \tau)$, the intensity-intensity correlation function⁴.

The two x-ray pulses can be created in a variety of ways. For example, they can be obtained from splitting a single coherent pulse by mean of an optical setup⁵, splitting the electron pulse by slicing before the x-rays are generated⁶, or generating the two at the source with two separate laser systems to create two electron pulses using the two-bucket mode⁷. These different methods cover the range from tens of fs to or order one microsecond, bridging the gap with the timescales available at high-brightness synchrotron sources. XPFS opens up a new realm of energy ranges corresponding to collective excitations in condensed matter.

While it is the unique combination of coherent and ultra-short pulse durations of photons produced by X-FELs that foster this technique with a time resolution orders of magnitude faster than that available with current detector read-out capabilities, the high fluence that results can potentially induce out-of-equilibrium dynamics in the system being probed. In order to preserve the system from this effect, the beam intensity has to be drastically reduced to the weak photon counting limit³. In this sparse

photon regime, the intrinsic sample dynamics are deduced by invoking a photon statistical tool to analyze single shot detector images. The fitting statistics to reliably extract $C(q, \tau)$ is established through hundreds or thousands of frames using the formalism of quantum optics. Mathematically, this is carried out via a negative binomial distribution⁸ which contains the contribution from the expected scattering (Gamma distribution) and the shot-noise (Poisson distribution), convoluted within a single function:

$$P(k, \bar{k}, M) = \frac{\Gamma(k+M)}{k! \Gamma(M)} \left(\frac{\bar{k}}{\bar{k}+M} \right)^k \left(\frac{M}{\bar{k}+M} \right)^M, \quad (1)$$

where $\Gamma(x)$ is the gamma function, k is an integer number of photons per speckle, \bar{k} the average number of photons per speckle, and M is the number of modes (degrees of freedom) in the speckle patterned, inversely related to the contrast by $C = 1/\sqrt{M}$. For a given contrast and mean intensity, this function gives the probability of counting k photons per speckle. Experimentally, the starting zero delay point, $C_0(q, \tau = 0)$ is expected to be near unity, but can differ ($M > 1$) depending on the transverse and longitudinal coherence. This is especially true for large scattering angles, where the longitudinal coherence is reduced, and can contain an increased number of modes. For $\tau > 0$, the contrast becomes a measure of the sample dynamics. The delay time τ is increased until it is larger than the correlation time, at which point the two speckle pattern reduce to a contrast of $C_0/\sqrt{2}$, for full decorrelation⁸. The speckle size is given by $s = \lambda d/l$ where λ is the wavelength, l is the distance to detector and d the beam size. Since λ is usually dictated by the sample of interest in the soft x-ray range for resonant scattering, given a finite pixel size, oversampling of the speckles is achieved by varying the beam size or the distance from the sample to the detector.

One of the critical issues when using pixelated detectors, is the electron cloud generated around each detected photon hit due to the charge sharing. Depending on the x-ray energy and the size of pixels, the clouds can overlap and eventually coalesce to form islands. Furthermore, readout noise can extend the blurring in a non-trivial way as well. Since a wrong registration of the photon within the correct speckle will have a strong impact on the determination of the contrast, a "deblurring" method called a droplet algorithm has been applied to decipher the effect of charge sharing on every image. Though other authors have studied optimization of droplet algorithms⁹, the prescriptions mainly apply to the situation where charge clouds are small compared to the pixels. The opposite situation however was met in the experiment by Seaberg et al.³, the first proof-of-principle of XPFS using the double pulse summed-contrast method carried out at the LCLS¹⁰. In that case, an Andor CCD camera bearing a small pixel size of 13 μm was employed to achieve sufficient oversampling. This leaves a strenuous task for the droplet algorithm in the disentangling of large charge clouds that overflow into neighboring pixels, and with it the many algorithm parameters that can compromise a full restoration of the photon map.

In this paper, we demonstrate a solution to this problem for incorporating a variational droplet algorithm for high resolution pixelated detectors used in XPFS. By modeling this experiment with a detailed account of the detector characteristics, extensive simulations have been carried out in order to determine the precise algorithmic parameters that extract the true absolute contrast. The limit of the average photon intensity in which the algorithm operates best is also examined. Finally, additional photon energies are analyzed in view of future experiments to study fluctuations at XFELs.

Model

Experimental data are modeled by simulating pixelated grids with weak intensity speckle patterns using photon statistics following a negative binomial distribution (Eq.1), and onto which Gaussian shaped charge clouds are convolved with each photon, to account for charge sharing effects on the CCD detector:

$$g(x, y) = \frac{1}{2\pi\sigma_G^2} \exp\left(-\frac{(x-x_c)^2}{2\sigma_G^2} - \frac{(y-y_c)^2}{2\sigma_G^2}\right), \quad (2)$$

where σ_G is the radius of the charge cloud, also referred to as the strength of the charge sharing parameter. The variables x and y are the pixel coordinate and (x_c, y_c) is the random location center of the charge cloud within that pixel. In addition, pixels receive a dose of readout noise along with a small gain non-uniformity.

A similar model was already successfully applied in the work of Lehmkuhler et al.¹¹ with which they were able to reproduce an experimental histogram from data taken at the SACLA free-electron laser and thereby deduce the spatial coherence, and the resultant contrast, by simply adjusting the model parameters and fitting. However, for larger charge cloud probability (or smaller pixels) as found in the experiment discussed here, a simple contrast interpretation from the histogram can become ambiguous and calls for a more rigorous analysis of the charge cloud maps using a droplet algorithm.

To this aim, photon maps of known contrast are generated from a random sampling of the negative binomial distribution (Eq.1) in the weak scattered photon limit. Speckle patterns are first simulated with an average photon count rate on the interval

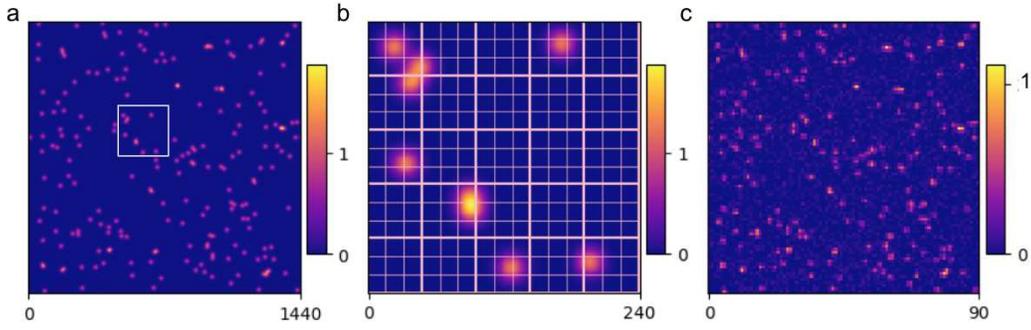


Figure 1. CCD charge sharing model of speckle patterns. a. Sparse distribution of weakly scattered photons on a 3 pixels/speckle basis, generated via the negative binomial distribution (Eq.1) within the range $\bar{k} = [0.025 - 0.25]$ photons/pixel/speckle. The selected contrast is 0.7. The map is of high-resolution since every pixel is sub-divided by 16x16 to obtain a sub-pixel placement of Gaussian shaped charge clouds (FWHM ~ 1 pixel). b. Enlarged portion of the high resolution photon map (white square box in a), displaying a 5x5 speckles area (bold lines boundaries), each speckle containing 3x3 pixels (light lines boundaries). Photons can sprawl over both pixel and speckle boundaries. c. Final ccd image (or speckle pattern) obtained after binning the high resolution photon map A to a coarser 90x90 grid and adding readout noise ($\sigma_N = 0.05$ photon). A single photon readout is 350 ADU.

$\bar{k} = [0.025, 0.25]$ photon/speckle, on a 90 x 90 pixel basis and with a nominal speckle size of $\sim 3 \times 3$ pixels. This particular \bar{k} range accounts for the inexorable stochastic shot-to-shot intensity fluctuation of the XFEL beam, while also being within the sparse photon limit. The beam fluctuation distribution would likely follow a Gamma distribution without undergoing scattering from the sample. Each pixel is then sub-divided into a 16x16 array to attain sub-pixel resolution in the placement of the random photon location within each speckle. Once a photon is allocated to a speckle, it is randomly placed within the corresponding 3x3-pixel/speckle area, according to a homogeneous distribution within each speckle (Fig. 1a). By convolving these high-resolution maps with a Gaussian distributed charge cloud (Eq.2) of size $\sigma_G = 0.45$ pixels, strong charge sharing is enabled (Fig. 1b). The intensity of a single charge cloud is set to an ADU of 350. Both values were taken from the experimental data in Seaberg et al.³, which used a photon energy of 1.19 keV. Finally, to emulate the experimental data, the high-resolution map is binned back by 16x16 and Gaussian noise of width $\sigma_N = 0.05$ photon, centered around zero, is added to every pixel (Fig. 1c).

Because of the statistical nature of the process, acceptable accuracy can only be obtained by gathering a sufficiently large number of two-pulse speckle patterns. Sets of 1000 patterns are thus simulated for each configuration. Once the patterns are available, the goal is to recover the initial contrast value, using the same analysis tools as for the experimental data. We start by assessing the accuracy in generating sets of 1000 photon maps with the given contrasts (Eq.1) from the fitting of the photon statistics $P(k)$ for $k=1,2$, and 3. An example of a fit for a simulated contrast $C=0.7$ is given in Fig. 2a. The respective uncertainties for $k=1,2$, and 3 are used for the final contrast using a weighted average, though it has recently been shown this can also be accomplished by using the ratio, $R_k = P(k+1)/P(k)$, to solve for a closed-form analytic solution (Shen et al., accepted).

To obtain the normalised RMS error reported Fig. 2b, the sets of 1000 photon maps were generated and fitted 50 times for every contrast value in the range [0.1-1] for 0.1 increments. It was shown that only the lowest contrast value of 0.1 had a deviation larger than 10%.

A cardinal issue with the extraction of the photon statistics is the charge cloud size. In previously published numerical models, the charge cloud size variation was estimated to be null¹¹ or as small as 15% of its average size¹². Based on the premise that this variation has a major impact on our results, and that droplet algorithms rely on an intimate knowledge of the individual photons, we focus on the investigation of single photon events, comparing experimental and simulated data.

For the simplest scenario, droplets with a total intensity that matches the single photon ADU are isolated and fitted with a 2D Gaussian shaped cloud model via a least squares optimization of Eq.2. In the real experiment however, the variation in the location of the detected photon with respect to the pixel can cause the droplet structure to become complicated rapidly. In the following, the cylindrical symmetry is relaxed and independent widths are considered along the vertical (σ_x) and horizontal (σ_y) directions, and the generated radius is elliptical, i.e. $\sigma_G = \sqrt{\sigma_x + \sigma_y}$.

The success rate and accuracy of the radius σ_G fitting optimization was tested first by generating 50 noiseless charge cloud maps with ($\sigma_G = 0.45$). The convergence success rate obtained was $\sim 90\%$, with a slightly overestimated average radius $\sigma_G \sim 0.5$ along with a narrow distribution that corresponds to the "digitization" of the under-sampled cloud. The accuracy in

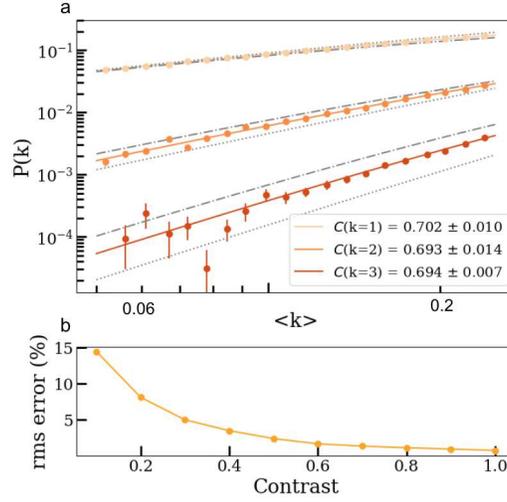


Figure 2. Fitting of observed probabilities $P(k)$ for $k = 1$ to 3 photons/speckle as a function mean photon density \bar{k} . a. Fits from a set of 1000 speckled patterns generated within $\bar{k} = [0.025 - 0.025]$ and contrast $C = 0.7$ with Eq.1. The thin curves represents the contrast boundaries for $C = 0$ and $C = 1$ respectively. Eq.1 is also used for fitting C is calculated from a weighted average of the fitted $k = 1 - 3$ distribution. The generated and fitted contrasts agrees with each other well. b. Plot of the normalised rms error obtained after repeating 50 times the fitting procedure of a. over newly generated batches of speckled patterns.

retrieving the sub-pixel center location (x_c, y_c) is not the focus of this work. After injection of both a readout noise ($\sigma_N = 0.05$ photon), and a variation of 15% in the cloud radius (composed from the independent $\sigma_{x,y}$ widths), the fitting convergence rate drops to $\sim 80\%$. While the recovered average radius remains the same with $\sigma_G \sim 0.5$, its distribution becomes very broad, as can be seen in model 1 of Fig.4b). The "zeroing threshold" was set approximately to the level of noise $\sigma_G = 0.05 \text{ photon}$. However, it is to be noted that the fitted distribution shape and width is dictated by the level of thresholding.

The experimental distribution of σ_G is reported in Fig. 4b. This knowledge alone does not allow one to further advance the numerical models. Therefore, an additional and essential piece of information is derived from this fit, what we call the single photon pixel spectral width, i.e. the probability for single charge cloud radius pixel spread. This new information is reported in Fig. 4a and while as expected, a single photon mainly occupies 4 to 6 pixels, with a non-negligible amount of single photons being bounded to a single pixel ($\sim 20\%$). This contradicts the original assumptions for model 1, hence this knowledge is reincorporated into a more advanced model. To do so, a cloud radius is assigned to each of the probability density peaks (experimental model Fig. 4a respectively with $(\sigma_G = [0.1, 0.25, 0.35, 0.45, 0.55, 0.6, \dots])$, e.g., a small radius of 0.1 is appointed to the first peak, with a weighted probability of 20 %, and inversely, a large radius $\sigma_G > 0.6$ for the last peak with a low weighted probability. As can be observed, the output of the new simulation, i.e., model 2, qualitatively recreates the experimental data. Another way to reach a more robust output is an optimization loop. However, the stochastic nature of the charge splitting among CCD pixels¹³ could be an obstacle in that scenario.

An additional way to discriminate between the models, is provided from the direct comparison of pixel readout histograms with the experimental data (Fig. 5a. Model 2 clearly reproduces the amplitude of the first photon peak at a nominal readout $r = 340$ ADU. The low amplitude of the second photon peak is due to the dilution of the speckles to 3×3 pixel together with the sub-pixel placements, making events containing 2 full photons within a single pixel statistically insignificant. Since readout noise and gain non-uniformity can only modestly influence the fits, it is apparent that the original model 1 is not suitable for this type of experiment.

By repeating the photon droplet exploration without constraining the signal to a single photon, the full extent of the droplet total readout is identified, as seen in the histogram of Fig. 5b. While this last point does not differentiates the models, it gives valuable information to run droplet algorithms. The multiple peaks of the nominal photon readout (~ 340 ADU, or 1.19 keV), corresponds to the energy calibration of the detector while the width corresponds to the energy resolution. The photon boundaries denoted by the gray vertical lines determine the integer number of photons within each droplet. The maximum of 4 photons/droplet mainly results from selecting speckle patterns within a limited range of average photon counts per speckle $\bar{k} = [0.025, 0.25]$. Extending the range of \bar{k} results in very large droplets extending over multiple speckles, leading to droplet percolation. Possessing a fully tunable model that reproduces the experimental data, we now demonstrate the absolute contrast retrieving abilities of the advanced droplet algorithm.

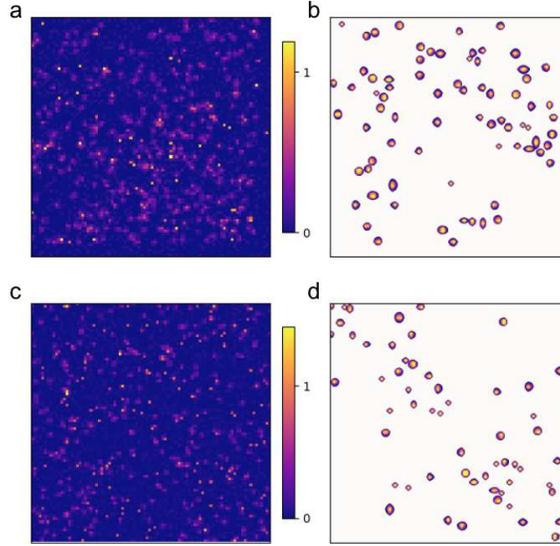


Figure 3. Single charge cloud's size analysis from experimental and simulated data. a. Experimental data. b. The single photon events are isolated from a by selecting connected islands containing a photon readout ($0.5 < \text{photon} < 1.5$) and which shape are fitted with 2d Gaussian shaped cloud model optimization, Eq.2. c. Simulated speckled pattern with variable charge cloud's sizes. The Gaussian fit shows a variability in single photon's size (distribution reported in Fig. 4.) d. Single photon events are isolated and fitted from c.

Contrast evaluation: results

In this section, the results on contrast extraction using the most straightforward of the droplet algorithms are presented. These aim at converting the raw detector intensity map into a photon count map, accounting for the detector signal-to-noise and charge sharing effects. These algorithms have been thoroughly discussed in the recent work of Sun et al.¹², where the authors observed that the droplet algorithms commonly used only differ by a constant bias, and hence can be corrected with a linear model. Having a real-time online analysis in mind, we naturally chose the fastest algorithm, the so called Greedy Guess (GG). The Least Squares Fit (LSF)⁹ which rectifies the photon assigned by the GG algorithm is much slower and is usually adapted for an offline analysis scenario.

The GG algorithm processes every speckle pattern in the following manner. In step one, the effects of readout noise are suppressed by "zeroing" the pixels containing values below a given threshold. This operation leaves islands of connected pixels with non-zero signal, identified as droplets by the algorithm. In our study, the connected direction are not discriminated, i.e., diagonal pixels are included to accommodate for the large photons clouds. In the second step, the signal within each identified droplet is summed and ascribed an integer number of photons, according to the readout histograms found in Fig. 5b. Then, the highest pixel value of the droplet is assigned a photon, and a Gaussian cloud centered on that pixel is subtracted from the droplet. The intensity and width of the Gaussian are assigned the mean single photon ADU and the estimated charge cloud size respectively, from which location a Gaussian charge cloud is subtracted whose total intensity is equivalent to the readout of one single photon (Eq.2). This operation is repeated until the total integer photon count of the droplet reaches zero. Since our GG algorithm operates on the basis of a Gaussian charge cloud model, we name it the Gaussian Greedy Guess (GGG) algorithm. The average photon count/speckle \bar{k} and the photon probabilities $P(k)$ for $k = 1, 2$, and 3 are then computed from the photon maps and fitted with the negative binomial distribution to extract the contrast $C(q, \tau)$ of the speckle pattern.

The single photon ADU, the pixel intensity threshold, and the charge cloud size are thus critical hyper-parameters of the analysis process that can dramatically alter the final result. While the former can often be unambiguously defined from the droplet intensity histogram (Fig. 5), there is not a clear criterion on how to set the other two. This becomes particularly problematic for low signal-to-noise cases, where small variations can have significant effects. With our model, this phase parameter space can now be explored with prior knowledge of the true contrast value, allowing us to define precise rules on how to fix these hyper-parameters for realistic experimental conditions.

Examples of a contrast fit for a pixel intensity threshold of ~ 0.1 photon and charge cloud size ranging from $\sigma_G = 0.2$ to 0.6 pixels is shown in Fig. 6a. A linear response of the greedy guess is observed, with a slope that tends to 1 as the cloud radius approaches $\sigma_G = 0.4$. These two values locate the lowest region of the surface error displayed in Fig. 6b. For every

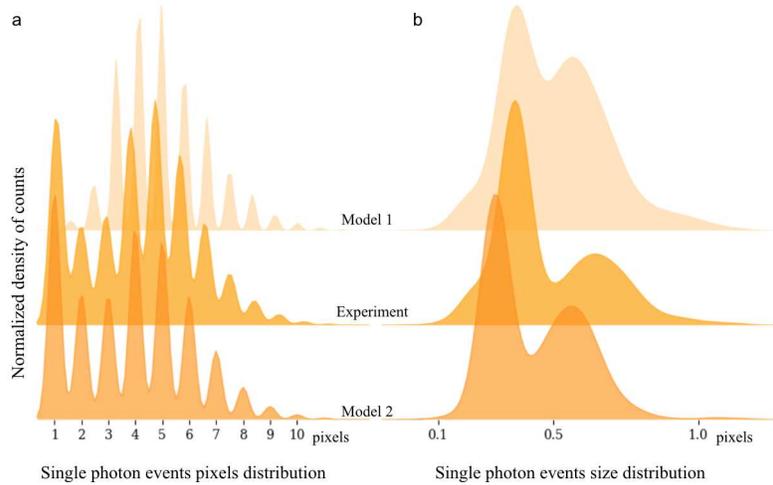


Figure 4. Single charge cloud's pixel and size distribution of experimental and simulated data. The analysis described Fig. 3 is repeated over 50 speckled patterns with an applied threshold of 0.05 photon, comparable to experimental readout noise level. a. Two models of charge clouds size distribution (Fixed and variable charge cloud size) are compared to experimental data. b. Single photon events's size distribution from Gaussian model Eq.2. The simulation based on a weighted photon size distribution reproduces the experimental findings.

pair of hyper-parameters, the error metric measures the r.m.s distance between the extracted contrasts to the ideal contrasts (dashed magenta line). For error metric values < 0.025 (r.m.s), the extracted contrast is deemed absolute (even though the lowest contrast point (0.1) does not fall within the error bars of the ideal reconstructions).

It is not surprising to find that the optimal radius is actually close to the center of mass of the distribution reported in Fig. 4b. Moreover, it is even closer considering the fitting overestimation described for obtaining Fig. 4b. As the thresholding grows too large (> 0.125 photon), the tails of the charge cloud severely peel off and push the signal distribution towards its center, thus creating a tendency to overestimate the occurrences of multiple photons falling into the same pixel⁵. In addition, to assert that the nominal 16x16 sub-pixel grids provided sufficient resolution of the photon placement, simulations with a larger 32x32 sampling were performed. No change over the accuracy of the contrast extraction was noted. Lowering the resolution further however presumably will affect the results.

Importantly, the degree of average photon count rate per speckle \bar{k} up to which the contrast can be accurately extracted was also investigated. For the sets of simulated data, it is found to be as high as $\bar{k} = 0.25$ photon/speckle, a value that approaches the low photon count limit where the speckle statistics formalism has been shown to operate. This is expedient, as accepting a large bandwidth of scattering intensities coming from the untamed SASE pulses, is tantamount to shortening the time to collect a data set at every delay point. Pushing the \bar{k} envelope further is a task worthy of the LSF optimization and its subtle assignment of the photon positions, however it is not guaranteed, as it suffers a separate set of drawbacks⁵.

To further generalize the reconstruction parameters, the simulations were extrapolated to higher photon energies with 1.68 and 2.24 keV (525, 700 ADU), respectively. Even though experimental data are not available at these energies yet, the improved SNR can in principle only facilitate this exercise using the GGG droplet algorithm. Conversely, analysis at a lowered photon energy is expected to be far more critical. The CCD pixel readout width along with the gain non-uniformity were thus increased up to $\sigma_N \sim 0.065$ photon and $\sigma_r = 0.05$, shown in Fig. 7. The range of optimal thresholding is found to be at the same fraction of the photon energy, $\sim 0.08 + \sigma_N/2$ corresponding to the lowest areas in the error metrics. σ_G is common to all the results with a value of ~ 0.4 . The charge cloud radius distribution is hypothesized to be about the same for all the simulated photon energies.

Conclusion

In conclusion, numerical tools were developed to emulate the very distinctive response of a CCD detector to soft x-ray single photon speckle patterns for ultrafast fluctuations studies, which can produce strong charge sharing. The ensuing obstacles were accounted for with a modified fast droplet algorithm, with its set of parameters enabling a linear and non-biased contrast extraction determined for a few soft x-ray photon energies, simulated within the low photon count limit. For a given readout noise, the optimized parameters are energy invariant, with a threshold starting at $\sim 10\%$ of the photon energy and that slightly increases with the width of the readout noise. The optimal charge cloud radius was found to be located at the adjusted center

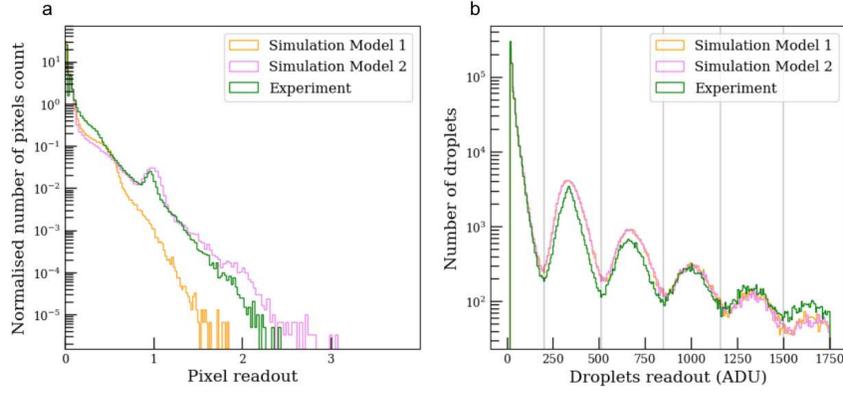


Figure 5. Pixel and droplet histogram for a 1000 experimental and simulated data within same photon average $\bar{k} = [0.025 - 0.025]$ and noise readout. a. CCD pixel histogram. Orange and pink are simulation with a fixed and variable photon's cloud size. b. Droplet histogram in ADU with peaks corresponding to multiples of a single photon. The vertical lines defines the boundaries of the integer number of photons within the droplets, with a maximum of 4 photons per droplet observed.

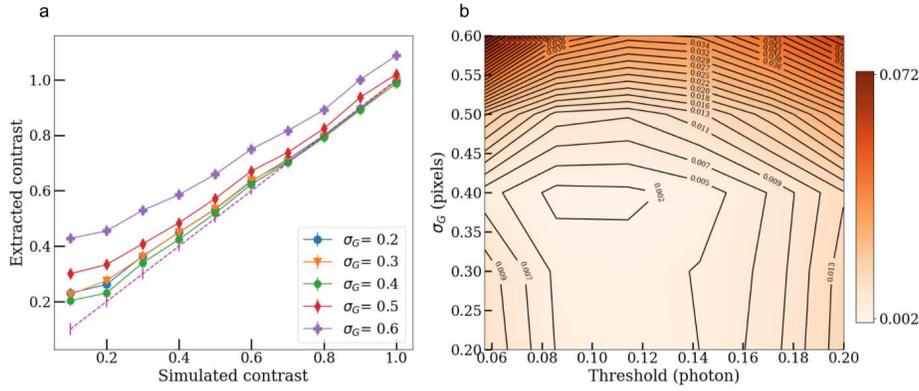


Figure 6. Extracted contrasts from Greedy Guess (GG) droplet algorithm for a 1000 simulated shots. A. Plot at a selected threshold of 0.1 photon with different size of charge cloud used in GG. The violet dashed line indicates the ideal/simulated contrasts with error bars based on the r.m.s value calculated in Fig. 2b. B. Surface error metric plot describing what are the optimal hyper-parameters that minimizes the distance to the ideal contrasts (r.m.s). The optimal $\sigma = 0.4$ is somehow smaller than the average ($\sigma = 0.46$) calculated in Fig. 4b

of mass of its distribution. However, the photon energies were extrapolated up to only twice the experimental photon energy of ~ 1.12 keV due to a lack of data to compare to at a higher energy range, at which point the charge cloud model may start to substantially deviate. Using an advanced Gaussian model for creating the simulations and performing the droplet analysis enables absolute contrast extraction. Notably, the intrinsic pulse energy fluctuations due to the SASE process at the XFEL were considered and shown not to affect the photon determination accuracy up to a large fraction of the low photon limit (where the sample dynamics is not influenced). Furthermore, this motivates the possibility for a real-time analysis of XPFS data at XFELs in the future.

Acknowledgements

We would like to thank Y. Sun for helpful comments with the manuscript. The use of the Linac Coherent Light Source (LCLS), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-76SF00515. We gratefully acknowledge support for work at SIMES through the Department of Energy, Office of Science, Basic Energy Sciences as well, through the Materials Sciences and Engineering Division, also under Contract DE-AC02-76SF00515. J.J Turner acknowledges support from the US Department of Energy, Office of Science, Basic Energy Sciences through the Early Career Research Program.

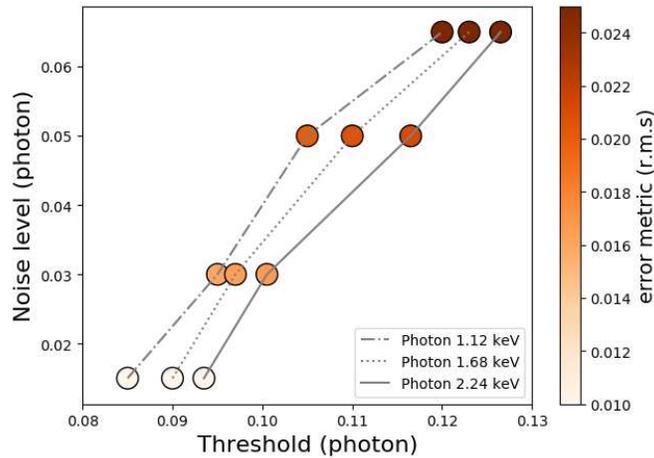


Figure 7. Absolute contrast extraction for different energies (1.12, 1.68, 2.24 keV) and level of noise readouts. Only optimal threshold is reported for the different conditions since its counterpart σ_G is common to all the results with a value of 0.4. The error metrics minimum value are color coded for each of the absolute contrast results.

References

1. Sutton, M. A review of x-ray intensity fluctuation spectroscopy. *C. R. Physique* **9**, 657 (2008).
2. Sinha, S. K., Jiang, Z. & Lurio, L. B. X-ray photon correlation spectroscopy studies of surfaces and thin films. *Adv. Mater.* **26**, 7764–7785, DOI: [10.1002/adma.201401094](https://doi.org/10.1002/adma.201401094) (2014).
3. Seaberg, M. H. *et al.* Nanosecond x-ray photon correlation spectroscopy on magnetic skyrmions. *Phys. Rev. Lett.* **119**, 067403, DOI: [10.1103/PhysRevLett.119.067403](https://doi.org/10.1103/PhysRevLett.119.067403) (2017).
4. Gutt, C. *et al.* Measuring temporal speckle correlations at ultrafast x-ray sources. *Opt. Express* **17**, 55–61 (2009).
5. Sun, Y. *et al.* Characterization of the lcls “nanosecond two-bunch” mode for x-ray speckle visibility spectroscopy experiments. *Proc. SPIE* **10237**, 102370N–102370N–10, DOI: [10.1117/12.2265454](https://doi.org/10.1117/12.2265454) (2017).
6. Lutman, A. A. *et al.* Experimental demonstration of femtosecond two-color x-ray free-electron lasers. *Phys. Rev. Lett.* **110**, 134801 (2013).
7. Decker, F. *et al.* Two Bunches with ns-separation with LCLS. In *Proceedings of FEL2015*, 634–638 (2015).
8. Goodman, J. *Speckle Phenomena in Optics: Theory and Applications* (Roberts & Company, 2007).
9. Hruszkewycz, S. O. *et al.* High contrast x-ray speckle from atomic-scale order in liquids and glasses. *Phys. Rev. Lett.* **109**, 185502 (2012).
10. Bostedt, C. *et al.* Linac coherent light source: The first five years. *Rev. Mod. Phys.* **88**, 015007, DOI: [10.1103/RevModPhys.88.015007](https://doi.org/10.1103/RevModPhys.88.015007) (2016).
11. Lehmkuhler, F. *et al.* Single shot coherence properties of the free-electron laser sacla in the hard x-ray regime. *Sci. Rep.* **4**, DOI: [10.1038/srep05234](https://doi.org/10.1038/srep05234) (2014).
12. Sun, Y., Montana-Lopez, J., Fuoss, P., Sutton, M. & Zhu, D. Accurate contrast determination for X-ray speckle visibility spectroscopy. *J. Synchrotron Radiat.* **27**, 999–1007, DOI: [10.1107/S1600577520006773](https://doi.org/10.1107/S1600577520006773) (2020).
13. Amorese, A. *et al.* Enhanced spatial resolution of commercial soft x-ray ccd detectors by single-photon centroid reconstruction: Technical note. *Nucl. Instruments Methods Phys. Res. Sect. A: Accel. Spectrometers, Detect. Assoc. Equip.* **935**, DOI: [10.1016/j.nima.2019.02.044](https://doi.org/10.1016/j.nima.2019.02.044) (2019).

Figures

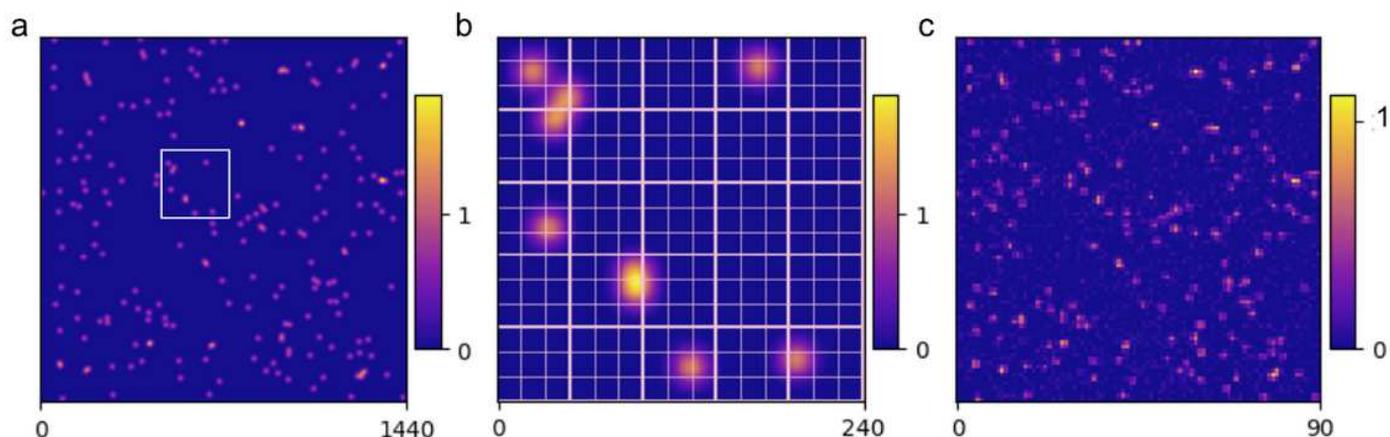


Figure 1

CCD charge sharing model of speckle patterns. a. Sparse distribution of weakly scattered photons on a 3 pixels/speckle basis, generated via the negative binomial distribution (Eq.1) within the range $k = [0:0.25-0:25]$ photons/pixel/speckle. The selected contrast is 0.7. The map is of high-resolution since every pixel is sub-divided by 16x16 to obtain a sub-pixel placement of Gaussian shaped charge clouds (FWHM ~ 1 pixel). b. Enlarged portion of the high resolution photon map (white square box in a), displaying a 5x5 speckles area (bold lines boundaries), each speckle containing 3x3 pixels (light lines boundaries). Photons can sprawl over both pixel and speckle boundaries. c. Final ccd image (or speckle pattern) obtained after binning the high resolution photon map A to a coarser 90x90 grid and adding readout noise ($\sigma_N = 0:05$ photon). A single photon readout is 350 ADU.

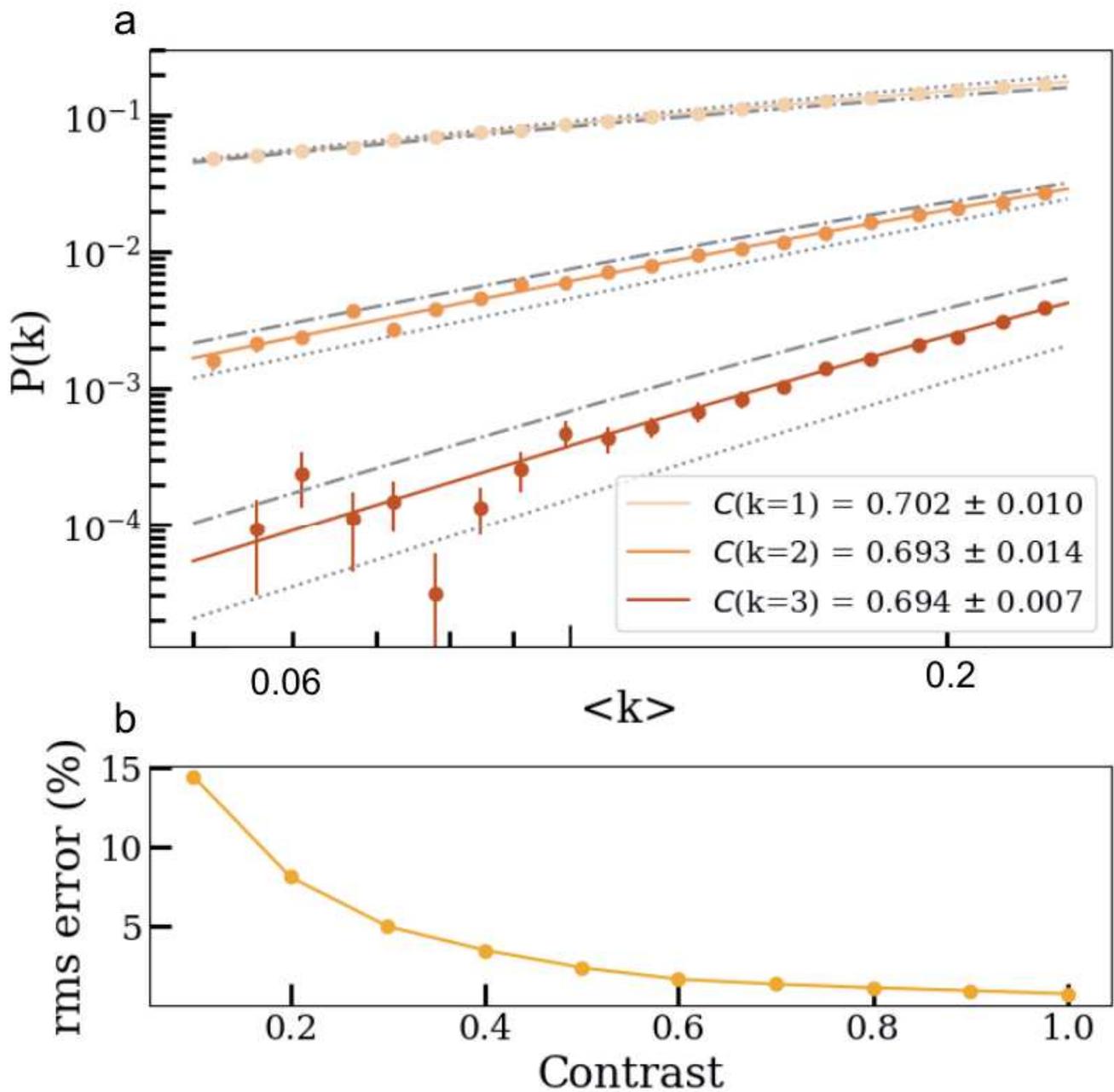


Figure 2

Fitting of observed probabilities $P(k)$ for $k = 1$ to 3 photons/speckle as a function of mean photon density $\langle k \rangle$. a. Fits from a set of 1000 speckled patterns generated within $\langle k \rangle = [0:0.025-0:0.025]$ and contrast $C = 0:7$ with Eq.1. The thin curves represent the contrast boundaries for $C = 0$ and $C = 1$ respectively. Eq.1 is also used for fitting C is calculated from a weighted average of the fitted $k = 1-3$ distribution. The generated and fitted contrasts agree with each other well. b. Plot of the normalized rms error obtained after repeating 50 times the fitting procedure of a. over newly generated batches of speckled patterns.

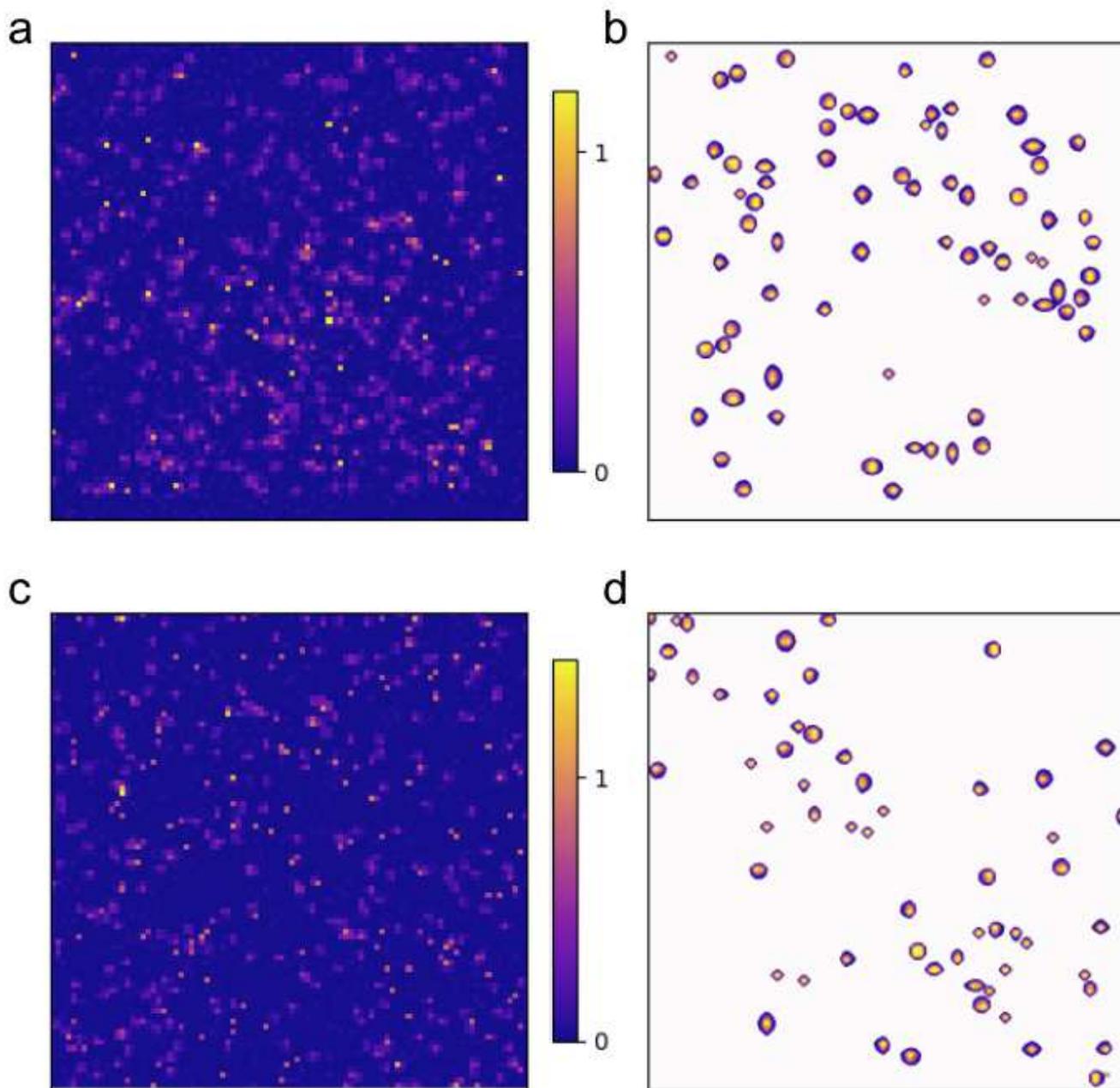


Figure 3

Single charge cloud's size analysis from experimental and simulated data. a. Experimental data. b. The single photon events are isolated from a by selecting connected islands containing a photon readout ($0.5 < \text{photon} < 1.5$) and which shape are fitted with 2d Gaussian shaped cloud model optimization, Eq.2. c. Simulated speckled pattern with variable charge cloud's sizes. The Gaussian fit shows a variability in single photon's size (distribution reported in Fig. 4.) d. Single photon events are isolated and fitted from c.

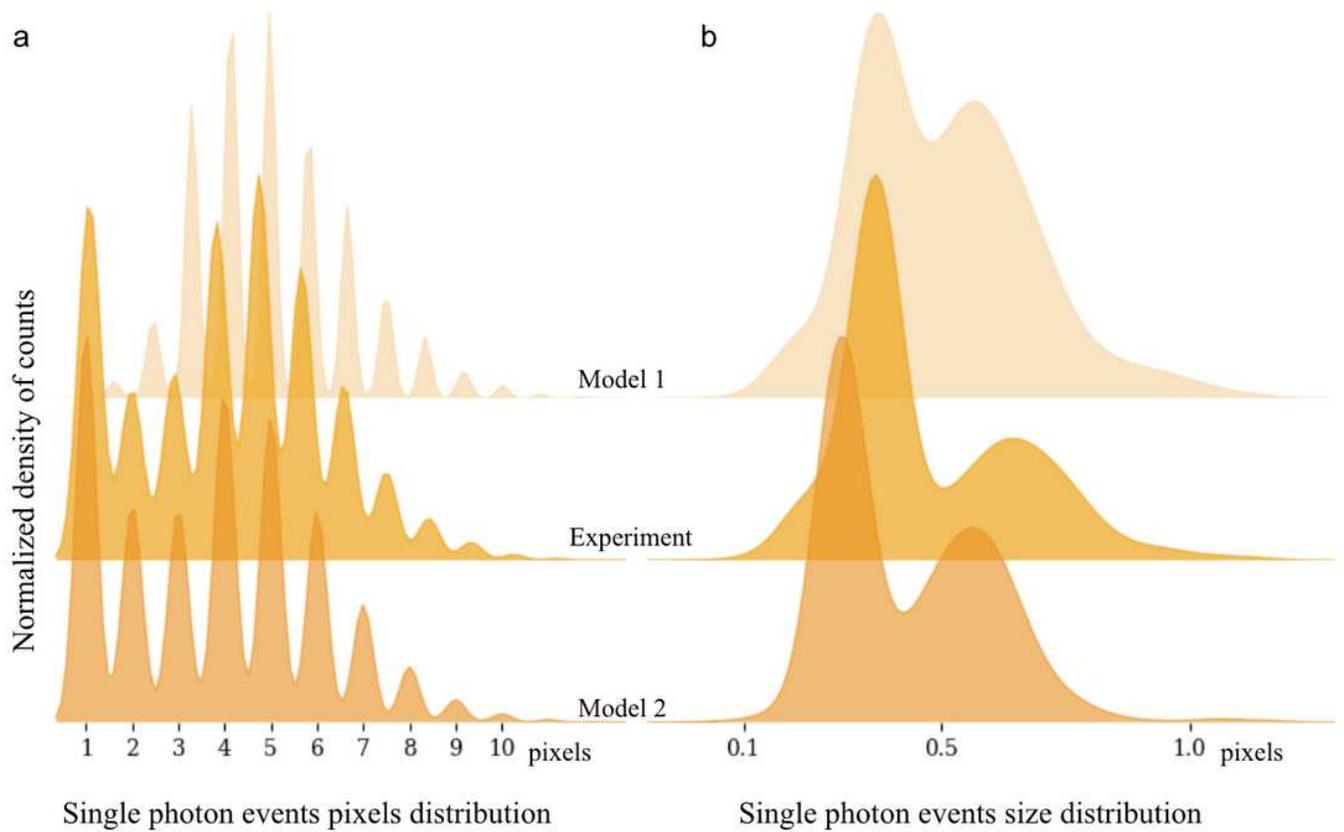


Figure 4

Single charge cloud's pixel and size distribution of experimental and simulated data. The analysis described Fig. 3 is repeated over 50 speckled patterns with an applied threshold of 0.05 photon, comparable to experimental readout noise level. a. Two models of charge clouds size distribution (Fixed and variable charge cloud size) are compared to experimental data. b. Single photon events's size distribution from Gaussian model Eq.2. The simulation based on a weighted photon size distribution reproduces the experimental findings.

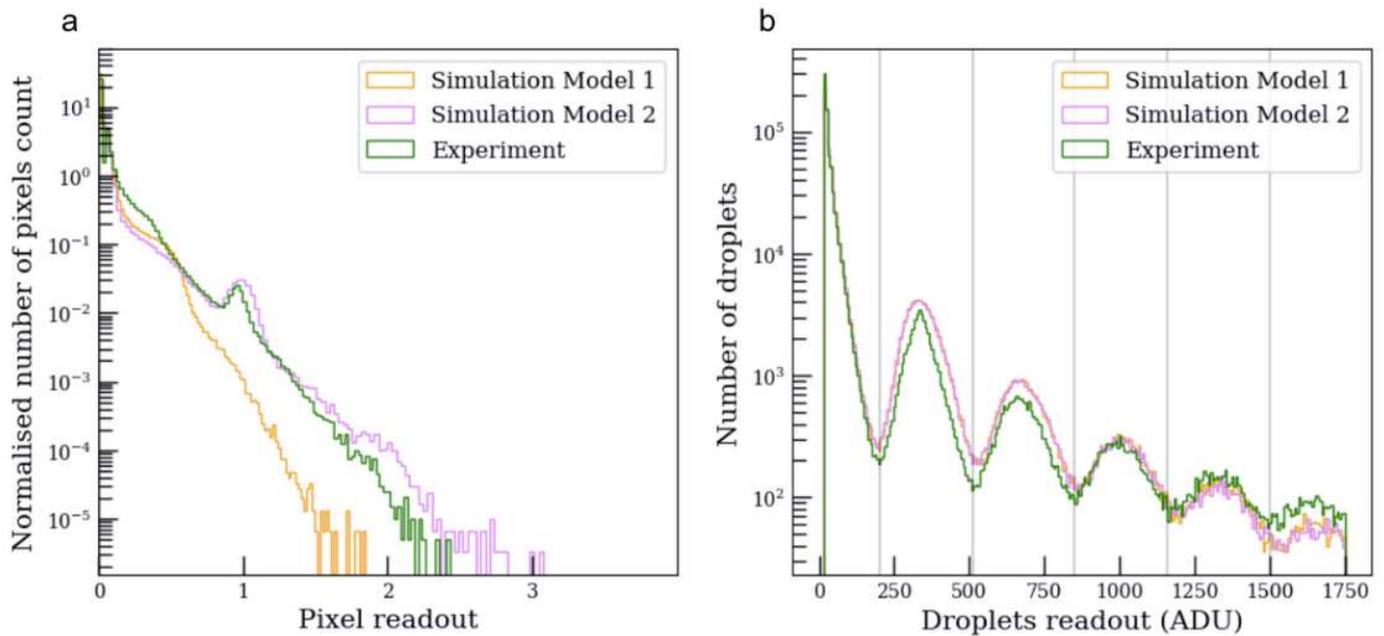


Figure 5

Pixel and droplet histogram for a 1000 experimental and simulated data within same photon average $\bar{k} = [0:0.025-0:0.025]$ and noise readout. a. CCD pixel histogram. Orange and pink are simulation with a fixed and variable photon's cloud size. b. Droplet histogram in ADU with peaks corresponding to multiples of a single photon. The vertical lines defines the boundaries of the integer number of photons within the droplets, with a maximum of 4 photons per droplet observed.

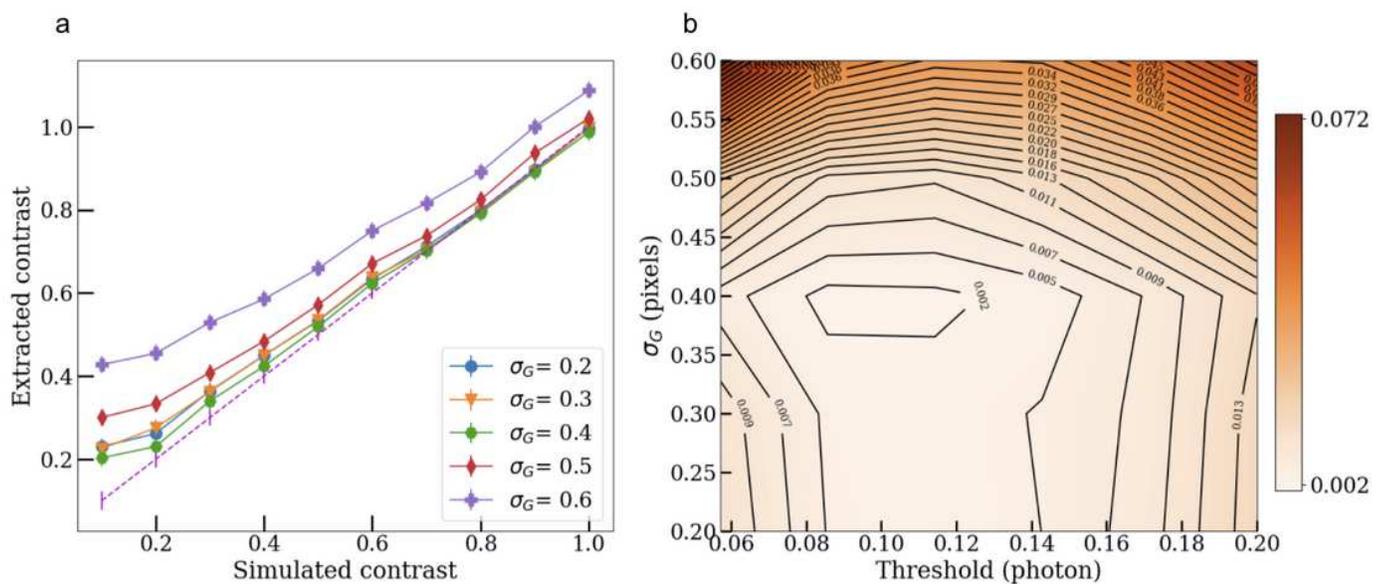


Figure 6

Extracted contrasts from Greedy Guess (GG) droplet algorithm for a 1000 simulated shots. A. Plot at a selected threshold of 0.1 photon with different size of charge cloud used in GG. The violet dashed line

indicates the ideal/simulated contrasts with error bars based on the r.m.s value calculated in Fig. 2b. B. Surface error metric plot describing what are the optimal hyper-parameters that minimizes the distance to the ideal contrasts (r.m.s). The optimal $\sigma = 0.4$ is somehow smaller than the average ($\sigma = 0.46$) calculated in Fig. 4b

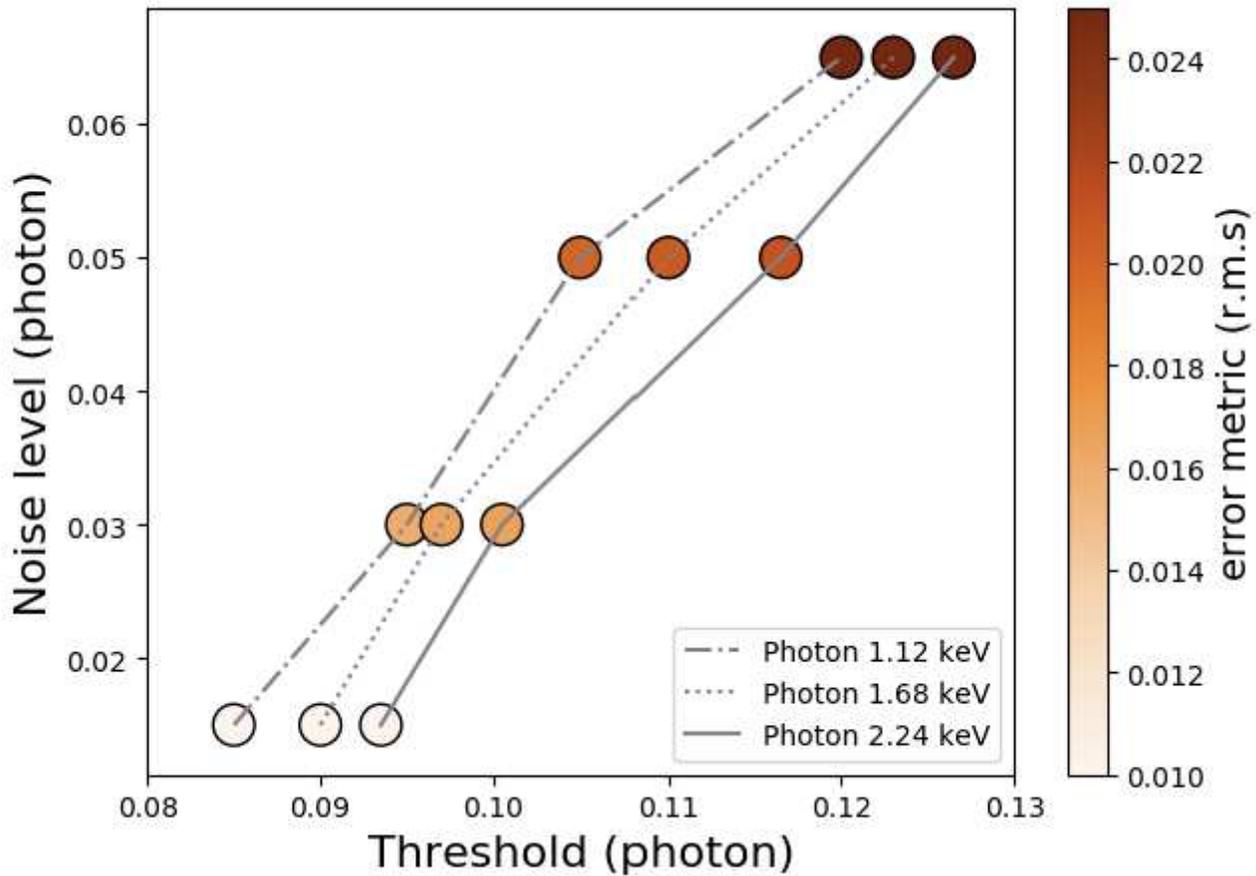


Figure 7

Absolute contrast extraction for different energies (1.12, 1.68, 2.24 keV) and level of noise readouts. Only optimal threshold is reported for the different conditions since its counterpart σG is common to all the results with a value of 0.4. The error metrics minimum value are color coded for each of the absolute contrast results