Optimization of multi-point phase retrieval in edge illumination X-ray imaging

Lorenzo Massimi[1], Tom Partridge[1], Alberto Astolfo[2], Marco Endrizzi[1], Charlotte K. Hagen[1], Peter R. T. Munro[1], David Bate[2], and Alessandro Olivo[1]

[1]Department of Medical Physics and Biomedical Engineering, University College London, Gower St, London WC1E 6BT, UK

[2]Nikon X-Tek Systems Ltd., Tring Business Centre, Icknield Way, Tring, Hertfordshire HP23 4JX, UK

email: l.massimiphd@gmail.com

Abstract

Purpose: In this work, an analytical model describing the noise in the retrieved three contrast channels, transmission, refraction and ultra-small angle scattering, obtained with edge illumination X-ray phase-based imaging system is presented and compared to experimental data.

Methods: In edge illumination, images acquired at different displacements of the pre-sample mask (i.e. different illumination levels referred to as points on the "illumination curve"), followed by pixel-wise curve fitting, are exploited to quantitatively retrieve the three contrast channels. Therefore, the noise in the final image will depend on the error associated with the fitting process. We use a model based on the derivation of the standard error on fitted parameters, which relies on the calculation of the covariance matrix, to estimate the noise and the cross-channel correlation as a function of the position of the sampling points. In particular, we investigated the most common cases of three and five sampling points. In addition, simulations have been used to better understand the role of the integration time for each sampling point. Finally, the model is validated by comparison with the experimental data acquired with an edge illumination setup based on a tungsten rotating anode X-ray source and a photon counting detector.

Results: We found a good match between the predictions of the model and the experimental data. In particular, for the investigated cases, an arrangement of the sampling points leading to minimum noise and cross-channel correlation can be found. Simulations revealed that, given a fixed overall scanning time, its distribution into the smallest possible number of sampling points needed for phase retrieval leads to minimum noise thanks to higher statistics per point.

Conclusions: This work presents an analytical model describing the noise in the edge illumination contrast channels as function of the illumination curve sampling. In particular, an optimal sampling scheme leading to minimum noise has been determined when three or five sampling points are used, which represents the most common acquisition scheme. In addition, the correlation between noise in the different channels and the role of the number of points and exposure time have been also investigated. In general, our results suggest a series of procedures that should be followed in order to optimize the experimental acquisitions.

Introduction

In conventional X-ray imaging, contrast arises from differences in the absorption coefficients, which can be very low when imaging soft tissue specimens, leading to poor signal-to-noise ratio (SNR) and excessive dose. Furthermore, staining protocols are often required. Phase contrast imaging may provide a viable alternative, through the exploitation of phase variations encountered by the X-ray beam when traversing a specimen[1,2]. In particular, since the real part of the complex refractive index ($n = 1 - \delta + i\beta$) is up to three orders of magnitude larger than the absorption one at x-ray energies relevant for medical imaging (above 10 keV), phase imaging can provide greater contrast and better SNR at the same or even reduced dose , especially for high resolution applications[3,4,5]. Currently, the combination of phase contrast and tomography at synchrotron radiation facilities delivers high-contrast images of soft tissues with micron and sub-micron resolution, which allows volumetric quantitative analyses; the ability to do this non-destructively makes the same specimen available for further investigations such as conventional histology[6,7]. Therefore, X-ray phase contrast imaging is becoming increasingly important in the pre-clinical investigation of pathological conditions[8,9,10]. The limited access to synchrotrons currently represents the main limit to the widespread application of this technique. For this reason, phase imaging techniques based on conventional X-ray sources have been developed, which are typically based on the use of optical elements such as absorption and phase gratings[11,12]. In addition to transmission and phase imaging, these techniques provide access to the ultra-small angle scattering (or dark-field) signal, which has proven to be useful both for material and medical imaging applications[13,14]. Edge illumination is one of these techniques. It is based on the use of two absorption masks to shape the beam into a series of beamlets, and detect a change in their propagation direction as a consequence of refraction[12,15]. Changes in the beamlets' width and intensity are a consequence of ultra-small angle scattering and transmission, respectively. The relatively simple implementation and versatility in terms of scanning modes and accessible spatial resolution levels make edge illumination a promising phase detection scheme for clinical applications allowing also single-image retrival approaches[16,17,18]. The phase sensitivity is fully described by the illumination curve (IC), which expresses the quantitative relationship between the change in beamlet direction and recorded change of intensity on the detector[19]. It is usually Gaussian shaped, and represents the basis for the phase retrieval which is performed by quantifying the perturbation that the IC undergoes when a sample is placed into the beam path[20]. Quantification is usually achieved by a pixel-wise Gaussian fit of the intensity values obtained by displacing one of the masks in a series of positions, with and without the sample. Since phase retrieval relies on curve fitting, the noise in the final image will depend on the error associated with the fitting process. In this work, we present an analysis of the noise in the retrieved transmission, refraction and dark field contrast channels obtained with an edge illumination setup when using multi-point retrieval. A theoretical model based on the derivation of the standard error on fitted parameters is developed and compared to the experimental data, showing a very good agreement. Different experimental conditions, involving the acquisition of three or five input images, have been considered, and in each case the positions of the IC sampling points leading to the minimum noise and cross-channel correlation have been determined for each contrast channel. We also used simulations to investigate whether it is preferable to distribute the same overall statistics in more or fewer sampling points, a typical question when the overall acquisition time is limited. The cross-channel correlation has also been investigated. This work will help determine the acquisition scheme in a multi-point scan that optimises the subsequent phase retrieval. In addition, it presents a noise model that can be adapted to different experimental techniques based on curve fitting, as well as to different IC shapes.

Materials and Methods

Edge illumination

Edge illumination is a phase gradient method particularly well suited to laboratory applications since it is achromatic and does not require a coherent X-ray beam[15,21,22]. A schematic view of an edge illumination system is shown in Fig.1(a). This method is based on the use of two absorption gratings, usually referred to as masks. The first (sample) mask is placed before the sample, and splits the main X-ray beam into a series of beamlets. The second (detector) mask is positioned in front of the detector so as to intercept a portion of each beamlet. When an object is inserted into the beam path, refraction causes a shift of the beamlets away from or towards the corresponding aperture in the detector mask, leading to a change in the recorded intensity. A quantitative relationship exists between the recorded in-

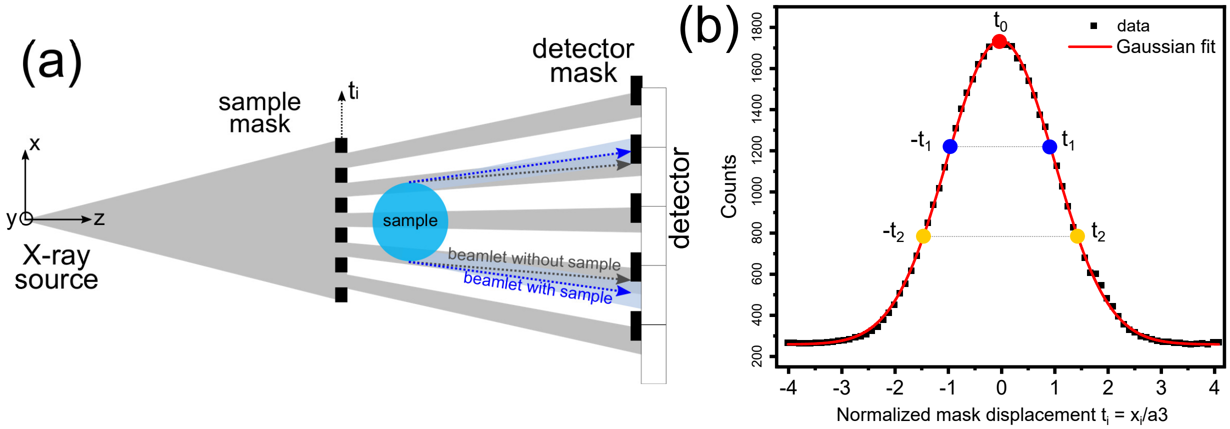

Figure 1: Panel (a) shows a schematic view of a typical edge illumination system. Panel (b) reports the illumination curve obtained with the specific edge illumination system used and a fit using a Gaussian model. The investigated arrangement of the sampling points is also indicated.

tensity change and the refraction angle; this is expressed through the IC, which characterizes the phase sensitivity of an edge illumination system. It can be measured experimentally by moving the masks relative to each other each other and recording the transmitted intensity; usually the sample mask is scanned, while the detector mask is kept still, see Fig.1(a). It can be expressed mathematically as:

$$IC(x) = (A_1 * S * A_2)(x) \tag{1}$$

where A_1 and A_2 are the sample and detector mask transmission functions, and S is the source shape projected onto the detector plane. The $*$ symbol denotes the convolution operator. Since the focal spot is usually Gaussian shaped, the IC is well described by a Gaussian function as shown by the fitting of a real dataset in Fig.1(b). The IC is also the starting point for a quantitative phase retrieval algorithm since transmission, refraction and ultra small-angle scattering have different effects on the curve. Specifically, transmission reduces the intensity of each beamlet depending on the imaginary part of the refractive index β, and refraction shifts each beamlet according to the first derivative of the phase shift with respect to the transverse coordinate x[20]. In addition, ultra small-angle scattering is responsible for a change in the width of the IC[23]. Combining the effects of these three processes on a beamlet, the intensity recorded by each pixel at position x of detector column y for a relative mask position t can be expressed as:

$$I(x, y, t) = T(x, y)[O(x, y) * IC(x, y, t - \Delta t)] + d \tag{2}$$

where $T(x, y)$ is the sample transmission function and $IC(x, y, t - \Delta t)$ is the illumination curve, shifted because of refraction and convolved with the object scattering function O. An offset d has been introduced to take into account that usually the IC does not go to zero because of residual beam transmission through the masks. The quantities $T(x, y)$ and Δx are quantitatively related to the imaginary and unit decrement of the real part of the refractive index (n=1-δ+iβ), respectively. In particular, $T(x, y) = e^{-\int \mu(x,y,z)dz}$, where $\mu = (4\pi/\lambda)\beta$ and λ is the wavelength of the incident radiation and $\Delta t \sim z_{od} \nabla_x \int \delta(x, y, z)dz$, where z_{od} is the sample to detector distance and ∇_x is the gradient in the sample mask plane and perpendicular to the direction of the apertures. Assuming a Gaussian approximation for both the IC and the scattering function, Eq.2 can be written as:

$$I(t) = TA_{IC}\, e^{\frac{-(t-\Delta t)^2}{(2\sigma^2)}} + d \tag{3}$$

where the x, y dependency has been neglected for simplicity and A_{IC} corresponds to the maximum value of the IC without the offset, and the scattering function $O(x, y)$ has been described by a Gaussian with unit amplitude and same centre as the IC, and width σ_O[20,23]. Therefore, $\sigma = \sqrt{\sigma_{IC}^2 + \sigma_O^2}$, where σ_{IC} is the width parameter of the IC without the object. The acquisition of at least three images at different IC positions allows to solve eq.3 for each detector pixel and retrieve $T(x, y)$, Δx and σ (and therefore σ_O) which are related to physical properties of the investigated sample. A straightforward way to proceed is to perform pixel-wise Gaussian curve fitting of the form:

$$G(t, \mathbf{a}_{f,s}) = a_{1\,f,s} e^{\frac{-(t-a_{2\,f,s})^2}{(2a_{3\,f,s}^2)}} + d_{f,s} \tag{4}$$

where the subscripts f, s refer to the fit parameters obtained without and with the sample, respectively. If the offset d exists, it can be assumed a-priori or determined by on an iterative basis[20]. The extracted fit parameters are related to the physical quantities in eq.3 by:

$$T = a_{1\,s}/a_{1\,f} \qquad \Delta x = a_{2\,s} - a_{2\,f} \qquad \sigma_O^2 = a_{3\,s}^2 - a_{3\,f}^2 \tag{5}$$

where all relations apply on a pixel-wise basis. If the offset d is assumed to be the same with and without the object, it is sufficient to sample the IC at three positions to fit the $G(x, \mathbf{a}_{f,s})$ model to the experimental data. If this assumption is violated, as in the case of a sample causing non-negligible beam hardening, the offset must be taken into account as a parameter in the fit model and at least four sampling point are needed. The Gaussian model of eq.4 is referred to as "non-normalized". Similarly, a normalized model can be defined by dividing the Gaussian by its area:

$$G(t, \mathbf{a}_{f,s}) = \frac{a_{1\,f,s}}{\sqrt{2\pi}a_{3\,f,s}} e^{\frac{-(t-a_{2\,f,s})^2}{(2a_{3\,f,s}^2)}} + d_{f,s} \tag{6}$$

While there is no physical reason to prefer one of these formulations, as we will show in the Results section, the normalization factor has an impact in terms of correlation between the extracted parameters.

Noise model

As indicated by eq.5, the retrieved physical sample parametersare obtained from the pixel-wise estimated Gaussian fit parameters. The conventional way to find the optimal set of fit parameters is by minimization of the normalized residuals, which means finding the set of parameters $\mathbf{a^0}$ that solve the system of equations:

$$\frac{\partial R}{\partial a_j} = \frac{\partial}{\partial a_j} \sum_i^n \frac{1}{\sigma_i^2} \left[y_i - f(x_i, a_j^0) \right]^2 = 0 \tag{7}$$

where f is the fitted model and j and i run from 1 to the number of model parameters m and to the number n of experimental data points y_i, respectively. σ_i is the uncertainty on each of the measured y_i points.To solve this equation, a set of parameters $\mathbf{a}^k$ is chosen at the beginning as initial guess. In the most used fitting algorithms, assuming the chosen $\mathbf{a}^k$ is reasonably close to $\mathbf{a}^0$, the function f can be linearized by Taylor expansion[24]:

$$f(x_i, \mathbf{a}^0) \sim f(x_i, \mathbf{a}^k) + \sum_j^m \frac{\partial f(x_i, a_j)}{\partial a_j}(a_j^0 - a_j^k) = f(x_i, \mathbf{a}^k) + \sum_j^m J_{ij} \Delta a_j \tag{8}$$

where the index k is indicating the iteration number, $\mathbf{J_{ij}}$ are the elements of the Jacobian matrix $\mathbf{J}$ of f, and Δa_j are the distances between the set of parameters at iteration k and the target one minimizing the residuals. Substituting the eq.8 into eq.5 and rearranging the terms[24,25], we obtain the matrix equation:

$$(\mathbf{J}^T \mathbf{W} \mathbf{J}) \mathbf{\Delta a} = \mathbf{J}^T \mathbf{W} \mathbf{\Delta y} \tag{9}$$

where $\Delta y_i = y_i - f(x_i, a_j^k)$ and $\mathbf{W}$ is an $n \times n$ matrix with entries $1/\sigma_i^2$ for each of the n experimental points along the diagonal. From eq.9 the distance of the parameters set at iteration k from the target one can be obtained as:

$$\mathbf{\Delta a} = (\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1}\mathbf{J}^T\mathbf{W}\mathbf{\Delta y} = \mathbf{C}\mathbf{J}^T\mathbf{W}\mathbf{\Delta y} \tag{10}$$

where $\mathbf{C}$ is a symmetric $m \times m$ matrix defined as $\mathbf{C} = \mathbf{H}^{-1} = (\mathbf{J}^T\mathbf{W}\mathbf{J})^{-1}$, which is referred to as the covariance matrix. It is worth noting that, in the linear least square case, eq.10 represents the exact solution, while in the non-linear case discussed here it represents the distance of the current parameters set from the target one. Therefore, the parameters set can now be updated as $\mathbf{a}^{k+1} = \mathbf{a}^k + \mathbf{\Delta a}$ and the entire process is repeated until some convergence criteria are met. Eq.10 provides also the basis to calculate the error on the fitted parameters, which is indicated by δa_j and can be written as:

$$\delta a_j = \sum_i^n \frac{\partial a_j}{\partial y_i}\delta y_i \tag{11}$$

The calculation of the derivatives by means of eq.10 leads to the matrix equation:

$$\delta\mathbf{a} = \mathbf{C}\mathbf{J}^T\mathbf{W}\delta\mathbf{y} \tag{12}$$

It is now possible to calculate the variance and the covariance for the variables $\mathbf{a}$ as:

$$\sigma_{\mathbf{a}}^2 = \langle\delta\mathbf{a}\delta\mathbf{a}^T\rangle = \langle\mathbf{C}\mathbf{J}^T\mathbf{W}\delta\mathbf{y}\delta\mathbf{y}^T\mathbf{W}\mathbf{J}\mathbf{C}^T\rangle = \mathbf{C}\mathbf{J}^T\mathbf{W}\langle\delta\mathbf{y}\delta\mathbf{y}^T\rangle\mathbf{W}\mathbf{J}\mathbf{C}^T \tag{13}$$

where $\langle . \rangle$ denotes the average over the errors on the experimental data points. Since these can be assumed to be uncorrelated, the covariance $\langle\delta y_i \delta y_j\rangle$ is always zero except when $i = j$, which represents the variance σ_i^2. Therefore, $\mathbf{W}\langle\delta\mathbf{y}\delta\mathbf{y}^T\rangle$ is the identity matrix, and eq.13 becomes:

$$\sigma_{\mathbf{a}}^2 = \mathbf{C}(\mathbf{J}^T\mathbf{W}\mathbf{J})\mathbf{C}^T = \mathbf{C} \tag{14}$$

which shows that the diagonal elements of $\mathbf{C}$ represent the variance on the fitted parameters, while the off-diagonal terms are their covariances, i.e.:

$$\sigma_{a_j}^2 = C_{jj} \qquad \sigma_{a_j - a_k}^2 = C_{jk} \tag{15}$$

This result shows that, in general, the errors on the coefficients are correlated, which means that the off-diagonal terms in the covariance matrix do not vanish. Eq.15 is the starting point for the noise analysis performed in this work. Following the definition of $\mathbf{C}$, the elements of the covariance matrix can be obtained from the inversion of $\mathbf{H}$, whose elements are of the form:

$$h_{jk} = \sum_{i=1}^n \frac{1}{\sigma_i^2}\frac{\partial f(x_i, \mathbf{a})}{\partial a_j}\frac{\partial f(x_i, \mathbf{a})}{\partial a_k} \tag{16}$$

Three-point retrieval

In order to proceed further, we assume that f is well represented by a Gaussian model, which justifies the use of eq.5 to fit the experimental intensity distribution described by eq.3. The offset is assumed not to vary following the introduction of the sample, so that only three IC sampling points are needed. We also assume that the uncertainty σ_i on the measured value y_i is a function of the value of the point itself and of the set of parameters $\mathbf{a}$, i.e. $\sigma_i = \sigma(x_i, \mathbf{a})$, that from an experimental point of view corresponds to use the standard deviation of a series of measurements obtained from a Poissonian distribution as a noisy estimate of the true noise value. Under these assumptions $\mathbf{H}$ becomes a 3×3 matrix the independent elements of which are:

$$\begin{aligned} h_{11} &= \sum_{i=1}^{n} \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \quad & h_{12} &= \frac{a_1}{a_3} \sum_{i=1}^{n} t_i \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \quad & h_{13} &= \frac{a_1}{a_3} \sum_{i=1}^{n} t_i^2 \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \\ h_{22} &= \sum_{i=1}^{n} t_i^2 \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \quad & h_{23} &= \frac{a_1^2}{a_3^2} \sum_{i=1}^{n} t_i^3 \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \quad & h_{33} &= \frac{a_1^2}{a_3^2} \sum_{i=1}^{n} t_i^4 \frac{e^{-t_i^2}}{\sigma(x_i,\mathbf{a})} \end{aligned} \tag{17}$$

where $t_i = (x_i - a_2)/a_3$. In general, all the h_{ij} elements are different from zero. However, without loss of generality, the origin of the x-axis can be set in a_2 (i.e. $a_2 = 0$), and the n data points can be assumed to be symmetrically arranged on either side of the IC peak (see plot in Fig.1(b)) which represents the most typical acquisition scheme for edge illumination. With this choice, all the elements with an odd power of t vanish, namely $h_{12} = h_{23} = 0$. Moreover, $\mathbf{H}$ is a symmetric-defined matrix, therefore:

$$\mathbf{H} = \begin{pmatrix} h_{11} & 0 & h_{13} \\ 0 & h_{22} & 0 \\ h_{13} & 0 & h_{33} \end{pmatrix} \tag{18}$$

which can be analytically inverted, leading to the following expressions for the variances of the fitted Gaussian amplitude, centre and width:

$$\sigma_{a_1}^2 = C_{11} = \frac{h_{33}}{h_{11}h_{33} - h_{13}^2} \qquad \sigma_{a_2}^2 = C_{22} = \frac{1}{h_{22}} \qquad \sigma_{a_3}^2 = C_{33} = \frac{h_{11}}{h_{11}h_{33} - h_{13}^2} \tag{19}$$

while only the off-diagonal element $C_{13} = C_{31}$ are different from zero and equal to:

$$\sigma_{a_1-a_3}^2 = C_{13} = -\frac{h_{13}}{h_{11}h_{33} - h_{13}^2} \tag{20}$$

indicating that a degree of correlation exists between the fitted amplitude and width. In order to obtain a theoretical model which can be directly compared to the experimental data regardless of the sample, we further restrict the analysis to noise in the background, the reduction of which will boost the SNR of the image[5]. This allows to assume the same expected value for the fitted parameters with and without the sample and constant offset. Therefore, the error on the transmission, refraction and dark-field contrast channels can be obtained from the variance on fitted parameters simply by applying error propagation to eq.5:

$$\sigma^2_{Transmission} = 2\,\frac{\sigma^2_{a_1}}{a_1^2} \qquad \sigma^2_{Refraction} = 2\,\sigma^2_{a_2} \qquad \sigma^2_{Dark-field} = 2\,\sigma^2_{a_3} \tag{21}$$

Assuming that the three data points are of the form $(-x_1, 0, x_1)$ where as said 0 corresponds to the peak of the IC, and a Poisson-like noise of the form $\sigma(x_i, \mathbf{a}) = \sqrt{A\,G(x_i, \mathbf{a}) + B}$, which agrees with the behaviour of the normalized data in use (see supplementary materials), the h_{ij} elements in eq.17 can be calculated analytically leading to the following expressions for the noise in the background of each contrast channel:

$$\begin{gathered} \sigma^2_{Transmission} = \frac{2}{a_1^2}[A(a_1 + d) + B] \qquad\qquad \sigma^2_{Refraction} = \frac{a_3^4}{x_1^2 a_1^2} e^{t_i^2} \left(A a_1 e^{-t_i^2/2} + Ad + B \right) \\ \sigma^2_{Darkfield} = \frac{a_3^6}{x_1^4 a_1^2} \left[\left(2 + e^{t_i^2}\right)(B + 2dA) + A a_1 \left(2 + e^{t_i^2/2}\right) \right] \end{gathered} \tag{22}$$

As well as on the system parameters (through the shape of the IC, which is a function of the source and mask parameters), the noise in the refraction and dark-field channels depends on the sampling points, suggesting that the noise in the retrieved images can be optimized by a careful choice of their position. Remarkably, the noise in the transmission channel is found not to depend on the IC width nor to the sampling points position. It mainly depends on the term $a_1 + d$ which represents the top value of IC, confirming the empirical observation that the only way to improve the conventional transmission contrast is to increase the photon statistic[16] It is worth noting that the noise variances can be obtained in the case of a perfect Poissonian response setting $A = 1$ and $B = 0$. On the other hands, if the noise of the system is characterized by a Gaussian distributed noise which is independent from the position of the sampling points, which can be the case of the dark current for an integrating detector, the same analysis can be repeated simply this new component to the definition of σ, obtaining new equations for the noise in each channel.

Five-point retrievals

In the previous section, restricting the analysis to the background noise allowed to assume that the offset in the fitting model did not vary with the introduction of the sample, and therefore that only three sampling points were required for the retrieval. However, four or more sampling points are needed when imaging specimens for which the offset cannot be assumed constant, such as those causing a significant degree of beam hardening. In this case, we will use five sampling points, which is compatible with our approach where a point on the top of the IC is accompanied by an equal number of additional points placed symmetrically on each side, and leads to a more robust retrieval of the four sample parameters. Assuming again sampling points symmetric with respect to the IC top, i.e. $(-x_2, -x_1, 0, x_1, x_2)$ and $a_2 = 0$, the matrix $\mathbf{H}$ becomes a 4×4 matrix of the form:

$$\mathbf{H} = \begin{pmatrix} h_{11} & 0 & h_{13} & h_{14} \\ 0 & h_{22} & 0 & 0 \\ h_{13} & 0 & h_{33} & h_{34} \\ h_{14} & 0 & h_{34} & h_{44} \end{pmatrix} \tag{23}$$

which can be inverted to obtain $\mathbf{C}$. From $\mathbf{C}$ the same considerations leading to eq.22 can be followed to obtain explicit expressions for the background noise of the three contrast channels and their correlations, which are again a function of the fitted IC parameters $\mathbf{a}$ and of the positions of the sampling points. Explicit expressions are reported in the supplementary materials due to their increased length and complexity.

Data acquisition

The imaging system uses a tungsten anode COMET MXR-160HP/11 x-ray source (Comet, Wünnewil-Flamatt, Switzerland), which, for collecting these data, was operated at 90 kVp and 7.7 mA with a nominal focal spot of 0.4 mm. To increase the intensity of the phase signal, this was reduced to 70 μm along the x direction with a Huber slit (Huber Diffraktionstechnik GmbH & Co. KG, Rimsting, Germany) placed against the output window. The detector is a single photon counting Cd-Te CMOS (XCounter XC-FLITE FX2, Direct Conversion, Danderyd, Sweden) with 2048×128 square pixels 100 μm in side, placed at approximately 2.1 m from the x-ray source. Pre-sample and detector masks were placed at 1.60 m and 2.06 m from the source, respectively. The detector mask was 20 cm tall and featured 28 μm

wide apertures (one per detector pixel), with a regular period of 98 μm. The pre-sample mask was 15 cm tall and featured 21 μm wide apertures, with a regular 75 μm period, offset along the z-axis, by -22 mm to create a Moire fringe pattern at the detector. The masks were fabricated by electroplating approximately 300 μm of gold on a 1 mm thick graphite substrate by Microworks GmbH (Karlsruhe, Germany). They were mounted on pairs of linear translators for movement along and across the optical axis (Newport, Irvine, CA), and on a cradle for rotation around the optical axis (Kohzu, Kawasaki, Japan). This system allows producing an IC spanning over many pixels, and tuning the number of sampling points by changing the pre-sample mask offset along the z-axis, and therefore the size of the Moire fringe. Since the noise analysis was restricted to the background, 50 flat-field images of 1 second exposure each have been used for the data analysis. The image size was 712 × 78 pixels, where the number of rows and columns correspond to independent IC profiles and to the number of IC sampling points, respectively. In this case, the latter is high enough to allow for an efficient sub-sampling (see supplementary materials).

Data analysis

As a preliminary step, all the 712 × 50 IC profiles were fitted with the Gaussian model reported in eq.5 to obtain the average amplitude, centre and offset required for theoretical model. The results are reported in table 1 in terms of mean value and standard deviation, where $a_{1,2,3,4}$ are the IC's amplitude, centre, width and offset, respectively. The x-axis for

Parameter	units	mean value	standard deviation
a_1	number of photons	1700	19
a_2	mm	$6 \cdot 10^{-5}$	$1 \cdot 10^{-3}$
a_3	mm	0.1	$1 \cdot 10^{-3}$
a_4	number of photons	300	6

Table 1: Average values of the experimental Gaussian profiles

the fit has been chosen symmetrically with respect to the top of the IC, in accordance to the $a_2 = 0$ assumption used to develop the theoretical model. To eliminate the effects of masks and beam inhomogeneities, the average of 35 images has been used as a flat-field to normalize the remaining 15 images, which have been used for the actual data analysis.

The noise of the so obtained data are described by a Poisson-like distribution of the form $\sqrt{A\,N + B}$, where N is the number of photons and $A = 1.15$, $B = 20$ (see supplementary materials). In order to preserve the average value of each column after normalization, these have been multiplied by the average counts of the corresponding columns in the flat-field; as a reminder, because of the moiré-style acquisition, columns correspond to different points on the IC (see above). For this reason, images have been rearranged in a set where the n-th image corresponds to the n-th column of all the 15 original images (see supplementary material for details). Therefore, the new dataset consists of a number of images equal to the number of columns in the original images, each corresponding to a subsequent point on the IC as would be acquired in a standard multi-point EI acquisition. This rich dataset with many available IC points allowed selecting multiple combinations corresponding to both three and five sampling point acquisitions. In the first case they are of the form ($-x_1$, 0, x_1) where x_1 is the distance from the sampling point located at 0 (the top of the IC). In the latter, they are of the form ($-x_2$,$-x_1$, 0, x_1, x_2), where both x_1 and x_2 are varied (with $x_2 \geq x_1$). According to the number and arrangement of the investigated sampling points, the corresponding "re-arranged" images are selected, and a pixel-wise weighted non-linear least square fit applied, with weights equal to the inverse of the value of the fitting function in agreement to eq.7. The fit was performed by means of Matlab® *lsqcurvefit* using default parameters. The result is a series of about 712×15 values for amplitude, centre and width. Two subsets of size 50×50 have been extracted at random for each parameter, and the corresponding transmission, refraction and dark-field signals have been calculated according to eq.5. This random extraction of 50×50 subsets for each parameter was repeated 100 times, and the standard deviation for the three contrast channels recorded each time. The final noise value was obtained as the average of all these standard deviation values, and its uncertainty as their standard deviation. A similar approach is used to calculate the covariance according to its definition:

$$cov(X, Y) = E[(X - E[X])(Y - E[Y])] \tag{24}$$

where E denotes the expected value, and X and Y are two random distributed variables. By definition the covariance is not bounded and its units change according to the meaning of X and Y. Therefore, to enable comparing different contrast channels, the correlation has been used instead, which is defined as:

$$corr(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}} \tag{25}$$

where var indicates the variance. Unlike the covariance, the correlation has no dimension and is limited to the interval $[-1,1]$ where 1 and -1 indicate total positive or negative correlations.

Finally, to investigate also the relationship between background noise and exposure time per point, Monte Carlo simulations have been used. In order to reduce complexity when increasing the number of points they have been evenly distributed in the $\pm 3a_3$ range. For each investigated number of points $\mathbf{X} = (x_1...x_n)$ with $n = (3,5,7,9,11)$, the intensity values $G(\mathbf{X}, \mathbf{a})$ were extracted, where $\mathbf{a}$ is the set of parameters reported in table 1 describing the experimental IC parameters. A Poisson-like noise like the one observed in the normalized data (see supplementary materials) was added. The set of intensity values obtained in this way was fitted as previously described, considering a constant offset. This procedure has been repeated 10^6 times, resulting in a large set of values for Gaussian amplitude, centre and width. In order to simulate the retrieval of the individual contrast channels, a subset of 800 values was randomly chosen. Half of the values were used as a sample images and half as flat-field images, and the different contrast channels calculated by pixel-wise application of eq.5, assuming the standard deviation over the 400 values as the noise. Finally, the average value of the standard deviation over 10^3 repetitions was considered as the final noise value.

Results

In Fig.2 the noise values expected for the three retrieved contrast channels are compared to the corresponding experimental values for the three points retrieval case. The sampling points are of the form $t_0 = 0$ and $t_1 = \pm x_1/a_3$, so that the entire triplet is determined by varying a single parameter. Examples of the experimental retrieved images are also shown, with the choice of the three sampling points shown as red dots in the inset. A very good agreement between predicted and experimental values is observed for all three channels. In particular, the error on the retrieved transmission, see Fig.2(a) and(d), is confirmed to be independent from the choice of the sampling points, in agreement with eq.22 which shows that it depends only on the top value of the IC. Therefore, the SNR in the transmission

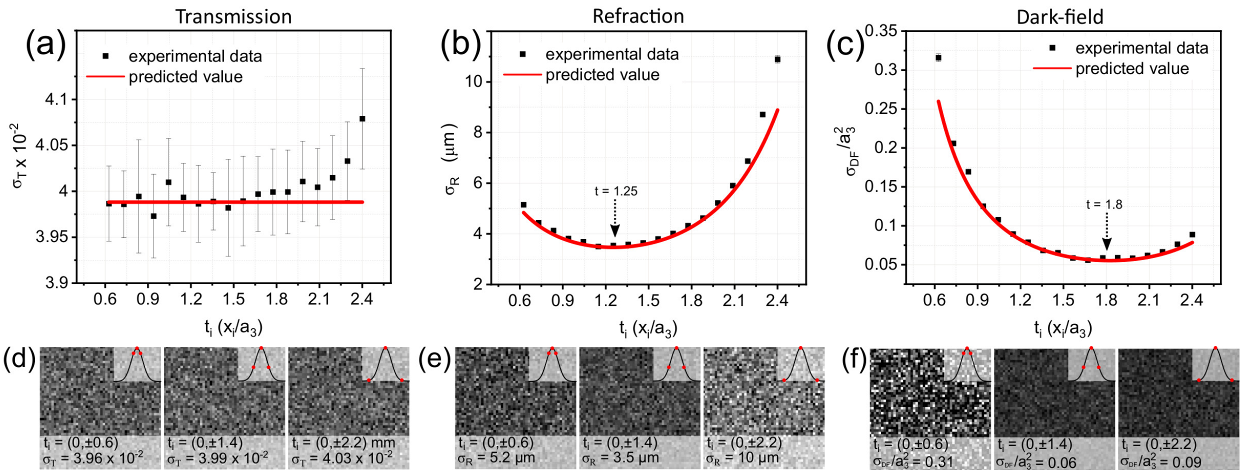

Figure 2: Panel (a) to (c) show the comparison between the experimental values for the standard deviation in the background (black dots) and the values predicted by the model (red lines) for the transmission, refraction and dark-field contrast channels, respectively. Panels (d) to (f) show example images corresponding to the retrieved contrast channels at different positions of the sampling points. Images are shown on the same gray level scale for each contrast channels. Arrows in panels (b) and (c) point at the minimum of the curve for the refraction and dark-field channels.

contrast can be improved only by increasing the X-ray flux. On the other hand, noise in the refraction and dark-field channels has a more complex behaviour. In both cases a minimum can be found, located at $t = 1.25$ and $t = 1.8$ for refraction and dark-field, respectively, as shown in Fig.2(b) and (c). In refraction, noise is minimised by choosing the off-centre frames further away from the IC's maximum slope position ($t = 1$), for which phase sensitivity is highest when acquiring two points only (see Fig.2(e))[26]. Moving down the IC beyond the $t = 1.25$ point, noise starts to rise fast due to the increased uncertainty on the IC centre estimation. This can be explained by an overall reduced phase sensitivity, since only sampling points for which the IC derivative is approaching zero (top and tails) are now being considered[27]. For the same reason, an increase in the noise is found when the two additional sampling points approach the top of the IC. In the dark-field channel, the minimum is located further away from the IC's maximum slope position, at approximately twice the standard deviation, approaching the tails of the curve as show in Fig.2(c) and (f). In this case the noise rises rapidly when the sampling points approach the top, since the fit becomes less sensitive to changes in width. Finally, a deviation from the model starts to be visible when t_i approaches the tails of the curve, which can be explained by a non perfect

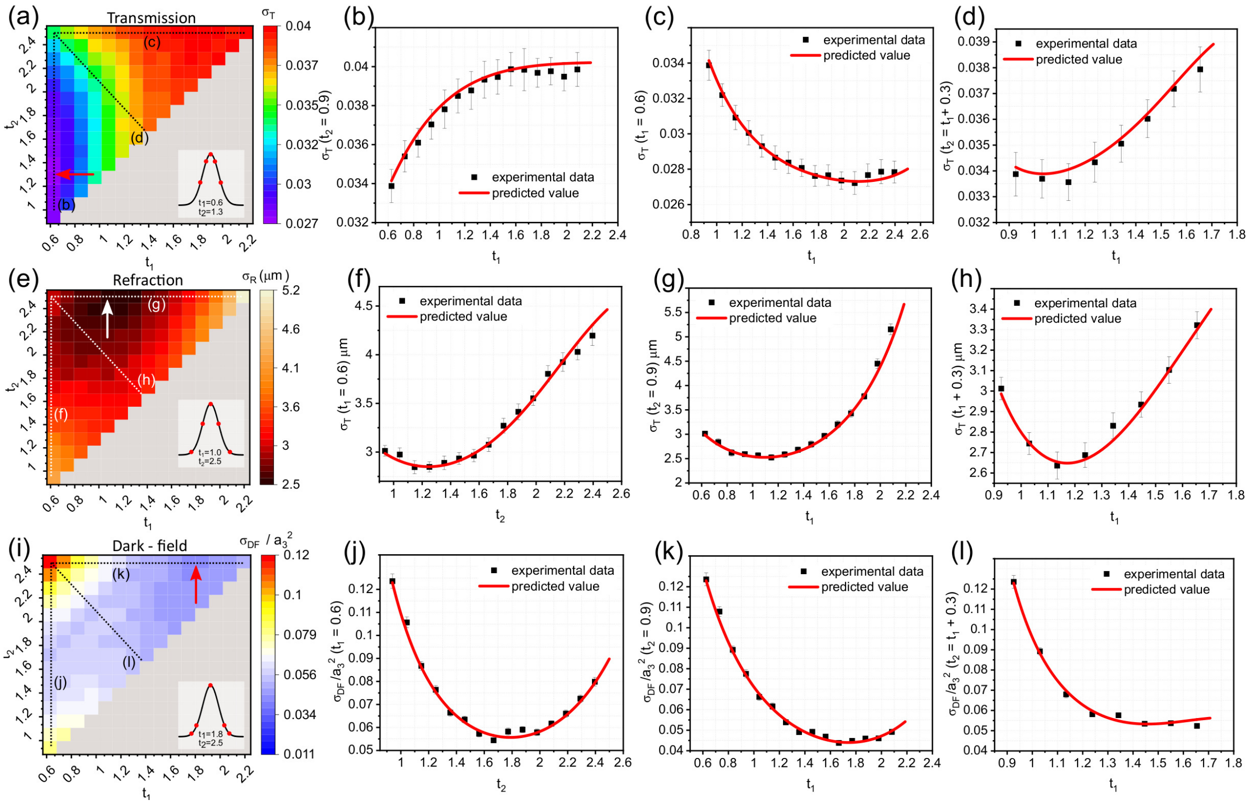

Figure 3: Panels (a), (e) and (i) show the experimental standard deviation values for transmission, refraction and dark-field, respectively, as a function of the IC sampling points t1 and t2. For each contrast channel, line profiles along the dotted lines shown in the 2D plots are extracted and compared to the model predictions, as shown in the panels to the right of each 2D plot. Coloured arrows and the ICs in the insets indicate the combination of sampling points leading to the minimum standard deviation.

convergence of the fitting algorithm since all the points in the fitting model have now a null derivative.

A similar analysis is shown in Fig.3 for the five-point retrieval case. In this case, a 2D plot is needed to show the noise as a function of both sampling point pairs $\pm t_1$ and $\pm t_2$, while t_0 is again kept constant at the top of the IC. Results for transmission, refraction and dark-field are shown in Fig.3(a),(e) and (i), respectively. While the acquisition of at least four sampling points is necessary to fit also the offset, the fact that we are restricting our analysis to the background justifies the assumption of constant offset, since no beam hardening is expected. When comparing the results for the three contrast channels, some differences can be observed. For transmission, the arrangement leading to minimum noise is obtained for $t_1 = 0.6$, $t_2 = 1, 3$, as shown in Fig.3(a). This corresponds to placing two points very close to the top of the IC, and the other two close to the maximum slope position. For refraction, the minimum noise arrangement is obtained two points at the maximum slope and the other two close to the tails, namely for $t_1 = 1$ and $t_2 = 2.5$, as shown in Fig.3(e). For dark-field, also the first two points are shifted towards the tails, with minimum noise achieved for $t_1 = 1.8$ and $t_2 = 2.5$, as shown in Fig.3(i). For all contrast channels, line profiles extracted from the 2D plots along three different directions are compared to the model's predictions, revealing a very good match, as shown in Fig.3(b) to (d), Fig.3(f) to (h) and Fig.3(j) to (l) for transmission, refraction and dark-field, respectively.

Another interesting aspect to be analysed is the correlation between the noise of the fitted parameters which may translate into a correlation between the contrast channels. According to eq.15 this can be addressed by analysing the off-diagonal elements of the covariance matrix or, equivalently, the correlation coefficient described in eq.25. The results are shown in Fig.4 for both three and five sampling points. Panels (a) to (c) show the correlation between fitted parameters when three sampling points are acquired (and therefore a constant offset assumed). A negative correlation between amplitude and width is observed when a non-normalized Gaussian function is used for fitting (see eq.4). As the distance t_1 of the sampling points from the top increases, the correlation decreases, approaching 0.2 when the extreme position $t = 2.4$ is considered. While this position leads to the lowest correlation, it is not convenient in terms of noise for both the refraction and dark-field channels, as shown in Fig.3. However, when a normalized Gaussian profile is used for the fitting (see eq.6), a positive correlation is found (blue line in Fig.4(a)), which reaches a minimum correlation value of 0.2 at a more convenient location, $t = 1.6$. In both cases, the match between model predictions and experimental data is very good. Conversely, no correlation is expected between amplitude and centre and centre and width, as shown in Fig.4(b) and (c), respectively. The expected zero correlation is also confirmed by the experimental data, which fluctuate around zero with a slight increase for the a_1-a_2 correlation at high t values. A similar analysis was performed using five sampling points. The correlation is investigated as a function of the position of both pairs of sampling points $\pm t_1$ and $\pm t_2$, while $t_0 = 0$ is kept constant. A normalized Gaussian model was used which the analysis of the three-point case revealed to be preferable. Offset is included in the fitting parameters, since qualitative experimental observations indicated it introduces a non-negligible degree

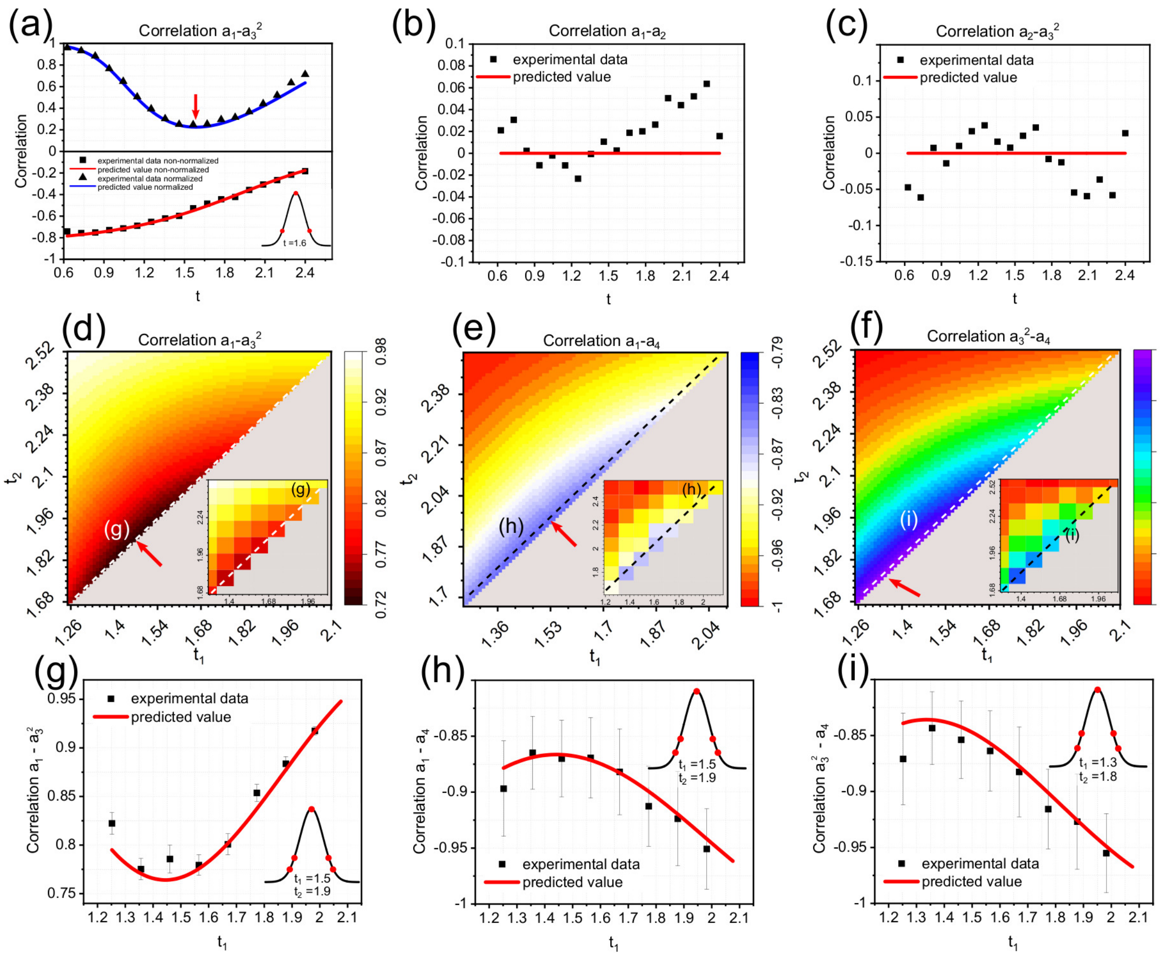

Figure 4: Panels (a) to (c) show the comparison between expected and experimentally measured correlation between the fitting parameters when three sampling points are used. For the correlation between amplitude and width (panel (a)), the different results obtained by using normalized (blue line) and non-normalized (red line) Gaussians are shown. Panels (d) to (f) show the same comparison between theoretical 2D plots obtained when using 5 sampling points, with the corresponding experimental plots shown in the insets. In this case, the offset has been considered as a free parameter, and the correlations between transmission, dark field and offset (taken in pairs) are reported. Panels (g) to (i) show line profile extracted from the 2D plots across the dashed lines indicated. In all panels (d) to (f), the position leading to minimum correlation is indicated by red arrows, and the corresponding positions of the sampling points is shown as red dots in the ICs in the insets of panels (g) to (i).

of correlation. The model predictions are presented as a three 2D plot in Fig.4(d) to (f), with the corresponding experimental plots shown as insets. Line profiles extracted along the indicated dashed lines for both model and experiment are reported in panels (g) to (i), showing a good match in all cases. A minimum can be found in all three cases. This is located at $t_1 = 1.5$ and $t_2 = 1.9$ for amplitude-width and amplitude-offset correlations, and at slightly higher IC positions $t_1 = 1.3$ and $t_2 = 1.8$ for the width-offset correlation. The use of a normalized amplitude coefficient (see eq.6) leads to a positive correlation between amplitude and width, which can be expected since an increase in amplitude is compensated for by an increase in width to keep their ratio constant. Conversely, amplitude-offset and width-offset exhibit an anti-correlated behaviour. This can be explained by considering that an increase in the offset 'cuts out' the bottom part of the IC, leading to a decrease in its width as well as in its amplitude. The introduction of the offset leads, in general, to a higher degree of correlation between parameters, which is in the $|0.75|$ to $|0.9|$ range for all the three contrast channels.

To understand the impact of increasing the sampling points above five as well as of the reduction in the exposure time per point, a Monte Carlo simulation has been performed with three to eleven sampling points and a fixed imaging time, ranging from 1 to 20 seconds. To keep the overall imaging time constant, the exposure time per point was adjusted as a function of the number of sampling points, which corresponds to the common experimental situation when a constraint exists on scanning time. As the number of points increase, the equations are still valid, however, the level of complexity makes unfeasible to explore the entire parameter space. Therefore, the points have been chosen to be evenly spaced in the $\pm 3a_3$ range, even though such configuration may not be the one resulting in minimum noise. The results are shown in Fig.5(a) to (c), with the arrangement of the corresponding sampling points schematised in Fig.5(d). For a given overall scanning duration, noise increases with the number of points for both transmission and dark-field as the time per point decreases. In particular, a sudden increase in the noise is found when going from three to five points, while for higher numbers of points the rate at which noise increases slows down. Conversely, refraction seems to be less affected, with noise exhibiting a slight increase with the number of points only when the total scanning time is low (1-2 s), and a flat behaviour at higher scanning times. This is due to the higher stability of the fit in determining the curve's centre, which can be reliably estimated even if the points are significantly affected by noise.

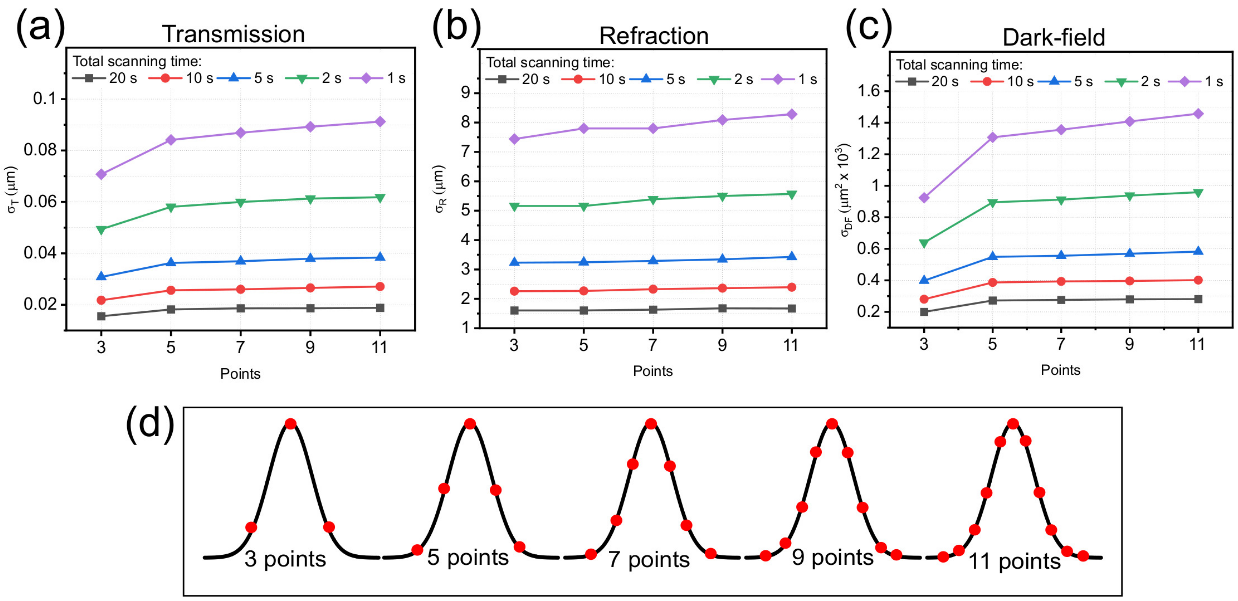

Figure 5: Panels (a) to (c) show the results of Monte Carlo simulations of the noise in retrieved transmission, refraction and dark-field images for a varying number of sampling points equally distributed within a pre-determined total scanning time. The points have been chosen to be evenly spaced in the ±3a3 range, as per the schematic in panel (d).

Discussion

Phase sensitivity in edge illumination is expressed by the IC, which is well described by a Gaussian function when using laboratory sources[28]. The IC is also the basis of phase retrieval, which is performed by means of pixel-wise fitting of images captured at different illumination levels[20]. Therefore, the noise in the retrieved contrast channels depends on the error on the fit, which can be calculated by means of the covariance matrix. Good agreement between the noise calculated according to this model and directly extracted from experimental data was found, which allowed to select optimal IC positions when acquiring data for multi-point phase retrieval (see supplementary materials for an analysis of the validity limits of the presented model).

In common experimental conditions, the total scanning time is usually fixed. Therefore, the exposure time per point must be adjusted according to the acquired number of points. The best performance in terms of noise is achieved by using the lowest number of points needed to fit the IC, namely three points if the offset can be considered constant, and four otherwise even if we investigated the five-point case which is suitable for the developed model requiring an odd number of points. This is due to the increased relative error on each single measurements, which increases when the scanning time per point is decreased to keep the overall scanning time constant.

It was also found that, both when using three and five sampling points, a configuration leading to minimum noise exists, which is a function of the parameters of the edge illumination setup in use.

The analysis of the correlation between channels showed that, when using three sampling points and a constant offset, the correlation between amplitude and width can be minimised but not eliminated, suggesting that a certain degree of correlation between transmission and dark-field can not be avoided when using curve fitting. Interestingly, a different behaviour is observed when using non-normalized vs normalized Gaussian profiles. In both cases, the same minimum correlation value can be achieved, but in very different ways. In the non-normalized case, the minimum is obtained with a set of sampling points far from the ideal (i.e. noise-minimising) positions for both refraction and dark-field . Conversely, when a normalized Gaussian profile is used, the position of minimum correlation in found at a more convenient location, close to minimum noise configuration for both refraction and dark field. We also investigated the role of the offset in terms of its correlation with other channels when five sampling points are acquired, and indeed observed a very high correlation of the offset with both transmission and dark-field. While this correlation may not be a concern when low absorbing or highly scattering samples are imaged, it can become a problem when using high energy X-rays for which both signals are reduced. Therefore, the use of normalized Gaussian is always recommended when performing quantitative analysis.

It is worth noting that this analysis is limited to the background noise, since samples introduce a level of complexity which is difficult to model. The error propagation equations (see eq.21) which relate the variance of the contrast channels to the one of the fitted Gaussian parameters, do not hold anymore within the sample; moreover, offset cannot usually be assumed to be constant. The change in offset is strictly dependent on the specific sample, and can vary on a pixel-by-pixel basis, which makes it extremely difficult to predict. In addition, the use of the covariance to estimate errors requires knowledge of the noise on the experimentally measured intensity. If a photon counter detector is used, a Poisson noise behaviour can be assumed *a priori*; however, if this is not the case, an experimental measurement of the intensity variance at each sampling point is required. Finally, the noise analysis reported here has been performed on normalized images. Therefore, this noise values must be considered as a lower limit, since mask inhomogeneities may introduce an additional random noise component which is not accounted for in our analytical model. However, the good agreement with experimental data obtained with real, imperfect masks would suggest this is a lesser concern, well addressed by the flat field correction procedure. Finally, it is worth noting that changes in the illumination curve position across the field of view due to misalignment of the system and mask unevenness, have a reduced impact on the noise optimization since these are in the range of a few microns even for large masks while the mask period is usual dozens of microns (see supplemental material).

Conclusions

In this work, an analytical model describing the noise background in retrieved multimodal edge illumination images has been developed and compared to experimental data. Since phase retrieval in edge illumination is related to curve fitting, the model was based on the analysis of the standard error on fitted parameters. The good match observed with experimental data demonstrates the validity of the proposed model, which was then used to tackle some common questions occurring when acquiring multi-points edge illumination images. In particular, the arrangement of IC points leading to the best noise performance for each contrast channel, and to the minimum correlation between parameters, was determined. Finally, it was found that, given a fixed overall scanning time, its distribution into a smaller number of sampling points with higher statistics leads to minimum noise, suggesting the use of the smallest possible number of sampling points for phase retrieval. Overall, these results indicate a series of optimised procedures which should be followed in order to optimize experimental acquisitions in edge illumination.

Acknowledgments

This work is supported by EPSRC (Grant EP/T005408/1). CKH and ME are supported by the Royal Academy of Engineering under the RAEng Research Fellowships scheme. PRTM is supported by a Royal Society University Research Fellowship. AO is supported by the

Royal Academy of Engineering under their Chairs in Emerging Technologies scheme.

References

[1] A. Snigirev, I. Snigireva, V. Kohn, S. Kuznetsov, and I. Schelokov, On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation, Review of scientific instruments **66**, 5486–5492 (1995).

[2] A. Momose and J. Fukuda, Phase-contrast radiographs of nonstained rat cerebellar specimen, Medical physics **22**, 375–379 (1995).

[3] B. L. Henke, E. M. Gullikson, and J. C. Davis, X-ray interactions: photoabsorption, scattering, transmission, and reflection at E= 50-30,000 eV, Z= 1-92, Atomic data and nuclear data tables **54**, 181–342 (1993).

[4] C. Hagen, P. Munro, M. Endrizzi, P. Diemoz, and A. Olivo, Low-dose phase contrast tomography with conventional x-ray sources, Medical physics **41**, 070701 (2014).

[5] C. K. Hagen, P. C. Diemoz, and A. Olivo, On the relative performance of edge illumination x-ray phase-contrast CT and conventional, attenuation-based CT, Medical physics **44**, 1876–1885 (2017).

[6] L. Massimi et al., Exploring Alzheimer's disease mouse brain through X-ray phase contrast tomography: From the cell to the organ, NeuroImage **184**, 490–495 (2019).

[7] M. Fratini et al., Simultaneous submicrometric 3D imaging of the micro-vascular network and the neuronal system in a mouse spinal cord, Scientific reports **5**, 1–7 (2015).

[8] L. Massimi et al., Assessment of plaque morphology in Alzheimer's mouse cerebellum using three-dimensional X-ray phase-based virtual histology, Scientific Reports **10**, 1–10 (2020).

[9] F. Palermo et al., X-ray Phase Contrast Tomography Serves Preclinical Investigation of Neurodegenerative Diseases, Frontiers in neuroscience **14**, 1137 (2020).

[10] A. Cedola et al., X-ray phase contrast tomography reveals early vascular alterations and neuronal loss in a multiple sclerosis model, Scientific reports **7**, 1–11 (2017).

[11] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources, Nature physics **2**, 258–261 (2006).

[12] A. Olivo, K. Ignatyev, P. Munro, and R. Speller, Design and realization of a coded-aperture based X-ray phase contrast imaging for homeland security applications, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **610**, 604–614 (2009).

[13] K. Scherer et al., X-ray dark-field radiography-in-vivo diagnosis of lung cancer in mice, Scientific reports **7**, 1–9 (2017).

[14] D. Shoukroun, L. Massimi, F. Iacoviello, M. Endrizzi, D. Bate, A. Olivo, and P. Fromme, Enhanced composite plate impact damage detection and characterisation using X-Ray refraction and scattering contrast combined with ultrasonic imaging, Composites Part B: Engineering **181**, 107579 (2020).

[15] A. Olivo and R. Speller, A coded-aperture technique allowing x-ray phase contrast imaging with conventional sources, Applied Physics Letters **91**, 074106 (2007).

[16] L. Massimi et al., Detection of involved margins in breast specimens with X-ray phase-contrast computed tomography, Scientific reports **11**, 1–9 (2021).

[17] L. Massimi et al., Laboratory-based x-ray phase contrast CT technology for clinical intra-operative specimen imaging, in Medical Imaging 2019: Physics of Medical Imaging, volume 10948, page 109481R, International Society for Optics and Photonics, 2019.

[18] C. K. Hagen, O. Roche i Morgó, and A. Olivo, Predicting the noise in hybrid (phase and attenuation) x-ray images acquired with the edge illumination technique, Medical Physics **47**, 4439–4449 (2020).

[19] M. Endrizzi, F. A. Vittoria, G. Kallon, D. Basta, P. C. Diemoz, A. Vincenzi, P. Delogu, R. Bellazzini, and A. Olivo, Achromatic approach to phase-based multi-modal imaging with conventional x-ray sources, Optics express **23**, 16473–16480 (2015).

[20] M. Endrizzi and A. Olivo, Absorption, refraction and scattering retrieval with an edge-illumination-based imaging setup, Journal of Physics D: Applied Physics **47**, 505102 (2014).

[21] M. Endrizzi, F. Vittoria, L. Rigon, D. Dreossi, F. Iacoviello, P. Shearing, and A. Olivo, X-ray phase-contrast radiography and tomography with a multiaperture analyzer, Physical review letters **118**, 243902 (2017).

[22] M. Endrizzi, F. A. Vittoria, G. Kallon, D. Basta, P. C. Diemoz, A. Vincenzi, P. Delogu, R. Bellazzini, and A. Olivo, Achromatic approach to phase-based multi-modal imaging with conventional x-ray sources, Optics express **23**, 16473–16480 (2015).

[23] M. Endrizzi, P. C. Diemoz, T. P. Millard, J. Louise Jones, R. D. Speller, I. K. Robinson, and A. Olivo, Hard X-ray dark-field imaging with incoherent sample illumination, Applied Physics Letters **104**, 024106 (2014).

[24] P. R. Bevington, D. K. Robinson, J. M. Blair, A. J. Mallinckrodt, and S. McKay, Data reduction and error analysis for the physical sciences, Computers in Physics **7**, 415–416 (1993).

[25] P. Richter, Estimating errors in least-squares fitting, Telecommun. Data Acquisition Prog. Rep **42**, 107–137 (1995).

[26] P. R. Munro, C. K. Hagen, M. B. Szafraniec, and A. Olivo, A simplified approach to quantitative coded aperture x-ray phase imaging, Optics express **21**, 11187–11201 (2013).

[27] G. Kallon, F. Vittoria, I. Buchanan, M. Endrizzi, and A. Olivo, An experimental approach to optimising refraction sensitivity for lab-based edge illumination phase contrast set-ups, Journal of Physics D: Applied Physics **53**, 195404 (2020).

[28] P. Diemoz, C. Hagen, M. Endrizzi, and A. Olivo, Sensitivity of laboratory based implementations of edge illumination x-ray phase-contrast imaging, Applied Physics Letters **103**, 244104 (2013).