X-Ray Ghost-Tomography: Artefacts, Dose Distribution, and Mask Considerations

Andrew Maurice Kingston, Glenn R. Myers, Daniele Pelliccia, Imants D. Svalbe, and David M. Paganin

***Abstract*—Ghost imaging has recently been successfully achieved in the X-ray regime. Due to the penetrating power of X-rays this immediately opens up the possibility of ghost-tomography. No research into this topic currently exists in the literature. Here, we present adaptations of conventional X-ray tomography techniques to this new ghost-imaging scheme. Several numerical implementations for tomography through X-ray ghost-imaging are considered. Specific attention is paid to schemes for reducing the noise-like artefacts of the resulting tomographic reconstruction, issues related to dose fractionation, and considerations regarding the ensemble of illuminating masks used for ghost-imaging. Each theme is explored through a series of numerical simulations, and several suggestions offered for practical realisations of ghost-tomography.**

***Index Terms*—Tomography, x-rays, reconstruction algorithms, radiation imaging, biomedical imaging, geometrical optics, optical signal processing, inverse problems.**

I. Introduction

GHOST imaging is an indirect-imaging technique that utilises intensity correlations to enable spatial information of an object to be determined using a spatially-insensitive detector. The concept originated in the domain of visible-light optics [1]–[4] and has two key features: (i) only photons (or other imaging quanta, e.g., atoms [5]) that do *not* interact with the object are registered by a position-sensitive detector, and (ii) imaging quanta that do interact with the object are not registered by a position sensitive detector but rather with a large single-pixel detector commonly called a *bucket* [6]. Neither of these two sets of measurements is sufficient to form an image of the object, however, through the correlations between them an image can be computed. The computer thus forms an intrinsic component of the technology and ghost imaging is a quintessential example of computational imaging.

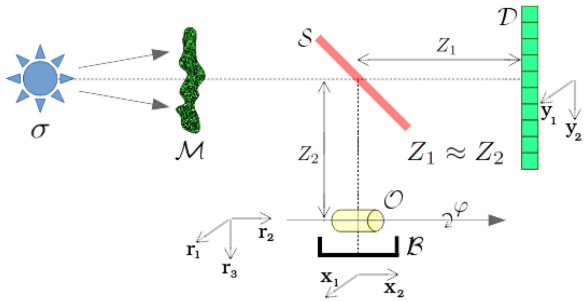

Fig. 1. Generic experimental setup for X-ray ghost imaging and tomography with X-ray source, σ, mask for producing structured illumination, $\mathcal{M}$, beam splitter, $\mathcal{S}$, position-sensitive detector, $\mathcal{D}$, bucket detector, $\mathcal{B}$, and object, $\mathcal{O}$. Note that the axes labelled r_1, x_1 and y_1 are out-of-the-page (see Section II).

Interaction with the object in ghost imaging (GI) has traditionally been via the mechanism of reflection, but with the recent translation to the X-ray regime, transmission ghost imaging is now possible. Ghost-tomography then becomes accessible, however, no research on this currently exists in the literature. Here we explore numerically some techniques to adapt conventional X-ray tomography approaches to GI. In particular, we explore schemes for (i) reducing the noise-like artefacts of the resulting tomographic reconstruction, (ii) issues related to dose fractionation, and (iii) considerations regarding the ensemble of illuminating masks used for GI. We also offer several suggestions towards achieving a practical realisation of ghost-tomography.

Initial developments of ghost imaging (GI) were quantum-mechanical in nature, utilising "spooky action" of entangled photon pairs [7]. Under such formulations, one photon from an entangled pair would be recorded by a position sensitive detector, while the other photon might, or might not, be recorded by a bucket detector depending on its interaction with an object. Since that time *classical* GI has been realised using intensity correlation between pairs of coherent wavefields with spatially highly-structured patterns [8], [9]. Paired coherent wavefields can be generated through use of a beam-splitter as depicted in the schematic of an experimental X-ray GI set-up in Fig. 1. The bucket detector records a single number (the *bucket signal*) for each illuminating pattern. This bucket signal is proportional to the total number of imaging quanta transmitted (or reflected) by the object. The patterned structure may be random (e.g., through speckle fields or photon shot noise [10]) or can be engineered into the illumination through the use of coded apertures [11]

Manuscript received August 13, 2018; revised October 30, 2018; accepted November 2, 2018. Date of publication November 9, 2018; date of current version February 7, 2019. The work of A. M. Kingston and G. R. Myers was supported by the Australian Research Council and FEI-Thermo Fisher Scientific through Linkage Project LP150101040. The work of D. M. Paganin was supported by the Experiment Division of the ESRF. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Kees Joost Batenburg. *(Corresponding author: Andrew Maurice Kingston.)*

A. M. Kingston and G. R. Myers are with the Department of Applied Mathematics, Research School of Physics and Engineering, CTLab: National Laboratory for Micro Computed-Tomography, and Advanced Imaging Precinct, The Australian National University, Canberra, ACT 2601, Australia (e-mail: Andrew.Kingston@anu.edu.au; glenn.myers@anu.edu.au).

D. Pelliccia is with Instruments & Data Tools Pty Ltd., Rowville, VIC 3178, Australia (e-mail: daniele@idtools.com.au).

I. D. Svalbe and D. M. Paganin are with the School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia (e-mail: Imants.Svalbe@monash.edu; david.paganin@monash.edu).

Digital Object Identifier 10.1109/TCI.2018.2880337

or spatial light modulators [8], [9]. More recently, a technique known as *computational* GI has arisen from these engineered methods [12]. Rather than the wavefields being paired in time through a beam-splitter, here wavefields are paired in space by pre-recording a set of structured illuminations with high-fidelity and subsequently recording the effects of these illuminations when interacting with an object using a bucket detector. For further information we refer the reader to a recent GI review by Moreau *et al.* [13].

A. Advantages of Ghost Imaging

To date ghost imaging has been shown to have possible advantages over conventional imaging, in terms of both resolution and reduced-data/dose as well as imaging in turbulent environments. In terms of resolution, a traditional GI set-up is limited by the point-spread function of the position sensitive (or two-dimensional (2D)) detector or the probe wavelength. However, in computational GI coded-apertures that are fabricated with structure known accurately below the resolution of a 2D detector, or patterned illuminations modulated with sub-wavelength precision, can enable super-resolution imaging. For example, resolution below the probe wavelength has been achieved using GI in this fashion with THz radiation [14]. In terms of imaging the object with a reduced number of measurements (or compressed sensing), GI is very similar to the single-pixel camera (SPC) which is ideal for this scenario [15]. In SPC, the illumination resulting from interaction with the object is structured by a programmable mask at the camera prior to measurement; the total intensity along with the mask structure is recorded. A set of these SPC measurements, combined with the appropriate compressed sensing (CS) denoising scheme, can obtain an image from far fewer measurements than conventional imaging [15]. This efficiency is achieved by leveraging the knowledge that the objective function (or image of the object) can be represented in some transform space with a sparse set of coefficients [13]. In contrast to SPC, the GI technique involves structuring the illumination incident on the object. Typically, these scenarios are equivalent mathematically, and the same CS properties apply to GI. However, two key differences arise: (i) when using ionising radiation GI reduces dose while SPC does not, and (ii) structuring the incident illumination field in GI enables one to probe the object using optical effects such as refraction [16]–[18], diffraction [16], [19], and polarisation [20]–[22] which is not possible when filtering at the camera in SPC. The success of CS schemes in SPC to obtain an image from fewer measurements is surely one of the keys to reducing dose in X-ray ghost imaging (and tomography). A complementary technique for image recovery recently introduced in GI is that of deep-learning [23]; this is not explored here.

B. X-Ray Ghost Imaging and Ghost Tomography

Here we focus on the translation of GI concepts to the X-ray domain, a relatively recent development for which there are only several currently-published papers as of this writing [24]–[29]. Due to the penetrating power of X-ray radiation, these techniques produced transmission images rather than the traditional reflection images generated by GI with visible-light. The ability to perform *transmission* GI opens up the possibility for three-dimensional (3D) X-ray GI or X-ray ghost-tomography (XGT) of optically opaque objects. The majority of published X-ray ghost-imaging reconstructions to date are in essence one-dimensional (1D). 2D stencil (i.e., binary) images have been obtained with relative accuracy [27], [28]. We show in this paper that the method used to compute the 2D transmission images, in these papers, is not sufficient for performing tomography.

A practical realisation of 3D X-ray GI is not a straightforward progression from 2D X-ray GI: further exploration is required. To identify and overcome the key practical obstacles to 3D X-ray GI, we conduct a case study by numerical simulation. Our objective here is not to tackle the general XGT problem in a mathematically rigorous fashion. We instead opt for a thorough exploration of the impact that experimental factors have on the problem, and design a feasible experimental protocol for 3D X-ray GI. We therefore restrict ourselves to a very simple object (that one would consider for an initial experimental attempt) and less than 100,000 measurements in total (since we recorded 10,000 per 24 hr period in the 2D X-ray GI experiment reported in Pelliccia *et al.* [28], at the European Synchrotron Radiation Facility). Having established the feasibility of conducting a multi-day 3D X-ray GI synchrotron experiment, we then develop algorithms that offer improved tomographic reconstruction quality, when compared to our initial naive implementation. The understanding gained from this study, along with these techniques developed, formed the foundation underpinning the first experimental realisation of X-ray ghost tomography [30].

Computed tomography (CT) of optically opaque objects using X-ray GI has not been considered in the literature, however, the concept of ghost-tomography and 3D ghost-imaging does indeed exist. 2D ghost-tomography of simple objects has been explored using THz radiation [31], however, refraction effects typically dominate transmission imaging with THz photons (as for infra-red and optical photons) and compensating for these effects is the focus of the paper. Direct 3D imaging has been achieved using GI in an optical coherence tomography (OCT) context [32]. While limited to optically transparent objects, this quantum-OCT system has advantages in dealing with group-velocity dispersion, enabling imaging with a higher spectral bandwidth, and resolution given a Gaussian spectrum. A method for 3D imaging of a surface by GI, i.e., ghost topography, has also been developed using time-of-flight with a single-pixel camera, in the context of remote sensing [33].

C. Comparison of Ghost Tomography and Compressive Tomography

3D CT using illumination masks (known as compressive tomography, see e.g., [34], [35]) also has many similarities to the current work. In compressive tomography the *bucket* signal is measured with a position sensitive detector and there is only one mask applied per view-angle. This can be viewed as a subset of 3D XGT since each unobstructed pixel-value measured can be produced as a bucket signal with a pin-hole mask translated to the said pixel. Thus the image/volume reconstruction techniques developed in the context of compressive tomography are

generally applicable here and the properties and advantages of XGT are a super-set of those determined for compressed tomography. Compressive tomography has been shown to be a far more effective way to reduce dose in CT compared with reducing the number of viewing angles [36], and we expect XGT to be even more effective. The principal limitation in compressive tomography, is that reduced sampling (or measurements) applies to transmission image space, while in XGT the sampling can occur in any transform space of the transmission images, and is determined by the structure of the illumination patterns.

D. Paper Outline

The methods used to simulate radiographic volume projection, illumination patterns, and the corresponding bucket signals are presented in Section II. Section III treats X-ray GI in 2D to obtain projection images. Four approaches are compared, and recommendations made for future experiments in 2D X-ray GI. Section IV considers a naive two-step approach to ghost-tomography, which first separately reconstructs a series of two-dimensional projection X-ray ghost images and then combines that series into a 3D ghost tomogram. For the GI step, the same suite of algorithms is considered as was used in the preceding section; for the subsequent tomography step, two standard methods are considered: (i) filtered back-projection (FBP) [37], and (ii) the simultaneous iterative reconstruction technique (SIRT) [38]. We also consider regularisation techniques such as CS that exploits sparsity to remove noise-like artefacts from the reconstructions, as well as studying dose fractionation. Section V develops a single-step X-ray ghost tomographic reconstruction that does not require the intermediate step of reconstructing 2D ghost projection images. We compare spatially random speckle masks with coded masks, for both 2D and 3D X-ray ghost imaging, in Section VI. Section VII discusses possibilities for reduced dose in X-ray GI compared to direct X-ray imaging, gives general remarks regarding random versus coded masks, and gives a way to avoid ring artefacts in ghost-tomography. Section VII outlines future research followed by some concluding remarks in Section VIII.

II. Method

In this section we describe the simulation process to generate data for the following ghost-tomography study. We first outline the experimental set-up used for the simulations as well as the synthetic object (or phantom) used for this case study. To simulate a ghost tomography experiment, we are required to produce (i) *transmission functions* of a 3D volume containing a simple phantom, (ii) a set of *illumination masks*, and (iii) the corresponding *bucket values* by applying these masks to (i). We conclude this section by observing some important assumptions made herein and establish a quality benchmark for subsequent ghost imaging and ghost tomography reconstruction results.

A. Experimental Set-Up

The experimental geometry used here follows the schematic in Fig. 1, similar to that used in previous experimental X-ray studies for 1D and 2D GI [24], [25], [28]. Here, a synchrotron X-ray source, σ, illuminates a thin transmissive mask, $\mathcal{M}$, with a parallel beam. A beam-splitter, $\mathcal{S}$, creates two arms in the ghost-imaging setup. The transmitted beam has its transverse spatial intensity structure registered by a position-sensitive detector, $\mathcal{D}$. The reflected beam passes through an object, $\mathcal{O}$, before having the total intensity, transmitted by the object, recorded by the *bucket* detector, $\mathcal{B}$. The distance Z_1 is made similar to the distance Z_2 since the transverse spatial distribution of intensity transmitted by the mask will in general exhibit Fresnel diffraction, and it is important that the illuminating patterns measured at $\mathcal{D}$ be equal to the intensity field illuminating the object, up to a multiplicative constant. We choose Z_1 (and thus Z_2) to be sufficiently small that Fresnel diffraction can be ignored.

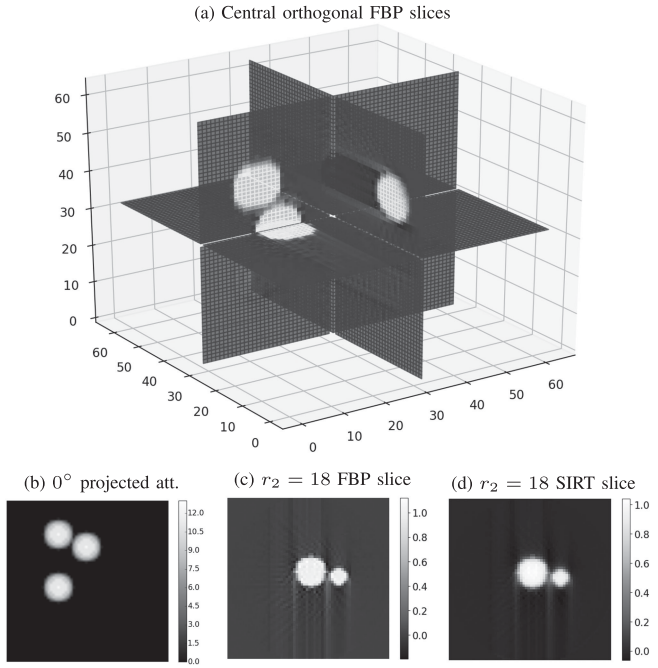

Fig. 2. (a) Orthogonal 2D slices through the 64^3 voxel tomogram reconstructed with FBP. (b) The 64^2 pixel projection image of the phantom at 0°. Slices through the reconstructed volume at $r_2 = 18$ reconstructed by (c) FBP, and (d) 32 iterations of SIRT. All reconstruction is based on noise-free simulated conventional imaging of each projection image.

B. Phantom (or Simulated Object)

The phantom considered here is comprised of three non-overlapping, identical spheres arranged in a 3D volume parameterised by Cartesian coordinates $\mathbf{r} = (r_1, r_2, r_3)$. The 3D X-ray linear attenuation coefficient $\mu(\mathbf{r})$, is modelled using $\mathcal{N}^3$ cubic voxels with arbitrary dimension ϵ and $\mathcal{N} = 64$. The spheres have a diameter of 12 voxels and an attenuation of 1.0 per voxel that can also be arbitrarily scaled. A depiction of this volume can be found in Fig. 2a.

C. Simulating Transmission Functions

Since we have a parallel X-ray beam we use square pixels, again with dimensions ϵ, to represent the $\mathcal{N} \times \mathcal{N} = 64 \times 64$ pixel transmission and projected attenuation images of the

phantom, $\mu(\mathbf{r})$. These are denoted $T(\mathbf{x})$ and $A(\mathbf{x})$ respectively where the 2D Cartesian coordinates $\mathbf{x} = (x_1, x_2)$ parameterise the transverse position in the contact plane: the plane normal to the direction of X-ray propagation through the object, contacting the *downstream* side of the object. Parallel beam geometry implies rotation by azimuthal angle φ, about a single axis, and provides sufficient information for exact 3D tomographic reconstruction. φ denotes rotation of $\mu(\mathbf{r})$ about the r_2-axis (which corresponds to the x_2-axis of the imaging system), and describes the angle between the r_1-axis and the x_1-axis. Projected attenuation, $A_\varphi(\mathbf{x})$, of the volume representing the object, $\mu(\mathbf{r})$, at angle φ is simulated by applying the X-ray projection transform $\mathcal{P}$ to each horizontal slice of the volume independently as follows:

$$\begin{aligned} A_\varphi(\mathbf{x}) &= \mathcal{P}_\varphi[\mu(\mathbf{r})] \\ &= \langle \mu(r_1, x_2, r_3) | \delta(r_1 \cos\varphi + r_3 \sin\varphi - x_1) \rangle_{(r_1, r_3)}, \end{aligned} \tag{1}$$

where $\langle | \rangle_{(r_1, r_3)}$ denotes the inner product spanned by Cartesian coordinate (r_1, r_3). Observe in Fig. 1 that co-ordinate systems $\mathbf{x}$ and $\mathbf{r}$ share a common axis $x_2 = r_2$, with the position of axis r_1 (and axis r_3) relative to axis x_1 being defined by the object rotation angle, φ, about the r_2 axis. Projection (as defined in Eqn. 1) can be achieved numerically by rotating the volume, $\mu(\mathbf{r})$, about the r_2-axis, perpendicular to the X-ray beam direction, by φ and then summing rows along the beam axis. An example of a projected attenuation image at $\varphi = 0$ is given in Fig. 2. To satisfy Nyquist angular-sampling requirements, $\pi\mathcal{N}/2$ azimuthal angles are required [39]; we have thus generated 90 projected attenuation images with φ distributed evenly over π radians. The transmission function of the object is found as

$$T_\varphi(\mathbf{x}) \equiv \exp[-A_\varphi(\mathbf{x})]. \tag{2}$$

D. Simulating Illumination Patterns

Let there be an ensemble of N illumination masks, $\mathcal{M} = \{M_n(\mathbf{y}) : n \in \mathbb{Z}_N\}$, used to generate N illumination patterns registered by the position sensitive detector $\mathcal{D}$, $\{I_n(\mathbf{y}) : n \in \mathbb{Z}_N\}$. Here, $\mathbf{y} = (y_1, y_2)$ denotes Cartesian coordinates in planes orthogonal to the optic axis from source to detector. The masks, M_n, and corresponding illuminating patterns, I_n, are again discretised using square pixels with arbitrary physical dimensions ϵ. Two classes of mask will be of particular interest here: (i) random masks whose transmitted intensity distribution is a spatially random speckle field. Each realisation has the same characteristic transverse length scale (speckle size) and root-mean-square (RMS) intensity at every point in the field of view of the mask. Here, the illuminating $\mathcal{N} \times \mathcal{N}$-pixel patterns constitute an ensemble of realisations of an $\mathcal{N} \times \mathcal{N}$ pseudo-random binary matrix with a mean of 0.5 [40], [41]; (ii) coded masks whose transmitted intensity distribution generates a linearly independent set of illuminating intensity maps, according to a deterministic algorithm such as is used to calculate uniformly redundant arrays [42], so-called perfect arrays [43], etc. A spatially uniform flux of X-rays with arbitrary intensity of 1.0 is assumed such that the illuminating intensity patterns, $I_n(\mathbf{y})$, generated by each mask are simply $I_n(\mathbf{y}) = M_n(\mathbf{y})$.

E. Computing Bucket Values

The jth XGT measurement, from a total of J measurements, is comprised of three elements: the estimated intensity distribution, the object orientation, and the bucket signal, i.e., $\{I_j(\mathbf{x}), \varphi_j, B_j\}$. The intensity distribution (measured at $\mathcal{D}$) $I_j(\mathbf{y})$ is generated by one of the N masks such that $I_j(\mathbf{y}) \in \{I_n(\mathbf{y}) : n \in \mathbb{Z}_N\}$ and $I_j(\mathbf{x})$ is estimated from this. In a system of the type shown in Fig. 1, x is the 2D coordinate system normal to the bucket-arm of the system, and y is the 2D coordinate system normal to the detector $\mathcal{D}$ arm. The object orientation φ_j is one of the L azimuthal angles such that $\varphi_j \in \{\varphi_l : l \in \mathbb{Z}_L\}$. The bucket value or signal, B_j, is the integrated intensity incident on the bucket detector, $\mathcal{B}$. As Fresnel diffraction has been assumed negligible, the transmitted X-ray intensity in the contact plane $I'_n(\mathbf{x})$, is

$$I'_j(\mathbf{x}) = I_j(\mathbf{x}) T_{\varphi_j}(\mathbf{x}), \tag{3}$$

and the bucket signal, B_j, is then approximated via

$$B_j \approx \langle I_j(\mathbf{x}) | T_{\varphi_j}(\mathbf{x}) \rangle_{\mathbf{x}}. \tag{4}$$

F. Fundamental Assumptions

There are several important points to note here: (i) Our simulations all assume zero noise in the photon detection process, i.e., the effects of noise (e.g., photon shot noise, detector noise) and other detector imperfections (e.g., hot detector pixels, dead detector pixels, non-uniformities in detector gain, imperfections and losses of the beam-splitter) are not considered; (ii) The physical pixel/voxel size, ϵ, does not need to be specified; (iii) The projected attenuation also has no scale. (iv) To linearise the problem we assume both that the phantom is *weakly absorbing*, and that the mean illuminating intensity, $\langle I_j(\mathbf{x}) | 1 \rangle_{\mathbf{x}}$, is constant. This enables us to work directly with projected attenuation (as expanded below). An example phantom may be composed of three poly(methyl-methacrylate), or PMMA, spheres of 2 mm diameter imaged with 30 keV X-rays. PMMA has a density of 1.18 g cm^{-3} giving a linear attenuation coefficient of 0.072 cm^{-1} or 93% transmission. Transmission bucket measurements, B_j can then be found directly from the projected attenuation bucket measurements, $B_j^A \approx \langle I_n(\mathbf{x}) | A_{\varphi_l}(\mathbf{x}) \rangle_{\mathbf{x}}$ by approximating the exponential term with the first order Maclaurin expansion:

$$\begin{aligned} B_j &= \langle I_j(\mathbf{x}) | \exp[-A_{\varphi_j}(\mathbf{x})] \rangle_{\mathbf{x}} \\ &\approx \langle I_j(\mathbf{x}) | 1 - A_{\varphi_j}(\mathbf{x}) \rangle_{\mathbf{x}} \\ &= \langle I_j(\mathbf{x}) | 1 \rangle_{\mathbf{x}} - B_j^A. \end{aligned} \tag{5}$$

This approximation greatly simplifies the following numerical study, since we can work directly with attenuation data, without compromising the validity and implications of the results. A significant difference does occur when tomographic reconstruction is performed directly from measured bucket values (Section V); the determination of error in projected attenuation

from intensity bucket residuals will be presented for both cases (weak X-ray absorption and otherwise).

G. Quality Standard for Comparison

As a first benchmark for quality of tomographic reconstruction, the result of conventional (i.e., direct imaging) X-ray tomography using filtered back-projection is presented in Fig. 2. Central 2D slices in the r_1-, r_2-, and r_3- directions are presented in Fig. 2a, with the $r_2 = 18$ slice shown in Fig. 2c. Henceforth only this slice will be presented for comparison. Note that the mild streaks in the r_2-slice are due to a half-pixel offset introduced between the actual rotation axis, and the location of the axis used in the FBP reconstruction. This offset was kept in all remaining simulations. As a second benchmark, Fig. 2d shows the result of applying SIRT to the original projections, with 32 iterations. Regularisation can be achieved by reducing the number of iterations performed. In this initial ideal case, 32 iterations reconstructs the objective function with reasonable fidelity. Since this work is predominantly an exploration of techniques in the new field of ghost-tomography, we are as much interested in the type/appearance of image quality degradation as in the amount of degradation. Therefore, we principally rely on visual comparison with these benchmarks for quality inspection rather than simple quantitative metrics. Quality metrics will be used to indicate trends where full suites of images could not be presented.

III. X-Ray Ghost Projection Imaging

A naive method for XGT is an obvious two-step approach: first perform ghost imaging (GI) to recover each projection image; second, apply the conventional tomographic reconstruction technique of filtered back-projection. In this section we explore the first step of this process, namely how best to obtain projection images. We start by outlining the four different GI approaches considered in the present paper, then present and compare the results from employing each method to recover the projection image at angle 0°.

A. Ghost Imaging Techniques

1) Cross-Correlation: The conventional method for ghost image reconstruction, here termed cross correlation (XC) [9] and written as the operator $\mathcal{C}_{\varphi_l}$, approximates the intensity transmission function $T_{\varphi_l}(\mathbf{x})$ of the object using

$$T'_{\varphi_l}(\mathbf{x}) = \mathcal{C}_{\varphi_l}(B_j) \tag{6}$$

$$\equiv \frac{1}{N}\sum_{j=1}^{J}(B_j - \overline{B})I_j(\mathbf{x})\delta(\varphi_l - \varphi_j), \tag{7}$$

where $\overline{B} = \frac{1}{J}\sum_{j=1}^{J} B_j$ is the average bucket value, N is the number of measurements at angle $\varphi_l = \varphi_j$, and δ denotes the Kronecker delta.

2) Iterative Cross-Correlation: Given a current estimate $T^k_{\varphi_l}(\mathbf{x})$, which may be obtained by XC, an improved estimate is obtained via the update scheme

$$T^{k+1}_{\varphi_l}(\mathbf{x}) = T^k_{\varphi_l}(\mathbf{x}) + \gamma\mathcal{C}_{\varphi_l}[B_j - \langle I_j(\mathbf{x})|T^k_{\varphi_l}(\mathbf{x})\rangle_{\mathbf{x}}]. \tag{8}$$

Identifying the GI operator as the adjoint of the XC operator

$$\mathcal{C}^*_{\varphi_j}[T^k_{\varphi_l}(\mathbf{x})] \equiv \langle I_j(\mathbf{x})|T^k_{\varphi_l}(\mathbf{x})\rangle_{\mathbf{x}} \tag{9}$$

this can be written as

$$T^{k+1}_{\varphi_l}(\mathbf{x}) = T^k_{\varphi_l}(\mathbf{x}) + \gamma\mathcal{C}_{\varphi_l}\{B_j - \mathcal{C}^*_{\varphi_j}[T^k_{\varphi_l}(\mathbf{x})]\}, \tag{10}$$

suggesting this is a gradient descent method where γ is a Landweber relaxation factor with the objective function:

$$\arg\min_{T_{\varphi_l}(\mathbf{x})} \sum_j (B_j - \langle I_j(\mathbf{x})|T_{\varphi_l}(\mathbf{x})\rangle_{\mathbf{x}})^2. \tag{11}$$

We use $\gamma = \alpha/\sigma^2$ where σ^2 is the spatially averaged variance of $I_j(\mathbf{x})$. Given complete data, we set $\alpha = 0.25(J/\mathcal{N}^2)^2$ since this adds more relaxation as the problem becomes under-constrained ($J < \mathcal{N}^2$). This iterative cross-correlation (IXC) process is iterated until a suitable convergence criterion is achieved [28], [44].

3) Conjugate-Gradient Method: The IXC concepts can be adapted to utilise the conjugate-gradient method [45], i.e., conjugate gradient cross-correlation (CGXC), to iteratively improve the reconstruction [46].

4) Regularisation Methods: Regularisation of IXC is also considered (RXC), utilising various forms of *a posteriori* assertions about the objective function [9], [15], [35], [47], [48]. The objective function for IXC, Eq. (11) can be regularised as follows

$$\arg\min_{T_{\varphi_l}(\mathbf{x})} \sum_j (B_j - \langle I_j(\mathbf{x})|T_{\varphi_l}(\mathbf{x})\rangle_{\mathbf{x}})^2 + \lambda\mathcal{R}(T_{\varphi_l}(\mathbf{x})), \tag{12}$$

where $\mathcal{R}$ is a function that penalises the deviation of $T_\varphi(\mathbf{x})$ from the assertions, and λ weights the importance of this penalty. Three typical assertions, relevant in the present context, are (a) image-space sparsity, where $T_{\varphi_l}(\mathbf{x})$ is assumed to be negligible for most pixels, (b) gradient sparsity, for which $|\nabla_\perp T_{\varphi_l}(\mathbf{x})|$ is negligible for most pixels, $\nabla_\perp$ being the gradient operator in the $\mathbf{x}$ plane, and (c) smoothness, where the values of $\mathcal{F}[T_{\varphi_l}(\mathbf{x})]$ are negligible for high-spatial frequencies, with $\mathcal{F}$ denoting Fourier transformation.

B. Results and Discussion

1) Cross-Correlation: The XC GI method was applied to reconstruct the object transmission function in Fig. 2b from $J = 1000$, 2000, 3000, and 4000 masks. The corresponding normalised mean absolute deviation (MAD) values – namely the mean absolute value of the difference between the reconstructed and input projection image (see Fig. 2), scaled by the maximum image value – are 0.317, 0.230, 0.187, and 0.164. This error metric decreases as J is increased. Fig. 3a, b, c shows the resulting ghost projection images for $J = 1000, 2000$, and 4000.

2) Iterative Cross-Correlation: The images recovered by XC GI can be improved using IXC, as shown in Fig. 3d, e, f for $J = 1000$, 2000, and 4000. The normalised MAD values

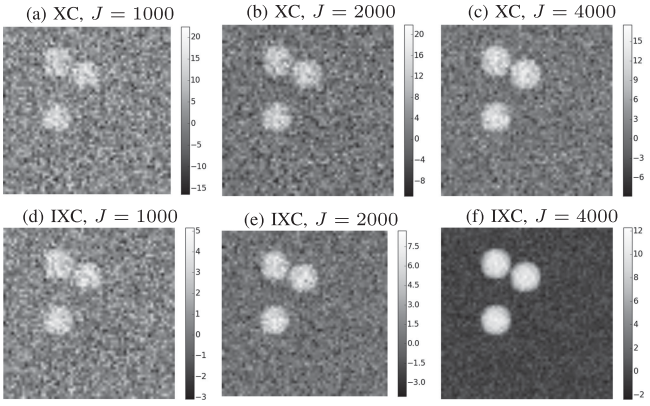

Fig. 3. The 0° projection ghost-images recovered from J measured bucket values generated with random binary masks with a mean value ≈ 0.5. Recovery performed using either the XC method or 100 IXC iterations (as specified). (a) XC, $J = 1000$, (b) XC, $J = 2000$, (c) XC, $J = 4000$, (d) IXC, $J = 1000$, (e) IXC, $J = 2000$, and (f) IXC, $J = 4000$.

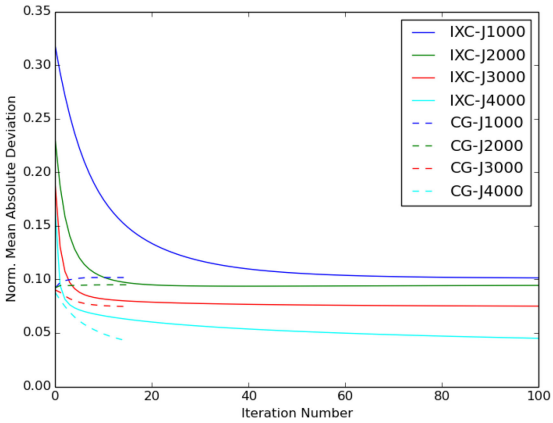

Fig. 4. Convergence plots for 100 iterations of IXC and 16 iterations of CGXC for $J = 1000$, 2000, 3000, and 4000.

following 100 IXC iterations corresponding to $J = 1000$, 2000, 3000, and 4000 masks are 0.101, 0.0943, 0.0750, and 0.0450. Convergence plots are shown in Fig. 4.

3) Conjugate-Gradient Method: Faster convergence than that of IXC can be achieved, for a moderate number of measurements, using the third means for 2D X-ray GI mentioned here, i.e., CGXC. However, for highly under-constrained data, stability becomes a problem. The normalised MAD values following 16 CGXC iterations corresponding to $J = 1000$, 2000, 3000, and 4000 masks are 0.102, 0.0949, 0.0747, 0.0423. The MAD values for $J = 3000$, and 4000 are already better than for 100 iterations of IXC after 16 iterations and continue to improve (e.g., MAD = 0.0238 for $J = 4000$ after 100 CGXC iterations). However, for $J = 1000$ and 2000 the MAD values are worse than for IXC and start to diverge after 16 CGXC iterations: for this reason the comparison plot stops at 16 iterations. These robustness issues suggest that CGXC would not be the method of choice given noisy and under-constrained experimental data.

4) Regularisation Methods: One can further improve on the IXC results for highly under-constrained problems by regularising the problem using *a priori* knowledge of the objective

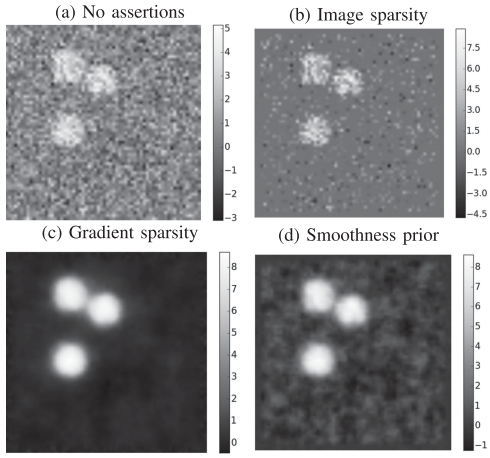

Fig. 5. The 0° ghost projection images recovered from $J = 1000$ measured bucket values generated with random binary masks with a mean value ≈ 0.5. Recovery using 1000 IXC iterations with *a posteriori* assertions. (a) None. (b) Image-space sparsity. (c) Gradient sparsity. (d) Smoothness. MAD: (a) 0.101, (b) 0.0547, (c) 0.0363, (d) 0.0487.

function, cf. Yao *et al.* 2014 [49]. This is the fourth and final 2D GI method considered here. The three relevant assertions are image-space sparsity, gradient sparsity and smoothness. The results of applying these to the 1000-bucket-measurement problem are shown in Fig. 5. We have used 1000 iterations of IXC since for best results these priors are enforced very lightly and gradually, e.g., $\alpha = 0.025(J/\mathcal{N}^2)^2$. All sparsity assumptions yield significant reduction in the noise-like artefacts, with gradient sparsity seeming to be the most appropriate.

From this set of simulations we observe the following: XC GI alone can be used provided the number of measurements, J, approaches the number of pixels, $\mathcal{N}^2$ (which is unlikely on a real experiment). IXC GI, although very slow, is robust and could improve on XC results for all images (although improvement decreases with J). Faster convergence was achieved by CGXC, however, this became unstable in the highly under-constrained cases, namely where $J \ll \mathcal{N}^2$. Since, for the object considered here, *a posteriori* properties can be asserted, regularisation techniques such as compressed sensing and Bayesian approaches were favourable when $J \ll \mathcal{N}^2$. This is an important observation, especially considering the limited number of total measurements expected in our experimental tomography scenario.

IV. Ghost-Tomography: Two-Step Approach

From this point we begin investigating XGT. We first explore performance of the naive two-step approach: (1) 2D ghost projection images are first recovered at each object orientation. These projections can be determined using one of the four 2D GI methods introduced earlier (XC, IXC, CGXC and RXC). (2) Conventional tomographic reconstruction, using these recovered X-ray ghost projection images, is then performed to compute the volume. Here the reconstruction is performed by employing two different standard tomographic approaches, namely (i) the analytic approach of filtered back-projection (FBP), or

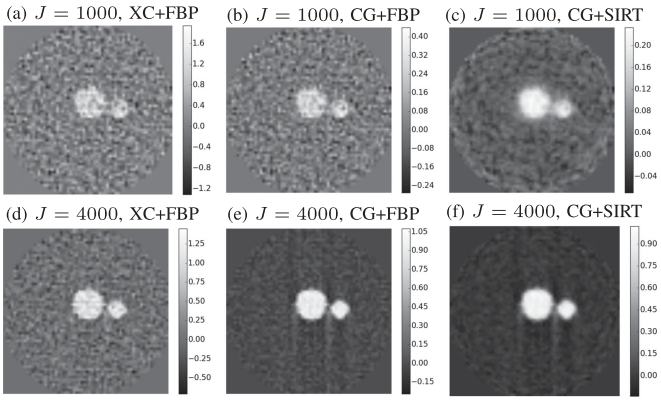

Fig. 6. The $r_2 = 18$, slice through the 64^3 voxel tomogram reconstructed from ghost projections recovered by (a) and (d) XC, (b) and (c), (e) and (f) 16 CGXC iterations from J measured bucket values using FBP or 100 SIRT iterations (as specified).

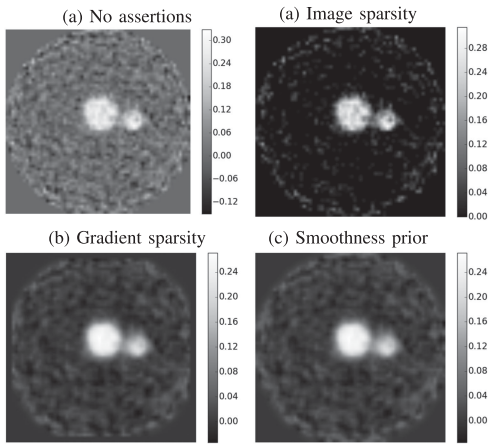

Fig. 7. Tomogram slice $r_2 = 18$ reconstructed from ghost projections recovered by 16 CGXC iterations from $N = 1000$ measured bucket values per azimuthal angle φ_l using 256 iterations of SIRT with regularisation. (a) None. (b) Image-space sparsity. (c) Gradient sparsity. (d) Smoothness.

(ii) iterative refinement using the simultaneous iterative reconstruction technique (SIRT) [38].

There are assumed to be $J = NL$ total recorded measurements, with N measurements performed at each of L object orientations φ. Here, $\{\varphi_l\}$ samples from an equally spaced subset of the 90 equally spaced azimuthal angles for which projection images are simulated $\{\varphi_l : l \in \mathbb{Z}_{90}\}$. In order to design a first XGT synchrotron experiment, we need to gain an appreciation of the effects of reducing N compared to reducing L, and which regularisation technique is most useful for this type of under-constrained system.

A. Number of Bucket Measurements Per Projection, N

We investigate the effect of altering the image quality of each X-ray ghost projection by changing the number of bucket value measurements N per projection to $N = 1000$, 2000, 3000, and 4000 with the number of viewing angles, L, set to 90. The results from applying FBP to projections generated by XC from 1000 and 4000 measurements are presented in Fig. 6a, d (XC+FBP). As expected, image quality improves with increasing N. Similar results from FBP applied to projections generated by 16 CGXC iterations from 1000 and 4000 measurements are presented in Fig. 6 b, e (CG+FBP). Here we observe the improvements that can be realised using iterative refinement to XC given $N \simeq \mathcal{N}^2$. A certain degree of regularisation can be achieved by exiting SIRT before convergence. The result of applying 100 SIRT iterations, with a relaxation scaling of 0.5 N/$\mathcal{N}^2$ applied to updates, to the $N = 1000$ and 4000 CGXC cases (CG+SIRT) are given in Fig. 6c, f. We observe that for $N = 4000$ this seems sufficient regularisation, however, significant artefacts still appear for the $N = 1000$ case. The latter could benefit from some regularisation.

B. Regularisation Techniques

Here we briefly investigate the noise-like high-frequency artefacts that arise above for the $N = 1000$ measurements seen in Fig. 6a, b, c. Results in Section III suggest that regularisation, where applicable, can achieve superior results compared to the marginal improvements obtained by IXC GI compared to XC GI for $N \ll \mathcal{N}^2$. The same three regularisation assertions are explored here, now in tomogram space during the iterative tomographic reconstruction stage. A limit of 256 iterations has been adopted, since each such iteration is computationally costly in 3D GI when compared to 2D GI.

Fig. 7 shows regularisation to be quite effective at suppressing noise-like artefacts resulting from insufficient measurements per viewing angle, as was the case for 2D GI in Section III. Combining image-space and gradient sparsity we expect results approaching that for $N = 4000$ can be achieved with four times less experiment time (at the expense of computation time). This is an important and encouraging result when designing synchrotron experiments. The next question to address is, given a fixed experiment time (or X-ray dose) how should this best be distributed amongst view-angles?

C. Dose Distribution

In conventional tomography, the dose fractionation theorem states that it is better to utilise many viewing angles, L, with low-dose than fewer viewing angles with high-dose [50]. Recall that in XGT, the total number of bucket measurements is $J = LN$, with N illumination patterns being used at L viewing angles. Assuming time per measurement is fixed, we can explore the effect of dose fractionation by keeping J constant, while varying the number L of azimuthal orientations equally-spaced over π radians.

Initially, we limit the numerical experiment to $J \approx 90,000$ bucket measurements. We investigated scenarios with $N = 1000$, 2000, 3000, and 4000 and corresponding $L = 90$, 45, 30, and 22. The tomographic slices generated by a naive XC+FBP two-step approach (i.e., XC recovery of ghost projection images, followed by FBP) in Fig. 8a, d show the results for $N = 1000$ and 4000 respectively. We have not used any iterative refinement here, to emphasise the type of artefacts present; we observe noise-like artefacts appear similar, with more streaking apparent as L reduces. The degree of these noise-like artefacts

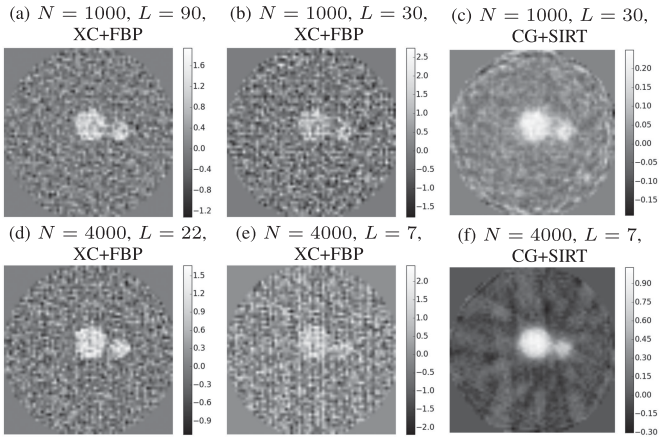

Fig. 8. Tomogram slice $r_2 = 18$ reconstructed from ghost projections recovered by XC or 16 CGXC iterations (as specified) from $J = LN$ measured bucket values using FBP or 256 SIRT iterations (as specified). (a) and (d) Show $J \approx 90,000$ measurements while (b), (c), (e), and (f) show $J \approx 30,000$.

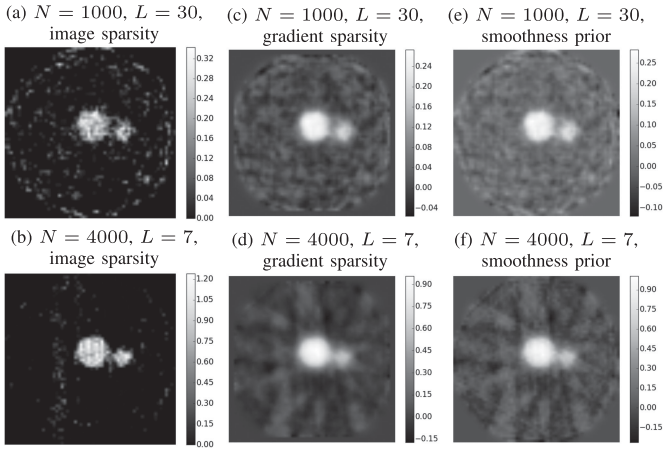

Fig. 9. Tomogram slice $r_2 = 18$ reconstructed from ghost projections recovered by 10 IXC iterations from $J = LN \approx 30,000$ measured bucket values using 256 SIRT iterations with *a posteriori* assertions. (a) and (d) Image-space sparsity. (b) and (e) Gradient sparsity. (c) and (f) Smoothness prior.

(from noise-free simulated data) again demonstrates the need for more sophisticated means of realising 3D X-ray GI in a practical manner, given the current experimental limitation that $J \lesssim 100,000$.

Further reducing the total number of bucket measurements to $J = 30,000$ (more realistic experimentally), we used the same set of four values for N with the corresponding number of viewing angles of $L = 30$, 15, 10, and 7. The results from two-step XGT for $N = 1000$ and 4000 are given in Fig. 8 b, c, e, f. The tomograms reconstructed by XC+FBP are again presented to emphasise artefacts; they are too degraded for all but the most crude tomographic requirements. The backgrounds are dominated by noise-like artefacts making it difficult to see the spheres. The streaking artefacts that arise when too few viewing angles, L, are utilised is more evident in this scenario (Fig. 8e) than for $J = 90,000$. 100 SIRT iterations applied to ghost projection images recovered by 16 CGXC iterations indicate what can be achieved without asserting priors. The difference in artefacts becomes quite clear here (Fig. 9c compared to 9f); tomograms from many low-dose projections are dominated by high-frequency artefacts, while tomograms from few high-dose projections are dominated by low-frequency streak artefacts. This affects suitability for regularisation as we explore next.

Given $J \approx 30,000$ ghost projection images recovered by 16 CGXC iterations, we investigated the effectiveness of employing 256 SIRT iterations with *a posteriori* regularisation. Selected results are presented in Fig. 9. We observe that for the 7×4000 measurements, gradient-space sparsity and smoothness priors have little effect on the low-frequency streak artefacts; in this case image sparsity most effectively suppresses artefacts since the streaks are low-intensity. For the case of 30×1000 measurements, all three methods have a similar degree of effectiveness and perhaps combined priors would prove useful.

Ghost-tomography by the two-step approach does not necessarily follow the dose fractionation theorem [50]. Since each projection image is recovered separately, there is a minimum amount of information required for reasonable 2D GI and the nature of the artefacts changes with the number of view-angles, L. Regularisation techniques, particularly in *compressive sensing*, are predominantly designed for denoising, thus in ghost-tomography these schemes are better suited to many low-dose projections. Of course this study used noise-free simulations and we expect a compromise, e.g., $N = 2000$, would be preferable experimentally when restricted to $J \simeq 30,000$.

V. Ghost-Tomography: Direct Approach

The naive two-step approach to 3D X-ray GI explored in the previous section only makes limited use of the information available in the measured data. The first step, performing 2D GI to recover projection images at each view-angle, treats each object orientation separately; the iterative procedure considers only data from measurements taken at the corresponding view-angle. Tomographic reconstruction then proceeds from these limited data 2D GI results. However, all measured data is of the same object and should be used in each iteration in a direct reconstruction from bucket signals to the 3D tomogram. A direct approach is derived and explored in this section.

To motivate the direct XGT approach, consider the question from an elementary perspective based on back-projection. A XC form of XGT can be achieved by back-projecting

$$\mathcal{C}(B_j) \equiv [B_j - \overline{B}]I_j(\mathbf{x}) \tag{13}$$

along the direction φ_j, and ensemble averaging over all of the back-projections corresponding to all object orientations, φ_j (see Fig. 1; cf. Eq. (6)). Note especially that the operator $\mathcal{C}$ is a generalised form of $\mathcal{C}_\varphi$ which is restricted to the sum over buckets at angle φ. Also note that we have used the weak-absorption approximation here. This XC XGT formulation permits as few as one bucket signal per object orientation, e.g., if every successive orientation over a sphere, (θ_j, φ_j), or over a cylindrical surface for cone beam radiation, (z_j, φ_j), is chosen randomly with a uniform distribution over the sphere/cylinder. This limiting case, of one bucket measurement per object orientation, exemplifies the fact that it is not necessary to proceed via an intermediate step of calculating 2D X-ray GI projections when

undertaking 3D XGT. Of course, practically speaking, to avoid rotating the object every time a subsequent bucket measurement is taken, and also in light of simulations presented below, it will often be advantageous to take a series of bucket measurements for each object orientation.

Here we derive an iterative tomographic reconstruction algorithm that operates directly from the measured bucket signal. Starting from Eq. (1) for X-ray projection, optimisation by gradient descent given a current estimate $\mu^k(\mathbf{r})$, for an improved estimate $\mu^{k+1}(\mathbf{r})$ is then found as:

$$\mu^{k+1}(\mathbf{r}) = \mu^k(\mathbf{r}) + \beta\mathcal{P}^*[A_\varphi(\mathbf{x}) - A^k_\varphi(\mathbf{x})]. \tag{14}$$

Here, β is again a Landweber relaxation factor (typically set to $(L\mathcal{N})^{-1}$) and $\mathcal{P}^*$ is the back-projection operator (the adjoint of the projection operator, $\mathcal{P}$).

Determination of the residual in projected attenuation, $A_\varphi(\mathbf{x}) - A^k_\varphi(\mathbf{x})$, depends on the relevant assumptions and simplifications. In our simulations, we have utilised the weak-absorption approximation, i.e., $T_\varphi(\mathbf{x}) = \exp[-A_\varphi(\mathbf{x})] \approx 1 - A_\varphi(\mathbf{x})$, yielding:

$$A_{\varphi_l}(\mathbf{x}) - A^k_{\varphi_l}(\mathbf{x}) = \mathcal{C}_{\varphi_l}\{\mathcal{C}^*_{\varphi_j}[1 - \mathcal{P}_{\varphi_l}\mu(\mathbf{r})] - B_j\}. \tag{15}$$

Noting that the Radon transform is linear, we can re-write Eq. (14) in the form discussed at the start of this section, following Eq. (13):

$$\mu^{k+1}(\mathbf{r}) = \mu^k(\mathbf{r}) + \beta\mathcal{P}^*\mathcal{C}\{\mathcal{C}^*_{\varphi_j}[1 - \mathcal{P}_{\varphi_l}\mu(\mathbf{r})] - B_j\}. \tag{16}$$

Without the weakly-attenuating approximation, the required correction in projected attenuation is determined as follows:

$$A_{\varphi_l}(\mathbf{x}) - A^k_{\varphi_l}(\mathbf{x}) = \log(\mathcal{C}_{\varphi_l}B_j) - \log\{\mathcal{C}_{\varphi_l}\mathcal{C}^*_{\varphi_j}\exp[-\mathcal{P}_{\varphi_l}\mu(\mathbf{r})]\}. \tag{17}$$

We will first compare the performance of this approach to tomography with that of the two-step approach from the previous section. We will then explore the effects of dose-fractionation for direct XGT from bucket values.

A. Comparison With Naive Two-Step Ghost-Tomography

256 iterations of the above direct tomography algorithm were applied to the simulated data with $L = 90$ azimuthal viewing angles and $N = 1000$, 2000, 3000, and 4000 measured bucket values per view-angle. The corresponding normalised root mean squared error (RMSE) in the estimated bucket values after these 256 iterations is 4.63, 5.72, 6.24, and 6.44 ($\times 10^{-3}$). Normalisation in this case means scaled by the maximum bucket signal ($3 \times \frac{4}{3}\pi 6^3$). Each image appeared very similar, as indicated by the similar residuals. The result for $N = 1000$ is presented in Fig. 10c. For comparison, 16 CGXC iterations followed by 256 SIRT iterations was also performed for each scenario with and without image-space sparsity regularisation. The results for $N = 1000$ are presented in Fig. 10a, b for inspection. The power of direct tomography from the measured bucket values, with regard to the suppression of noise-like artefacts, can be observed in this figure: there are less artefacts observed in the under-constrained cases ($N = \{1000, 2000\}$) without any *a posteriori* assertions. However, there is a trade-off with computation time: the higher RMSE in the more well constrained problems

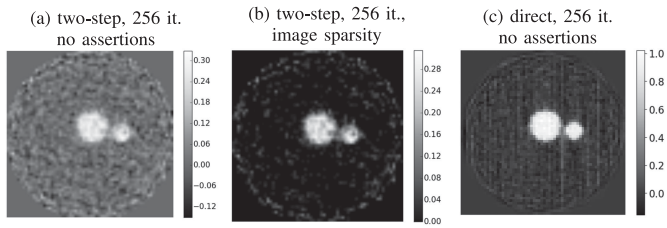

Fig. 10. Tomogram slice $r_2 = 18$ generated from measured bucket values with 90 equally spaced azimuthal viewing angles with 1000 bucket values per view-angle. Reconstruction performed using (a) the two-step approach with 16 CGXC followed by 256 SIRT iterations, (b) as in (a) including regularisation asserting image-space sparsity, and (c) 256 iterations of the direct approach.

($N = \{3000, 4000\}$) is due to the direct algorithm converging more slowly; high-frequency terms are still converging after 256 iterations. Each step of the direct algorithm must satisfy $L\times$ more constraints and thus resulting step sizes are smaller and convergence is slower. An ordered-subset form of direct tomography should converge more quickly since it has exactly the opposite effect.

Improved results could also be achieved in fewer tomographic reconstruction iterations by performing multiple IXC iterations before estimating the residual to be back-projected. However, it is IXC that dominates the computational complexity and not tomographic reconstruction. IXC requires $O(\mathcal{N}^5)$ operations since it involves correlating $N = O(\mathcal{N}^2)$ bucket values with $\mathcal{N}^2$ pixel illumination images over $L = O(\mathcal{N})$ angles. The (back-) projection operations involved in tomographic reconstruction require only $O(\mathcal{N}^4)$ operations with each of $\mathcal{N}^3$ voxels being (back-)projected over L angles. Thus there is little difference in computational cost between 256 iterations of the direct method and the two-step method with 256 IXC iterations (e.g., with regularisation) with FBP.

B. Dose Fractionation

Let us revisit the dose fractionation question as in Section IV. We limit the total number of bucket measurements to $J = 30,000$ with $L = 90$, 30, 15, 10, and 7 (and corresponding measurements per view-angle $N = 333$, 1000, 2000, 3000, and 4000). We have included 90 viewing angles (with $N = 333$) to explore the limits of dose fractionation since a minimum amount of information per view-angle is no longer required (and ideally $L = \pi\mathcal{N}/2 \approx 100$). We see in Fig. 11a, b that the $N = 333$ case is dominated by artefacts using the two-step approach. (Note that here IXC and CGXC could not improve on the results of XC). The corresponding normalised RMSE residuals after 256 direct tomography iterations were 2.65, 2.38, 3.29, 5.10, and 7.17 ($\times 10^{-3}$). The results for $N = 333$, 1000, and 4000 are presented in Fig. 11c, d, e.

Comparing the quality of these results with the two-step results in Fig. 11b and Fig. 8c, f as well as for regularised SIRT in Fig. 9, the power of direct ghost-tomography, which utilises all measured data simultaneously, is again evident since superior results are produced with no *a posteriori* assertions.

With regard to dose fractionation, we again observed that using too few bucket-values per view-angle resulted in noise-like

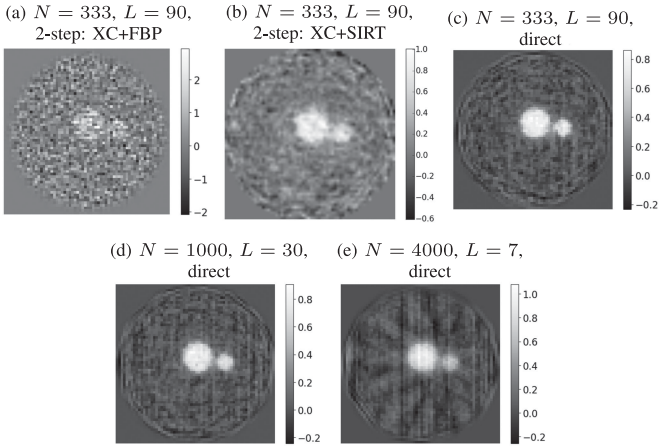

Fig. 11. Tomogram slice $r_2 = 18$ generated from $J = LN \approx 30,000$ measured bucket values. Reconstruction performed from N bucket values per view-angle using (a) two-step tomography: XC GI with FBP, (b) two-step tomography: XC GI with 256 SIRT iterations, and (c)–(e) 256 direct tomography iterations.

artefacts, and conversely, too few view-angles results in streaking artefacts. The best results in this study, limiting the total number of bucket measurements to 30,000, correspond to $L \geq 30$ equally-spaced azimuthal views, each of which have $N \leq 1000$ bucket measurements. However, we recall that these are noise-free simulations; when performing actual experiments, a trade-off would be recommended of $N = 2000$ with $L = 15$. The regularisation techniques explored in Section IV would also help in the case of noise.

A virtue of the direct ghost-tomography approach developed here is its simultaneous use of the whole dataset, rather than first analysing separate subsets of the data for each object orientation as in the two-step approach. We expect such whole-of-dataset algorithms to be more effective when dealing with real experimental data and in future XGT developments.

VI. Mask Type and Mask Scanning Considerations

Thus far we have utilised spatially random masks in all simulations. In certain contexts, such as when using x-ray shot noise from individual charged-particle bunches to generate speckle fields in a GI setting [25], randomness is assured. However, in many experimental circumstances one will have the choice between random and coded masks and they will be varied with a scanning procedure. We suspect that, in analogy with what is done with spatial light modulators [8], [9], appropriately designed coded apertures will typically perform better than random masks, on account of the fact that the corresponding coded-aperture illumination patterns can be designed to form a strictly orthogonal basis [11], [41].

The random masks used thus far correspond to illuminating intensity distributions given by an ensemble of $J = LN$ random $\mathcal{N} \times \mathcal{N}$ matrices, each element being an independent deviate drawn from the same uniform probability distribution. Generalisations include different probability distributions such as normal or Poisson processes, and/or introducing coupling between adjacent pixels via convolution with a suitable discrete convolution kernel [41].

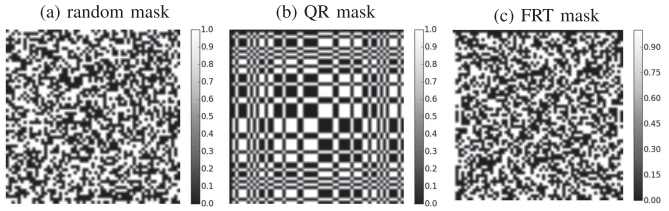

Fig. 12. Examples of types of binary mask explored here. (a) Random mask. (b) Quadratic residue (QR) mask. (c) Finite Radon transform (FRT) based mask.

It is important, in the context of an investigation of the functional form adopted for the illuminating masks, to note that the standard XC ghost formula in Eq. (6) considers the set of linearly independent illumination patterns to form a complete or near-complete (indeed, over-complete if $J > \mathcal{N}^2$) mathematical basis for $\mathcal{N}^2$-pixel images. Moreover, there is an implicit assumption, in this formula, that the set of background-subtracted illumination patterns $\{I_j(\mathbf{x}) - \overline{I}\}$ approximate an orthogonal basis [28], [41]; here, $\overline{I}$ denotes the spatial average of the intensity for each illuminating pattern, with this spatial average assumed to be approximately the same for all illuminating patterns. This observation leads one to consider non-random orthogonal (or near-orthogonal) masks [11]. This is the topic of the present section. We consider only orthogonality per view-angle; global orthogonality (over the set of all viewing angles) is a topic for future research.

An infinite multiplicity of orthogonal masks could be devised for reconstructing pixelated arrays of a specified size [11]. Since such masks are deterministic, if their associated transmission functions $I_j(\mathbf{x})$ are sufficiently well known then the detector $\mathcal{D}$ in Fig. 1 may be eliminated altogether. In the simulations presented below, we compare the previously considered random masks, with two classes of non-random mask (coded aperture) that have near-perfect autocorrelation modified using a scanning procedure. We then examine their relative performance in (2D) X-ray GI and (3D) XGT. We examine an issue that arises, pertaining to ring artefacts, and consider a means to avoid such artefacts in practice.

A. Coded-Aperture Generation

Existing practical applications of X-ray GI do not use a random set of masks, each of which is independently generated for each bucket reading. Rather, a single mask is scanned over the sample [26]–[28]. This single mask, be it random or deterministic, is transversely displaced between bucket readings. Here we assume the masks are periodic in y_1/y_2, with respective periods of $\mathcal{N}\Delta_{y_1}$ and $\mathcal{N}\Delta_{y_2}$, hence for an $\mathcal{N} \times \mathcal{N}$ mask there are $\mathcal{N}^2$ possible unique bucket measurements. A mask with perfect auto-correlation [43], i.e., zero at all values except the origin, produces an orthogonal basis in this case. There are many ways to produce such masks. We investigate two methods: (i) modified uniformly redundant array constructed using quadratic residues [51] (Fig. 12b); (ii) a method based on the finite Radon transform (FRT) [43] (Fig. 12c). We compare their performance against a random mask (Fig. 12a), in both GI and XGT, in the following subsections.

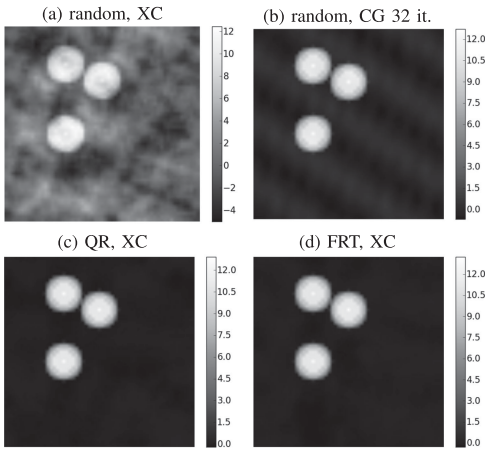

Fig. 13. Ghost image recovery using all 3481 mask positions for (a) random mask by XC, (b) random mask by 32 CGXC iterations, (c) quadratic residue mask by XC, and (d) finite Radon transform based mask by XC.

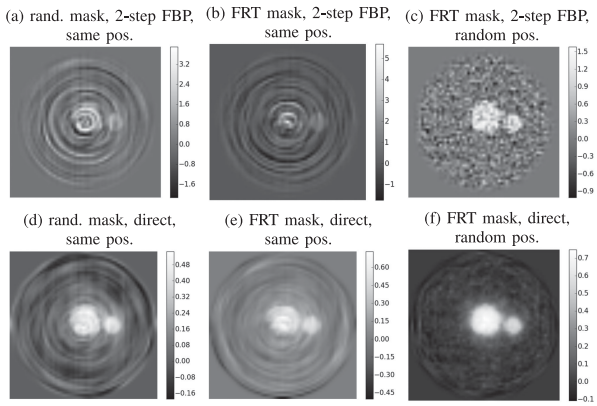

Fig. 14. Tomogram slice $r_2 = 16$ measured with scanning mask of type (a) and (d) random, or (b) and (e) FRT-based, and translated to the *same* random set of $N = 870$ positions per view-angle. (a) and (b) Generated by the two-step method using XC and FBP. (d) and (e) Generated directly from bucket values using 100 iterations of Eq. 14. (c) and (f) are the same as (b) and (d) but with a *different* set of random positions per view-angle.

These masks are constructed for $\mathcal{N} = 59$, giving 3481 elements. This number is odd, so we cannot have $\{-1, 1\}$ cancelling to precisely give zeros in the auto-correlation, hence we will have *near*-perfect auto-correlations equal to ± 1 away from the origin. In fact, since intensity maps must be non-negative and intensity transmission masks are restricted to values between zero and unity, we will not have zero cross-correlations at all, but rather $\mu\sigma^2\mathcal{N}^2$ where μ is the mean, σ^2 is the variance, and $\mathcal{N}^2$ is the mask size. They can be made orthogonal if the mean is subtracted, an important requirement in the context of GI [11], [25], [28], [41].

Auto-correlations generated by the masks in Fig. 12 appear similar with a peak value of 1741. The principal difference is that the random mask off-origin correlations have a range of 100 while the coded-apertures only have a range of 1. These auto-correlations are proportional to the point-spread function of each mask set [11], [28], [52]. They are not perfect since they are not precisely proportional to a spatial Kronecker delta at the origin, however deviation is mild.

B. Performance of Coded and Random Masks in Ghost Imaging

1) Ghost Imaging Performance: Given masks with an orthogonal basis (after average-subtraction is applied), XC should give the exact inverse from the full complement of $\mathcal{N}^2$ noise-free bucket values. The results of XC GI are presented in Fig. 13 for the random mask, quadratic-residue mask and FRT-based coded mask. Being nearly orthogonal, coded masks perform significantly better than random masks which exhibit artefacts with long range correlation. The result of applying 32 CGXC iterations to the random mask data is also presented in Fig. 13b; image quality approaches that using XC with coded apertures, at significant computational cost, however the MAD is still $5\times$ larger. Since the coded aperture masks are nearly orthogonal, IXC achieves little over XC alone.

A similar trend is observed for reduced input data, i.e., $N < \mathcal{N}^2$. We investigated performance of each mask type given $N =$ 870, 1740, and 2610 bucket measurements. After 32 CGXC iterations applied to the random mask data, both the residual and MAD is reduced, however, the errors are still larger than that for XC, given either of the coded-aperture data, and has more low spatial-frequency components.

We have observed that greater efficiency is achieved when the background-subtracted illuminating intensity maps are strictly orthogonal rather than merely orthogonal in expectation value [11] (as is the case for random masks [41]). These noise-free simulations indicate a slight advantage in using specific coded-apertures over randomly generated apertures; artefacts contain no long-range correlations, and XC is sufficient for image recovery. This is consistent with the previously cited theoretical studies [11], [41]. We expect this advantage to be more significant given noisy data from experiment, since data collected with an orthogonal coded-aperture is more well-conditioned and ghost recovery is more stable.

2) Ghost Tomography Performance: Recall that we have generated a set of masks that are orthogonal per view-angle, i.e., orthogonal for radiography, not tomography. The advantages discovered in Section VI-B1 for ghost radiography will not necessarily apply to XGT in general; we explore that in this section. Tomographic reconstructions were performed in two ways: (i) FBP from XC recovered projections, and (ii) 100 iterations of the direct tomography algorithm developed in Section V. Reconstructions were performed on simulated data with $L = 90$ viewing angles and $N = 870$, 1740, 2610, and 3481 bucket measurements. The results for $N = 870$ are in Fig. 14. Note that here we have only presented results for the FRT-based mask (the quadratic residue mask gave similar results) and we have included results from a random mask for comparison.

The tomogram slices in Fig. 14a, b are dominated by ring artefacts. Tomogram quality degradation due to these rings increases with reduced number of measurements, N. Apart from these artefacts we see that, as for ghost imaging in Section VI-B1, errors from the random mask have a lower spatial-frequency

component in Fig. 14a (cf. tomograms obtained using the perfect coded arrays, in which such low-frequency artefacts are absent, Fig. 14b). Moreover, Fig. 14c, d shows that the direct tomography algorithm, regularised using 100 iterations, can reduce these errors and the ring artefacts.

C. Ring Artefacts

Ring artefacts arise since the simulated data has utilised the same set of mask positions for each projection angle. In *compressive tomography*, this would equate to using the same mask at each viewing-angle, and is an immediately obvious cause of ring-artefacts, since the same set of pixels record no information about the sample at all angles [36]. In XGT however, we use many masks per viewing-angle and every *pixel* of the ghost projection-image, that is to be recovered, has been measured for a large fraction of those masks (although it is integrated with the contribution from many other *pixels*). Instead of having *pixels* with no information, we have *pixels* with varying degrees of statistical certainty. It is not intuitive that ring artefacts would ensue from this, but clearly they can (particularly using XC for projection GI).

To demonstrate this cause of the ring artefacts, we generated datasets with the same number of bucket measurements per view-angle (as for the above demonstration), but with the mask positions selected *randomly* per viewing-angle. Results are presented in Fig. 14 c,f for the FRT-based mask using both the XC+FBP two-step, and the XC-SIRT direct reconstruction methods given the same number of measurements as for Fig. 14 b,e, i.e., $L = 90$ and $N = 870$.

The superiority of these simulations suggest that, experimentally, a different set of masks should be used per viewing-angle if possible; it is an effective means for suppressing ring artefacts in XGT. If a regular scanning grid pattern per view-angle is used in the experiment protocol, a randomised start position per view-angle is recommended. In the longer-term progress of the new field of ghost-tomography, development of more sophisticated ring-artefact-removal algorithms may alter this recommendation, particularly in contexts where *a priori* knowledge about the class of imaged samples may be employed. A (near) orthogonal coded aperture causes GI to be more robust since XC is the (well-conditioned) inverse.

VII. Discussion and Future Research

One clear, albeit unsurprising, message of the current study is the fact that considerable improvements may be obtained in X-ray GI, by going beyond the conventional XC reconstruction formula given in Eq. (6). The XC method is conceptually appealing, since it can be trivially derived from first principles by considering the ensemble of background-subtracted illuminating fields to form an (approximately) orthogonal set [28]. For cases where orthogonal coded-aperture masks are used, XC is sufficient. However, in all other cases XC is incomplete and requires iterative refinement methods such as IXC, CGXC, RXC, etc. We suspect that, in the future, refinement methods based on machine learning and artificial intelligence [53], [54] will become of progressively greater importance for X-ray GI, together with the strides being made in the burgeoning field of compressive sensing [48]. This will likely be a key avenue for future research, driven primarily by the quest for improved reconstructions using a minimal number of probe photons.

This leads to the question of whether or not X-ray GI may enable reduced dose relative to competing protocols [27]. The answer at this stage is *maybe*, with inequalities having been developed which, if violated, imply GI can achieve reduced dose relative to its direct-imaging counterpart [11], [41]. The logical possibility that GI may reduce dose is demonstrated by the following example: Suppose one performs an X-ray GI experiment of a sample whose transmission function is a greyscale map of Leonardo da Vinci's Mona Lisa, using a *single* weak illuminating intensity field that is also a greyscale map of the Mona Lisa. The resulting reconstruction (proportional to the illuminating intensity multiplied by the single bucket measurement) is generated with minimal dose. This contrived example is indicative of the more general result that X-ray GI may reduce dose, e.g., in cases where "the class of imaged objects is strongly correlated with a small number of illumination patterns" [11]. It remains unclear whether, in the longer term, X-ray GI can reduce dose in scenarios where sufficiently low doses are important (e.g., radiation damage to living tissues or other biological/radiation-sensitive materials, efficiently small acquisition time in an industrial-testing context, etc.). Relevant to this context is the ability of a low-resolution low-dose ghost reconstruction to locate a region of interest (ROI) in a sample, so that subsequent illumination patterns can be adapted by confining them to this ROI [55].

The extent to which one can use polychromatic radiation (allowing higher flux) for X-ray GI, forms another avenue for future research. Nowhere in our development did we explicitly refer to the coherence of the beam. For example, computational X-ray GI with a laboratory source has already been reported experimentally with a fully polychromatic beam [26], [27]. For 3D XGT in this context, one could first perform 2D X-ray GI for each orientation of the object, and then correct for the beam hardening [56] (naive two-step tomography approach). Beam-hardening effects can also be modelled in the forward process of the direct tomographic reconstruction approach, with an estimate of the polychromatic attenuation refined during iteration, e.g., [57], [58].

Another avenue for future work is prompted by the advances made in X-ray imaging in recent decades, due to X-ray phase contrast [59]. Through various incarnations including propagation-based X-ray phase contrast, analyzer-crystal phase contrast, grating-based phase contrast, speckle-tracking X-ray phase contrast, etc. [60], the harnessing of phase contrast in an X-ray imaging setting has led to significant increases in image contrast and resolution, as well as protocols for the (often rather substantial) reduction of sample dose [61]. The method devised by Paganin *et al.* [62] has been particularly successful in this regard, with acquisition-time-reductions in the tens of thousands being typical [61], [63]–[66]. While phase contrast has been incorporated into visible-light GI protocols through the use of interferometers [18], [67]–[69], these are difficult to translate to an X-ray setting. Indeed, the standard X-ray

GI setup is totally insensitive to any refractive effects (phase shifts) imparted by the sample since such effects have no influence on the total number of photons registered in each bucket measurement. A protocol has recently been proposed for X-ray phase contrast ghost imaging [41], which belongs to the class of signal-to-noise-ratio (SNR)-boosting X-ray phase retrieval algorithms [61], [70] epitomised by the Paganin method. Such phase-contrast X-ray GI protocols, particularly in light of the SNR-boosting property, may form an interesting avenue for future developments of phase-contrast XGT. Such developments would also be relevant to the preceding discussion regarding dose reduction.

VIII. Conclusion

Some key considerations regarding experimental realisation of X-ray ghost-tomography have been presented. Experimental protocols and data reconstruction techniques have been explored by numerical simulation using a very simple case study (as would be considered for a first experiment). Results suggest that coded-aperture masks that form an orthogonal basis are preferable to random masks for generating patterned illumination. The set of masks used per view-angle should be different if possible to avoid ring artefacts. The conventional cross-correlation method for ghost-imaging can be insufficient for ghost projection imaging (given a set of random masks); iterative refinement may be required for effective image recovery; regularisation, or *a posteriori* assertions, are essential for highly under-constrained situations. Optimal dose fractionation seems similar to that for tomography by conventional imaging. A tomographic reconstruction scheme that applies directly to the bucket measurements was designed and shown to make more effective use of the data, than the naive two-step method, due to its holistic approach. This method combined with regularisation, although more computationally costly, is recommended for experimental realisation of XGT. Possibilities regarding dose reduction were also discussed.

Acknowledgment

The authors would like to thank Dr. W. M. (Steve) Lee and Dr. G. Scarcelli for assistance with the literature survey. They acknowledge useful discussions with D. Ceddia, M. P. Olbinado, T. C. Petersen, A. Rack, and T. P. Simula.

References

[1] D. N. Klyshko, "Effect of focusing on photon correlation in parametric light scattering," *Sov. Phys. J.E.T.P.*, vol. 67, no. 6, pp. 1131–1135, 1988.
[2] A. V. Belinskiĭ and D. N. Klyshko, "Two-photon optics: diffraction, holography, and transformation of two-dimensional signals," *Sov. Phys. J.E.T.P.*, vol. 78, no. 3, pp. 259–262, 1994.
[3] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. Sergienko, "Optical imaging by means of two-photon quantum entanglement," *Phys. Rev. A*, vol. 52, no. 5, pp. R3429–R3432, 1995.
[4] D. V. Strekalov, A. V. Sergienko, D. N. Klyshko, and Y. H. Shih, "Observation of two-photon 'ghost' interference and diffraction," *Phys. Rev. Lett.*, vol. 74, no. 18, pp. 3600–3603, 1995.
[5] R. I. Khakimov *et al.*, "Ghost imaging with atoms," *Nature*, vol. 540, pp. 100–103, 2017.
[6] Y. Shih, "The physics of ghost imaging," in *Classical, Semi-Classical and Quantum Noise*. New York, NY, USA: Springer, 2012, pp. 169–222.
[7] M. J. Padgett and R. W. Boyd, "An introduction to ghost imaging: Quantum and classical," *Phil. Trans. R. Soc. A*, vol. 375, no. 2099, 2017, Art. no. 20160233.
[8] Y. Bromberg, O. Katz, and Y. Silberberg, "Ghost imaging with a single detector," *Phys. Rev. A*, vol. 79, 2009, Art. no. 053840.
[9] O. Katz, Y. Bromberg, and Y. Silberberg, "Compressive ghost imaging," *Appl. Phys. Lett.*, vol. 95, no. 13, 2009, Art. no. 131110.
[10] B. I. Erkmen and J. H. Shapiro, "Ghost imaging: From quantum to classical to computational," *Adv. Opt. Photon.*, vol. 2, no. 4, pp. 405–450, 2010.
[11] T. E. Gureyev, D. M. Paganin, A. Kozlov, Y. I. Nesterets, and H. M. Quiney, "On the efficiency of computational imaging with structured illumination," *Phys. Rev. A*, vol. 97, 2018, Art. no. 053819.
[12] J. H. Shapiro, "Computational ghost imaging," *Phys. Rev. A*, vol. 78, no. 6, 2008, Art. no. 061802.
[13] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, "Ghost imaging using optical correlations," *Laser Photon. Rev.*, vol. 12, 2018, Art. no. 1700143.
[14] R. I. Stantchev *et al.*, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," *Sci. Adv.*, vol. 2, no. 6, 2016, Art. no. e1600190.
[15] M. F. Duarte *et al.*, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.
[16] R. Borghi, F. Gori, and M. Santarsiero, "Phase and amplitude retrieval in ghost diffraction from field-correlation measurements," *Phys. Rev. Lett.*, vol. 96, 2006, Art. no. 183901.
[17] W. Gong and S. Han, "Phase-retrieval ghost imaging of complex-valued objects," *Phys. Rev. A*, vol. 82, 2010, Art. no. 023828.
[18] T. Shirai, T. Setälä, and A. Friberg, "Ghost imaging of phase objects with classical incoherent light," *Phys. Rev. A*, vol. 84, no. 4, 2011, Art. no. 041801.
[19] F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla, and L. A. Lugiato, "High-resolution ghost image and ghost diffraction experiments with thermal light," *Phys. Rev. Lett.*, vol. 94, 2005, Art. no. 183602.
[20] H. Kellock, T. Setälä, T. Shirai, and A. T. Friberg, "Higher-order ghost imaging with partially polarized classical light," *Proc. SPIE*, vol. 8171, 2011, Art. no. 81710Q.
[21] Y. Zhu, J. Shi, Y. Yang, and G. Zeng, "Polarization difference ghost imaging," *Appl. Opt.*, vol. 54, no. 6, pp. 1279–1284, 2015.
[22] Y. Liu, J. Shi, and G. Zeng, "Single-photon-counting polarization ghost imaging," *Appl. Opt.*, vol. 55, no. 36, pp. 10 347–10 351, 2016.
[23] M. Lyu *et al.*, "Deep-learning-based ghost imaging," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 17865.
[24] H. Yu *et al.*, "Fourier-transform ghost imaging with hard X rays," *Phys. Rev. Lett.*, vol. 117, no. 11, 2016, Art. no. 113901.
[25] D. Pelliccia, A. Rack, M. Scheel, V. Cantelli, and D. M. Paganin, "Experimental X-ray ghost imaging," *Phys. Rev. Lett.*, vol. 117, no. 11, 2016, Art. no. 113902.
[26] A. Schori and S. Schwartz, "X-ray ghost imaging with a laboratory source," *Opt. Express*, vol. 25, no. 13, pp. 14822–14828, 2017.
[27] A.-X. Zhang, Y.-H. He, L.-A. Wu, L.-M. Chen, and B.-B. Wang, "Tabletop X-ray ghost imaging with ultra-low radiation," *Optica*, vol. 5, pp. 374–377, 2018.
[28] D. Pelliccia, M. P. Olbinado, A. Rack, A. M. Kingston, G. R. Myers, and D. M. Paganin, "Towards a practical implementation of X-ray ghost imaging with synchrotron light," *IUCrJ*, vol. 5, no. 4, pp. 428–438, 2018.
[29] A. Schori, D. Borodin, K. Tamasaku, and S. Shwartz, "Ghost imaging with paired X-ray photons," *Phys. Rev. A*, vol. 97, Jun. 2018, Art. no. 063804.
[30] A. M. Kingston *et al.*, "Ghost tomography," *Optica*, vol. 5, no. 12, pp. 1516–1520, 2018.
[31] T. Mohr, A. Herdt, and W. Elsässer, "2D tomographic terahertz imaging using a single pixel detector," *Opt. Express*, vol. 26, no. 3, pp. 3353–3367, 2018.
[32] M. B. Nasr, B. E. A. Saleh, A. V. Sergienko, and M. C. Teich, "Demonstration of dispersion-canceled quantum-optical coherence tomography," *Phys. Rev. Lett.*, vol. 91, no. 8, 2003, Art. no. 083601.
[33] M.-J. Sun *et al.*, "Single-pixel three-dimensional imaging with time-based depth resolution," *Nat. Commun.*, vol. 7, 2016, Art. no. 12010.
[34] K. Choi and D. Brady, "Coded aperture computed tomography," *Proc. SPIE*, vol. 7468, no. 2009, 2009, Art. no. 74 680B.
[35] D. J. Brady, A. Mrozack, K. MacCabe, and P. Llull, "Compressive tomography," *Adv. Opt. Photon.*, vol. 7, no. 4, pp. 756–813, 2015.
[36] Y. Kaganovsky *et al.*, "Compressed sampling strategies for tomography," *J. Opt. Soc. Amer. A*, vol. 31, no. 7, pp. 1369–1394, 2014.
[37] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. Piscataway, NJ, USA: IEEE Press, 1988.

[38] P. Gilbert, "Iterative methods for the three-dimensional reconstruction of an object from projections," *J. Theor. Biol.*, vol. 36, no. 1, pp. 105–117, 1972.
[39] R. A. Crowther, D. J. DeRosier, and A. Klug, "The reconstruction of a three-dimensional structure from projections and its application to electron microscopy," *Proc. Roy. Soc. London A*, vol. 317, no. 1530, pp. 319–340, Jun. 1970.
[40] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[41] D. Ceddia and D. M. Paganin, "Random-matrix bases, ghost imaging and X-ray phase contrast computational ghost imaging," *Phys. Rev. A*, vol. 97, 2018, Art. no. 062119.
[42] E. E. Fenimore and T. M. Cannon, "Coded aperture imaging with uniformly redundant arrays," *Appl. Opt.*, vol. 17, no. 3, pp. 337–347, 1978.
[43] B. Cavy and I. Svalbe, "Construction of perfect auto-correlation arrays and zero cross-correlation arrays from discrete projections," in *Combinatorial Image Analysis: 17th International Workshop, IWCIA 2015 Kolkata, India, November 24-27 2015 Proceedings* (Lecture Notes in Computer Science 9448), R. Barneva, B. Bhattacharya, and V. Brimkov, Eds. New York, NY, USA: Springer, 2015, pp. 232–243.
[44] H. Huang, C. Zhou, T. Tian, D. Liu, and L. Song, "High-quality compressive ghost imaging," *Opt. Commun.*, vol. 412, pp. 60–65, 2018.
[45] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*. 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[46] L. Bian, J. Suo, Q. Dai, and F. Chen, "Experimental comparison of single-pixel imaging algorithms," *J. Opt. Soc. Amer. A*, vol. 35, no. 1, pp. 78–87, 2018.
[47] V. Katkovnik and J. Astola, "Compressive sensing computational ghost imaging," *J. Opt. Soc. Amer. A*, vol. 29, no. 8, pp. 1556–1567, 2012.
[48] S. Qaisar, R. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: From theory to applications, a survey," *J. Commun. Netw.*, vol. 15, no. 5, pp. 443–456, 2013.
[49] X.-R. Yao *et al.*, "Iterative denoising of ghost imaging," *Opt. Express*, vol. 22, no. 20, pp. 24 268–24 275, 2014.
[50] R. Hegerl and W. Hoppe, "Influence of electron noise on three-dimensional image reconstruction," *Z. Naturforsch.*, vol. 31, pp. 1717–1721, 1976.
[51] S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Appl. Opt.*, vol. 28, no. 20, pp. 4344–4352, Oct. 1989.
[52] F. Ferri, D. Magatti, L. A. Lugiato, and A. Gatti, "Differential ghost imaging," *Phys. Rev. Lett.*, vol. 104, 2010, Art. no. 253603.
[53] Z. D. C. Kemp, "Propagation based phase retrieval of simulated intensity measurements using artificial neural networks," *J. Opt.*, vol. 20, no. 4, 2018, Art. no. 045606.
[54] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.*, vol. 7, 2018, Art. no. 17141.
[55] S. Sun *et al.*, "Multi-scale adaptive computational ghost imaging," *Sci. Rep.*, vol. 6, 2016, Art. no. 37013.
[56] R. A. Brooks and G. Di Chiro, "Beam hardening in X-ray reconstructive tomography," *Phys. Med. Biol.*, vol. 21, no. 3, pp. 390–398, 1976.
[57] M. Krumm, S. Kasperl, and M. Franz, "Reducing non-linear artifacts of multi-material objects in industrial 3D computed tomography," *N.D.T. E. Int.*, vol. 41, no. 4, pp. 242–251, 2008.
[58] G. Van Gompel *et al.*, "Iterative correction of beam hardening artifacts in CT," *Med. Phys.*, vol. 38, no. 7, pp. S36–S49, 2011.
[59] S. W. Wilkins, Y. I. Nesterets, T. E. Gureyev, S. C. Mayo, A. Pogany, and A. W. Stevenson, "On the evolution and relative merits of hard X-ray phase-contrast imaging methods," *Phil. Trans. R. Soc. A*, vol. 372, no. 2010, 2014, Art. no. 20130021.
[60] D. M. Paganin, *Coherent X-ray Optics*. London, U.K.: Oxford Univ. Press, 2006.
[61] M. J. Kitchen *et al.*, "CT dose reduction factors in the thousands using X-ray phase contrast," *Sci. Rep.*, vol. 7, 2017, Art. no. 15953.
[62] D. Paganin, S. C. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins, "Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object," *J. Microsc.*, vol. 206, no. 1, pp. 33–40, 2002.
[63] M. A. Beltran, D. M. Paganin, K. Uesugi, and M. J. Kitchen, "2D and 3D X-ray phase retrieval of multi-material objects using a single defocus distance," *Opt. Express*, vol. 18, no. 7, pp. 6423–6436, 2010.
[64] M. A. Beltran *et al.*, "Interface-specific X-ray phase retrieval tomography of complex biological organs," *Phys. Med. Biol.*, vol. 56, no. 23, pp. 7353–7369, 2011.
[65] Y. I. Nesterets and T. E. Gureyev, "Noise propagation in X-ray phase-contrast imaging and computed tomography," *J. Phys. D, Appl. Phys.*, vol. 47, no. 10, 2014, Art. no. 105402.
[66] T. E. Gureyev *et al.*, "Investigation of the imaging quality of synchrotron-based phase-contrast mammographic tomography," *J. Phys. D, Appl. Phys.*, vol. 47, no. 36, 2014, Art. no. 365401.
[67] B. Jack *et al.*, "Holographic ghost imaging and the violation of a Bell inequality," *Phys. Rev. Lett.*, vol. 103, no. 8, 2009, Art. no. 083602.
[68] W. Gong and S. Han, "Phase-retrieval ghost imaging of complex-valued objects," *Phys. Rev. A*, vol. 82, no. 2, 2010, Art. no. 023828.
[69] L. Martínez-León, P. Clemente, Y. Mori, V. Climent, J. Lancis, and E. Tajahuerce, "Single-pixel digital holography with phase-encoded illumination," *Opt. Express*, vol. 25, no. 5, pp. 4975–4984, 2017.
[70] T. E. Gureyev, Y. I. Nesterets, A. Kozlov, D. M. Paganin, and H. M. Quiney, "On the 'unreasonable' effectiveness of transport of intensity imaging and optical deconvolution," *J. Opt. Soc. Amer. A*, vol. 34, no. 12, pp. 2251–2260, 2017.

Authors' photographs and biographies not available at the time of publication.