

# Random-matrix bases, ghost imaging, and x-ray phase contrast computational ghost imaging

David Ceddia\* and David M. Paganin

*School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia*



(Received 14 February 2018; published 19 June 2018)

A theory of random-matrix bases is presented, including expressions for orthogonality, completeness, and the random-matrix synthesis of arbitrary matrices. This is applied to ghost imaging as the realization of a random-basis reconstruction, including an expression for the resulting signal-to-noise ratio. Analysis of conventional direct imaging and ghost imaging leads to a criterion which, when satisfied, implies a reduced dose for computational ghost imaging. We also propose an experiment for x-ray phase contrast computational ghost imaging, which enables differential phase contrast to be achieved in an x-ray ghost-imaging context. We give a numerically robust solution to the associated inverse problem of decoding differential phase contrast x-ray ghost images, to yield a quantitative map of the projected thickness of the sample.

DOI: [10.1103/PhysRevA.97.062119](https://doi.org/10.1103/PhysRevA.97.062119)

## I. INTRODUCTION

Conventional direct imaging is typically realized using position-sensitive detectors such as those that are currently used in digital photography, x-ray radiography, neutron imaging, electron microscopy, etc. Ghost imaging [1,2] is an emerging alternative in which the conventional direct-imaging methodology is replaced with an indirect-imaging process which synthesizes an image given a series of known illuminations and associated position-insensitive “bucket” detector measurements.

In its simplest form, ghost imaging synthesizes an image from a set of temporally acquired single-pixel detector (bucket reading) outputs which are each correlated with a random intensity pattern. Explicitly,

$$\mathcal{I}(i, j) = \frac{1}{N} \sum_{k=1}^N (B_k - E[B]) I_k(i, j), \quad (1)$$

where  $\mathcal{I}(i, j)$  denotes the synthesized image,  $N$  is the number of spatially random intensity patterns  $I_k(i, j)$  which is indexed by  $k$ , and  $E[B]$  is the expectation value of the bucket reading

$$B_k = \sum_{i=1}^n \sum_{j=1}^m f(i, j) I_k(i, j) = \langle f, I_k \rangle. \quad (2)$$

Here  $\langle \cdot \rangle = \sum_{i=1}^n \sum_{j=1}^m f(i, j)$  is the sample intensity pattern, and  $(i, j)$  denote the pixel coordinates [3].

Ghost imaging originated in visible-light studies [4–11]. The field has been advanced by incorporating compressive sensing ideas [12–15] and studies of turbulent environmental robustness such as lidar systems [16], atmospheric turbulence in thermal imaging [17], and underwater fluctuations in optical imaging [18]. Visible-light ghost imaging has been augmented to achieve absorption-contrast ghost imaging using x rays [19–23] and atoms [24]. The potential for reduced dose

[22], increased resolution, compressive sensing capabilities [12], and turbulence robustness [25] are all under ongoing investigation.

The underlying concept of ghost imaging has been pointedly distilled to “a vector projection of the [sample] transmission function over  $[N]$  different random vectors” [12] or “a form of orthogonal [random] function expansion” [23]. We build on these ideas to develop a mathematical theory of random-matrix bases, with particular focus on the inherent noise associated with such a reconstruction. We develop an inequality which, if satisfied, implies that computational ghost imaging exhibits reduced dose when compared to its direct imaging counterpart. If this inequality is not satisfied, this implies that alternative reconstruction strategies, including but not limited to those based on the concept of compressive sensing, are necessary for ghost imaging to reduce dose in comparison to direct imaging.

Owing to the potential for reduced dose [22], a promising application of ghost imaging is in x-ray medical imaging. As previously mentioned, already realized are experimental setups for x-ray-absorption contrast ghost imaging. Outstanding, however, is a means to realize x-ray phase contrast ghost imaging. In seeking a means to achieve this, we are motivated by the significant advances in x-ray direct imaging that have been enabled through the use of phase contrast [26]. There exist means for phase contrast ghost imaging in the optical regime [27], although the experimental setups are not readily transferable to the x-ray regime. We propose an experimental setup to achieve x-ray differential phase contrast computational ghost imaging. We also consider the associated inverse problem, of inverting the differential phase contrast x-ray ghost image to obtain a quantitative map of the projected thickness of the sample which led to such an image.

We close this introduction with an outline of the remainder of the paper. Section II develops the formalism of using a set of real random  $n \times m$  matrices as a basis over which arbitrary  $n \times m$  matrices may be decomposed. While we view such decompositions in the context of imaging, where the matrix to be decomposed is a pixelated image, the formalism may be applied to a broader range of discretized fields.

\*david.ceddia@gmail.com

Loosely, the procedure may be viewed as synthesizing signals by superposing noise. The statistical orthogonality of the random matrices is treated, together with completeness and an error metric for the statistical errors in the random-matrix decomposition. A comparison is also drawn between directly decomposing a given pixelated image using a random-matrix basis and decomposition using a basis that has been orthonormalized using the Gram-Schmidt process. Section III gives a comparison between direct and computational ghost imaging with respect to fluence or dose. An inequality is developed which, if satisfied, implies a particular ghost-imaging scenario to have reduced dose relative to its direct imaging counterpart. Section IV proposes a means for x-ray phase contrast computational ghost imaging, namely, a means for x-ray computational ghost imaging that is sensitive to the phase of the coherent x-ray field that impinges on the specified x-ray ghost-imaging system. This gives a technique for x-ray ghost imaging that is sensitive to the refractive properties of a sample. A solution is developed to the associated inverse problem of quantitative phase-amplitude reconstruction of the projected thickness of a single-material sample, using x-ray phase contrast ghost imaging. Section V discusses the broader significance of some of our results, together with some future avenues for research. We conclude with Sec. VI.

## II. RANDOM-MATRIX BASES

The infinite monkey theorem [28] states that an infinite random sequence almost surely, i.e., with a probability that is infinitesimally close to unity, contains every possible finite subsequence. In principle, this permits synthesis of an arbitrary sequence via one realization of a random sequence. To say that such a process is extremely inefficient would be an understatement, owing to the impracticably large mean times one needs to wait for a given desired sequence to occur. In an alternative, practical approach, a given discrete random output can be cut into pieces of the size of the desired sequence to form a random basis. Each random basis member can then be superposed in a weighted sum to synthesize a desired signal (or some approximation of it). Given that the present paper works in an imaging context, with particular emphasis on computational imaging and ghost imaging, we will discuss the topic of random bases in the context of a matrix basis. Here each element of the matrix may be associated with a pixel in a Cartesian lattice.

Consider the standard matrix basis set  $\{e_{ij}\}$  that spans  $M_{nm}$  (i.e.,  $nm$  distinct matrices of size  $n \times m$ , each of which has a single element equaling unity, with all other elements being zero). We will define a random basis member as corresponding to the randomly weighted sum of the standard set

$$R_k = R(1,1)e_{11} + \dots + R(n,m)e_{nm}. \quad (3)$$

Here  $k \in [1, N]$ ,  $i \in [1, n]$ ,  $j \in [1, m]$ , and  $R(i, j)$  is the realization of a zero-centered random variable  $R$  drawn from the probability density function  $\text{Pr}(R)$ . We reserve defining an explicit probability density function for  $R$  (e.g., uniform, Gaussian, Poissonian, etc.) and instead leave our results in terms of parameters of that distribution (i.e., expectation value  $E[R]$  and variance  $\text{Var}[R]$ ).

Alternatively, we can work in reverse from the random bases via the definitions

$$\begin{aligned} E[R] &= \lim_{n,m \rightarrow \infty} \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m R_k(i, j) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N R_k(i, j) \end{aligned} \quad (4)$$

and

$$\begin{aligned} \text{Var}[R] &= \lim_{n,m \rightarrow \infty} \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m \{R_k(i, j) - E[R]\}^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \{R_k(i, j) - E[R]\}^2, \end{aligned} \quad (5)$$

where these equalities do not necessarily hold for any finite choice of  $n, m, N$  but are increasingly well approximated with increasing  $n, m, N$ . Random bases defined in this way will display ergodic behavior in expectation. Ergodicity is usually defined as the equality of ensemble and time averages [29]. The related but different concept of ergodicity in expectation refers to the expectation value over a matrix set being equal to the expectation value over the spatial distribution, where the statistics of each will be governed by  $\text{Pr}(R)$ . That does not mean, however, that the spatial average of each realized basis member will be the same, nor will each set average for a given coordinate  $(i, j)$  be the same. In each case these are finite sums that are ultimately subject to variance.

We consider two related cases: (i) synthesis of a desired matrix  $f(i, j)$  using a nonorthogonal discrete random-matrix basis and (ii) synthesis using an orthonormal discrete random basis. The first case is relevant, e.g., to ghost imaging using spatially random shot-noise fields, with the latter case being relevant, e.g., for computational imaging using specifically designed pseudorandom masks. Owing to the generation of a particular  $N$ -member random-matrix basis being a stochastic process, a probabilistic approach is employed. For the nonorthogonal discrete random-basis case, the exact function is retrieved in expectation and noise is quantified in the variance. The orthonormal random basis is achieved via applying the Gram-Schmidt process [30,31] to a given nonorthogonal set of random matrices and forms a more efficient reconstruction method when compared to the first case listed above. It is important to note regarding a particular realized set of  $N$  random matrices, each real element of which is by assumption generated by the same specified probability distribution, that subsequent Gram-Schmidt orthogonalization yields a sequence of random matrices whose statistics in general change with position in the matrix sequence. This implies a slightly different approach for the two cases.

### A. Orthogonality

We assess orthogonality of two different random matrices by considering their Frobenius inner product [32], namely, the discrete version of the inner product between two real functions

of two real variables

$$\langle R_k, R_{k'} \rangle = \sum_{i=1}^n \sum_{j=1}^m R_k(i, j) R_{k'}(i, j). \quad (6)$$

We restrict ourselves to the case  $E[R] = 0$  (i.e., the distribution has been zero centered, which can be obtained from an arbitrary random variable via  $R \equiv X - E[X]$ ). Being a random process, the probability of these basis members being strictly orthogonal is typically of zero measure. Despite not being able to make deterministic statements regarding orthogonality, we can make probabilistic statements regarding expectation value and variance of orthogonality. To calculate the expected value, consider

$$\begin{aligned} E[\langle R_k, R_{k'} \rangle] &= nm E[R_k(i, j) R_{k'}(i, j)] \\ &= nm \text{Var}[R] \delta_{kk'}, \end{aligned} \quad (7)$$

where  $\delta_{kk'}$  is an ensemble Kronecker delta (1 for  $k = k'$  and 0 for  $k \neq k'$ ). From this result, we can see that regardless of the particular distribution of  $R$ , so long as  $R$  is zero centered, two random basis members will be orthogonal in expectation. We emphasize the difference between orthogonal and orthogonal in expectation: The former label refers to two random matrices having zero Frobenius product, with the latter stating that the Frobenius product is a random variable with zero expectation value.

We turn to the deviations from orthogonality that we can expect as quantified by the variance. The variance of the Frobenius product expressed in Eq. (6) is

$$\begin{aligned} \text{Var}[\langle R_k, R_{k'} \rangle] &= nm \text{Var}[R_k R_{k'}] \\ &= \begin{cases} nm \text{Var}[R]^2 & \text{for } k \neq k' \\ nm \text{Var}[R^2] & \text{for } k = k', \end{cases} \end{aligned} \quad (8)$$

where we have used the independence of each random number realization to set the covariance to zero. Note also that although  $R_k$  and  $R_{k'}$  are different random numbers, their expectation value and variance will be the same since they are drawn from the same distribution. The larger the matrix size  $n \times m$ , the larger the variance from orthogonality. Although, if we were to normalize the inner product with  $1/nm \text{Var}[R]$ , the right-hand side of Eq. (8) would decrease with increasing matrix size. We note, in this context, that our normalization differs from that which is typically used in the theory of random matrices [33].

### B. Completeness relation

We turn attention to the completeness relation

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N \text{Var}[R]} \sum_{k=1}^N R_k(i, j) R_k(i', j') \\ = \delta(i - i', j - j'), \end{aligned} \quad (9)$$

where  $\delta(i - i', j - j')$  is a spatial Kronecker delta and all other symbols are as previously defined. We are particularly interested in the finite- $N$  version, namely,

$$\frac{1}{N \text{Var}[R]} \sum_{k=1}^N R_k(i, j) R_k(i', j') \approx \delta(i - i', j - j'), \quad (10)$$

where of interest is the rate at which the approximation will converge to a Kronecker  $\delta$ , as quantified by the variance of the above sum. Confirming we indeed obtain the Kronecker  $\delta$  as claimed, consider the expectation

$$\begin{aligned} E \left[ \frac{1}{N \text{Var}[R]} \sum_{k=1}^N R_k(i, j) R_k(i', j') \right] \\ = \frac{1}{\text{Var}[R]} E[R(i, j) R(i', j')] = \delta(i - i', j - j'), \end{aligned} \quad (11)$$

where we have used the fact  $R$  is a zero-centered random variable, which implies that  $E[R^2] = \text{Var}[R]$ . So we indeed obtain a spatial Kronecker  $\delta$  as an expectation.

Moving onto the variance, this can be calculated via

$$\begin{aligned} \text{Var} \left[ \frac{1}{N \text{Var}[R]} \sum_{k=1}^N R_k(i, j) R_k(i', j') \right] \\ = \frac{1}{N^2 \text{Var}[R]^2} \sum_{k=1}^N \text{Var}[R(i, j) R(i', j')] \\ = \begin{cases} \frac{1}{N} & \text{for } (i, j) \neq (i', j') \\ \frac{1}{N} \frac{\text{Var}[R^2]}{\text{Var}[R]^2} & \text{for } (i, j) = (i', j'), \end{cases} \end{aligned} \quad (12)$$

where we have again used the fact that each of the covariances is zero since each  $R_k(i, j)$  is independent. As the number  $N$  of basis members is increased, the closer we converge to the expectation value of the spatial Kronecker  $\delta$ . As  $N \rightarrow \infty$ , we achieve a complete (indeed, overcomplete) set of nonorthogonal basis members.

### C. Synthesis of arbitrary matrices using a random-matrix basis

Suppose one wishes to express an  $n \times m$  arbitrary discretized function or matrix  $f(i, j)$  as a linear combination of realizations of random  $n \times m$  matrices, each element of each random matrix being an independent deviate drawn from the same arbitrary real probability distribution. To synthesize  $f(i, j)$ , take the completeness relation expressed in Eq. (9), multiply both sides by  $f(i, j)$ , and then sum over all spatial points to obtain

$$\begin{aligned} f(i, j) &= \sum_{i'=1}^n \sum_{j'=1}^m f(i', j') \delta(i - i', j - j') \\ &= \lim_{N \rightarrow \infty} \frac{1}{N \text{Var}[R]} \sum_{i'=1}^n \sum_{j'=1}^m f(i', j') \sum_{k=1}^N R_k(i, j) R_k(i', j'). \end{aligned} \quad (13)$$

Interchanging the order of the sums and approximating the infinite sum with a finite sum, we see that  $f(i, j)$  is approximated with finitely many terms

$$f(i, j) \approx \frac{1}{N \text{Var}[R]} \sum_{k=1}^N w_k R_k(i, j) \equiv f_{(N)}(i, j), \quad (14)$$

where  $w_k \equiv \langle f, R_k \rangle$  is the weighting coefficient of the random matrix  $R_k(i, j)$  and  $f_{(N)}(i, j)$  is the  $N$ th-order approximation to the exact expression  $f(i, j)$ . This is a key expression, for it demonstrates that a signal  $f(i, j)$  can indeed be approximated by a linear combination of noise maps. A decomposition of

this nature is commonly employed in computational imaging and ghost imaging [3,34], although the preceding analysis is not tied to either particular context.

To quantify the noise in a reconstruction made with a finite number of basis elements, we calculate the variance

$$\begin{aligned} \text{Var}[f_{(N)}(i, j)] &= \frac{1}{N} \sum_{i'(\neq i)=1}^n \sum_{j'(\neq j)=1}^m f^2(i', j') \\ &\quad + \frac{f^2(i, j)}{N} \left( \frac{\text{Var}[R^2]}{\text{Var}[R]^2} \right), \end{aligned} \quad (15)$$

where we have twice used the facts that each  $R_k(i, j)$  is independent and the covariance of each term with respect to the others is zero. For  $\text{Var}[R^2] \approx \text{Var}[R]^2$ , or large  $N$  such that the first term is sufficiently dominant, we can make the simplifying approximation

$$\text{Var}[f_{(N)}(i, j)] \approx \frac{1}{N} \sum_{i'=1}^n \sum_{j'=1}^m f^2(i', j') = \frac{\langle f^2 \rangle}{N}. \quad (16)$$

The variance in a finite reconstruction is independent of the distribution of  $R$ , depending only on the spatial sum of the square of the desired function  $f$  and the number of basis members  $N$ .

The local signal-to-noise ratio (SNR)  $S$  is

$$S(i, j) \equiv \frac{E[f_{(N)}(i, j)]}{\sqrt{\text{Var}[f_{(N)}(i, j)]}} \approx \frac{f(i, j)}{\sqrt{\langle f^2 \rangle / N}}, \quad (17)$$

with the corresponding global SNR being the root-mean-square average of (17):

$$S \approx \sqrt{\frac{N}{nm}}. \quad (18)$$

This leads to a form of noise-resolution uncertainty principle

$$S^2 \times (nm) \approx N, \quad (19)$$

where  $nm$  is a measure of the resolution and  $N$  is the number of basis members or, in a sense, the number of measurements required. Hence there is a direct tradeoff between resolution and SNR for a given number of basis members or measurements. Note that this statement refers to the noise inherent to the process of image reconstruction from a nonorthogonal random basis as outlined above and in no way includes the noise contributions and physical limitations of experimental realization which will have an additional noise-resolution uncertainty principle [35].

#### D. Random-matrix basis with Gram-Schmidt orthogonalization

A finite realized set of random-matrix basis members will not be strictly orthogonal, as seen in Sec. II A. They can, however, be made strictly orthonormal via the Gram-Schmidt process [30]. Further, supposing we have a set of weighting coefficients already generated for our original random basis set, we can transform both objects to achieve a more efficient

representation via

$$\begin{aligned} w'_k &= w_k - \sum_{\ell=1}^{k-1} \frac{\langle R_\ell, R_k \rangle}{\langle R_\ell, R_\ell \rangle} w_\ell, \quad \tilde{w}_k = \frac{w'_k}{\langle R'_k, R'_k \rangle}, \\ R'_k &= R_k - \sum_{\ell=1}^{k-1} \frac{\langle R_\ell, R_k \rangle}{\langle R_\ell, R_\ell \rangle} R_\ell, \quad \tilde{R}_k = \frac{R'_k}{\langle R'_k, R'_k \rangle}. \end{aligned} \quad (20)$$

Here the primed quantities are intermediate steps and the tildes represent the orthonormalized basis quantities.

The implementation of the Gram-Schmidt algorithm can be computationally expensive and even unstable. To improve the stability and efficiency of the algorithm, it can be implemented via QR decomposition and variants thereof [23,36]. We make particular note of the use of the Householder transformation in this context [31]. Irrespective of how orthonormalization is achieved, the corresponding function reconstruction is

$$\begin{aligned} f(i, j) &= \sum_{k=1}^{nm} \tilde{w}_k \tilde{R}_k(i, j) \\ &\approx \sum_{k=1}^N \tilde{w}_k \tilde{R}_k(i, j) \equiv f^{(N)}(i, j), \end{aligned} \quad (21)$$

where  $f^{(N)}$  is the approximate version of  $f(i, j)$  projected onto the incomplete  $N$ -member basis set ( $N < nm$ ). Note that we have set  $N$  to be less than  $nm$  above since this directly implies the basis to have insufficient members to be complete; the set would be complete with  $N$  members and overcomplete with greater than  $N$  members. Note also that the subscript  $N$  in the nonorthogonal case is swapped for a superscript in the orthonormal case, to distinguish the two. Finally, we augment the basis via inclusion of one constant basis member  $R_1(i, j) = 1/\sqrt{nm} \forall i, j$ , to improve convergence. That is,  $R_1(i, j)$  will efficiently achieve the spatial average of the desired matrix  $f(i, j)$  and the random basis will then only be required to capture deviations from this spatial average.

As expressed in Eq. (21), an exact reconstruction of  $f(i, j)$  is achieved with  $nm$  basis members (i.e., a complete set). Seeking a more explicit statement regarding the noise present in a reconstruction using an incomplete random basis with Gram-Schmidt utilization, consider the variance of the special case consisting of a single constant value basis member

$$\text{Var}[f^{(N)}(i, j)] = \frac{1}{nm} \langle [\bar{f} - f(i, j)]^2 \rangle \equiv \text{Var}[f], \quad (22)$$

where  $\bar{f}$  is the spatial average of  $f(i, j)$ . Now consider the case with a complete basis set and hence an exact reconstruction. In this latter case the variance per pixel vanishes. Finally, consider an intermediate case using an incomplete basis set with between two and  $nm - 1$  basis members inclusive. Given that the basis matrices are inherently random, each basis member will on average contribute equally to the reconstruction. This is similar in sentiment to the idea of equally likely microstates in statistical mechanics [37]. Adding independent deviations in quadrature implies the variance of an incomplete basis set is given by linearly interpolating between the one special-constant basis case and the  $nm$  basis-set case, hence

$$\text{Var}[f^{(N)}(i, j)] = \text{Var}[f] \left( 1 - \frac{N}{nm} \right). \quad (23)$$

It is emphasized that the dependence on  $f$  only arises due to the inclusion of the initial privileged basis member  $R_1(i, j) = 1/\sqrt{nm} \forall i, j$ , to boost convergence. The expression (23) was seen to be consistent with the results of numerical simulation, for several different probability distributions (Gaussian, Poissonian, and uniform).

We adopt the same global SNR definition as in Sec. II C. Hence

$$S = \sqrt{\frac{\langle f^2 \rangle}{\text{Var}[f]} \left( \frac{1}{nm - N} \right)}. \quad (24)$$

Note the scaling of SNR here compared to the nonorthogonal case. Previously we had  $S \approx \sqrt{N/nm}$ , whereas now we have  $S \propto \sqrt{1/(nm - N)}$ . Note also that “noise” in this context refers to spatially random errors in the synthesis of  $f(i, j)$  that is obtained by superposing random matrices in an idealized setting devoid of experimental noise. Each will reach an infinite SNR, the Gram-Schmidt case in  $N = nm$  orthonormal random matrices, while the nonorthogonal case requires  $N \rightarrow \infty$  basis matrices. The above comparison is independent of the particular random distribution employed to develop the random basis. Further, in the case of an incomplete basis set for the Gram-Schmidt case, the representation of the reconstruction is not unique and depends on the order of the basis elements being orthonormalized. Hence a further reduction in noise can be achieved by averaging over different Gram-Schmidt-basis reconstruction results [23].

### III. COMPUTATIONAL GHOST-IMAGING SHOT-NOISE ANALYSIS

To compare conventional direct imaging and computational ghost imaging with respect to fluence or dose, we can make a preliminary assessment by counting “balls” in “bins.” Here the bins correspond to a pixel detector and the balls to the photons. For the direct image, suppose we have  $n \times m$  bins with  $nm\tilde{\lambda}$  balls to make the image. We will distribute  $\tilde{\lambda}$  balls into each bin and let  $f(i, j)$  be the proportion of balls that will go in at each  $(i, j)$  bin [i.e.,  $0 \leq f(i, j) \leq 1$  is the transmission value of the sample]. We describe the influence of noise by the Poisson distribution given its relevance in describing shot noise present in pixel photon measurements [29, 38]. This gives the normalized bin value  $U(i, j)$  in the direct imaging case as

$$U(i, j) = \frac{P(f(i, j)\tilde{\lambda})}{\tilde{\lambda}}, \quad (25)$$

where  $P$  represents the Poisson distribution with a mean value of  $f(i, j)\tilde{\lambda}$ . The expected value and variance of the normalized bin value are representative of the noise-free value and the expected noise, respectively. That is, we expect each normalized bin value to converge to

$$\text{E}[U(i, j)] = \text{E}\left[\frac{P(f(i, j)\tilde{\lambda})}{\tilde{\lambda}}\right] = f(i, j), \quad (26)$$

where we have used  $\text{E}[P(f(i, j)\tilde{\lambda})] = f(i, j)\tilde{\lambda}$ . The expected noise is given by

$$\text{Var}[U(i, j)] = \text{Var}\left[\frac{P(f(i, j)\tilde{\lambda})}{\tilde{\lambda}}\right] = \frac{f(i, j)}{\tilde{\lambda}}, \quad (27)$$

where we have used  $\text{Var}[P(f(i, j)\tilde{\lambda})] = f(i, j)\tilde{\lambda}$ . The total expected noise in the entire direct image is the spatial sum of these values:

$$\text{Var}[\mathcal{I}_{\text{DI}}] = \sum_{i=1}^n \sum_{j=1}^m \frac{f(i, j)}{\tilde{\lambda}} = \frac{\langle f \rangle}{\tilde{\lambda}}. \quad (28)$$

Moving onto the ghost-imaging case, we now perform this comparison for computational ghost imaging. We wish to employ a scheme that will converge to the exact image in a finite number of measurements. This precludes simply using randomly generated, nonorthogonal masks which require infinitely many basis members to reach an overcomplete basis set. Nor will we consider taking  $nm$  nonorthogonal masks and seek to orthonormalize them (via the Householder transformation or otherwise), whereby experimental uncertainties will propagate in a relatively complex fashion. Instead, as a baseline analysis, we will consider a computer-generated complete set of  $nm$  orthonormal random matrices  $\{\tilde{R}_k(i, j)\}$  of size  $n \times m$  which includes one constant member  $\tilde{R}_1(i, j) = 1/\sqrt{nm} \forall i, j$ . In general, these can be mapped to non-negative ghost-imaging masks  $\{I_k(i, j)\}$  via the transformation

$$I_k(i, j) = \frac{1}{\xi} [\tilde{R}_k(i, j) - \min], \quad (29)$$

where  $\{\tilde{R}_k(i, j)\} \in [\min, \max]$ ,  $\xi \equiv (\max - \min)$ , and  $\{I_k(i, j)\} \in [0, 1]$ . Further, we will consider the order: mask, then sample (which is the conventional setup for ghost imaging but not necessarily a requirement; see Sec. VI below for further information on this point). To conserve the dose that the sample is exposed to, we will enforce that we have  $\tilde{\lambda}$  balls downstream of the mask (which is consistent with  $nm\tilde{\lambda}$  total balls reaching the sample). This gives the uniform incident illumination  $x$  prior to the mask as

$$\langle I_k \rangle \frac{x}{nm} = \tilde{\lambda}, \quad \Rightarrow x = \frac{nm}{\langle I_k \rangle} \tilde{\lambda}. \quad (30)$$

This construction has the associated image or normalized bin value

$$U(i, j) = \sum_{k=1}^{nm} (B_k + \eta B_1) \tilde{R}_k(i, j), \quad (31)$$

where  $B_k$  is the  $k$ th bucket reading and  $\eta B_1$  is min multiplied by the spatially integrated sample transmission function  $\langle f \rangle$ , which can be determined from the constant basis member via  $\eta \equiv \min/(1/\sqrt{nm} - \min)$ . Explicitly, the measured bucket reading  $B_k$  is given by

$$B_k = \xi \frac{\langle I_k \rangle}{\tilde{\lambda}} P\left(\frac{\tilde{\lambda}}{\langle I_k \rangle} \langle I_k, f \rangle\right). \quad (32)$$

We now calculate the expectation and variance of the effect of Poisson noise on the ghost-imaging case. The expectation is

$$\begin{aligned} \text{E}[U(i, j)] &= \sum_{k=1}^{nm} (\text{E}[B_k] + \eta \text{E}[B_1]) \tilde{R}_k(i, j) \\ &= \sum_{k=1}^{nm} (\xi \langle I_k, f \rangle + \min\langle f \rangle) \tilde{R}_k(i, j) \\ &= \sum_{k=1}^{nm} \langle \tilde{R}_k, f \rangle \tilde{R}_k(i, j) = f(i, j). \end{aligned} \quad (33)$$

For the variance

$$\begin{aligned} \text{Var}[U(i, j)] &= \left( \tilde{R}_1(i, j) + \eta \sum_{k=1}^{nm} \tilde{R}_k(i, j) \right)^2 \text{Var}[B_1] \\ &+ \sum_{k=2}^{nm} \text{Var}[B_k] \tilde{R}_k^2(i, j), \end{aligned} \quad (34)$$

where we have once again set the covariance in different bucket readings to zero. The bucket reading variance is

$$\begin{aligned} \text{Var}[B_k] &= \xi^2 \text{Var} \left[ \frac{\langle I_k \rangle}{\tilde{\lambda}} P \left( \frac{\tilde{\lambda}}{\langle I_k \rangle} \langle I_k, f \rangle \right) \right] \\ &= \xi^2 \frac{\langle I_k \rangle}{\tilde{\lambda}} \langle I_k, f \rangle, \end{aligned} \quad (35)$$

which implies that the variance of a normalized bin filled with balls via the ghost-imaging scheme is

$$\begin{aligned} \text{Var}[U(i, j)] &= \left( \tilde{R}_1(i, j) + \eta \sum_{k=1}^{nm} \tilde{R}_k(i, j) \right)^2 \xi^2 \frac{\langle I_1 \rangle}{\tilde{\lambda}} \langle I_1, f \rangle \\ &+ \sum_{k=2}^{nm} \xi^2 \frac{\langle I_k \rangle}{\tilde{\lambda}} \langle I_k, f \rangle \tilde{R}_k^2(i, j). \end{aligned} \quad (36)$$

This in turn implies that the variance of the ghost image, being the spatial sum of bins, is

$$\begin{aligned} \text{Var}[\mathcal{I}_{\text{GI}}] &= \left\langle \left( \frac{1}{\sqrt{nm}} + \eta \sum_{k=1}^{nm} \tilde{R}_k(i, j) \right)^2 \right\rangle \xi^2 \frac{\langle I_1 \rangle}{\tilde{\lambda}} \langle I_1, f \rangle \\ &+ \sum_{k=2}^{nm} \xi^2 \frac{\langle I_k \rangle}{\tilde{\lambda}} \langle I_k, f \rangle, \end{aligned} \quad (37)$$

where we have used  $\tilde{R}_1(i, j) = 1/\sqrt{nm} \forall i, j$  and  $\langle R_k^2 \rangle = 1$ . This implies that computational ghost imaging may reduce dose over direct imaging by boosting the SNR [22] if a basis set exists such that for the same image, the variance of the ghost image is less than that of the direct image. Such dose reduction therefore requires the inequality

$$\Omega \xi^2 \langle I_1 \rangle \langle I_1, f \rangle + \xi^2 \sum_{k=2}^{nm} \langle I_k \rangle \langle I_k, f \rangle < \langle f \rangle \quad (38)$$

to be satisfied, where  $\Omega \equiv \langle [1/\sqrt{nm} + \eta \sum_{k=1}^{nm} \tilde{R}_k(i, j)]^2 \rangle$ . Conversely, if a basis set is such that the above inequality is violated, the ghost-imaging case will be less dose efficient than the corresponding direct-imaging case (cf. a similar inequality obtained in a different context, recently reported by Gureyev *et al.* [39]).

It is reasonable in some cases to make the approximation  $\sum_{k=1}^{nm} \tilde{R}_k \approx 0 \forall i, j$ , which reduces Eq. (37) to

$$\text{Var}[\mathcal{I}_{\text{GI}}] \approx \sum_{k=1}^{nm} \xi^2 \frac{\langle I_k \rangle}{\tilde{\lambda}} \langle I_k, f \rangle. \quad (39)$$

Ignoring  $\xi^2$  and  $\langle I_k \rangle$ , which scale approximately as  $1/nm$  and  $nm$ , respectively, suggests that in order for ghost imaging to reduce dose, the sum of bucket readings needs to be less than the spatial sum of the transmission function. That is, we want the mask-sample combination to be more absorbing than the

sample alone which corresponds to fewer balls being detected, which when normalized, produces a lower uncertainty.

As an illustration of the logical possibility for dose reduction on account of the above inequalities, consider the following extreme limiting case. Suppose one wishes to perform computational ghost imaging on a sample whose two-dimensional intensity transmission function is a pixelated grayscale image of a given highly textured structure. Suppose too that a very small number of pseudorandom illuminating masks is used, each of which happen to have an unusually high degree of correlation with the illuminated sample. The resulting computational ghost image will then be obtainable with an unusually small number of bucket measurements and a correspondingly reduced dose. While this scenario is of course very highly improbable when genuinely random masks are used, the possibility of such a scenario becomes significantly more likely when one utilizes feedback computational ghost imaging, a point which is very briefly explored later in the present paper.

It is physically reasonable that the dose-reduction inequalities should be both object and mask dependent. Notwithstanding this fact, if one wishes to image a particular class of object, it may be useful to average the above dose-reduction inequalities over a statistical ensemble of possible objects.

The preceding analysis regarding dose is conservative as it does not consider incorporation of compressive sensing ideas [40]. That is, compressive sensing allows one to perform ghost imaging in less than  $nm$  measurements, which is beneficial regarding reduced dose. Moreover, in cases where the above inequality is violated, the necessity for a more sophisticated approach to ghost reconstruction, such as given by compressive ghost-imaging methods [12–15], is made apparent as a necessary condition for ghost imaging to reduce dose when compared to its direct-imaging counterpart [39].

## IV. X-RAY PHASE CONTRAST COMPUTATIONAL GHOST IMAGING

### A. Theory

Here we apply the random-matrix-basis concept to x-ray phase contrast computational ghost imaging. In this context, phase contrast may be defined as any imaging modality yielding intensity maps that are a function of the phase of the field that is input into the associated imaging system [41]. In an x-ray setting, phase contrast permits visualization of x-ray-transparent structures. While many means for realizing x-ray phase contrast exist in a direct-imaging setting [41], there is a dearth of such methods for x-ray ghost imaging. Indeed, there are only a few papers on x-ray ghost imaging [19–23], all of which demonstrate absorption-contrast x-ray ghost imaging, which is only sensitive to the magnitude of the field visualized by the associated ghost-imaging system. While means exist for realizing phase contrast in non-x-ray ghost-imaging settings [27,42–44], these are not readily transferred to x-ray imaging.

The proposed experimental setup for computational x-ray phase contrast ghost imaging utilizes a quasimonochromatic [29] x-ray source, an analyzer crystal [45], an ensemble of known speckle masks, and a bucket detector (see Fig. 1). The x-ray phase information imparted by the refractive properties of the sample will be encoded as transverse spatial intensity

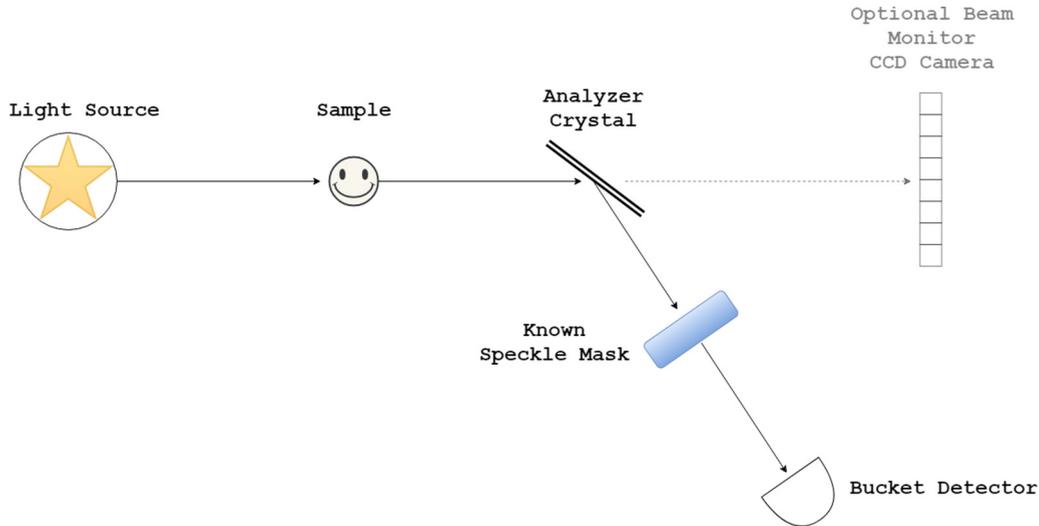


FIG. 1. Schematic of computational x-ray phase contrast ghost imaging.

variations via the Fourier-space filtration provided by reflection from the analyzer crystal [41,45]. Indeed, the angle-dependent reflectivity function (rocking curve) of the analyzer crystal is key in this setup to achieving bucket signals which are sensitive to the phase gradients of the impinging x-ray radiation. The analyzer-crystal-filtered intensity field is then incident on a known speckle mask and the combined effect on the intensity is integrated up in the bucket detector. The process is repeated for each speckle mask in the set of known masks. Note that in an experiment one can obtain such a set of speckle masks by transversely scanning a single larger mask [21]. Note also that for each bucket-signal acquisition the beam fully illuminates the sample. One may optionally include position-sensitive beam monitors at either or both of the indicated positions, to adapt to inherent temporal fluctuations of the source from measurement to measurement.

Consider the intensity  $I(i, j)$  and phase  $\phi(i, j)$  of a beam measured in a plane perpendicular to the beam propagation ( $z$  axis) on a pixel grid  $(i, j)$ . We consider the case that the sample and mask are each made from a single and in general different uniform material. This single-material assumption is not particularly restrictive regarding the mask, since the sole purpose of the mask is to generate a set of linearly independent spatially random intensity patterns. Further, this assumption is far less restrictive than it might appear, regarding the class of objects that may be imaged, for two reasons: (i) The same assumption has been successfully employed in several hundred papers in a propagation-based phase contrast x-ray tomography setting, since many samples of interest may be locally (i.e., in three spatial dimensions) described as consisting of a single material of interest [46–48], and (ii) for sufficiently high x-ray energies, low-atomic-number materials such as those in soft biological tissues behave approximately as a single material, namely, electrons [49], with the atomic nuclei having only a small effect on both the absorptive and refractive properties of the stated soft-tissue materials in this regime [50].

Under both the single-material assumption and the projection approximation, the phase shift due to the sample is [41]

$$\phi_{\text{sample}}(i, j) = -k\delta T(i, j), \quad (40)$$

where  $k = 2\pi/\lambda$  is the wave number,  $\lambda$  is the wavelength,  $\delta \equiv 1 - \text{Re}\{n\}$  (with  $\text{Re}\{n\}$  the real part of the refractive index of the sample material), and  $T(i, j)$  is the projected thickness of the sample. The angular deviation from the  $z$  axis, which is taken to be the optic axis, post the sample is

$$\Delta\theta(i, j) = -\frac{1}{k} \frac{\partial}{\partial x} [\phi_{\text{sample}}(i, j)] = \delta \frac{\partial T}{\partial x}. \quad (41)$$

From here we can calculate the intensity of the beam post the analyzer crystal interaction. We use the Beer-Lambert law of absorption [41] to model the sample interaction (and mask interaction). The reflectivity of the crystal as a function of deflection angle, namely, the rocking curve [45], is taken to be known. Assuming a linear approximation to the rocking curve [51] gives

$$\begin{aligned} I_{\text{sample}}(i, j) &= I_0 \exp[-\mu T(i, j)] [\alpha + \beta \Delta\theta(i, j)] \\ &= I_0 \exp[-\mu T(i, j)] \left( \alpha + \beta \delta \frac{\partial T}{\partial x} \right), \end{aligned} \quad (42)$$

where  $I_0$  is the uniform intensity input,  $\mu$  is the linear attenuation coefficient of the sample,  $\alpha$  is the zero-angle deflection reflectivity, and  $\beta$  is a linear approximation to the rocking-curve slope. We then subject the intensity field to absorption from the speckle mask, to give the bucket detector reading  $B_k$  as

$$B_k = \left\langle I_0 \exp[-\mu T(i, j) - \mu_0 R_k(i, j)] \left( \alpha + \beta \delta \frac{\partial T}{\partial x} \right) \right\rangle, \quad (43)$$

where  $\mu_0$  and  $R_k(i, j)$  are the absorption coefficient and projected thickness of the mask respectively. Note, in this context and for the remainder of the paper, that  $R_k(i, j)$  is a random value greater than zero which is slightly different than in the previous two sections where it was a zero-centered random value.

We can compute the x-ray phase contrast computational ghost image  $\mathcal{P}$  using the standard ghost-imaging formula (1) or via the Gram-Schmidt process. For the standard ghost-imaging case, we will require a normalization of  $\text{Var}[I_0 \exp(-\mu_0 R)]$  to

obtain the expected image

$$E[\mathcal{P}(i, j)] = \alpha \left( 1 - \frac{\beta\delta}{\mu\alpha} \frac{\partial}{\partial x} \right) \exp[-\mu T]. \quad (44)$$

The expected noise in the image will be given by the variance

$$\text{Var}[\mathcal{P}_{(N)}(i, j)] \approx \frac{\langle [\alpha(1 - \frac{\beta\delta}{\mu\alpha} \frac{\partial}{\partial x}) \exp(-\mu T)]^2 \rangle}{N} \quad (45)$$

or with Gram-Schmidt utilization

$$\begin{aligned} & \text{Var}[\mathcal{P}^{(N)}(i, j)] \\ & \approx \text{Var} \left[ \alpha \left( 1 - \frac{\beta\delta}{\mu\alpha} \frac{\partial}{\partial x} \right) \exp(-\mu T) \right] \left( 1 - \frac{N}{nm} \right), \end{aligned} \quad (46)$$

where the variance of  $\alpha(1 - \frac{\beta\delta}{\mu\alpha} \frac{\partial}{\partial x}) \exp(-\mu T)$  is relative to the spatial average.

The expected value of the phase contrast ghost image has the form

$$\mathcal{P} = C \left( 1 - G \frac{\partial}{\partial x} \right) \exp[-\mu T(x, y)], \quad (47)$$

where  $C = \alpha$  and  $G = \beta\delta/\mu\alpha$  are known real constants. This x-ray phase contrast ghost image can be viewed as encoding quantitative information regarding the projected thickness  $T$  of the sample, thereby inviting solution to the associated inverse problem [52] of obtaining  $T$  from  $\mathcal{P}$ . The situation is somewhat analogous to Gabor's original conception of (inline holographic) imaging as a two-step process, namely, recording followed by reconstruction [53]. To this end, Fourier transform the above differential equation with respect to  $x$ , use the Fourier derivative theorem, and then solve the resulting algebraic equation for the Fourier transform of  $\exp(-\mu T)$ . Inverse Fourier transformation gives an expression mathematically identical to that previously obtained by Paganin *et al.* in a direct-imaging x-ray phase-contrast context unrelated to ghost imaging [54–56]:

$$T(x, y) = -\frac{1}{\mu} \ln \left\{ \mathfrak{F}_x^{-1} \left[ \frac{\mathfrak{F}_x(\mathcal{P})}{C(1 - iGk_x)} \right] \right\}. \quad (48)$$

Here  $\mathfrak{F}_x$  and  $\mathfrak{F}_x^{-1}$  denote the Fourier transform and inverse Fourier transform in the  $x$  direction, respectively, and  $k_x$  denotes the Fourier-space coordinate dual to  $x$ . We use a Fourier-transform convention in which the Fourier derivative theorem takes the form  $\partial/\partial x = \mathfrak{F}_x^{-1}(ik_x)\mathfrak{F}_x$  [41]. For  $G \neq 0$ , the Fourier-space filter  $1/(1 - iGk_x)$  has modulus  $(1 + G^2k_x^2)^{-1/2}$  that never vanishes, never has division-by-zero instability, and never exceeds unity. This class of inverse-problem solution belongs to the category of SNR-boosting algorithms recently investigated by Gureyev *et al.* [57]. A related member of this algorithm class [46] has been shown to boost the SNR by factors  $F$  on the order of hundreds (or, equivalently, enable dose and data acquisition time to be reduced by factors  $F^2$  on the order of tens of thousands) in the process of reconstruction [47,48,58–60].

TABLE I. Material parameters for  $E \approx 40$  keV.

Element	Z	$M_A$ (g/mol)	$\mu/\rho$ (cm <sup>2</sup> /g)	$\rho$ (g/cm <sup>3</sup> )	$f_1$
carbon	6	12.011	0.2076	1.700	6.00115
aluminium	13	26.982	0.5685	2.699	13.0206
copper	29	63.546	4.862	8.960	29.2497
gold	79	196.966	12.98	19.32	79.1108

## B. Simulations

### 1. Parameters and rocking-curve model

In our simulations, the analyzer-crystal rocking curve is approximated by the Pearson VII distribution [61]

$$R(\theta) = R_0 \left( 1 + \frac{\theta^2}{\mathfrak{M}a^2} \right)^{-\mathfrak{M}}. \quad (49)$$

Here  $R$  is reflectivity,  $R/R_0$  is relative reflectivity, and the independent variable  $\theta$  is the deviation from the Bragg angle associated with the operative analyzer-crystal reflection;  $R_0$ ,  $\mathfrak{M}$ , and  $a$  are parameters of the crystal and the x-ray energy. Note that the same symbol is used for both reflectivity and a random variable; it will always be clear from the context which meaning is intended. Taking a Taylor series about  $\theta_0$ , corresponding to a single chosen angular detuning of the analyzer crystal from the Bragg condition for which the entirety of the images is taken, we obtain an expression for the linearization parameters that approximate the rocking curve ( $\alpha$  and  $\beta$ ) in terms of  $R_0$ ,  $\theta_0$ ,  $\mathfrak{M}$ , and  $a$ :

$$\alpha = R(\theta_0) = R_0 \left( 1 + \frac{\theta_0^2}{\mathfrak{M}a^2} \right)^{-\mathfrak{M}}, \quad (50)$$

$$\beta = \left. \frac{dR}{d\theta} \right|_{\theta_0} = -\frac{2\theta_0 R_0}{a^2} \left( 1 + \frac{\theta_0^2}{\mathfrak{M}a^2} \right)^{-\mathfrak{M}-1}. \quad (51)$$

The numerical parameters for the Pearson VII curve correspond to a Si(333) analyzer crystal at illuminating energy  $E = 40$  keV and were taken from Majidi *et al.* [62] as  $a = 0.7146$ ,  $\mathfrak{M} = 2.3737$ , and  $R_0 = 1$ . The uniform initial intensity was assumed to be normalized so that  $I_0 = 1$ . Material values were obtained from the NIST database [63] and are listed in Table I.

Note that form factor  $f_1$  values were obtained at  $E = 39.19543$  keV. We hence computed the refractive index decrement  $\delta$  via [64]

$$\delta = \frac{n_a r_0 \lambda^2}{2\pi} f_1, \quad (52)$$

where  $r_0$  is the classical electron radius,  $\lambda$  is the x-ray wavelength,  $n_a$  is

$$n_a \approx \frac{\rho N_A}{M_A}, \quad (53)$$

$\rho$  is the material density,  $N_A$  is Avogadro's number, and  $M_A$  is the molar mass.

### 2. Idealized x-ray phase contrast computational ghost-imaging simulations

In the present section we make two natural simplifying assumptions, both of which will be dropped in the next section.

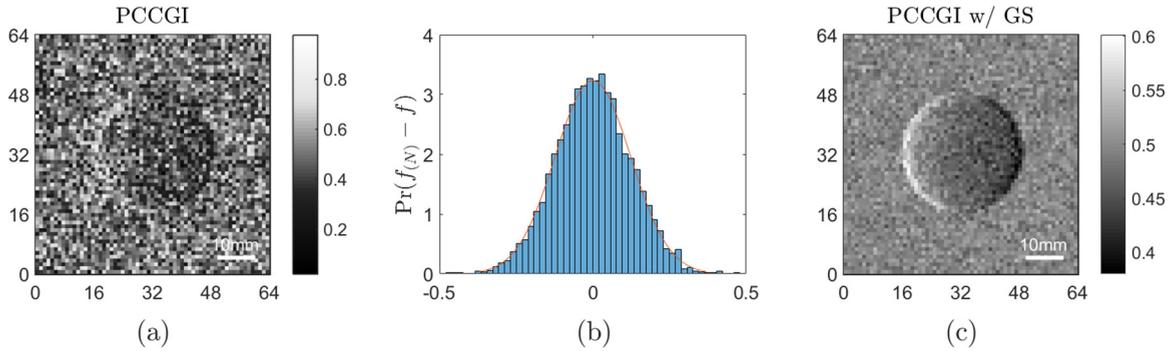


FIG. 2. Idealized x-ray phase contrast computational ghost-imaging (PCCGI) simulation of a carbon ellipsoid on a  $64 \times 64 \text{ mm}^2$  pixel array. (a) Reconstruction performed with standard ghost-imaging formula with  $N = 16 \text{ nm}$ . (b) Pixel variation from the expected value in (a) as a probability distribution, overlaid with the predicted distribution using the predicted variance and assumed Gaussian statistics. (c) Reconstruction performed with Gram-Schmidt (GS) utilization for  $N = 0.5 \text{ nm}$ .

(i) The idealized ghost-imaging case we consider here has a 1:1 pixel-to-speckle correspondence, giving a unique random illuminating intensity at each pixel. This amounts to seeking a ghost reconstruction with resolution equal to the mask speckle size, a natural restriction that is consistent with the fact that the point-spread function of the reconstruction is equal to the autocovariance of the ensemble of speckle masks [23,39,65]. Indeed, the 1:1 pixel-to-speckle correspondence investigated here is optimal, since sampling the speckles any more finely does not improve spatial resolution (in the ghost reconstruction) beyond the maximum resolution given by the speckle size. (ii) The effects of detector noise are not considered in the present section. Any deleterious effect of detector noise in the present context is typically rather mild, since bucket signals are rather insensitive to noise on account of the fact that they integrate over the entire transverse extent of the beam.

The simulated sample being imaged is an amorphous carbon ellipsoid with projected thickness

$$T(x, y) = 2 \text{Re}[\sqrt{(r/\aleph)^2 - (y/\aleph)^2 - (x/\aleph)^2}], \quad (54)$$

where  $\aleph = 8$ ,  $r = 16 \text{ mm}$ , and  $\text{Re}$  denotes the real part. For this simulated object,  $R/R_0$  was approximately within the range (0.3,0.65), consistent with the linear approximation to the rocking curve that has been adopted. Each pixel is  $1 \times 1 \text{ mm}^2$  in size, corresponding to an overall field of view on the order of several centimeters, as is often used, e.g., in medical-imaging x-ray synchrotron beamlines [66] and laboratory-source x-ray imaging [56].

We will compare image reconstructions of the projected thickness performed with the standard ghost-imaging formula (1) and with Gram-Schmidt utilization (21). The Pearson VII rocking-curve linearization is performed about the 50% relative reflectivity point to obtain a maximum range of the approximately linear region, which was taken to span from 0.2 to 0.8. The linear-approximation parameters were calculated to be  $\alpha = 0.5011$  and  $\beta = 9.3875 \times 10^5 \text{ } \mu\text{rad}^{-1}$ .

The results of this first simulation are given in Fig. 2. Figure 2(a) exhibits x-ray ghost differential phase contrast, as the white edge bounding the left side of the ellipsoid and the corresponding dark edge on the right side of the ellipsoid. This corresponds to the derivative term in Eq. (47), which in turn arises from the refractive properties of the sample. Note

that, while these images [e.g., Fig. 2(a)] closely resemble those obtained using analyzer-crystal x-ray phase contrast imaging [45], we emphasize that Fig. 2(a) is an x-ray *ghost* phase contrast image whose phase contrast is never directly measured with a position-sensitive detector. Note also that, while the eye typically perceives some absorption within the boundaries of the ellipsoid in Fig. 2(a), this is an optical illusion related to the physiology of the human eye [67,68]. Thus the dominant contrast mechanism in Fig. 2(a) is differential phase contrast, with absorption contrast being negligible. The histogram of noise present in Fig. 2(a) is shown in Fig. 2(b) with the numerically obtained distribution (blue bars) overlaid with a red curve based on the theory developed above. Finally, Fig. 2(c) shows the differential x-ray phase contrast ghost image obtained when the Gram-Schmidt process is used to orthogonalize the speckle fields prior to the ghost reconstruction. The significant improvement in efficiency of the reconstruction with Gram-Schmidt utilization is consistent with previous studies on absorptive x-ray ghost imaging [23].

### 3. Comparison of computational x-ray absorption and phase contrast ghost-imaging simulations

To perform a more realistic x-ray ghost-imaging simulation, for the case of both absorption contrast and phase contrast, we drop both of the simplifying assumptions listed at the beginning of the preceding section. We thus expand considerations to include spatially smoothed speckles with Poisson shot noise [29,38] in the bucket reading acquisition. For the Poisson noise associated with the bucket detector as a single-pixel detector, we use the model

$$B_k'' = \frac{P(B_k \tilde{\lambda}')}{\tilde{\lambda}'}, \quad (55)$$

where  $B_k''$  and  $B_k$  are the measured and ideal bucket readings, respectively,  $\tilde{\lambda}' = \tilde{\lambda}(I_k)$ ,  $\tilde{\lambda}$  is the number of balls or imaging quanta in the uniform beam,  $P$  is the Poisson distribution, and  $I_k(i, j) = I_0 \exp[-\mu_0 R_k(i, j)]$  is a speckle intensity pattern. The speckle mask in this case is generated by taking a single large array of uniformly and independently distributed random numbers (0 and  $250 \text{ } \mu\text{m}$ ) and filtering it with a normalized two-dimensional rotationally symmetric Gaussian kernel having a standard deviation of  $2/3$  of a pixel. Each speckle

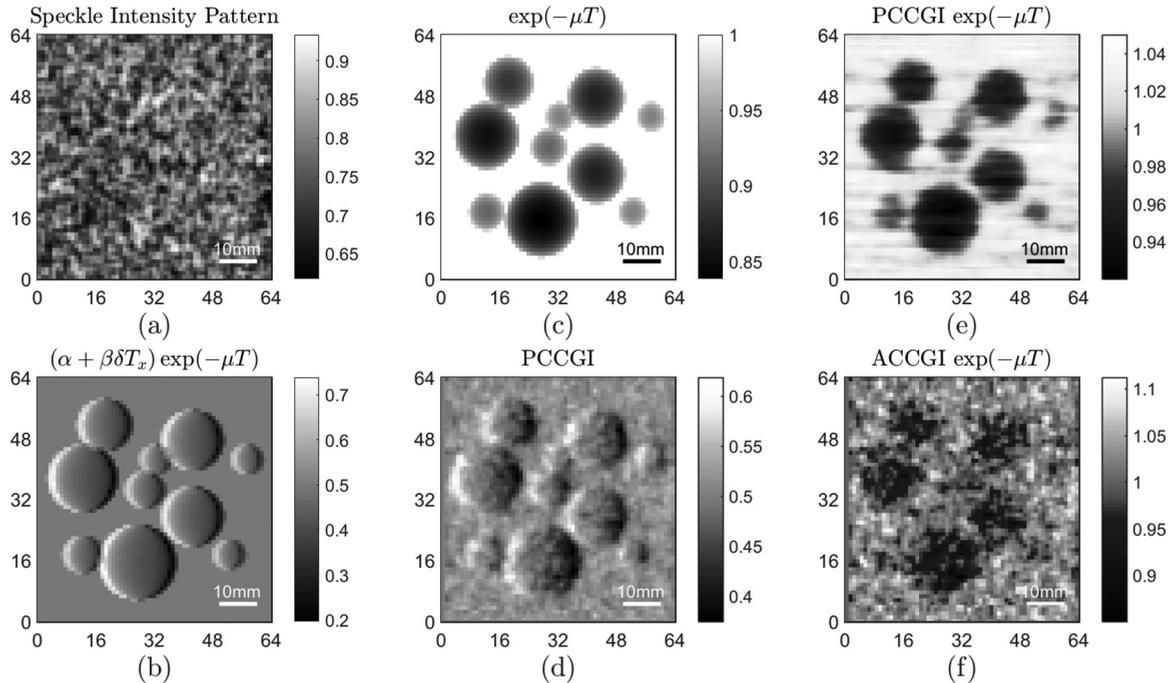


FIG. 3. Comparison of x-ray phase contrast and absorption contrast computational ghost-imaging simulation made on a  $64 \times 64 \text{ mm}^2$  pixel array with smoothed speckle intensity patterns and shot-noise inclusion, performed with  $N = 64 \text{ nm}$  and  $\tilde{\lambda} = 10^8$ . (a) Typical speckle intensity pattern. (b) Expected phase contrast image, with  $T_x \equiv \partial T / \partial x$ . (c) Expected absorption contrast image. (d) Obtained phase contrast image. (e) Reconstructed absorption image from phase contrast. (f) Obtained absorption contrast computational ghost image (ACCGI). Note the boost in SNR in passing from (f) to (e).

intensity pattern is then taken to correspond to a random transverse translation of this single larger pattern. A typical speckle intensity pattern created via this method is given in Fig. 3(a). The sample being imaged here is a collection of ten carbon ellipsoids, each described by Eq. (54), with  $\aleph = 4$  and  $r = \{4, 4, 4, 5, 5, 7, 8, 8, 9, 10\} \text{ mm}$ .

For the case that we have more pixels than speckles, the resolution and reconstruction noise of the ghost image will not be determined by the number of pixels but rather the number of speckles. Note that this limitation, which is always the case for the means of ghost imaging that has been considered here, may be removed or reduced if, e.g., additional *a priori* knowledge such as sparsity constraints are able to be employed through compressive sensing methods. Note also that the noise mentioned here is that inherent to the reconstruction process utilizing spatially random masks, which is distinct from noise obtained from experimental measurement.

The results of this more realistic simulation are in Fig. 3. A typical illuminating speckle field is given [Fig. 3(a)], with the expected phase-contrast ghost image [see Eq. (47)] [Fig. 3(b)] and the expected absorption-contrast image [Fig. 3(c)]. The simulated x-ray differential phase contrast ghost image shows that, for weakly absorbing samples, the phase contrast image can still consist of quite prominent features dominated by the edge contrast associated with the spatial derivative in Eq. (47) [Fig. 3(d)]. This x-ray ghost phase contrast leads to better visibility through the noise in Fig. 3(d) and ultimately retrieves [using Eq. (48)] a better absorption image [Fig. 3(e)] compared to Fig. 3(f).

While the simulation results required a large number  $N$  of speckle intensity patterns, recall that this is a base case

employing the standard ghost-imaging formula (1). Such a reconstruction only improves in SNR proportional to  $\sqrt{N}$ . More sophisticated reconstruction algorithms, such as those that utilize compressive sensing, machine learning, basis orthogonalization, etc., may achieve much better efficiency. Such improvements are beyond the scope of the simulated x-ray study given in the present paper, whose core focus is on establishing proof of concept for x-ray phase contrast ghost imaging.

## V. DISCUSSION

By considering a random-matrix reconstruction of a desired two-dimensional discretized function, we have obtained  $\mathcal{S} \approx \sqrt{N}/nm$ , which is consistent with the findings of ghost-imaging studies [3,69,70]. This suggests ghost imaging to be the experimental realization of the mathematical concept of a random basis reconstruction and is not inherently reliant on any underlying physical phenomena.

A key theme of this paper is employing a set of random matrices as a mathematical basis that can be used to expand discretized functions of interest. While the focus has been on expanding discretized two-dimensional functions using two-dimensional random matrices, the idea can be implemented in any integer number of dimensions. Indeed, the concept of a random-matrix basis can be considered in a purely mathematical setting, if desired, to sit alongside the Fourier series, Taylor series, and other classic series as a useful tool that allows one to approximate signals as a superposition of simpler functions.

A peculiarity of the random basis, when compared to deterministic bases, is that the spatially random basis

synthesizes signals by superposing noise. This fact facilitates applications where the requisite random matrices are spontaneously realized through (i) classical processes (e.g., via transmission of plane waves through spatially random screens in conventional x-ray ghost imaging [19,21–23] or atmospheric shifts in the case of ground-based telescopes which would build on the atmospheric turbulence ghost-imaging work of [17]) and (ii) quantum processes (e.g., via photon shot noise obtained via synchrotron radiation from individual electron bunches [20]). Another peculiarity of random bases is that every basis member is in some sense statistically equivalent, implying that the basis elements cannot be ordered in a manner analogous to the ordering of deterministic bases in terms of increasing spatial frequency, rapidity of oscillation, etc. The latter point invites the previously suggested parallel between random-matrix basis members, and the thermodynamic concept of a microstate. Indeed, averaging over microstates in thermodynamic ensembles is conceptually rather close to the concept of ensemble averaging over a random-matrix basis. Further exploration, of the connection between thermodynamic concepts and the concept of random-matrix bases, would form an interesting avenue for further investigation.

We do not claim that the particular random-matrix synthesis and decomposition methods in the present paper are the most efficient. For example, as previously stated, the efficiency and robustness of our proof-of-concept x-ray phase contrast computational ghost-imaging procedure could certainly be improved via the use of concepts and techniques from compressed sensing [40]. Another avenue for improvement, inspired by some recent works on inverse imaging problems, is the utilization of machine learning and artificial neural networks (see, e.g., [71]) to realize more robust and efficient ghost-imaging algorithms that are able to make fuller use of all available *a priori* knowledge. Indeed, the improvement afforded by such extensions could be quantified by comparison with the analytical base case results developed in the present paper.

With respect to phase contrast or holographic x-ray ghost imaging, there exists the additional hurdle of potentially unwanted phase information being imparted by the mask. Recall the setup for x-ray phase contrast computational ghost imaging (Sec. IV A); the reordering of the mask and sample, with respect to the usual sequence in which the mask occurs upstream of the sample, was enacted for this reason. Supposing we had the reverse order for mask and sample, the bucket reading would be

$$B_k = \sum_{i=1}^n \sum_{j=1}^m I_0 \exp[-\mu T(i, j) - \mu_0 R_k(i, j)] \times \left[ \alpha + \beta \delta \frac{\partial T(i, j)}{\partial x} + \beta \delta_0 \frac{\partial R_k(i, j)}{\partial x} \right], \quad (56)$$

where symbols with subscript 0 are mask parameters. The spatial gradient of the mask in this case presents an unwanted contribution that corrupts the bucket reading and ghost-imaging process in a manner problematic to remedy. Hence we recommend that for x-ray phase contrast and holographic ghost-imaging investigations in general, moving the mask or

spatial light modulator (SLM) to downstream of the sample may produce a more practical implementation.

In a similar fashion, consider now environmental or turbulence robustness [17,25,72–74]. The spatial distribution of the mask intensity pattern must be approximately unchanged when illuminating the sample (or vice versa) to avoid corruption of the ghost-imaging process, but need not necessarily be preserved for the entire path from sample to detector (as is consistent with the findings of [18]). That is, from the perspective of a series expansion, the weighting coefficient or bucket reading must be of the form of a known basis, multiplied by the sample’s transmission function, and spatially integrated. This permits arbitrary scrambling of the information post the mask-sample interaction.

In the case of direct imaging, the minimum resolvable area is limited by the pixel size. In the case of ghost imaging, the limiting factor is the mask speckles, which presents a potentially easier manufacturing obstacle to overcome. That is, a potential avenue of future work could be to nanomanufacture the mask, allowing ghost imaging to be performed on the submicrometer scale. This may yield a means to produce geometrical superresolution [75] images (cf. [76]).

Finally, we consider the concept of feedback computational ghost imaging. This utilizes information gained from each speckle-field interaction with the sample to inform future intensity field inputs [77,78]. Thus the  $n$ th illuminating speckle field is adaptively chosen based on all preceding  $n - 1$  speckle fields and their associated bucket signals, in a manner that seeks maximal new information in each subsequent measurement. The aim of this method is to produce a rapidly converging imaging system that seeks to reduce dose by reducing the number of measurements required; this is on-the-fly compressive acquisition, as suggested by ABmann and Bayer [77]. For example, we could use the gained information to essentially inform a guess of what the exact image is. Alternatively, since SNR can have a spatial distribution, we might choose to focus future intensity patterns based on real-time SNR distribution results. Regarding implementation of this concept, and further to the recent results of Sun *et al.* [78], suppose one wanted to converge to the exact image via a global optimization algorithm. A goal could be to meet the condition that the SLM and the sample provide the same attenuation

$$\frac{B'_k}{\langle I_k^2 \rangle} = 1, \quad (57)$$

where  $B'_k = \langle I_k, f \rangle$  subject to appropriate normalization. This condition could be augmented into something suitable for improvement via simulated annealing [79]. Further, we might hope to utilize deep learning algorithms and artificial intelligence (AI) to improve convergence. Work in the area of AI imaging has dramatically progressed in recent times [71], including in the ghost-imaging area [80].

## VI. CONCLUSION

Synthesis of functions sampled on a discrete Cartesian lattice, via a linear combination of random-matrix basis elements, was considered. This may be viewed as superposing discrete noise maps in order to approximate arbitrary discrete signals.

While this concept has broad applicability in physics, focus was given to functions of two spatial variables. In this context, statements were made regarding orthogonality, completeness, arbitrary function reconstruction, and SNR.

Particular focus was given to applications in ghost imaging and computational imaging. This aligns with the idea that ghost imaging is the experimental realization of the mathematical process of expressing an image as a weighted sum of known intensity patterns (random or otherwise). We constructed an inequality that must be satisfied if the conventional noncompressive form of ghost imaging, which may be directly derived from the random-matrix-basis concept, is to yield reduced dose as compared to direct imaging. This inequality stems from comparing the expected noise or variance of each imaging process for a given dose. That is, e.g., if a ghost image yields a lower expected noise for a given dose than the direct image, that implies the ghost image could have achieved the same noise as the direct image with a lower dose. Cases in which this inequality is not satisfied may be viewed as motivating the need for compressive sensing or related concepts, as a necessary condition for dose reduction in a ghost-imaging or computational-imaging context.

We proposed an experimental setup to achieve computational x-ray phase contrast ghost imaging. This was motivated by the dearth of x-ray ghost-imaging experiments able to yield phase contrast, the infancy of the field of x-ray ghost imaging, and the renaissance in coherent direct x-ray imaging that is currently under way due to phase contrast. Both the forward and inverse problems were considered. The former demonstrated the feasibility of x-ray differential phase contrast ghost imaging. The latter gave a solution to the inverse problem of obtaining quantitative projected-thickness information regarding a sample, using a single x-ray differential phase contrast ghost image of that sample.

#### ACKNOWLEDGMENTS

We acknowledge useful discussions with Mario Beltran, Carsten Detlefs, Jean-Pierre Guigay, Timur Gureyev, Andrew Kingston, Alex Kozlov, Kieran Larkin, Anthony Mays, Kavan Modi, Glenn Myers, Margie Olbinado, Timothy Petersen, Daniele Pelliccia, Alexander Rack, Tapio Simula, and Imants Svalbe. D.M.P. acknowledges financial support from the European Synchrotron Radiation Facility.

- 
- [1] B. I. Erkmén and J. H. Shapiro, Ghost imaging: From quantum to classical to computational, *Adv. Opt. Photon.* **2**, 405 (2010).
  - [2] M. J. Padgett and R. W. Boyd, An introduction to ghost imaging: Quantum and classical, *Philos. Trans. R. Soc. A* **375**, 20160233 (2017).
  - [3] Y. Bromberg, O. Katz, and Y. Silberberg, Ghost imaging with a single detector, *Phys. Rev. A* **79**, 053840 (2009).
  - [4] D. N. Klyshko, Effect of focusing on photon correlation in parametric light scattering, *J. Exp. Theor. Phys.* **67**, 1131 (1988).
  - [5] A. V. Belinskii and D. N. Klyshko, Two-photon optics: diffraction, holography, and transformation of two-dimensional signals, *J. Exp. Theor. Phys.* **78**, 259 (1994).
  - [6] D. V. Strekalov, A. V. Sergienko, D. N. Klyshko, and Y. H. Shih, Observation of Two-Photon “Ghost” Interference and Diffraction, *Phys. Rev. Lett.* **74**, 3600 (1995).
  - [7] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, Optical imaging by means of two-photon quantum entanglement, *Phys. Rev. A* **52**, R3429 (1995).
  - [8] R. S. Bennink, S. J. Bentley, and R. W. Boyd, “Two-Photon” Coincidence Imaging with a Classical Source, *Phys. Rev. Lett.* **89**, 113601 (2002).
  - [9] G. Scarcelli, V. Berardi, and Y. Shih, Can Two-Photon Correlation of Chaotic Light be Considered as Correlation of Intensity Fluctuations? *Phys. Rev. Lett.* **96**, 063602 (2006).
  - [10] M. I. Kolobov, *Quantum Imaging* (Springer Science + Business Media, New York, 2007).
  - [11] Y. Shih, in *Classical, Semi-classical and Quantum Noise* (Springer, Berlin, 2012), pp. 169–222.
  - [12] O. Katz, Y. Bromberg, and Y. Silberberg, Compressive ghost imaging, *Appl. Phys. Lett.* **95**, 131110 (2009).
  - [13] P. Zerom, K. W. C. Chan, J. C. Howell, and R. W. Boyd, Entangled-photon compressive ghost imaging, *Phys. Rev. A* **84**, 061804 (2011).
  - [14] V. Katkovnik and J. Astola, Compressive sensing computational ghost imaging, *J. Opt. Soc. Am. A* **29**, 1556 (2012).
  - [15] W.-K. Yu, M.-F. Li, X.-R. Yao, X.-F. Liu, L.-A. Wu, and G.-J. Zhai, Adaptive compressive ghost imaging based on wavelet trees and sparse representation, *Opt. Express* **22**, 7133 (2014).
  - [16] C. Zhao, W. Gong, M. Chen, E. Li, H. Wang, W. Xu, and S. Han, Ghost imaging lidar via sparsity constraints, *Appl. Phys. Lett.* **101**, 141123 (2012).
  - [17] D. Shi, C. Fan, P. Zhang, J. Zhang, H. Shen, C. Qiao, and Y. Wang, Adaptive optical ghost imaging through atmospheric turbulence, *Opt. Express* **20**, 27992 (2012).
  - [18] M. Le, G. Wang, H. Zheng, J. Liu, Y. Zhou, and Z. Xu, Underwater computational ghost imaging, *Opt. Express* **25**, 22859 (2017).
  - [19] H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, Fourier-Transform Ghost Imaging with Hard X-Rays, *Phys. Rev. Lett.* **117**, 113901 (2016).
  - [20] D. Pelliccia, A. Rack, M. Scheel, V. Cantelli, and D. M. Paganin, Experimental X-Ray Ghost Imaging, *Phys. Rev. Lett.* **117**, 113902 (2016).
  - [21] A. Schori and S. Schwartz, X-ray ghost imaging with a laboratory source, *Opt. Express* **25**, 14822 (2017).
  - [22] 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* **5**, 374 (2018).
  - [23] D. Pelliccia, M. P. Olbinado, A. Rack, and D. M. Paganin, Practical x-ray ghost imaging with synchrotron light, [arXiv:1710.00727](https://arxiv.org/abs/1710.00727) [IUCrJ (to be published)].
  - [24] R. I. Khakimov, B. M. Henson, D. K. Shin, S. S. Hodgman, R. G. Dall, K. G. H. Baldwin, and A. G. Truscott, Ghost imaging with atoms, *Nature (London)* **540**, 100 (2016).
  - [25] R. Meyers, K. Deacon, and Y. Shih, A new two-photon ghost imaging experiment with distortion study, *J. Mod. Opt.* **54**, 2381 (2007).

- [26] 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, *Philos. Trans. R. Soc. A* **372**, 20130021 (2014).
- [27] T. Shirai, T. Setälä, and A. T. Friberg, Ghost imaging of phase objects with classical incoherent light, *Phys. Rev. A* **84**, 041801 (2011).
- [28] A. Eddington, *The Nature of the Physical World: The Gifford Lectures* (Macmillan, New York, 1928).
- [29] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).
- [30] H. Anton and C. Rorres, *Elementary Linear Algebra with Applications* (Wiley, New York, 1987).
- [31] P. Diaconis, What is... a random matrix? *Not. Am. Math. Soc.* **52**, 1348 (2005).
- [32] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1990).
- [33] G. W. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices* (Cambridge University Press, Cambridge, 2010).
- [34] J. H. Shapiro, Computational ghost imaging, *Phys. Rev. A* **78**, 061802 (2008).
- [35] T. Gureyev, Y. Nesterets, and F. de Hoog, Spatial resolution, signal-to-noise and information capacity of linear imaging systems, *Opt. Express* **24**, 17168 (2016).
- [36] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart, Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization, *Math. Comput.* **30**, 772 (1976).
- [37] F. W. Sears and G. L. Salinger, *Thermodynamics, Kinetic Theory, and Statistical Thermodynamics*, 3rd ed. (Addison-Wesley, Reading, 1975).
- [38] Y. M. Blanter and M. Büttiker, Shot noise in mesoscopic conductors, *Phys. Rep.* **336**, 1 (2000).
- [39] T. E. Gureyev, D. M. Paganin, A. Kozlov, Y. I. Nesterets, and H. M. Quiney, Complementary aspects of spatial resolution and signal-to-noise ratio in computational imaging, *Phys. Rev. A* **97**, 053819 (2018).
- [40] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, Compressive sensing: From theory to applications, A survey, *J. Commun. Netw.* **15**, 443 (2013).
- [41] D. M. Paganin, *Coherent X-Ray Optics* (Oxford University Press, Oxford, 2006).
- [42] B. Jack, N. Leach, J. Romero, S. Franke-Arnold, M. Ritsch-Marte, S. M. Barnett, and M. J. Padgett, Holographic Ghost Imaging and the Violation of a Bell Inequality, *Phys. Rev. Lett.* **103**, 083602 (2009).
- [43] W. Gong and S. Han, Phase-retrieval ghost imaging of complex-valued objects, *Phys. Rev. A* **82**, 023828 (2010).
- [44] L. Martínez-León, P. Clemente, Y. Mori, V. Climent, J. Lancis, and E. Tajahuerce, Single-pixel digital holography with phase-encoded illumination, *Opt. Express* **25**, 4975 (2017).
- [45] E. Förster, K. Goetz, and P. Zaumseil, Double crystal diffraction for the characterization of targets for laser fusion experiments, *Krist. Tech.* **15**, 937 (1980).
- [46] 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.* **206**, 33 (2002).
- [47] 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* **18**, 6423 (2010).
- [48] M. A. Beltran, D. M. Paganin, K. K. W. Siu, A. Fouras, S. B. Hooper, D. H. Reser, and M. J. Kitchen, Interface-specific x-ray phase retrieval tomography of complex biological organs, *Phys. Med. Biol.* **56**, 7353 (2011).
- [49] P. Cloetens (private communication).
- [50] X. Wu, H. Liu, and A. Yan, X-ray phase-attenuation duality and phase retrieval, *Opt. Lett.* **30**, 379 (2005).
- [51] D. Chapman, W. Thomlinson, R. E. Johnston, D. Washburn, E. Pisano, N. Gmür, Z. Zhong, R. Menk, F. Arfelli, and D. Sayers, Diffraction enhanced x-ray imaging, *Phys. Med. Biol.* **42**, 2015 (1997).
- [52] P. C. Sabatier, Past and future of inverse problems, *J. Math. Phys.* **41**, 4082 (2000).
- [53] D. Gabor, A new microscopic principle, *Nature (London)* **161**, 777 (1948).
- [54] D. Paganin, T. E. Gureyev, K. M. Pavlov, R. A. Lewis, and M. Kitchen, Phase retrieval using coherent imaging systems with linear transfer functions, *Opt. Commun.* **234**, 87 (2004).
- [55] D. Briedis, K. K. W. Siu, D. M. Paganin, K. M. Pavlov, and R. A. Lewis, Analyser-based mammography using single-image reconstruction, *Phys. Med. Biol.* **50**, 3599 (2005).
- [56] D. J. Vine, D. M. Paganin, K. M. Pavlov, J. Kräußlich, O. Wehrhan, I. Uschmann, and E. Förster, Analyser-based phase contrast imaging and phase retrieval using a rotating anode x-ray source, *Appl. Phys. Lett.* **91**, 254110 (2007).
- [57] 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. Am. A* **34**, 2251 (2017).
- [58] Y. I. Nesterets and T. E. Gureyev, Noise propagation in x-ray phase-contrast imaging and computed tomography, *J. Phys. D* **47**, 105402 (2014).
- [59] T. E. Gureyev, S. C. Mayo, Y. I. Nesterets, S. Mohammadi, D. Lockie, R. H. Menk, F. Arfelli, K. M. Pavlov, M. J. Kitchen, F. Zanconati, C. Dullin, and G. Tromba, Investigation of the imaging quality of synchrotron-based phase-contrast mammographic tomography, *J. Phys. D* **47**, 365401 (2014).
- [60] M. J. Kitchen, G. A. Buckley, T. E. Gureyev, M. J. Wallace, N. Andres-Thio, K. Uesugi, N. Yagi, and S. B. Hooper, CT dose reduction factors in the thousands using x-ray phase contrast, *Sci. Rep.* **7**, 15953 (2017).
- [61] K. Pearson, IX. Mathematical contributions to the theory of evolution. XIX. Second supplement to a memoir on skew variation, *Philos. Trans. R. Soc. A* **216**, 429 (1916).
- [62] K. Majidi, J. Li, C. Muehleman, and J. G. Brankov, Noise and analyzer-crystal angular position analysis for analyzer-based phase-contrast imaging, *Phys. Med. Biol.* **59**, 1877 (2014).
- [63] Attenuation coefficients were taken from <https://physics.nist.gov/PhysRefData/XrayMassCoef/tab3.html>; the form factor  $f_1$  was taken from <https://physics.nist.gov/PhysRefData/FFast/html/form.html>; and molar mass and density values were taken from <https://physics.nist.gov/PhysRefData/XrayMassCoef/tab1.html>
- [64] D. Attwood and A. Sakdinawat, *X-Rays and Extreme Ultraviolet Radiation: Principles and Applications* (Cambridge University Press, Cambridge, 2017).

- [65] F. Ferri, D. Magatti, L. A. Lugiato, and A. Gatti, Differential Ghost Imaging, *Phys. Rev. Lett.* **104**, 253603 (2010).
- [66] A. W. Stevenson, S. C. Mayo, D. Häusermann, A. Maksimenko, R. F. Garrett, C. J. Hall, S. W. Wilkins, R. A. Lewis, and D. E. Myers, First experiments on the Australian synchrotron imaging and medical beamline, including investigations of the effective source size in respect of x-ray imaging, *J. Synchrotron Radiat.* **17**, 75 (2010).
- [67] D. M. Eagleman, Visual illusions and neurobiology, *Nat. Rev. Neurosci.* **2**, 920 (2001).
- [68] S. A. Wallis and M. A. Georgeson, Mach bands and multiscale models of spatial vision: The role of first, second, and third derivative operators in encoding bars and edges, *J. Vision* **12**, 18 (2012).
- [69] B. I. Erkmen and J. H. Shapiro, Signal-to-noise ratio of Gaussian-state ghost imaging, *Phys. Rev. A* **79**, 023833 (2009).
- [70] P. Clemente, V. Durán, E. Tajahuerce, V. Torres-Company, and J. Lancis, Single-pixel digital ghost holography, *Phys. Rev. A* **86**, 041803(R) (2012).
- [71] Z. D. C. Kemp, Propagation based phase retrieval of simulated intensity measurements using artificial neural networks, *J. Opt.* **20**, 045606 (2018).
- [72] R. E. Meyers, K. S. Deacon, and Y. Shih, in *Quantum Communications and Quantum Imaging VI*, edited by R. E. Meyers, Y. Shih, and K. A. Deacon, SPIE Proc. Vol. 7092 art. 7092-15 (SPIE, Bellingham, 2008).
- [73] R. E. Meyers, K. S. Deacon, and Y. Shih, Turbulence-free ghost imaging, *Appl. Phys. Lett.* **98**, 111115 (2011).
- [74] R. E. Meyers, K. S. Deacon, and Y. Shih, Positive-negative turbulence-free ghost imaging, *Appl. Phys. Lett.* **100**, 131114 (2012).
- [75] P. Milanfar, *Super-Resolution Imaging* (CRC Press, Boca Raton, 2010).
- [76] 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.* **94**, 183602 (2005).
- [77] M. Aßmann and M. Bayer, Compressive adaptive computational ghost imaging, *Sci. Rep.* **3**, 1545 (2013).
- [78] S. Sun, W.-T. Liu, H.-Z. Lin, E.-F. Zhang, J.-Y. Liu, Q. Li, and P.-X. Chen, Multi-scale adaptive computational ghost imaging, *Sci. Rep.* **6**, 37013 (2016).
- [79] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science* **220**, 671 (1983).
- [80] T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, and T. Ito, Computational ghost imaging using deep learning, *Opt. Commun.* **413**, 147 (2018).