Optics Letters

Optimizing illumination for precise multi-parameter estimations in coherent diffractive imaging

DORIAN BOUCHET,[1,2,*] **JACOB SEIFERT,**[1] **AND ALLARD P. MOSK**[1]

[1]*Nanophotonics, Debye Institute for Nanomaterials Science, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands*
[2]*Present address: Université Grenoble Alpes, CNRS, LIPhy, 38000 Grenoble, France*
Corresponding author: dorian.bouchet@univ-grenoble-alpes.fr

Coherent diffractive imaging (CDI) is widely used to characterize structured samples from measurements of diffracting intensity patterns. We introduce a numerical framework to quantify the precision that can be achieved when estimating any given set of parameters characterizing the sample from measured data. The approach, based on the calculation of the Fisher information matrix, provides a clear benchmark to assess the performance of CDI methods. Moreover, by optimizing the Fisher information metric using deep learning optimization libraries, we demonstrate how to identify the optimal illumination scheme that minimizes the estimation error under specified experimental constraints. This work paves the way for an efficient characterization of structured samples at the sub-wavelength scale.

The fast and precise characterization of nanoscale devices is an essential aspect of advanced semiconductor manufacturing processes. It is thus crucial to ensure that optical measurements can reveal every important feature of nanostructured samples with an excellent precision. To achieve this goal, a common approach is to numerically reconstruct the permittivity distribution of the sample, either from interferometric measurements [1] or from intensity measurements via ptychography-like techniques [2]. In many cases of interest, some *a priori* knowledge of the sample is also available to the observer. For instance, in nanofabrication, the geometry of manufactured samples is usually known with high precision, and only a few critical parameters need to be monitored after the lithography process [3]. Typically, it is assumed that the sample can be described using a sparse representation in a known basis. Such an approach, referred to as sparsity-based coherent diffractive imaging (CDI), leads to a significant reduction in the number of parameters that need to be estimated from the measured diffraction patterns, therefore mitigating ill-posedness of the inverse problem that needs to be solved [4,5]. Furthermore, the resolution of reconstructed images is

not limited by Rayleigh's criterion, so that parameters can be estimated with sub-wavelength precision [6–9].

As for any imaging technique, an important aspect of sparsity-based CDI is to identify an optimized approach to illuminate the sample [10,11]. Formally, the estimation precision achievable with different incident fields can be compared using the Cramér–Rao lower bound (CRLB), which is a central concept in estimation theory. This concept is currently widely used in single-molecule localization microscopy [12,13] and in quantum metrology [14,15]. It has also been proposed as a new resolution measure for imaging systems [16,17], and the possibility to identify optimal incident fields that minimize the CRLB was recently investigated, for instance to localize a single particle in a complex environment [18] or to characterize a phase object hidden behind a scattering medium [19].

In this Letter, we describe a method to find illumination schemes that optimize the precision of parameter estimation in sparsity-based CDI. As an example, we present different approaches to characterize a parameterized sample composed of three vertical lines (Fig. 1), either by determining optimal positions for the incident field or by identifying the optimal design for a zone plate that shapes the incident field. In addition, we analyze the resulting CRLB in terms of contributions of the quantum fluctuations of coherent states, the absence of phase information in the measurements, and cross talk between
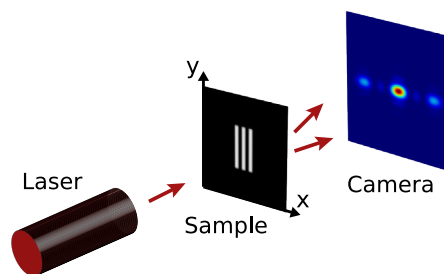

Fig. 1. Representation of a CDI setup used for the characterization of a parameterized sample composed of three vertical lines. A coherent light source illuminates the sample, and diffraction patterns are measured by a camera located in the detection plane.

parameters. These results offer new insights to improve the performance of methods based on CDI when the dose per acquisition may be limited, notably for the characterization of delicate samples or when high throughput is required.

In CDI, one seeks to characterize a sample by estimating a set of M parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_M)$ from measurements of one or several diffraction patterns that constitute the data $\boldsymbol{X}$. Noise fluctuations in the data impose a fundamental limit to the achievable precision on the determination of $\boldsymbol{\theta}$. Indeed, the covariance matrix Σ of any unbiased estimator of $\boldsymbol{\theta}$ must satisfy the Cramér–Rao inequality, which states that the matrix $(\Sigma - \mathcal{J}^{-1})$ is always nonnegative definite [20]. In this expression, the matrix $\mathcal{J}$ is known as the Fisher information matrix, defined by $\mathcal{J} = \langle [\nabla_{\boldsymbol{\theta}} \ln p(\boldsymbol{X}; \boldsymbol{\theta})][\nabla_{\boldsymbol{\theta}} \ln p(\boldsymbol{X}; \boldsymbol{\theta})]^T \rangle$, where $p(\boldsymbol{X}; \boldsymbol{\theta})$ is a joint probability density function, $\nabla_{\boldsymbol{\theta}}$ is a partial derivative operator defined by $\nabla_{\boldsymbol{\theta}} = (\partial/\partial\theta_1, \ldots, \partial/\partial\theta_M)^T$, and $\langle \cdots \rangle$ denotes the expectation operator acting over noise fluctuations. While the probability density function $p(\boldsymbol{X}; \boldsymbol{\theta})$ can describe any type of noise, we assume here that the values measured by the N_p pixels of the camera are statistically independent and follow a Poisson distribution, which corresponds to measurements limited only by shot noise. Considering a set of N_m diffraction patterns measured using different incident fields, the Fisher information matrix is then expressed by

$$[\mathcal{J}]_{ij} = \sum_{k,l} \frac{1}{I_{k,l}} \left(\frac{\partial I_{k,l}}{\partial \theta_i} \right) \left(\frac{\partial I_{k,l}}{\partial \theta_j} \right), \qquad (1)$$

where $I_{k,l}$ denotes the expected value of the intensity for the kth pixel and for the lth diffraction pattern. The resulting CRLB on the standard error on the estimated value of θ_i is given by

$$\mathcal{C}_i = \sqrt{[\mathcal{J}^{-1}]_{ii}} \, . \qquad (2)$$

This bound is asymptotically reached by maximum-likelihood (ML) estimators, which can be implemented by searching for the global maximum of the log-likelihood function [20,21].

In conventional CDI, it is impractical to calculate the CRLB due to the computational complexity of inverting the large Fisher information matrices that arise when samples are described by many parameters [22,23]. In contrast, the formalism is suitable to quantify the precision achievable with sparsity-based CDI, when samples can be described in sparse representations involving a reduced number of unknown parameters. In such cases, it is then possible to define an objective function that can be optimized to identify optimal illumination schemes tailored for the estimation of $\boldsymbol{\theta}$. For single-parameter estimations, the relevant objective function is simply given by the CRLB for the parameter [19]. For multi-parameter estimations, however, different relevant objective functions can be defined. As a possible objective function, one could choose the trace of $\mathcal{J}^{-1}$, which provides a measure of the average CRLB but does not guarantee that a controlled threshold value bounds the CRLB for every parameter (see Supplement 1, Section 1). For this reason, we use the spectral radius of $\mathcal{J}^{-1}$ as an objective function, which is defined as being the largest eigenvalue of $\mathcal{J}^{-1}$. The CRLB on the standard error on the estimated value of the first principal component is then expressed as follows:

$$\mathcal{C}_\rho = \sqrt{\rho\left(\mathcal{J}^{-1}\right)}, \qquad (3)$$

where $\rho(\mathcal{J}^{-1})$ denotes the spectral radius of $\mathcal{J}^{-1}$. The inequality $\mathcal{C}_i \leq \mathcal{C}_\rho$ holds for any parameter θ_i. Minimizing this objective function essentially leads to a reduction of the CRLB for the parameters that are the most difficult to estimate, a feature that is highly desirable for practical applications when the metrological specifications involve a single tolerance value that applies to all parameters.

To demonstrate the benefits of this approach in sparsity-based CDI, we consider a sample composed of three vertical lines (Fig. 1). These lines are separated from each other by a distance of 10 µm, each line being characterized by a width of 10 µm and a length of 100 µm. A sparse representation of the sample is obtained by describing these lines with 12 parameters $\boldsymbol{\theta} = (x_1, \ldots, x_6, y_1, \ldots, y_6)$, corresponding to the coordinates of the edges of the lines. We assume that the sample is illuminated with a coherent field at a wavelength $\lambda = 561$ nm. We choose a total number of photons incident on the sample of $n = 3 \times 10^6$; one can deduce the CRLB for other values of n by remarking that the CRLB for shot noise limited measurements scales with $1/\sqrt{n}$. Diffraction patterns are then calculated using a scalar diffraction approach by propagating the resulting field using the angular spectrum representation. This method allows us to calculate the expected value of the intensity $I_{k,l}$ that would be measured by a camera located at a distance $z = 10$ mm from the sample and, thus, to calculate the associated 12×12 Fisher information matrix using a finite-difference approximation of Eq. (1) (see Supplement 1, Section 2).

Tailoring the spatial distribution of the probe field provides us with degrees of freedom that can be tuned to minimize $\mathcal{C}_\rho$. In a constrained configuration, the shape of the distribution is fixed (e.g., a Gaussian beam), and it is desired to identify optimal values for the position of the probe field and its spatial extent. To solve this optimization problem, we employ the Adam optimizer, which is commonly used to train deep neural networks [24,25] and which is implemented in the open-source platform TensorFlow. We first consider the acquisition of four independent diffraction patterns, each of them obtained by illuminating the sample using a Gaussian beam with $n/4$ photons. The Adam optimizer is then used to identify the probe positions and the full width at half-maximum (FWHM) that minimize the CRLB for the first principal component $\mathcal{C}_\rho$. Note that such optimization procedure is especially effective when the *a priori* knowledge available on $\boldsymbol{\theta}$ is of the order of the FWHM of the probe field (see Supplement 1, Section 3). After the optimization process, the value of $\mathcal{C}_\rho$ is 44 nm [Fig. 2(a)], which is well below the wavelength of the incident light thanks to the sparse representation of the object. Optimal probe positions are identified at critical areas of the sample, with an optimized FWHM of 15 µm [Figs. 2(b) and 2(c) and Figs. 3(a)–3(h)]. This optimal illumination scheme can be interpreted as a trade-off between the necessity to illuminate all important areas of the object and the requirement to minimize the number of photons wasted by missing the object or the camera. For comparison, we performed the same analysis for a conventional ptychographic scheme. To ensure that the probes significantly overlap over the field of view [26], we chose a FWHM of 100 µm and four probe positions distributed in a square grid of side length 50 µm centered on the object. The value of $\mathcal{C}_\rho$ obtained with this conventional scheme is 127 nm, hence showing that $\mathcal{C}_\rho$ is reduced by a factor of 3 with the optimized scheme.
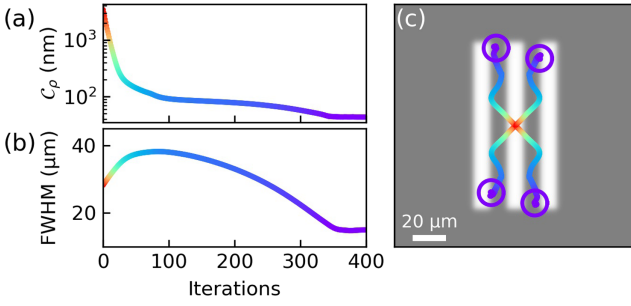
Fig. 2. Evolution of (a) the objective function $\mathcal{C}_\rho$ during the optimization process, as well as (b) the FWHM of the Gaussian probe field and (c) the probe positions represented in the sample plane. The purple circles shown in (c) are centered at the optimized positions, with a diameter equal to the optimized FWHM. The color of the curves represents the value of $\mathcal{C}_\rho$, ranging from 3.4 μm (red) to 44 nm (purple).

We can also use Eq. (2) to calculate the CRLB for each parameter after the minimization of $\mathcal{C}_\rho$ [Fig. 3(i)]. Interestingly, the formalism allows us to analyze the contribution of different error sources. Indeed, information is partly lost both because of the influence of parameter cross talk and because the phase of the field $\varphi_{k,l}$ is not captured by the measurements. When θ_i is to be estimated, other parameters can be considered as nuisance parameters that can increase the CRLB via cross talk [20]. Estimations of θ_i are the same regardless of whether other parameters are known or unknown only if $[\mathcal{J}]_{ij} = 0$ for $i \neq j$. We can thus assess the influence of parameter cross talk by calculating the lower bound on the standard error on the estimated value of θ_i as if the Fisher information matrix was diagonal. This bound is given by $\mathcal{C}_i' = 1/\sqrt{\mathcal{J}_i'}$, where

$$\mathcal{J}_i' = \sum_{k,l} \frac{1}{I_{k,l}} \left(\frac{\partial I_{k,l}}{\partial \theta_i} \right)^2. \tag{4}$$

In addition, the absence of phase measurements also leads to an increase of the CRLB. This can be assessed by calculating the lower bound on the standard error on the estimated value of θ_i assuming that both the intensity and the phase can be measured by the observer—the precision of estimations is then only limited by the quantum fluctuations of coherent states. This bound is expressed by $\mathcal{C}_i'' = 1/\sqrt{\mathcal{J}_i''}$, where $\mathcal{J}_i'' = 4 \sum |\partial E_{k,l}/\partial \theta_i|^2$ is the Fisher information corresponding to single-parameter estimation using an ideal homodyne detection scheme [19]. Note that $\mathcal{J}_i''$ is also equal to the quantum Fisher information associated with the estimation of a single parameter from uncorrelated coherent states [19,27]. Introducing $E_{k,l} = \sqrt{I_{k,l}} \exp(i\varphi_{k,l})$, we can decompose $\mathcal{J}_i''$ as follows:

$$\mathcal{J}_i'' = \sum_{k,l} \frac{1}{I_{k,l}} \left(\frac{\partial I_{k,l}}{\partial \theta_i} \right)^2 + 4 \sum_{k,l} I_{k,l} \left(\frac{\partial \varphi_{k,l}}{\partial \theta_i} \right)^2. \tag{5}$$

The two terms that appear in the second member of Eq. (5) can be interpreted as the Fisher information enclosed in the intensity and the phase of the detected field, respectively.

The different bounds that are introduced here satisfy the chain of inequalities $\mathcal{C}_i'' \leq \mathcal{C}_i' \leq \mathcal{C}_i \leq \mathcal{C}_\rho$, as can be seen in Fig. 3(i). The influence of parameter cross talk varies depending on the considered parameter, but we observe that parameters
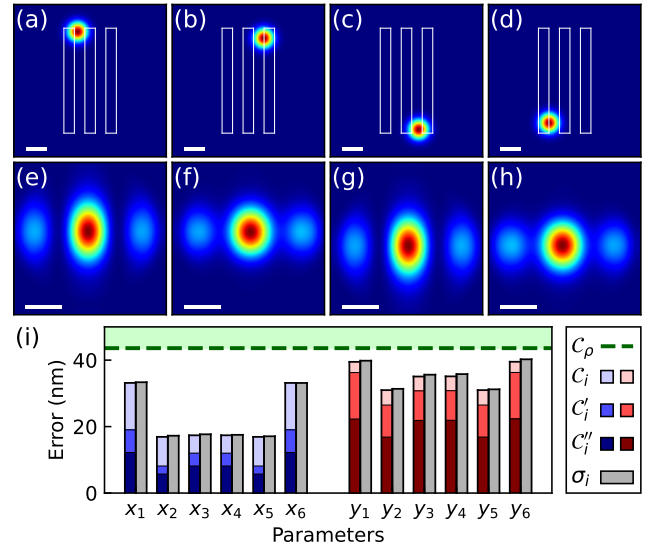

Fig. 3. (a)–(d) Spatial distributions of the excitation intensity in the sample plane for the optimal probe positions, assuming that four diffraction patterns are measured. The position of the sample is represented by white lines. Scale bars, 20 μm. (e)–(h) Spatial distributions of the intensity in the detection plane. Scale bars, 200 μm. (i) CRLB for each parameter after the minimization of $\mathcal{C}_\rho$, along with the RMS error obtained by performing ML estimations on 10^4 numerically generated diffraction patterns.

defining the x position of the line edges are more affected than those defining the y position of the line edges. Furthermore, after the propagation of the field to the detection plane, the Fisher information associated with intensity and phase measurements [first and second terms of the second member of Eq. (5), respectively] are approximately equal, which explains why the CRLB is then degraded by a factor close to $\sqrt{2}$ by the absence of phase information.

In order to show that the calculated CRLB can be approached with ML estimators, we numerically generate a set of 10^4 noisy diffraction patterns. For each pattern, we first randomly modify the value of each parameter according to a normal distribution, with a standard deviation of 0.5 μm. We then calculate the expected value of the intensity in the detection plane and use it to randomly generate noisy data with Poisson statistics. The value of all parameters is then estimated by maximizing the log-likelihood function with the Adam optimizer. The root mean square (RMS) error σ_i of the estimated values of each parameter is close to the fundamental limit $\mathcal{C}_i$ [Fig. 3(i)], which demonstrates here the efficiency of the ML estimator.

It is known that a structured illumination can improve the resolution of imaging techniques, which notably led to the development of randomized zone plates for use in ptychography [28,29]. Here, we can use our numerical framework to deterministically identify the design of the zone plate that is optimal for precisely characterizing the sample. To this end, we now consider a continuous transmission mask located at a distance of 10 mm upstream of the sample. The radius of the zone plate is set to 180 μm, so that the largest spatial frequency of the field in the sample plane is the same as for the Gaussian beams represented in Figs. 3(a)–3(d). Starting from a uniform initial guess, we run the Adam optimizer to find the design of the zone plate that minimizes $\mathcal{C}_\rho$ for a single-shot measurement
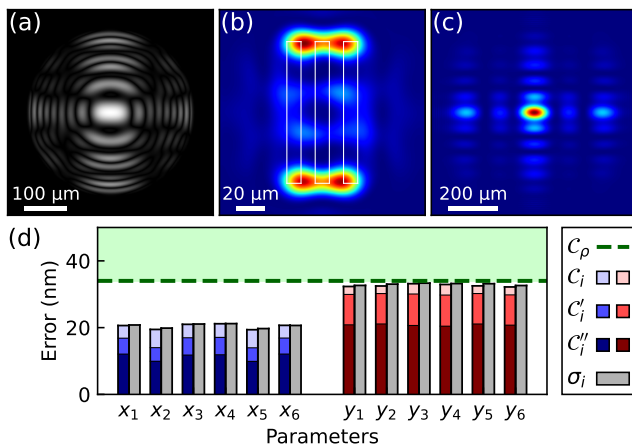
Fig. 4. (a) Optimal design of the zone plate that minimizes $\mathcal{C}_\rho$, assuming that a single diffraction pattern is measured. (b) Spatial distribution of the excitation intensity in the sample plane, as generated by the optimized zone plate. The position of the sample is represented by white lines. (c) Resulting spatial distribution of the intensity in the detection plane. (d) CRLB for each parameter after the minimization of $\mathcal{C}_\rho$, along with the RMS error obtained by performing ML estimations on 10^4 numerically generated diffraction patterns.

[Fig. 4(a)]. This zone plate generates an intensity in the sample plane that is high at all critical areas of the sample [Fig. 4(b)], producing a structured intensity pattern in the detection plane [Fig. 4(c)]. As shown in Fig. 4(d), the value of $\mathcal{C}_\rho$ resulting from the optimization process is 34 nm, which is significantly lower than the optimized value of 44 nm obtained in the case of the Gaussian beams. Thus, for a given total number of photons incident on the sample, a single-shot measurement using the optimized zone plate allows for a better precision on the estimation of θ as compared to what can be achieved with four measurements performed using a Gaussian beam illuminating the sample at different positions. This demonstrates the potential of optimized zone plates for the precise characterization of structured samples at high throughput, as often needed for industrial applications [3].

In summary, we calculated the CRLB to assess the precision achievable with sparsity-based CDI, and we used the formalism to identify optimal illumination schemes that allow all parameters to be precisely estimated while limiting the number of photons interacting with the sample. We envision that this strategy could be applied in future work by representing objects with different choices of basis functions, such as a wavelet basis or a basis of Gabor functions [30]. Implementing a Bayesian approach could also allow for more flexibility in the *a priori* knowledge that can be described using the formalism [20]. Furthermore, advanced numerical frameworks could be used to go beyond the first Born approximation and to characterize strongly scattering samples in two or three dimensions [2,31].

REFERENCES

1. O. Haeberlé, K. Belkebir, H. Giovaninni, and A. Sentenac, J. Mod. Opt. **57**, 686 (2010).
2. J. Rodenburg and A. Maiden, in *Springer Handbook of Microscopy* (Springer, 2019), pp. 819–904.
3. J. Alexander Liddle and G. M. Gallatin, Nanoscale **3**, 2679 (2011).
4. U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, IEEE Trans. Comput. Imaging **2**, 59 (2016).
5. H.-Y. Liu, D. Liu, H. Mansour, P. T. Boufounos, L. Waller, and U. S. Kamilov, IEEE Trans. Comput. Imaging **4**, 73 (2018).
6. A. Szameit, Y. Shechtman, E. Osherovich, E. Bullkich, P. Sidorenko, H. Dana, S. Steiner, E. B. Kley, S. Gazit, T. Cohen-Hyams, S. Shoham, M. Zibulevsky, I. Yavneh, Y. C. Eldar, O. Cohen, and M. Segev, Nat. Mater. **11**, 455 (2012).
7. P. Sidorenko, O. Kfir, Y. Shechtman, A. Fleischer, Y. C. Eldar, M. Segev, and O. Cohen, Nat. Commun. **6**, 8209 (2015).
8. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, R. G. Dixson, and M.-A. Henn, Light Sci. Appl. **5**, e16038 (2016).
9. T. Zhang, C. Godavarthi, P. C. Chaumet, G. Maire, H. Giovannini, A. Talneau, M. Allain, K. Belkebir, and A. Sentenac, Optica **3**, 609 (2016).
10. L. Bian, J. Suo, G. Situ, G. Zheng, F. Chen, and Q. Dai, Opt. Lett. **39**, 6648 (2014).
11. A. Muthumbi, A. Chaware, K. Kim, K. C. Zhou, P. C. Konda, R. Chen, B. Judkewitz, A. Erdmann, B. Kappes, and R. Horstmeyer, Biomed. Opt. Express **10**, 6351 (2019).
12. H. Deschout, F. C. Zanacchi, M. Mlodzianoski, A. Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, Nat. Methods **11**, 253 (2014).
13. Y. Shechtman, S. J. Sahl, A. S. Backer, and W. Moerner, Phys. Rev. Lett. **113**, 133902 (2014).
14. M. Szczykulska, T. Baumgratz, and A. Datta, Adv. Phys. X **1**, 621 (2016).
15. J. S. Sidhu and P. Kok, AVS Quantum Sci. **2**, 014701 (2020).
16. S. Ram, E. S. Ward, and R. J. Ober, Proc. Natl. Acad. Sci. **103**, 4457 (2006).
17. A. Sentenac, C.-A. Guérin, P. C. Chaumet, F. Drsek, H. Giovannini, N. Bertaux, and M. Holschneider, Opt. Express **15**, 1340 (2007).
18. D. Bouchet, R. Carminati, and A. P. Mosk, Phys. Rev. Lett. **124**, 133903 (2020).
19. D. Bouchet, S. Rotter, and A. P. Mosk, "Maximum information states for coherent scattering measurements," arXiv:2002.10388 (2020).
20. H. L. V. Trees, K. L. Bell, and Z. Tian, *Detection Estimation and Modulation Theory, Part I* (Wiley, 2013).
21. P. Thibault and M. Guizar-Sicairos, New J. Phys. **14**, 063004 (2012).
22. H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, J. Opt. Soc. Am. A **12**, 834 (1995).
23. X. Wei, H. P. Urbach, and W. M. J. Coene, Phys. Rev. A **102**, 043516 (2020).
24. D. P. Kingma and J. Ba, in *3rd International Conference on Learning Representations, San Diego* (2015).
25. G. Barbastathis, A. Ozcan, and G. Situ, Optica **6**, 921 (2019).
26. O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, Ultramicroscopy **108**, 481 (2008).
27. C. W. Helstrom, J. Stat. Phys. **1**, 231 (1969).
28. G. R. Morrison, F. Zhang, A. Gianoncelli, and I. K. Robinson, Opt. Express **26**, 14915 (2018).
29. M. Odstrčil, M. Lebugle, M. Guizar-Sicairos, C. David, and M. Holler, Opt. Express **27**, 14981 (2019).
30. H. H. Barrett and K. J. Myers, *Foundations of Image Science* (John Wiley & Sons, 2013).
31. R. J. Dilz and M. C. van Beurden, J. Comput. Phys. **345**, 528 (2017).