Complementary aspects of spatial resolution and signal-to-noise ratio in computational imaging

T. E. Gureyev,[1,2,3,4,*] D. M. Paganin,[2] A. Kozlov,[1] Ya. I. Nesterets,[5,3] and H. M. Quiney[1]
[1]*ARC Centre of Excellence in Advanced Molecular Imaging, The University of Melbourne, Parkville VIC 3010, Australia*
[2]*School of Physics and Astronomy, Monash University, Clayton VIC 3800, Australia*
[3]*School of Science and Technology, University of New England, Armidale NSW 2351, Australia*
[4]*Data61, Commonwealth Scientific and Industrial Research Organisation, Clayton VIC 3168, Australia*
[5]*Manufacturing, Commonwealth Scientific and Industrial Research Organisation, Clayton VIC 3168, Australia*

(Received 8 January 2018; published 15 May 2018)

A generic computational imaging setup is considered which assumes sequential illumination of a semitransparent object by an arbitrary set of structured coherent illumination patterns. For each incident illumination pattern, all transmitted light is collected by a photon-counting bucket (single-pixel) detector. The transmission coefficients measured in this way are then used to reconstruct the spatial distribution of the object's projected transmission. It is demonstrated that the square of the spatial resolution of such a setup is usually equal to the ratio of the image area to the number of linearly independent illumination patterns. If the noise in the measured transmission coefficients is dominated by photon shot noise, then the ratio of the square of the mean signal to the noise variance is proportional to the ratio of the mean number of registered photons to the number of illumination patterns. The signal-to-noise ratio in a reconstructed transmission distribution is always lower if the illumination patterns are nonorthogonal, because of spatial correlations in the measured data. Examples of imaging methods relevant to the presented analysis include conventional imaging with a pixelated detector, computational ghost imaging, compressive sensing, super-resolution imaging, and computed tomography.

DOI: 10.1103/PhysRevA.97.053819

I. INTRODUCTION

The duality between decomposition and synthesis is one of the most pervasive and fruitful ideas in mathematical physics. Key to rendering precise the decomposition-synthesis duality is the concept of expansions utilizing a superposition of basis objects drawn from a complete set. The process of decomposition involves breaking down an object belonging to a suitably wide class of possible objects, expressing it as a weighted superposition of objects in the complete basis set. The inverse process of synthesis takes the weighting coefficients referred to in the previous sentence, using them to synthesize the corresponding object from the said coefficients.

Geometrically, the decomposition-synthesis duality sets up an evident correspondence with linear algebra, a link that is developed further in the formalism of functional analysis [1]. This is particularly useful when the underpinning equations, such as the Maxwell equations for electromagnetic waves [2] or the parabolic equation of paraxial wave optics [3], are linear differential equations. The process of decomposition may thereby be viewed as determining all possible projections of a vector in a specified function space (the set of all possible vectors in the function space being associated with the set of all possible objects to be decomposed), with the process of synthesis corresponding to constructing a given vector (representing a particular object) from the knowledge of each of its projections [1].

A wide class of optical imaging scenarios may be viewed from this perspective. An obvious example is a pixelated charge-coupled device (CCD) camera, in which an arbitrary $L \times L$ pixel image may be synthesized as a superposition of one-pixel basis images, each of which has only one of the pixels uniformly illuminated [4]. Other examples relevant to optical imaging include the plane-wave (Fourier) basis [5], other complete-eigenfunction expansions such as Hermite-Gauss and Gauss-Laguerre bases [3], Floquet expansions [6], the Huygens construction as embodied in the Rayleigh-Sommerfeld diffraction integrals [7] and the convolution formulation of Fresnel diffraction theory [8], the Green's-function formalism for linear shift-variant (LSI) imaging systems and the convolution formalism for such systems [5], computational ghost imaging (CGI) [9,10], ghost imaging (GI) using random speckle bases [11,12], the wavelet decomposition of optical images [13], and the singular-value decomposition of computed tomography [14]. The basis elements in this list, which is far from exhaustive, range from being maximally localized (as in the pixel basis, together with its limit case given by the Dirac-delta basis) to being maximally delocalized (as in the plane-wave basis, many modal decomposition bases, and in convolution-type propagator formalisms embodying the Huygens construction). The wavelet basis corresponds to basis elements with a level of localization that is intermediate between the pixel basis or the Dirac-delta basis, and the delocalized bases mentioned above.

In a complete set utilized in the context of solving a given optical problem, the choice of basis elements is arbitrary. A corollary of this arbitrariness is the power it imparts to choose a basis in which the solution to a given optical problem assumes

*Corresponding author: timur.gureyev@unimelb.edu.au

a particularly convenient, transparent, tractable, or otherwise desirable form. Thus, for example, a plane-wave basis may be chosen in a free-space diffraction context on account of the analytical power of the associated machinery of Fourier analysis together with the extreme numerical efficiency of the fast Fourier transform algorithm [5], a spatially random speckle field basis may be chosen in a GI context on account of its amenability to compressive sensing concepts [15], a wavelet basis may be chosen when it is useful to have a sparse representation of an optical image, a modal basis may be useful in an optical communication context where only a small number of such modes is likely to be excited [16], and so forth. For example, in the case of parallel-beam computed tomography, if the incident illumination patterns are structured radially in the shape of suitable Zernike polynomials, then a simple and fast reconstruction of the object can be provided in terms of Chebyshev polynomials and the measured transmission coefficients [14].

A second corollary of this arbitrariness, of the choice of complete basis in the setting of the optical decomposition-synthesis duality, is the opportunity it affords to study optical imaging from a perspective that is not limited to a particular choice of complete basis. This is the perspective adopted by the present paper, in addressing the particular questions of spatial resolution and signal-to-noise ratio for intensity-linear optical imaging using arbitrary complete bases. While most of the results of the present study are generally applicable to arbitrary bases as applied to optical imaging, we also draw some conclusions specific to the comparison between the localized pixel basis associated with direct imaging, and less localized bases associated with indirect forms of imaging such as CGI [9].

There is an evident analogy between inline holography, viewed as the two-step process of image recording followed by reconstruction [17], and the generic means of indirect intensity imaging considered here. While this latter problem, which is the core topic of the present paper, also considers imaging as a two-step process of recording followed by reconstruction, in our case the object to be reconstructed is the projected transmission distribution of a semitransparent object which can be obtained by means of a linear transformation of the intensity registered by the detector, rather than both the intensity and phase of a complex disturbance.

We close this introduction with an outline of the remainder of the paper. Section II presents a formalism for describing a generic intensity-linear imaging setup whose ultimate purpose is to determine a spatially resolved estimate for an object transmission distribution obtained from a set of integral transmitted intensity measurements by a single-pixel detector when the object is illuminated by a sequence of distinct structured illumination patterns. Special cases that can be described by this model include, but are not limited to, direct imaging using a pixelated detector and indirect imaging of the impinging intensity distribution using computational imaging [9,10], GI [11,12], and coded aperture imaging [18]. The present formalism is developed utilizing GI terminology, particularly in utilizing the term bucket signal for what is in essence the total scattering cross section [11,12], but as emphasized earlier the domain of applicability of the formalism is much broader than GI. This section considers a means for determining the

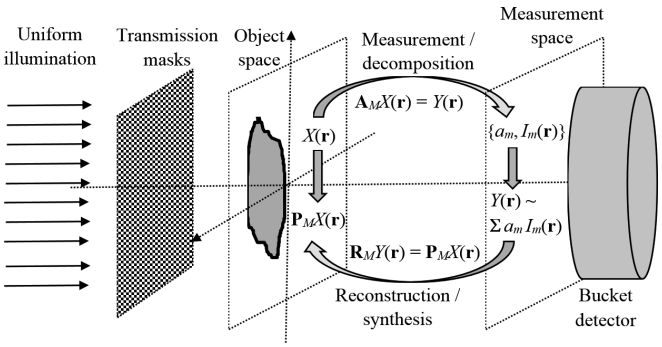

FIG. 1. Diagram of the imaging setup considered in the present paper.

point-spread function (PSF) associated with a given set of illumination patterns, both for the general case of a position-dependent PSF (Green's function), and the special case (where applicable) of a position-independent PSF. Section III then considers the associated question of spatial resolution for our rather general imaging setup, special cases of which include all of the previously mentioned forms of both direct and indirect imaging. The signal-to-noise ratio (SNR) in the reconstructed object transmission distribution is then considered in Sec. IV. The expression obtained for the SNR is factorized into a product of a function depending solely upon the photon statistics of the illuminating field and a function depending solely upon the object and basis functions. The analysis is then presented from the perspective of a recently introduced intrinsic quality characteristic [19,20]. This characteristic is invariant with respect to a broad class of intensity-linear shift-invariant transformations of imaging systems [20] and it effectively quantifies the optical quality of an imaging system (i.e., its efficiency of utilization of photons for imaging of the illuminated object). Section V draws specific comparison between the fully localized and nonlocalized bases, from the perspectives of SNR and spatial resolution of the reconstructed object transmission distribution, and the previously mentioned imaging quality metric. Technical details relevant to the examples considered in Sec. V can be found in the Appendixes. Section VI contains the conclusions, including a discussion of a possible classification of imaging systems on the basis of transformation of SNR and spatial resolution between the measurement (image) and the reconstruction (object) spaces.

II. A COMPUTATIONAL IMAGING SETUP

Consider Fig. 1, which shows an imaging setup where two-dimensional light intensity patterns $I_m(\mathbf{r})$, $\mathbf{r} = (x, y)$, $m = 1, 2, \ldots, M$, are used to illuminate, in sequence, a thin semitransparent object which is described by a deterministic dimensionless real non-negative transmission function $X(\mathbf{r})$. For simplicity, we consider here only the case of spatially coherent quasimonochromatic incident radiation. The illuminating patterns are assumed to have been implicitly integrated over some fixed exposure time, and therefore they are expressed in units of energy density (energy per unit area, $\mathrm{J/m^2}$). During each exposure, a bucket (single-pixel) photon-counting detector, with efficiency η (expressed in photons per joule) and a sufficiently large sensitive surface region Ω

collects all the transmitted light. The signal measured by the detector during the exposure with a given $I_m(\mathbf{r})$ is a single non-negative dimensionless number, $a_m = \eta \iint_\Omega X(\mathbf{r}) I_m(\mathbf{r}) d\mathbf{r}$, which corresponds to the number of registered photons. Such bucket signals may be measured at any distance from the object as long as the attenuation in the gap between the object and the detector is negligible. The goal of the imaging experiment is to reconstruct (or, at least, to approximate) the unknown object transmission function $X(\mathbf{r})$ from a set of measured bucket coefficients a_m, $m = 1,2,\ldots,M$.

The imaging scheme described by Fig. 1 may correspond to, for example, CGI [9–12], possibly in combination with compressive sensing [15,21], or to other imaging methods with structured illumination and a single-pixel detector. In the present paper, we are primarily interested in questions about the spatial resolution and signal-to-noise ratio that can be achieved in these types of imaging setup. Note, however, that our context is much broader than that of GI or CGI, notwithstanding our use of the term bucket for the photon detector.

In order to maximize the class of unknown transmission functions $X(\mathbf{r})$ that could be accurately reconstructed from the measurement of coefficients a_m, the illumination patterns $I_m(\mathbf{r})$ should be linearly independent, i.e., none of the functions $I_m(\mathbf{r})$ should be representable as a linear combination of other functions $I_{m'}(\mathbf{r})$, $m' \neq m$, with constant coefficients. If the set of patterns $I_m(\mathbf{r})$ is not linearly independent, the linear subspace spanned by the functions $I_m(\mathbf{r})$, $m = 1,2,\ldots,M$, will have a dimension $M' < M$, and the class of reconstructable transmission functions will be narrower compared to the linearly independent case. In some methods relevant to our study, the set of illumination patterns can be linearly dependent in practice, but that only represents a straightforward additional technicality for the model that we consider. For simplicity, in the present paper we assume that the set $I_m(\mathbf{r})$, $m = 1,2,\ldots,M$, is linearly independent.

In order to calculate the effect of photon shot noise on the coefficients a_m measured by the bucket detector, we assume that all $I_m(\mathbf{r})$ are obtained using the same incident illumination which is sufficiently monochromatic and has uniform intensity distribution, $I_{\text{in}}(\mathbf{r}) = I_{\text{in}}$, transmitted through different spatially varying transmission masks $T_m(\mathbf{r})$. So, for example, $I_m(\mathbf{r}) = I_{\text{in}} T_m(\mathbf{r})$ with $0 \leqslant T_m(\mathbf{r}) \leqslant 1$. The transmission distributions $T_m(\mathbf{r})$ are assumed to be deterministic (static) and dimensionless. A typical case is represented by the functions $T_m(\mathbf{r}) = \exp[-\mu L_m(\mathbf{r})]$, where μ is the linear attenuation coefficient of the mask material and $L_m(\mathbf{r})$ is the spatial distribution of the mask's projected thickness along the direction of the incident illumination. For simplicity we assume, where appropriate, that the integrals $t^2 \equiv \frac{1}{|\Omega|} \iint_\Omega T_m^2(\mathbf{r}) d\mathbf{r}$ are the same for all $m = 1,2,\ldots,M$, where $|\Omega|$ denotes the area of Ω. The illumination patterns are also assumed to be deterministic in our model, while the bucket coefficients are independent random variables with Poisson statistics, as discussed in detail below.

We will also consider an optional condition that the set $I_m(\mathbf{r})$ may satisfy.

Condition 1. The constant function, such that $f(\mathbf{r}) = 1$ everywhere in Ω, belongs to the linear space spanned by the functions $I_m(\mathbf{r})$, $m = 1,2,\ldots,M$, i.e., $1 = \sum_{m=1}^{M} \alpha_m I_m(\mathbf{r})$ for some constant coefficients α_m.

If condition 1 does not hold for a given set of illumination patterns $I_m(\mathbf{r})$, $m = 1,2,\ldots,M$, it means that the constant functions are linearly independent with respect to the set $I_m(\mathbf{r})$. In this case, it should usually be easy to add an extra illumination pattern $I_{M+1}(\mathbf{r}) = \text{const}$ [e.g., just a flat-field illumination with no mask, $I_{M+1}(\mathbf{r}) = 1$] to the set in order to satisfy condition 1.

We will see below that condition 1 expresses a form of an energy conservation law. In practical terms, condition 1 means that any uniform transmission function can be reconstructed from the measured bucket coefficients a_m, which is equivalent to the statement that the coefficients α_m in condition 1 can be expressed as linear combinations of the bucket coefficients $\eta \iint_\Omega I_m(\mathbf{r}) d\mathbf{r}$. The ability to reconstruct uniform transmission functions represents a desirable property expected from a well-designed imaging system. Condition 1 is usually satisfied in most real systems. This natural condition will be shown to lead to some convenient simplifications in our analysis of the spatial resolution and SNR below.

Mathematically, we will consider only square-integrable illumination patterns and transmission functions and view them as vectors from the linear space $L_2(\Omega)$ of square-integrable functions over Ω. Square-integrability of the illumination patterns is ensured by each illuminating pattern having a finite energy, while the square integrability of the transmission functions is ensured by the facts that they have finite transverse area and must vary between zero transmission and unity transmission. The scalar product of two vectors $\boldsymbol{V} \equiv V(\mathbf{r})$ and $\boldsymbol{U} \equiv U(\mathbf{r})$ is defined as

$$\langle \boldsymbol{V}, \boldsymbol{U} \rangle \equiv \frac{1}{|\Omega|} \iint_\Omega V(\mathbf{r}) U(\mathbf{r}) d\mathbf{r}. \tag{1}$$

We will denote the corresponding vector length as $\|\boldsymbol{V}\| = (\langle \boldsymbol{V}, \boldsymbol{V} \rangle)^{1/2}$. It will be convenient for us to work with dimensionless illumination vectors $\boldsymbol{W}_m \equiv W_m(\mathbf{r}) = \eta |\Omega| I_m(\mathbf{r})$, $m = 1,2,\ldots,M$. Note that the bucket coefficients can be represented as $a_m = \langle \boldsymbol{X}, \boldsymbol{W}_m \rangle$, where $\boldsymbol{X} \equiv X(\mathbf{r})$.

If the set of vectors $\boldsymbol{W}_m$, $m = 1,2,\ldots,M$, is linearly independent, they can always be orthonormalized [22,23]. Such orthonormalization is not unique. One can use, for example, the polar decomposition for this purpose. Let $\mathbf{W}$ be the matrix with elements $(W_{n\,m}) = \langle \boldsymbol{E}_n, \boldsymbol{W}_m \rangle$ in some orthonormal basis $\{\boldsymbol{E}_n\}$, i.e., the matrix that consists of vectors $\boldsymbol{W}_m = \mathbf{W}\boldsymbol{E}_m$ as its columns. The polar decomposition of $\mathbf{W}$ can be written as $\mathbf{W} = \mathbf{V}(\mathbf{W}^\dagger \mathbf{W})^{1/2}$, where the superscript "†" denotes transposition, $(\mathbf{W}^\dagger \mathbf{W})^{1/2}$ is a positive-definite matrix, and $\mathbf{V}$ is an orthogonal matrix. In particular, all columns of the matrix $\mathbf{V} = \mathbf{W}(\mathbf{W}^\dagger \mathbf{W})^{-1/2}$ are orthonormal with respect to the scalar product Eq. (1). Therefore, the set of vectors $\boldsymbol{V}_m = \mathbf{V}\boldsymbol{E}_m$ represents an orthonormalization of vectors $\boldsymbol{W}_m$, $m = 1,2,\ldots,M$:

$$\boldsymbol{V}_m = \sum_{m'=1}^{M} q_{m\,m'} \boldsymbol{W}_{m'}, \quad \text{and}$$
$$\langle \boldsymbol{V}_m, \boldsymbol{V}_{m'} \rangle = \delta_{mm'}, \quad m, m' = 1,2,\ldots,M, \tag{2}$$

where $\delta_{mm'}$ is the Kronecker delta. The coefficients $q_{m\,m'}$ in Eq. (2) correspond to the matrix $\mathbf{Q} = (\mathbf{W}^\dagger \mathbf{W})^{-1/2}$. The matrix

$\mathbf{Q}^2 = (\mathbf{W}^\dagger\mathbf{W})^{-1}$ with elements $q^{(2)}_{mm'} = \sum_{l=1}^{M} q_{ml}q_{lm'}$ can be used to construct a basis $\{\boldsymbol{U}_m\}$ that is biorthogonal to $\{\boldsymbol{W}_m\}$:

$$\begin{aligned} \boldsymbol{U}_m &= \sum_{m'=1}^{M} q_{m'm}\boldsymbol{V}_{m'} = \sum_{m'=1}^{M} q^{(2)}_{mm'}\boldsymbol{W}_{m'}, \\ \langle \boldsymbol{U}_m, \boldsymbol{W}_{m'}\rangle &= \delta_{mm'}, \\ m,m' &= 1,2,\ldots,M. \end{aligned} \tag{3}$$

The bases $\{\boldsymbol{V}_m\}$ and $\{\boldsymbol{U}_m\}$ are used below for reconstructing the object transmission function from the measured bucket coefficients.

In all realistic imaging systems only a finite number of elements in the previously mentioned bases will ever be used in a given experiment. However, by analogy to the utility of considering bases with infinitely many members (e.g., the infinite set of Fourier harmonics), it is useful to supplement the set of vectors $\boldsymbol{V}_m$, $m = 1,2,\ldots,M$ with suitable additional vectors to form a complete orthonormal basis $\{\boldsymbol{V}_m\}$, $m = 1,2,\ldots,\infty$. This complete basis spans the Hilbert space $L_2(\Omega)$ with the scalar product defined by Eq. (1). Obviously, the whole functional space $L_2(\Omega)$ can be represented as a sum of the subspace $V_M(\Omega)$, spanned by the first M basis vectors, and its orthogonal complement: $L_2(\Omega) = V_M(\Omega) \oplus V_M^{\perp}(\Omega)$.

The measured signal $\{a_m, \boldsymbol{W}_m\}$ can be represented as a vector $\boldsymbol{Y} = \sum_{m=1}^{M} a_m \boldsymbol{W}_m$ in the subspace $V_M(\Omega)$. When the vectors $\boldsymbol{W}_m$, $m = 1,2,\ldots,M$, are linearly independent, the coefficients a_m in this representation of the signal vector are unique, i.e., if $\boldsymbol{Y} = \sum_{m=1}^{M} a_m \boldsymbol{W}_m = \sum_{m=1}^{M} a'_m \boldsymbol{W}_m$, then $a_m = a'_m$ for all m. It is useful to introduce the (linear) measurement operator $\mathbf{A}_M$ which maps the object transmission function $X(\mathbf{r})$ into the measured signal vector: $\mathbf{A}_M X = \boldsymbol{Y} = \sum_{m=1}^{M} a_m \boldsymbol{W}_m$. By definition, the reconstruction (synthesis) operator $\mathbf{R}_M$ is equal to the inverse of the restriction of the measurement operator $\mathbf{A}_M$ to the subspace $V_M(\Omega)$. Let us show that $\mathbf{R}_M\boldsymbol{Y} = \sum_{m=1}^{M} a_m \boldsymbol{U}_m$, and that such reconstruction represents a projection of the object transmission function $\boldsymbol{X} = X(\mathbf{r})$ onto the subspace $V_M(\Omega)$.

Consider a projection operator from $L_2(\Omega)$ onto the subspace $V_M(\Omega)$. This operator acts on an arbitrary vector $\boldsymbol{X}$ from $L_2(\Omega)$ according to the expression $\mathbf{P}_M\boldsymbol{X} = \sum_{m=1}^{M} b_m \boldsymbol{V}_m$, $b_m \equiv \langle \boldsymbol{X}, \boldsymbol{V}_m\rangle$. Using Eq. (3), it is then easy to verify that

$$\begin{aligned} \mathbf{P}_M\boldsymbol{X} &= \sum_{m=1}^{M} b_m \boldsymbol{V}_m = \sum_{m=1}^{M} c_m \boldsymbol{W}_m = \sum_{m=1}^{M} a_m \boldsymbol{U}_m, \\ b_m &= \sum_{m'=1}^{M} q_{m\,m'}a_{m'}, \\ c_m &= \sum_{m'=1}^{M} q^{(2)}_{m\,m'}a_{m'}. \end{aligned} \tag{4}$$

In particular, Eq. (4) shows that $\mathbf{R}_M\boldsymbol{Y} = \mathbf{P}_M\boldsymbol{X}$, so that this reconstruction indeed corresponds to the projection of the object transmission function onto the vector subspace spanned by the illumination patterns. Note that, if the original illumination vectors $\boldsymbol{W}_m$, $m = 1,2,\ldots,M$, are themselves orthogonal and all have the same length, i.e., $\langle \boldsymbol{W}_m, \boldsymbol{W}_{m'}\rangle = w^2\delta_{mm'}$ with some constant $w > 0$, then we can take $\boldsymbol{U}_m = w^{-1}\boldsymbol{V}_m = w^{-2}\boldsymbol{W}_m$, $m = 1,2,\ldots,M$, $q_{mn} = w^{-1}\delta_{mn}$, and $c_m = w^{-1}b_m = w^{-2}a_m$, greatly simplifying Eq. (4).

Let us see under which conditions the reconstructed function $\mathbf{R}_M Y(\mathbf{r}) = \mathbf{P}_M X(\mathbf{r})$ represents a good approximation to the object transmission function $X(\mathbf{r})$. It is easy to verify that the operator $\mathbf{P}_M$ is a projector, i.e., $\mathbf{P}_M^2 = \mathbf{P}_M$ [24], and it can be represented as a linear integral operator

$$\mathbf{P}_M X(\mathbf{r}) = \frac{1}{|\Omega|}\iint_{\Omega} G_M(\mathbf{r},\mathbf{r}')X(\mathbf{r}')d\mathbf{r}' \tag{5}$$

with the kernel (Green's function)

$$G_M(\mathbf{r},\mathbf{r}') = \sum_{m=1}^{M} U_m(\mathbf{r})W_m(\mathbf{r}') = \sum_{m=1}^{M} V_m(\mathbf{r})V_m(\mathbf{r}'). \tag{6}$$

From a physical perspective, the operator $\mathbf{P}_M$ projects an arbitrary object $\boldsymbol{X}$ to its approximation as a linear combination of the M independent basis vectors. This generalized form of filtering, which is consistent with the concept of reconstructing an object to within a finite resolution, may naturally be considered in geometric terms as a projection since it discards any components of the object which are orthogonal to all of the basis vectors. This intuitively aligns with the formal definition of an operator $\mathbf{P}$ being a projector if $\mathbf{P}^2 = \mathbf{P}$, since applying a projection more than once has the same effect as applying it once.

In some imaging methods, such as CGI [9,10], the illumination patterns are often chosen in such a way that the following optional condition is satisfied.

Condition 2. The set of illumination patterns is such that the corresponding Green's function is shift invariant (spatially stationary), at least approximately, in the sense that $G_M(\mathbf{r}+\mathbf{h},\mathbf{r}'+\mathbf{h}) = G_M(\mathbf{r},\mathbf{r}')$ for any vector $\mathbf{h}$, provided that $\mathbf{r}$, $\mathbf{r}'$, $\mathbf{r}+\mathbf{h}$, and $\mathbf{r}'+\mathbf{h}$ lie within Ω.

If condition 2 holds, the Green's function can be represented as a function of the difference between the two coordinates: $G_M(\mathbf{r},\mathbf{r}') = P_M(\mathbf{r}-\mathbf{r}')$, where the corresponding function $P_M(\mathbf{r}) = \sum_{m=1}^{M} V_m(0)V_m(\mathbf{r})$ acts as a PSF [25,26]:

$$\mathbf{P}_M X(\mathbf{r}) = \frac{1}{|\Omega|}\iint_{\Omega} P_M(\mathbf{r}-\mathbf{r}')X(\mathbf{r}')d\mathbf{r}'. \tag{7}$$

Note that condition 1 implies that the spatial average of the PSF is equal to 1: $1 = \mathbf{P}_M 1(\mathbf{r}) = \frac{1}{|\Omega|}\iint_{\Omega} P_M(\mathbf{r}-\mathbf{r}')d\mathbf{r}'$. Hence, the convolution with the PSF does not change the total transmission through the object:

$$\begin{aligned} \iint_{\Omega} \mathbf{P}_M X(\mathbf{r})d\mathbf{r} &= \iint_{\Omega} \frac{1}{|\Omega|}\iint_{\Omega} P_M(\mathbf{r}-\mathbf{r}')X(\mathbf{r}')d\mathbf{r}'d\mathbf{r} \\ &= \iint_{\Omega} X(\mathbf{r}')d\mathbf{r}'. \end{aligned}$$

In this sense, condition 1 implies a natural conservation law which states that the projection onto subspace $V_M(\Omega)$ preserves the total transmission value of the object.

The measurement operator $\mathbf{A}_M$ can also be represented as a linear integral operator:

$$\mathbf{A}_M X(\mathbf{r}) = \frac{1}{|\Omega|}\iint_\Omega A_M(\mathbf{r},\mathbf{r}')X(\mathbf{r}')d\mathbf{r}' \tag{8}$$

with the kernel (Green's function)

$$A_M(\mathbf{r},\mathbf{r}') = \sum_{m=1}^{M} W_m(\mathbf{r})W_m(\mathbf{r}'). \tag{9}$$

Indeed, if we substitute Eq. (9) into Eq. (8), and take $X(\mathbf{r}) = \sum_{m=1}^{M} a_m U_m(\mathbf{r})$, we obtain

$$\begin{aligned}\frac{1}{|\Omega|}\iint_\Omega \sum_{m=1}^{M} W_m(\mathbf{r})W_m(\mathbf{r}')X(\mathbf{r}')d\mathbf{r}' &= \sum_{m=1}^{M} W_m(\mathbf{r})\left\langle \sum_{m'=1}^{M} a_{m'}\boldsymbol{U}_{m'}, \boldsymbol{W}_m\right\rangle \\ &= \sum_{m=1}^{M} a_m W_m(\mathbf{r}), \quad \text{i.e. } \mathbf{A}_M X(\mathbf{r}) = Y(\mathbf{r}),\end{aligned}$$

as required. Similarly, the reconstruction operator $\mathbf{R}_M$ can be represented as a linear integral operator with the kernel $R_M(\mathbf{r},\mathbf{r}') = \sum_{m=1}^{M} U_m(\mathbf{r})U_m(\mathbf{r}')$. When the illumination patterns are orthogonal, such that $\langle \boldsymbol{W}_m, \boldsymbol{W}_{m'}\rangle = w^2\delta_{mm'}$, both the measurement and the reconstruction operator are proportional to the identity operator in $V_M(\Omega)$: $\mathbf{A}_M\mathbf{P}_M X(\mathbf{r}) = w^2\mathbf{P}_M X(\mathbf{r})$ and $\mathbf{R}_M Y(\mathbf{r}) = w^{-2}Y(\mathbf{r})$.

III. SPATIAL RESOLUTION

In order for the reconstructed function $\mathbf{R}_M Y(\mathbf{r}) = \mathbf{P}_M X(\mathbf{r})$ to be a good approximation for $X(\mathbf{r})$, for an arbitrary imaged object $X(\mathbf{r})$, Eq. (7) indicates that the PSF, $P_M(\mathbf{r})$, should approximate Dirac's delta function. Equation (6) then becomes a completeness (closure) relation: $|\Omega|^{-1}\sum_{m=1}^{M} V_m(\mathbf{r})V_m(\mathbf{r}') \cong \delta(\mathbf{r}-\mathbf{r}')$ [24]. In general, however, the projection operator $\mathbf{P}_M$ effectively blurs any function, $X(\mathbf{r})$, on which it acts, degrading the spatial resolution as a result. We will see below that when the basis vectors correspond to the Fourier harmonics, the blurring due to the convolution with the relevant PSF corresponds to truncation of the Fourier decomposition of the object transmission function to the linear combination of the first M Fourier components (low-pass filtration). One can see that the degree of blurring in Eq. (5), as measured by the width of the reconstructed response to a delta-function-like input signal $X(\mathbf{r}')$, can be different for different points $\mathbf{r}$. However, when the Green's function is shift invariant, the width of the response to a localized input is the same at every point in Ω, being equal to the width of the PSF. This width corresponds to the spatial resolution of the imaging setup.

According to [20,27,28], a convenient measure of the width of a function $g(\mathbf{r})$ can be given by the expression

$$(\Delta_2 r)[g] \equiv \frac{\iint g(\mathbf{r})d\mathbf{r}}{\left(\iint g^2(\mathbf{r})d\mathbf{r}\right)^{1/2}}. \tag{10}$$

To motivate this definition, let H be the peak value of $g(\mathbf{r})$, and observe that the right-hand side of Eq. (10) may be approximated by the ratio of $H\Delta_2^2$ (height multiplied by area of base) to the square root of $H^2\Delta_2^2$ (square of the height multiplied by area of base), such a ratio therefore being equal to Δ_2 (width of the base). Unlike the more conventional measure of width, related to the variance of a function, the definition in Eq. (10) works well with the formalism of vector decomposition over an orthogonal basis, as demonstrated below. At the same time, the width defined according to Eq. (10) produces values which are fully consistent with the natural understanding of the width of a function in the case of Gaussians, Lorentzians, rectangular (uniform), and other popular distributions [20]. It is also important for the following to note that the definition of spatial resolution in Eq. (10) is optimistic, in the sense that for any integrable function $g(\mathbf{r})$ the following inequality holds: $(\Delta_2 r)[g] \leqslant (3\pi^{1/2}/2)(\Delta r)[g]$, where $\{(\Delta r)[g]\}^2 \equiv \iint |\mathbf{r}-\bar{\mathbf{r}}|^2|g(\mathbf{r})|d\mathbf{r}/\iint |g(\mathbf{r})|d\mathbf{r}$ [20]. In other words, the spatial resolution estimated in accordance with Eq. (10) is always finer than or equal to the more conventional spatial resolution, defined via the spatial variance of the PSF, multiplied by the constant $3\pi^{1/2}/2 \cong 2.66$. Details about the relationship between the two definitions of the spatial resolution (width of the PSF) can be found in Ref. [20].

First we consider the spatial resolution of the function $g_{\mathbf{r}'}(\mathbf{r}) = G(\mathbf{r}',\mathbf{r}) = G(\mathbf{r},\mathbf{r}')$ which depends on the argument $\mathbf{r}'$ as a parameter. We will assume for simplicity that condition 1 is satisfied. The corresponding results without condition 1 can be derived in the same way, but the expressions are more complex. When the constant function 1 lies in the space $V_M(\Omega)$, we have $\mathbf{P}_M 1 = 1$, and hence

$$\iint_\Omega g_{\mathbf{r}'}(\mathbf{r})d\mathbf{r} = \iint_\Omega G_M(\mathbf{r}',\mathbf{r})d\mathbf{r} = |\Omega|(\mathbf{P}_M 1)(\mathbf{r}') = |\Omega|.$$

Next, using the orthonormality of vectors $\boldsymbol{V}_m$, we obtain

$$\begin{aligned}\iint_\Omega g^2_{\mathbf{r}'}(\mathbf{r})d\mathbf{r} = \iint_\Omega G^2_M(\mathbf{r}',\mathbf{r})d\mathbf{r} &= \sum_{m=1}^{M} V_m^2(\mathbf{r}')\iint_\Omega V_m^2(\mathbf{r})d\mathbf{r} \\ &= |\Omega|\sum_{m=1}^{M} V_m^2(\mathbf{r}').\end{aligned}$$

Therefore,

$$(\Delta_2 r)[g_{\mathbf{r}}] = \left(|\Omega|/\sum_{m=1}^{M} V_m^2(\mathbf{r})\right)^{1/2}. \tag{11}$$

Obviously, for reasons discussed earlier, this resolution may in general be different at different points $\mathbf{r}$ in Ω. Note that the spatial average of the function $f_M(\mathbf{r}) = \sum_{m=1}^{M} V_m^2(\mathbf{r})$ over Ω is equal to M, due to the normalization of vectors $\boldsymbol{V}_m$:

$$\frac{1}{|\Omega|}\iint_\Omega \sum_{m=1}^{M} V_m^2(\mathbf{r})\,d\mathbf{r} = \sum_{m=1}^{M} \|\boldsymbol{V}_m\|^2 = M.$$

If the Green's function is shift invariant, the function $f_M(\mathbf{r})$ is constant in Ω, because $f_M(\mathbf{r}) = \sum_{m=1}^{M} V_m^2(\mathbf{r}) = G_M(\mathbf{r},\mathbf{r}) = G_M(0,0) = f_M(0)$, and hence $f_M(\mathbf{r}) = M$ for any $\mathbf{r}$. Therefore, in the shift-invariant case, $\iint_\Omega g^2_{\mathbf{r}'}(\mathbf{r})d\mathbf{r} = M|\Omega|$ and $(\Delta_2 r)[g_{\mathbf{r}}] = (|\Omega|/M)^{1/2}$. As mentioned previously, in the

shift-invariant case, the Green's function can be represented as a PSF, $G_M(\mathbf{r},\mathbf{r}') = P_M(\mathbf{r}-\mathbf{r}')$, and so the width of the Green's function in this case is equal everywhere to the width of the PSF, which determines the uniform spatial resolution:

$$(\Delta_2 r)[P_M] = (|\Omega|/M)^{1/2}. \tag{12}$$

The square of the width of the PSF can be interpreted as the effective pixel area. Equation (12) shows that the square of the spatial resolution of the considered computational imaging system using M illumination patterns $I_m(\mathbf{r})$, $m = 1,2,\ldots,M$, which are linearly independent and satisfy conditions 1 and 2, is always equal to the image area divided by the number of illumination patterns. This simple result is consistent, for example, with the expected spatial resolution of a 2D imaging system with a low-pass filter retaining only the first M Fourier harmonics. It is interesting that this intuitive result holds for the much more general class of bases considered here, than just the Fourier basis. We consider some relevant illustrative examples in more detail below.

By calculating the average width of the Green's function $A_M(\mathbf{r},\mathbf{r}') = \sum_{m=1}^{M} W_m(\mathbf{r})W_m(\mathbf{r}')$ using Eq. (10) as above, it is possible to verify that the spatial resolution of the measured signal $Y(\mathbf{r}) = \sum_{m=1}^{M} a_m\, W_m(\mathbf{r})$ is equal to

$$(\Delta_2 r)[A_M] = \left(|\Omega| \sum_m \sum_{m'} \langle \boldsymbol{W}_m \rangle \langle \boldsymbol{W}_{m'} \rangle \langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle / \sum_m \sum_{m'} \langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle^2 \right)^{1/2}.$$

In the orthogonal case, i.e., when $\langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle = w^2 \delta_{mm'}$, we have $\langle 1 \rangle = \langle \sum_{m=1}^{M} \langle W_m \rangle U_m(\mathbf{r}) \rangle = w^{-2} \sum_{m=1}^{M} \langle W_m \rangle^2$ as a consequence of condition 1. Using this fact it easy to verify that $(\Delta_2 r)[A_M]$ becomes equal to $(|\Omega|/M)^{1/2}$. Comparing this result with Eq. (12), we see that in the case of orthogonal illumination patterns the average spatial resolution is the same in the measurement and in the object spaces.

It is interesting to compare Eq. (12) with previously published results on spatial resolution in CGI. It is known that the spatial resolution in GI with thermal or pseudothermal light is limited primarily by the transverse coherence length in the detector plane, or, equivalently, by the speckle size [12,25,29–32]. The result is effectively the same in the case of CGI, where the spatial resolution is limited by the aperture of the spatial light modulator (SLM) in the case of far-field illumination conditions [9,33,34]. The imaging scheme considered in the present paper (Fig. 1) corresponds to CGI in the near-field regime corresponding to a very short distance between the SLM and the object. In that case, the spatial resolution is limited by the effective pitch (pixel) size h of the SLM [12,33], which affects the spatial resolution in the same way as the coherence length does in the case of far-field illumination. Comparing this with Eq. (12), we note that, while it is still true that in the case of a very large number of illumination patterns the spatial resolution is limited by the pitch size h of the transmission masks, Eq. (12) in fact offers a complementary result which holds in the case of an arbitrary (including very small) number of illumination patterns. The pitch limit replaces the result given by Eq. (12) only when the number of illumination patterns M becomes larger than $D = |\Omega|/h^2$, because in that case the set of illumination patterns can no longer be linearly independent. Note that D represents the dimensionality of the vector space of illumination patterns with area Ω and pitch size h. Therefore, any set with $M > D$ patterns must be linearly dependent. This can be clearly seen in the example with the pixelated masks considered in Sec. V.

IV. SIGNAL-TO-NOISE RATIO

Now let us consider the signal-to-noise ratio (SNR) of the reconstructed distribution described by Eq. (4) or, equivalently, Eq. (5). Here the signal at each point $\mathbf{r}$ is formally defined as an ensemble averaged mean value, $\overline{\mathbf{P}_M X}(\mathbf{r}) = \sum_{m=1}^{M} \bar{a}_m U_m(\mathbf{r})$, of the function $\mathbf{R}_M Y(\mathbf{r}) = \mathbf{P}_M X(\mathbf{r})$ reconstructed multiple times using Eq. (4) with coefficients a_m measured under identical experimental conditions; $\bar{a}_m$ are the mean values of the measured coefficients. The random character of the measurements of coefficients a_m is assumed to be determined by the typical behavior of a photon-counting detector, as discussed below. The corresponding noise is defined as the standard deviation of the reconstructed values $\mathbf{P}_M X(\mathbf{r})$ at a given point $\mathbf{r}$. The SNR is then

$$\mathrm{SNR}[\mathbf{P}_M X](\mathbf{r}) = \frac{\overline{\mathbf{P}_M X}(\mathbf{r})}{[\mathrm{Var}(\mathbf{P}_M X)(\mathbf{r})]^{1/2}}, \tag{13}$$

where $\mathrm{Var}(\mathbf{P}_M X)(\mathbf{r}) = \overline{[\mathbf{P}_M X(\mathbf{r}) - \overline{\mathbf{P}_M X}(\mathbf{r})]^2}$ is the corresponding noise variance. Using Eq. (4), we find that

$$\overline{[\mathbf{P}_M X(\mathbf{r}) - \overline{\mathbf{P}_M X}(\mathbf{r})]^2} = \overline{\left[\sum_{m=1}^{M} (a_m - \bar{a}_m) U_m(\mathbf{r})\right]^2} = \sum_{m=1}^{M} \sum_{m'=1}^{M} \mathrm{Cov}(a_m, a_{m'}) U_m(\mathbf{r}) U_{m'}(\mathbf{r}),$$

where $\mathrm{Cov}(a_m, a_{m'}) = \overline{a_m a_{m'}} - \bar{a}_m \bar{a}_{m'}$ are the covariances of the measured coefficients. We also assume that the measurements of different coefficients a_m are statistically independent and hence $\mathrm{Cov}(a_m, a_{m'}) = \delta_{mm'} \mathrm{Var}(a_m)$ and, therefore, $\mathrm{Var}(\mathbf{P}_M X)(\mathbf{r}) = \sum_{m=1}^{M} \mathrm{Var}(a_m) U_m^2(\mathbf{r})$. Hence,

$$\mathrm{SNR}[\mathbf{P}_M X](\mathbf{r}) = \sum_{m=1}^{M} \bar{a}_m U_m(\mathbf{r}) \Big/ \left[\sum_{m=1}^{M} \mathrm{Var}(a_m) U_m^2(\mathbf{r})\right]^{1/2}. \tag{14}$$

In order to estimate the noise in the measured coefficients a_m, we use the conventional semiclassical model of statistical optics [27]. According to this approach, the propagation of light is calculated for continuous classical deterministic electromagnetic fields and the quantization and randomness is considered in connection with light sources and photodetection. For most thermal and similar light sources, to a very good approximation, the dominant contribution to noise in the registered signals comes from photodetection shot noise and, possibly, also from other detector-related noise, such as electronic dark current noise [27]. For simplicity, we consider here only the case of a perfect photon-counting detector, for which only the photon shot noise is significant.

A photon-counting detector converts the incident radiant energy (integrated over the detector's entrance surface) into the corresponding photon numbers, with the efficiency of the process described by the previously introduced detection efficiency constant η which has the dimensionality of photons per Joule [27]. In this process of conversion, a deterministic flux of light energy entering the detector results in the stochastic process of photon counting satisfying Poisson statistics [27]. Therefore, the mean values of the measured coefficients $a_m = \eta I_{\rm in} \iint_\Omega X(\mathbf{r}) T_m(\mathbf{r}) d\mathbf{r}$ are equal to $\bar{a}_m = \bar{n}\, x_m$, where $\bar{n} = \eta\, I_{\rm in}\, |\Omega|$ is the mean number of photons used in the measurement of each coefficient, and $x_m = \frac{1}{|\Omega|} \iint_\Omega X(\mathbf{r}) T_m(\mathbf{r}) d\mathbf{r}$ are dimensionless transmission coefficients. Similarly, according to the properties of Poisson statistics of the shot noise, $\mathrm{Var}(a_m) = \bar{a}_m = \bar{n} x_m$.

Let $\bar{N} \equiv M\,\bar{n}$ be the mean total number of photons utilized in the whole experiment (this already takes into account the detection efficiency). Then, in view of Eq. (4), Eq. (14) can be rewritten as a product of two distinct factors:

$$\mathrm{SNR}[\mathbf{P}_M X](\mathbf{r}) = (\bar{N}/M)^{1/2} F_{M,X}(\mathbf{r}), \tag{15}$$

where the first factor, $(\bar{N}/M)^{1/2}$, reflects the effect of the photon statistics on the SNR, and the second term,

$$F_{M,X}(\mathbf{r}) \equiv \left(\sum_{m=1}^{M} x_m S_m(\mathbf{r})\right) \Bigg/ \left(\sum_{m=1}^{M} x_m S_m^2(\mathbf{r})\right)^{1/2}, \tag{16}$$

is a deterministic function of $\mathbf{r}$, where, by definition, the vectors $\boldsymbol{S}_m \equiv S_m(\mathbf{r})$ constitute a biorthogonal basis for $T_m(\mathbf{r})$: $\langle \boldsymbol{S}_m, \boldsymbol{T}_{m'} \rangle = \delta_{mm'}$. We will call $F_{M,X}(\mathbf{r})$ the form factor. This factor depends only on the set of transmission masks $T_m(\mathbf{r})$ and the object transmission function $X(\mathbf{r})$, but not on the number of photons. Note that while the vectors $\boldsymbol{T}_m$ correspond to the original illumination vectors normalized by the mean number of incident photons, such that $\boldsymbol{T}_m = \bar{n}^{-1} \boldsymbol{W}_m$, the vectors $\boldsymbol{S}_m$ represent the suitably normalized versions of the biorthogonal vectors $\boldsymbol{U}_m$: $\boldsymbol{S}_m = \bar{n}\, \boldsymbol{U}_m$.

Consider the case when the illumination vectors $\boldsymbol{W}_m \equiv \eta |\Omega| I_m(\mathbf{r})$, $m = 1,2,...,M$, are orthogonal and have the same length with respect to the scalar product defined in Eq. (1), so that $\langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle = w^2 \delta_{mm'}$, where $w = \bar{n} t$ and $t = \|T_m\|$ for all m. Here, we have $\boldsymbol{S}_m = t^{-2} \boldsymbol{T}_m$, $\langle \boldsymbol{S}_m, \boldsymbol{T}_{m'} \rangle = t^{-2} \bar{n}^{-2} \langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle = \delta_{mm'}$, which leads to an expression for the form factor in terms of the orthogonal transmission masks:

$$F_{M,X}(\mathbf{r}) = \left(\sum_{m=1}^{M} x_m T_m(\mathbf{r})\right) \Bigg/ \left(\sum_{m=1}^{M} x_m T_m^2(\mathbf{r})\right)^{1/2}. \tag{17}$$

The form factor in Eq. (17), and hence also the corresponding SNR, still generally depend on the object transmission function $X(\mathbf{r})$, as well as on the transmission masks $T_m(\mathbf{r})$, and can have different values at different points $\mathbf{r}$ inside Ω. In practice, when the performance of an imaging system is evaluated, the SNR is often measured in flat (uniform) areas of the reconstructed images, which are much larger than the area of the system's PSF. As the effect of uniform attenuation, $X(\mathbf{r}) = \mathrm{const}$, in a flat object is trivial, we shall only consider the flat SNR in the absence of the object, in which case $X(\mathbf{r}) = 1$. Considering the definition of the SNR in Eq. (13) with $X(\mathbf{r}) = 1$, we note first that, as a consequence of condition 1, $\overline{\mathbf{P}_M 1}(\mathbf{r}) = 1$. Since the noise variance in the reconstruction of a uniform object for arbitrary illumination patterns may still be position dependent, we shall spatially average the value of the noise variance, $\mathrm{Var}(\mathbf{P}_M 1)(\mathbf{r})$, in the denominator of Eq. (13). This leads to the notion of squared averaged flat SNR, SNR_a^2, defined as the ratio of the spatially averaged squared reconstructed signal to the spatially averaged noise variance, in the absence of an object:

$$\mathrm{SNR}_a^2 \equiv \frac{\langle [\overline{\mathbf{P}_M 1}(\mathbf{r})]^2 \rangle}{\langle \mathrm{Var}(\mathbf{P}_M 1)(\mathbf{r}) \rangle} = (\bar{N}/M) \big(F_{M,1}^a\big)^2. \tag{18}$$

In the derivation of Eq. (18), we used Eq. (15) and introduced the squared flat averaged form factor,

$$\big(F_{M,1}^a\big)^2 \equiv 1 \Big/ \left(\sum_{m=1}^{M} t_m \, \|\boldsymbol{S}_m\|^2\right), \tag{19}$$

which is defined by spatially averaging the squared denominator of Eq. (16), with $X(\mathbf{r}) = 1$ and $t_m \equiv \langle \boldsymbol{T}_m \rangle$. Note that the flat SNR defined in Eq. (18) does not depend on the imaged object, and hence is an intrinsic characteristic of the imaging system.

As the vectors $\boldsymbol{S}_m$ are biorthogonal to $\boldsymbol{T}_m$, and, in particular, $\langle \boldsymbol{S}_m, \boldsymbol{T}_m \rangle = 1$, the length of the vectors $\boldsymbol{S}_m$ must be larger than or equal to the inverse of the length of $\boldsymbol{T}_m$, so that $\|\boldsymbol{S}_m\| \geqslant t^{-1}$. Therefore $\sum_{m=1}^M t_m \|\boldsymbol{S}_m\|^2 \geqslant t^{-2} \sum_{m=1}^M t_m$, equality being achieved only in the case when the length of all vectors $\boldsymbol{S}_m$ is equal to t^{-1}. The latter is possible only when all vectors $\boldsymbol{S}_m$ are parallel to $\boldsymbol{T}_m$, which is the orthogonal case; the corresponding result can also be obtained directly from Eq. (17). Therefore, $\mathrm{SNR}_a^2 \leqslant (\bar{N}/M)\, \tilde{t}_M$, where the quantity $\tilde{t}_M \equiv t^2 / \sum_{m=1}^M t_m$ is a particular form of average transmission coefficient for a given set of transmission masks. Note that $\tilde{t}_M = \sum_{m=1}^M t_m^2 / \sum_{m=1}^M t_m$ in the orthogonal case, because $1 = \sum_{m=1}^M t_m \langle S_m \rangle = t^{-2} \sum_{m=1}^M t_m^2$ as a consequence of condition 1. When $\tilde{t}_M$ is multiplied by the mean number of incident photons, it makes the SNR_a^2 equal to the suitably averaged number of photons registered in each individual measurement, in the case of orthogonal illumination patterns and no object. This result agrees with the naturally expected behavior in the case of image noise dominated by photon shot noise. Note that this SNR does not increase with the number of measurements, as the measurements of individual bucket coefficients are independent. Instead, when the reconstruction makes use of all M individual measurements, the increased number of measured bucket coefficients is translated into a larger number of effective pixels in the reconstructed distribution, i.e., it results in improved spatial resolution, rather than in an increased SNR in each pixel.

We have shown that squared reconstructed flat SNR, SNR_a^2, is always equal to or smaller than the average number of photons registered in each individual measurement of a bucket coefficient, with equality achieved only in the case of orthogonal illumination patterns. The fact that, in the case of nonorthogonal illumination patterns, the SNR is smaller than the value expected in the case of uncorrelated Poisson statistics is related to the presence of effective spatial correlations between the data obtained with individual

illumination patterns, even though the measurements of individual bucket coefficients a_m are statistically independent. This phenomenon is studied further in Sec. V below.

It has previously been argued [19,20] that the ratio of SNR to spatial resolution, divided by the square root of the incident photon fluence, provides a good measure of the quality of an imaging system. The latter ratio, which was called the intrinsic quality characteristic (IQC) in Ref. [19], is also an invariant of the system which does not change under linear filtering [20]. Therefore, it is useful to estimate this characteristic in the case of computational imaging systems considered in the present paper. It follows from Eqs. (11) and (15) that, when the space $V_M(\Omega)$ contains constant functions (i.e., under condition 1), we have

$$Q_2(\mathbf{r}) \equiv \left(\frac{|\Omega|}{\bar{N}}\right)^{1/2} \frac{\mathrm{SNR}[\mathbf{P}_M X](\mathbf{r})}{(\Delta_2 r)[g_{\mathbf{r}}]} = F_{M,X}(\mathbf{r}) \left(\sum_{m=1}^{M} V_m^2(\mathbf{r})/M\right)^{1/2}. \tag{20}$$

The subscript index 2 in the notation $Q_2(\mathbf{r})$ reflects the fact that the spatial resolution is calculated according to Eq. (10), rather than on the basis of the spatial variance of the PSF [20]. When the Green's function is shift invariant (i.e., under condition 2), we have $\sum_{m=1}^{M} V_m^2(\mathbf{r}) = M$, and hence

$$Q_2(\mathbf{r}) = F_{M,X}(\mathbf{r}). \tag{21}$$

As we noted above, the form factor $F_{M,X}(\mathbf{r})$ is independent of the number of photons used in the imaging experiment, but it can exhibit distinct spatial distributions for different illumination masks and imaged objects.

An object-independent flat version of the IQC can be defined in terms of the flat SNR following Eqs. (18) and (19):

$$Q_{2,a} = F_{M,1}^{a} \leqslant (\tilde{t}_M)^{1/2}, \tag{22}$$

with $F_{M,1}^{a}$ defined in Eq. (19) and $\tilde{t}_M = t^2/\sum_{m=1}^{M} t_m$. When the illumination vectors satisfy the orthogonality condition $\langle \boldsymbol{W}_m, \boldsymbol{W}_{m'} \rangle = w^2 \delta_{mm'}$, it follows that $Q_{2,a} = (\tilde{t}_M)^{1/2}$.

When one is interested in estimating the efficiency of an imaging system with respect to the radiation dose delivered to the sample, which is proportional to the photon fluence incident on the object rather than on the masks, it can be useful to consider the quantity

$$\tilde{Q}_{2,a} \equiv (\bar{N}/\bar{N}_t)^{1/2} Q_{2,a} \leqslant M^{1/2} t / \sum_{m=1}^{M} t_m, \tag{23}$$

where $\bar{N}_t \equiv \bar{N} M^{-1} \sum_{m=1}^{M} t_m$ is the mean total number of photons incident on the object, after transmission through the masks, during the measurements of M bucket coefficients.

It is also possible to calculate the squared flat SNR in the measured signal $Y_{1,M}(\mathbf{r})$ corresponding to $X(\mathbf{r}) = 1$, by evaluating

$$\mathrm{SNR}_{a,\mathrm{in}}^2 \equiv \langle [\overline{Y_{1,M}}(\mathbf{r})]^2 \rangle / \langle \mathrm{Var}(Y_{1,M}\mathbf{r}) \rangle = (\bar{N}/M) \sum_m \sum_{m'} t_m t_{m'} \langle \boldsymbol{T}_m, \boldsymbol{T}_{m'} \rangle / \sum_m t_m \|\boldsymbol{T}_m\|^2.$$

This value becomes equal to $(\bar{N}/M)\tilde{t}_M$ in the orthogonal case, i.e., when $\langle \boldsymbol{T}_m, \boldsymbol{T}_{m'} \rangle = t^2 \delta_{mm'}$. This means, in particular, that in the case of orthonormal illumination patterns the flat SNR does not change in the reconstruction process. As we showed previously that the spatial resolution also does not change between the measurement and the object spaces in this case, it follows that the flat IQC $Q_{2,a} = (\tilde{t}_M)^{1/2}$ is invariant under the action of the reconstruction operator $\mathbf{R}_M$ in the case of orthogonal illumination patterns.

Similarly to the case of spatial resolution considered at the end of Sec. III, here we would like to compare the result given by Eq. (15) with previously published results on SNR in CGI. It has been widely accepted that the SNR in GI is proportional to the square root of the number of illumination patterns [25,33,35]. Comparing this with Eq. (15), we note that such behavior corresponds to the increase of the number of photons $\bar{N}$ in Eq. (15), which increases with each additional illumination pattern. The fact that Eq. (15) contains the number of illumination patterns, M, in the denominator instead of the numerator is due to the specific formulation of the problem in the present paper. Indeed, we do not assume that an increase in the number of illumination patterns is always associated with an increase in the total number of incident photons. Instead, we assume that a fixed mean number of incident photons $\bar{N}$ (the dose) can be distributed over a different number of illumination patterns M. In this formulation, when the number of illumination patterns increases, the mean number of photons in each pattern, $\bar{N}/M$, actually decreases. In addition to this difference in the formulation of the problem, the difference in the approach to the spatial resolution between our paper and most of the published works on CGI is also crucial to Eq. (15), as the SNR is usually calculated per pixel or per spatial resolution unit. As discussed at the end of Sec. III above, previously published papers on CGI typically assume that the spatial resolution is independent of the number of illumination patterns and is limited only by the transverse coherence length (in the case of far-field illumination) or the SLM pitch size (in the case of near-field illumination). Thus, the previously published results about SNR in CGI, compared to the results obtained in our paper, effectively correspond to the limit of a large number of illumination patterns, where the accumulation of additional illumination patterns does not affect the spatial resolution, but increases the number of incident photons, with the number of photons determining the SNR. In contrast, in the present paper we study the effect of a finite number of illumination patterns and their potential non-orthogonality on the SNR and spatial resolution of computational imaging systems. In order to illustrate the role of the key parameters of computational imaging systems affecting their SNR and spatial resolution, we consider several simple model systems below.

V. EXAMPLES

A. Localized illumination masks

Let us consider an example which is based on a conventional imaging setup with a pixelated detector with $L \times L$ pixels (for which the issues of spatial resolution, SNR, and IQC have been previously obtained using similar criteria [20]), but presented here in a form consistent with the imaging

setup with structured illumination and single-pixel detector as described in Sec. II of this paper. For this purpose, we define a set of $M = L^2$ single-pixel illumination patterns, $V_m(\mathbf{r}) = LT_m(\mathbf{r})$, where L is an integer and $T_m(\mathbf{r}) = T(\mathbf{r} - \mathbf{r}_m)$ are single-pixel transmission masks. The points $\mathbf{r}_m$ denote the positions of the centers of the detector pixels, each with area h^2, indexed by $m = 1,2,...,M$. As expected, the spatial resolution in this imaging system can be shown equal to (see details in Appendix A)

$$(\Delta_2 r)^2[g_{\mathbf{r}}] = |\Omega|/M = h^2. \tag{24}$$

The squared SNR is equal to

$$\mathrm{SNR}^2[\mathbf{P}_M X](\mathbf{r}) = (\bar{N}/M^2)\langle X(\mathbf{r})\rangle_m, \tag{25}$$

where $\langle X(\mathbf{r})\rangle_m = h^{-2}\iint\limits_{\Omega_{m(\mathbf{r})}} X(\mathbf{r}')d\mathbf{r}'$ is the average object transmission in the pixel with index $m(\mathbf{r})$, i.e., the pixel that contains the point $\mathbf{r}$. A similar result is obtained for the flat SNR: $\mathrm{SNR}_a^2 = (\bar{N}/M)(t^2/\sum_{m=1}^M t_m) = \bar{N}/M^2$, since $t^2 = t_m = \frac{1}{|\Omega|}\iint\limits_{\Omega_m} dr = 1/M$ in this case. The extra factor M^{-1} in the denominator of Eq. (25) reflects the fact that incident photons are used very inefficiently when single-aperture masks and a bucket detector are used instead of uniform illumination (no masks) and a position-sensitive detector (the masks block most of the incident photons). Note, however, that as far as the radiation dose delivered to the sample is concerned, the present method is as efficient as the direct imaging with a pixelated detector. Here, the average transmission coefficient of the masks is equal to $M^{-1}\sum_{m=1}^M t_m = M^{-1}$. Therefore, the SNR calculated with respect to the number of photons transmitted through the masks, i.e., the number of photons incident on the object, $\bar{N}_t = \bar{N}/M$, is equal to $\mathrm{SNR}_a^2 = \bar{N}_t/M$, which is the same as in the case of imaging with a pixelated detector, no masks, and the total number of incident photons equal to $\bar{N}_t$.

Equations (24) and (25) lead to the following result for the IQC:

$$Q_2^2(\mathbf{r}) = \frac{|\Omega|}{\bar{N}}\frac{\bar{N}\,x_{m(\mathbf{r})}}{M\,h^2} = x_{m(\mathbf{r})} = \frac{\langle X(\mathbf{r})\rangle_m}{M}. \tag{26}$$

If we recalculate this quantity with respect to the photons incident on the object by dividing $Q_2^2(\mathbf{r})$ by the average transmission coefficient of the masks, $M^{-1}\sum_{m=1}^M t_m = M^{-1}$, we arrive at the result $\tilde{Q}_2^2(\mathbf{r}) \equiv MQ_2^2(\mathbf{r}) = \langle X(\mathbf{r})\rangle_m$. In other words, the IQC of the system with single-pixel masks, evaluated with respect to the incident dose delivered to the sample, is equal to the square root of the absorption coefficient of the object. Considering the case $X(\mathbf{r}) = 1$ we obtain $Q_{2,a}^2 = M^{-1}$ and $\tilde{Q}_{2,a} = 1$. This coincides with the result previously derived in Ref. [20] by a different approach.

Consider now the imaging system with overlapping two-pixel transmission masks $T_m^{(2)}(\mathbf{r}) = T_m(\mathbf{r}) + T_{m+1}(\mathbf{r})$, $m = 1,2,\ldots,M$, where M is odd and the index m is formally considered cyclical, i.e., $T_{M+1}(\mathbf{r}) \equiv T_1(\mathbf{r})$. Such overlapping masks are no longer orthogonal, as $\langle \boldsymbol{T}_m^{(2)}, \boldsymbol{T}_{m+1}^{(2)}\rangle = 1/M$. It is straightforward to verify that the biorthogonal basis here consists of the vectors

$$S_m^{(2)}(\mathbf{r}) = (M/2)\left[\sum_{m'=1}^{m}(-1)^{m-m'}T_{m'}(\mathbf{r}) + \sum_{m'=m+1}^{M}(-1)^{m+1-m'}T_{m'}(\mathbf{r})\right].$$

Therefore, $\|\boldsymbol{S}_m^{(2)}\|^2 = (M^2/4)\sum_{m'=1}^M \|\boldsymbol{T}_{m'}\|^2 = M^2/4$. Evaluating the squared flat form factor according to Eq. (19) and substituting it into Eq. (22), we obtain that the squared flat IQC is equal to

$$Q_{2,a}^2 = 1/\left(\sum_{m=1}^{M} t_m^{(2)}\left\|\boldsymbol{S}_m^{(2)}\right\|^2\right) = 1/\left(\sum_{m=1}^{M}(2/M)(M^2/4)\right) = 2/M^2.$$

The corresponding IQC corrected for the average mask transmission coefficient $M^{-1}\sum_{m=1}^M t_m^{(2)} = 2/M$ is equal to $\tilde{Q}_{2,a}^2 = 1/M$. We see that both these squared IQCs are smaller, by factors proportional to $1/M$, than the corresponding characteristics for the single-mask system considered above. This is an instructive example of the effect of nonorthogonality of illumination patterns on the SNR and IQC of imaging systems. Note that the object-space spatial resolution of the system with overlapping two-pixel masks $T_m^{(2)}(\mathbf{r})$ is the same as the one for the system with single-pixel masks $T_m(\mathbf{r})$, as expected in the case of super-resolution imaging with subpixel shifting of a PSF. However, as we see from the above results, the SNR achieved in such super-resolution imaging will be considerably smaller compared to the SNR in the single-pixel imaging system, at the same incident photon fluence. In other words, in order to achieve the same image quality (i.e., the same SNR and the same spatial resolution) as in the system with the PSF of area h^2, by subpixel shifting of the PSF of area $2h^2$, one would have to increase the radiation dose delivered to the imaged object by a factor of M, where Mh^2 is the field of view.

B. Delocalized harmonic masks

Here we consider a computational imaging system defined as in Sec. II, but with the harmonic transmission masks $T_m(\mathbf{r}) = 1/2 + F_m(\mathbf{r})$, $F_m(\mathbf{r}) = f_{m_x}(x)f_{m_y}(y)$, $f_1(t) = 1/2$, $f_l(t) = (1/\sqrt{2})\sin(\pi lt/A)$ when $l = 2,4,\ldots$, is even, and $f_l(t) = (1/\sqrt{2})\cos[\pi(l-1)t/A]$, when $l = 3,5,\ldots$, is odd. The image domain is represented by the square $\Omega = \{-A/2 < x < A/2, -A/2 < y < A/2\}$. Note that we have normalized all Fourier harmonics in the above definition of transmission masks in a way that ensures that the transmission values are non-negative and do not exceed 1 at any point $\mathbf{r}$ in Ω.

The spatial resolution evaluated in accordance with Eq. (10) is equal to

$$(\Delta_2 r)^2[g_{\mathbf{r}}] = |\Omega|/M, \tag{27}$$

provided that we restrict the values of M to $M = 1 + 4M'$ for some positive integer M' (see Appendix B for details).

The SNR is described by Eq. (15) with

$$F_{M,X}(\mathbf{r}) = Q_2(\mathbf{r}) = \frac{x_1 + 4\sum_{m=2}^{M}(3x_m - 2x_1)F_m(\mathbf{r})}{\left\{x_1\left[1 - 16\sum_{m=2}^{M}F_m(\mathbf{r})\right] + 16\sum_{m=2}^{M}\sum_{m'=2}^{M}(9x_m\delta_{mm'} + 4x_1)F_m(\mathbf{r})F_{m'}(\mathbf{r})\right\}^{1/2}}. \tag{28}$$

Using Eq. (28) with $x_m = t_m$ and taking into account that $\langle \boldsymbol{F}_m \rangle = 0$, when $m > 1$, we obtain the following value for the flat IQC:

$$Q_{2,a}^2 = \left(F_{M,1}^a\right)^2 = t_1^2 \Big/ \left[t_1 + \sum_{m=2}^{M}(9t_m + 4t_1)\right] = 3/[4 + 40(M-1)]. \tag{29}$$

Dividing $Q_{2,a}^2$ by the average transmission coefficient of the masks, $M^{-1}\sum_{m=1}^{M} t_m = [(3/4 + (1/2)(M-1)]/M \cong 1/2$, when $M \gg 1$, we obtain that $\tilde{Q}_{2.a}^2 \cong 3/[2 + 20(M-1)]$ for large M.

One can see that here the value of the IQC, calculated with respect to the incident dose delivered to the sample, rapidly decreases when the number of illumination patterns M grows. This can be contrasted with the constant value $\tilde{Q}_{2,a} = 1$ obtained for general orthogonal illumination patterns at the end of Sec. V. This difference is a typical consequence of nonorthogonality of illumination patterns. The low value of the IQC here is related to the fact that the average signal in the considered computational imaging setup receives contribution from only one measurement, i.e., that of coefficient a_1, while the noise variances from measurements of all coefficients a_m, $m = 1, \ldots, M$, add up. In other words, the measurements of all nonzero coefficients a_m do not contribute to the average signal, but they do contribute to the noise. Note, however, that the measurements of nonzero coefficients a_m are still not useless, as far as the quality measure $Q_{2,a}$ is concerned, as these coefficients contribute to the spatial resolution, as is evident in Eqs. (11) and (12).

C. Pseudorandom delocalized masks

This example has a direct relationship to standard GI using random speckle fields [26], CGI [9–12] and compressive sensing [21]. Here we assume that the first illumination pattern is spatially uniform, $I_1(\mathbf{r}) \equiv I_1 = \text{const}$, and the other illumination patterns satisfy the equation [26]

$$\frac{1}{|\Omega|}\iint_{\Omega}[I_m(\mathbf{r}) - I_1][I_{m'}(\mathbf{r}) - I_1]d\mathbf{r} = \sigma^2\delta_{mm'}, \quad m,m' = 2,3,\ldots,M, \tag{30}$$

where the spatial average of the intensity patterns is assumed to be the same, $I_1 = \frac{1}{|\Omega|}\iint_{\Omega} I_m(\mathbf{r})d\mathbf{r}$, for all $m = 1, \ldots, M$, and $\sigma^2 = |\Omega|^{-1}\iint_{\Omega}[I_m(\mathbf{r}) - I_1]^2 d\mathbf{r}$ is the spatial variance of the intensity patterns, assumed to be the same for all $m > 1$ [the spatial variance of the pattern $I_1(\mathbf{r}) = \text{const}$ is obviously equal to zero]. Physically, Eq. (30) amounts to the natural condition that when distinct background-subtracted pseudorandom illumination patterns are multiplied together, they average to zero.

The spatial resolution here is described by Eq. (10), while the SNR is expressed by Eq. (15) with

$$F_{M,X}(\mathbf{r}) = Q_2(\mathbf{r}) = \frac{x_1 + \kappa^2\sum_{m=2}^{M}(x_m - x_1)F_m(\mathbf{r})}{\left\{x_1\left[1 - 2\kappa^2\sum_{m=2}^{M}F_m(\mathbf{r})\right] + \kappa^4\sum_{m=2}^{M}\sum_{m'=2}^{M}(x_m\delta_{mm'} + x_1)F_m(\mathbf{r})F_{m'}(\mathbf{r})\right\}^{1/2}}, \tag{31}$$

where $\kappa = I_1/\sigma \geqslant 1$ and $F_m(\mathbf{r}) = I_m(\mathbf{r})/I_1 - 1$, $m = 2,3,\ldots,M$ (see the details in Appendix C). As in the case of harmonic illumination masks, the form factor here can in general be different for different transmission functions and different points in the image domain. The squared flat IQC in this example is equal to

$$Q_{2,a}^2 = \left(F_{M,X}^a\right)^2 = t_1^2 \Big/ \left[t_1 + \kappa^2\sum_{m=2}^{M}(t_m + t_1)\right] = t_1/[1 + 2\kappa^2(M-1)], \tag{32}$$

where we have taken into account that the average transmission of all masks is assumed to be the same, i.e., $t_m = t_1$ for any m. Dividing $Q_{2,a}^2$ by the average transmission coefficient of the masks, which is equal to t_1, we obtain $\tilde{Q}_{2,a}^2 \cong 1/[1 + 2\kappa^2(M-1)]$. Therefore, the value of the IQC, calculated with respect to the incident dose delivered to the sample, rapidly decreases when the number of illumination patterns M grows. As in the case with overlapping two-pixel masks and harmonic delocalized masks considered above, this behavior is a consequence of nonorthogonality of illumination patterns.

Because the IQC of imaging systems with pseudorandom delocalized masks rapidly deteriorates as the number of illumination patterns increases, it appears that it can only remain reasonably high if the number of measurements M can be kept low. This effectively requires that the class of imaged objects is strongly correlated with a small number of illumination patterns, ensuring that the representation of any unknown imaged object $X(\mathbf{r})$ is sparse in the basis of vectors $I_m(\mathbf{r})$. As this is exactly the central premise of compressive sensing imaging [15,21], the method can be saved by this sparsity assumption. Indeed, the present argument can be viewed as motivating the necessity to use compressive sensing and/or related approaches in CGI of scenes with $M \gg 1$ resolution elements using $M \gg 1$ pseudorandom illumination patterns.

VI. DISCUSSION AND CONCLUSIONS

The major difference in the estimated behavior of the SNR in the case of computational imaging using orthogonal illumination patterns, as in the case of single-pixel masks, and the nonorthogonal delocalized harmonic masks or the pseudorandom illumination patterns satisfying Eq. (30), can be attributed to the nonorthogonality in the latter cases. The SNR significantly deteriorates when the coefficients c_m used for the reconstruction of the unknown object transmission function according to Eq. (4) are calculated from the experimentally measured coefficients a_m containing random noise. A decrease in SNR inevitably takes place when one subtracts two measured values, such as a_m, $m > 1$, and a_1, because the mean values of the coefficients are subtracted, reducing the magnitude of the resultant signal, but the variances add up, increasing the resultant noise. This unfavorable effect is well known in many forms of imaging that involve postdetection processing of experimentally registered light intensity distributions. For example, it is exactly this effect that is known to lead to low SNR and correspondingly large low-frequency artefacts in in-line phase-contrast imaging with the transport-of-intensity equation [36–38], when the signal of interest is obtained by subtraction of two images collected at different defocus distances.

The considerations presented in the previous paragraph can in fact be applied to a generalized model of imaging systems described by Eq. (4), where the signal of interest is reconstructed by means of an equation which is linear with respect to the measured values, a_m, which contain Poisson or other uncorrelated random noise. Note that the measured coefficients a_m can correspond to intensities measured at individual pixels of a detector or to integrated intensities collected with a structured illumination pattern and a single-pixel (bucket) detector, or to something else still. In fact, we can drop the reference to the method by which the input values a_m are measured, and also ignore the particular methods of image reconstruction, leaving just a general linear equation connecting the measured intensity values a_m and the reconstructed object coefficients c_m in the same vector basis:

$$c_m = \sum_{m'=1}^{M} r_{m\,m'} a_{m'}, \tag{33}$$

where $r_{mm'} \equiv q^{(2)}_{m\,m'}$ according to Eq. (4). The corresponding continuous version of Eq. (33) is given by a linear integral equation with the Green's function $R_M(\mathbf{r},\mathbf{r}') = \sum_{m=1}^{M} U_m(\mathbf{r}) U_m(\mathbf{r}')$. Depending on the qualitative manner in which the SNR and the spatial resolution are affected by Eq. (33), one can naturally divide such systems into three distinct classes:

(I) *Rotation and multiplicationlike class*. In this case the matrix $r_{mm'}$ is a product of an orthogonal matrix and a constant number, and the corresponding integral transform is proportional to a unitary operator. The simplest example can be given in the case $M = 2$ by the transformation $c_1 = a_1 + a_2$, $c_2 = a_2 - a_1$, which can be represented as a rotation by an angle equal to $\pi/4$ radians, followed by multiplication by the factor $\sqrt{2}$. In this case, the image reconstruction operation does not change the average SNR achieved in the measured signal (a decrease of SNR for some coefficients c_m, compared to the coefficients a_m, is exactly matched by the corresponding increase of other coefficients) and also does not change the spatial resolution (it preserves the width of the spatial Fourier spectrum).

(II) *Convolutionlike class*. In this case the matrix elements $r_{mm'}$ are typically all non-negative. The corresponding integral transform acts in a way similar to low-pass filtering. The simplest example can be given in the case $M = 2$ by the transformation $c_1 = a_1$, $c_2 = (a_1 + a_2)/2$. In this case, the image reconstruction operation increases the average SNR (reduces noise by means of spatial correlations), but worsens spatial resolution (shrinks the spatial Fourier spectrum), in the transition from the measurement space (corresponding to coefficients a_m) to the object space (corresponding to coefficients c_m).

(III) *Deconvolutionlike class*. In this case the matrix elements $r_{mm'}$ typically have alternating signs, and the corresponding integral transform acts somewhat similarly to high-pass filtering. The simplest example can be given in the case $M = 2$ by the transformation $c_1 = a_1$, $c_2 = 2a_2 - a_1$, which is inverse to the transformation given as an example for class II above. In the present case, the image reconstruction operation reduces the average SNR (increases noise by means of spatial decorrelations), but improves the spatial resolution (broadens the spatial Fourier spectrum).

The imaging system with single-pixel masks considered in Sec. V belongs to class I. In such systems, the spatial resolution and the average SNR do not change in the transition from the measurement space to the object space. Examples of systems from class II are represented by any low-pass filtering operation. Such systems improve the SNR (by means of spatial correlations), but lower the spatial resolution (blur the image). The imaging systems with overlapping two-pixel masks considered at the end of Sec. V A, the harmonic masks considered in Sec. V B, and the pseudorandom masks considered in Sec. V C all belong to class III. The average SNR in this case decreases in the transition from the measured bucket coefficients a_m to the object coefficients c_m. The spatial resolution does improve in the transition from the image (measurement) space to the object space. The latter fact was recently demonstrated in an experiment [26].

Interestingly, the situation with the imaging systems using nonorthogonal illumination patterns, as considered above, is exactly opposite to the situation with reconstructive imaging using the homogeneous transport-of-intensity equation (TIE-Hom) [38], as considered in our recent publication Ref. [39]. It has been shown previously that the TIE-Hom imaging is capable of increasing SNR by factors of up to two orders of magnitude, without sacrificing spatial resolution, in certain common imaging contexts [40–44]. We showed in Ref. [39] that the IQC can increase in the process of free-space propagation of a transmitted wave from the object space to the measurement space (i.e., the image or detector space), because here the spatial resolution can improve without an increase of the noise. The key to that fact is the behavior of photon noise, which, in the case of thermal light sources, is dominated by the photodetection shot noise. Given that the total number of photons is preserved in the process of free-space propagation, the average photon shot noise is the

same in the object and the image spaces (possibly, after the geometric magnification is taken into account). Thus the noise stays the same, while the spatial resolution can improve (the spatial Fourier spectrum can become broader) in the process of free-space propagation in the near-Fresnel region [39], giving one a net increase in the ratio of SNR to the spatial resolution. In the case of imaging with nonorthogonal illumination patterns, the situation is the opposite. As we have demonstrated above, in that case the ratio of the SNR to the spatial resolution (and the corresponding IQC) in the object space is lower in the case of nonorthogonal illumination patterns compared to the equivalent orthogonal case. This happens because the spatial resolution in the object space is the same for orthogonal and nonorthogonal (linearly independent) bases, as the spatial resolution in the reconstructed object is determined only by the dimensionality of the space spanned by the illumination vectors, as shown in Sec. III. At the same time, the SNR is lower in the case of nonorthogonal bases, compared to orthogonal bases with the same average transmission, as shown in Sec. IV, because in the nonorthogonal cases the SNR is drastically lowered (typically in proportion to the number of illumination patterns) by the reconstruction operator which effectively spatially decorrelates the noise in the measured signal [because Eq. (33) performs a deconvolution in this case].

Note that unlike the situation with the imaging setups considered in Ref. [20], in the setup studied in the present paper it was not possible to show that the IQC is the same in the measurement and object spaces, i.e., that it does not change in the process of object reconstruction. This is probably due to the fact that although the reconstruction operator $\mathbf{R}_M$ here is linear, it is not necessarily shift invariant when the illumination patterns are not orthogonal. The behavior of the spatial resolution and SNR in the case of measurements in nonorthogonal bases, and its consequences for the efficiency of the corresponding imaging systems, may be worth investigating further in the future.

ACKNOWLEDGMENT

The authors are grateful to D. Ceddia and F. de Hoog for useful discussions.

APPENDIX A: LOCALIZED ILLUMINATION MASKS

Let the points $\mathbf{r}_m = (m_x h, m_y h)$ denote the positions of the centers of the detector pixels with area h^2 each, indexed by integers $m = (m_x - 1)L + m_y$, $m_x, m_y = 1,2,...,L$, $m = 1,2,...,M$. The function $T(\mathbf{r})$ is the indicator function of a pixel, i.e., $T(x,y) = 1$, when $|x| < h/2$ and $|y| < h/2$, and $T(x,y) = 0$ is zero otherwise. Each shifted function $V_m(\mathbf{r}) = LT(\mathbf{r} - \mathbf{r}_m)$ is then equal to L inside the pixel with index m, represented by the domain $\Omega_m = \{|x - m_x h| < h/2, |y - m_y h| < h/2\}$, and is equal to zero elsewhere. The whole image domain is represented by the square

$$\Omega = \bigcup_{m=1}^{M} \Omega_m = \{h/2 < x < Lh + h/2, h/2 < y < Lh + h/2\},$$

consisting of $M = L^2$ pixels Ω_m, with the total area $|\Omega| = M|\Omega_m| = L^2 h^2$. The orthonormality relationship, $\langle V_m, V_{m'} \rangle = \delta_{mm'}$ for any $m, m' = 1,2,\ldots,M$, with the scalar product defined by Eq. (1), follows from the fact that the spatial supports of functions $V_m(\mathbf{r})$ with different indexes do not overlap, while for $m = m'$ we get $|\Omega|^{-1} \iint_\Omega V_m^2(\mathbf{r}) d\mathbf{r} = |\Omega|^{-1} L^2 |\Omega_m| = 1$.

The projector onto the space spanned by vectors $\boldsymbol{V}_m$, $m = 1,2,\ldots,M$, can be defined as in Eq. (5). The corresponding Green's function is not shift invariant in general (although it is invariant with respect to shifts by a whole number of pixels). The spatial resolution can be evaluated in accordance with Eq. (10):

$$(\Delta_2 r)^2[g_{\mathbf{r}}] = |\Omega| / \sum_{m=1}^{M} V_m^2(\mathbf{r}) = |\Omega|/L^2 = |\Omega|/M = h^2, \tag{A1}$$

where we have taken into account the fact that any point $\mathbf{r}$ in Ω lies within a single pixel, and hence there is one and only one basis function which is equal to L at this point (here we ignore the points $\mathbf{r}$ that lie on the boundaries of pixels, because such points constitute a set of measure zero). We see that the spatial resolution here is the same at any point $\mathbf{r}$. The result agrees with one's natural expectation that the spatial resolution of the direct imaging setup using a detector with the pixel size h should be equal to that pixel size.

The SNR can be calculated according to Eqs. (15) and (16). Because in this case the original illumination patterns $\boldsymbol{W}_m = \bar{n}\boldsymbol{T}_m = (\bar{n}/L)\boldsymbol{V}_m$ are orthogonal, the form factor $F_{M,X}(\mathbf{r})$ can be expressed by Eq. (17) with $x_m = \frac{1}{|\Omega|} \iint_\Omega X(\mathbf{r}) T(\mathbf{r} - \mathbf{r}_m) d\mathbf{r}$. Note also that because the supports of indicator functions of different pixels do not overlap, we have $T_m(\mathbf{r}) T_{m'}(\mathbf{r}) = \delta_{mm'} T_m^2(\mathbf{r})$ for any point $\mathbf{r}$ in Ω. Using this property, it is easy to see that

$$\left[\sum_{m=1}^{M} x_m T_m(\mathbf{r}) \right]^2 = \sum_{m=1}^{M} x_m^2 T_m^2(\mathbf{r}) = x_{m(\mathbf{r})}^2 \quad \text{and}$$

$$\sum_{m=1}^{M} x_m T_m^2(\mathbf{r}) = x_{m(\mathbf{r})},$$

where $m(\mathbf{r})$ is the index of the pixel containing the point $\mathbf{r}$ and

$$x_{m(\mathbf{r})} = \frac{1}{|\Omega|} \iint_\Omega X(\mathbf{r}') T_{m(\mathbf{r})}(\mathbf{r}') d\mathbf{r}'.$$

Substituting these expressions into Eq. (17), we obtain that $F_{M,X}^2(\mathbf{r}) = x_{m(\mathbf{r})}$, and therefore

$$\mathrm{SNR}^2[\mathbf{P}_M X](\mathbf{r}) = (\bar{N}/M) x_{m(\mathbf{r})}, \tag{A2}$$

which corresponds to the average number of (transmitted) photons registered in the pixel containing the point $\mathbf{r}$. Note that $x_{m(\mathbf{r})} = M^{-1} \langle X(\mathbf{r}) \rangle_m$, where

$$\langle X(\mathbf{r}) \rangle_m = \frac{1}{|\Omega_m|} \iint_\Omega X(\mathbf{r}') T_{m(\mathbf{r})}(\mathbf{r}') d\mathbf{r}'$$

is the average value of the object transmission function over the pixel containing the point $\mathbf{r}$.

APPENDIX B: DELOCALIZED HARMONIC MASKS

The illumination vectors here are indexed by $m = (m_x - 1)L + m_y$, $m_x, m_y = 1,2,...,L$, and $m = 1,2,...,M$, and are equal to $\boldsymbol{W}_m = \bar{n}\boldsymbol{T}_m$, where $\bar{n} = \eta\, I_{\rm in}\, |\Omega|$ is the average number of photons used in the measurement of each bucket coefficient. The orthonormalized basis can be chosen as follows: $\boldsymbol{V}_1 = (4/3)\bar{n}^{-1}\boldsymbol{W}_1 = (4/3)\boldsymbol{T}_1 = 1$ and $\boldsymbol{V}_m = 4\bar{n}^{-1}[\boldsymbol{W}_m - (2/3)\boldsymbol{W}_1] = 4[\boldsymbol{T}_m - (2/3)\boldsymbol{T}_1] = 4\boldsymbol{F}_m$. It is straightforward to verify that $\langle \boldsymbol{V}_m, \boldsymbol{V}_{m'}\rangle = \delta_{mm'}$, $m,m' = 1,2,\ldots,M$.

Note that $\sum_{m=1}^{M} V_m^2(\mathbf{r}) = M$, provided that $M = 1 + 4M'$ for some positive integer M', which is easy to show by grouping together terms with the same m_x and m_y. Restricting the values of M to integers of the form $1 + 4M'$, we obtain that the spatial resolution evaluated in accordance with Eq. (10) will be equal to

$$(\Delta_2 r)^2[g_{\mathbf{r}}] = |\Omega|/\sum_{m=1}^{M} V_m^2(\mathbf{r}) = |\Omega|/M. \tag{B1}$$

This shows, as one would naturally expect, that the spatial resolution here is equal to the image area divided by the number of Fourier harmonics used in the imaging system.

Using the simple relation $\langle \boldsymbol{F}_m, \boldsymbol{F}_{m'}\rangle = (1/16)\delta_{mm'}$, it is straightforward to verify by direct calculations that in this case the biorthogonal basis U_m, such that $\langle \boldsymbol{U}_m, \boldsymbol{W}_{m'}\rangle = \langle \boldsymbol{U}_m, \bar{n}(1/2 + \boldsymbol{F}_{m'})\rangle = \delta_{mm'}$, consists of functions $U_1(\mathbf{r}) = \bar{n}^{-1}(4/3)[1 - 8\sum_{m=2}^{M} F_m(\mathbf{r})]$, $U_m(\mathbf{r}) = \bar{n}^{-1}16F_m(\mathbf{r})$, $m = 2,3,...,M$. Taking into account that $S_m(\mathbf{r}) = \bar{n}\, U_m(\mathbf{r})$ and substituting the obtained expressions for $S_m(\mathbf{r})$ into Eqs. (16) and (21), we obtain Eq. (28).

APPENDIX C: PSEUDORANDOM DELOCALIZED MASKS

In view of Eq. (30), the orthonormalized vectors $\boldsymbol{V}_m$ can be naturally chosen here as

$$\begin{aligned} V_1(\mathbf{r}) &= \frac{I_1(\mathbf{r})}{I_1} = 1, \\ V_m(\mathbf{r}) &= \frac{I_m(\mathbf{r}) - I_1}{\sigma}, \\ m &= 2,3,\ldots,M. \end{aligned} \tag{C1}$$

If the set of illumination patterns is shift-invariant, i.e., it satisfies condition 2, then the spatial resolution of this imaging system is given by Eq. (12). In order to estimate the SNR using Eqs. (15)–(17), we need to find the biorthogonal basis $\boldsymbol{U}_m$ or the coefficients q_{ml}. Introducing the vectors $\boldsymbol{F}_m \equiv \boldsymbol{I}_m/I_1 - 1 = T_m/t_1 - 1$, $m = 2,3,\ldots,M$, $t_1 = \langle T_1\rangle = T_1$, we obtain from Eq. (30) that $\langle \boldsymbol{F}_m, \boldsymbol{F}_{m'}\rangle = \kappa^{-2}\delta_{mm'}$, where $\kappa = I_1/\sigma \geqslant 1$. It is then straightforward to verify by direct calculations that in this case the biorthogonal basis $\boldsymbol{U}_m$, such that $\langle \boldsymbol{U}_m, \boldsymbol{W}_{m'}\rangle = \langle \boldsymbol{U}_m, \bar{n}t_1(1 + \boldsymbol{F}_{m'})\rangle = \delta_{mm'}$, consists of functions $U_1(\mathbf{r}) = \bar{n}^{-1}t_1^{-1}[1 - \kappa^2\sum_{m=2}^{M} F_m(\mathbf{r})]$, $U_m(\mathbf{r}) = \bar{n}^{-1}t_1^{-1}\kappa^2 F_m(\mathbf{r})$, $m = 2,3,...,M$. Substituting $S_m(\mathbf{r}) = \bar{n}\, U_m(\mathbf{r})$ into Eqs. (16) and (21), we obtain Eq. (31).

[1] B. Z. Vulikh, *Introduction to Functional Analysis for Scientists and Technologists* (Pergamon, Oxford, 1963).

[2] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley and Sons, New York, 1999).

[3] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley and Sons, New York, 1991).

[4] J. Romberg, Imaging via compressive sampling, IEEE Signal Process. Mag. **25**, 14 (2008).

[5] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, 1968).

[6] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, New York, 2006).

[7] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (John Wiley and Sons, New York, 1991).

[8] J. T. Winthrop and C. R. Worthington, Convolution formulation of Fresnel diffraction, J. Opt. Soc. Am. **56**, 588 (1966).

[9] J. H. Shapiro, Computational ghost imaging, Phys. Rev. A **78**, 061802(R) (2008).

[10] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, 3D computational imaging with single-pixel detectors, Science **340**, 844 (2013).

[11] M. J. Padgett and R. W. Boyd, An introduction to ghost imaging: quantum and classical, Philos. Trans. R. Soc. London, Ser. A **375**, 20160233 (2017).

[12] B. I. Erkmen and J. H. Shapiro, Ghost imaging: from quantum to classical to computational, Adv. Opt. Photon. **2**, 405 (2010).

[13] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. (Academic, San Diego, 1999).

[14] F. Natterer, *The Mathematics of Computed Tomography* (SIAM, Philadelphia, 2001).

[15] O. Katz, Y. Bromberg, and Y. Silberberg, Compressive ghost imaging, Appl. Phys. Lett. **95**, 131110 (2009).

[16] P. Martinsson, H. Lajunen, and A. T. Friberg, Communication modes with partially coherent fields, J. Opt. Soc. Am. A **24**, 3336 (2007).

[17] D. Gabor, A new microscopic principle, Nature (London) **161**, 777 (1948).

[18] M. J. Cieślak, K. A. A. Gamage, and R. Glover, Coded-aperture imaging systems: Past, present and future development—A review, Radiat. Meas. **92**, 59 (2016)

[19] T. E. Gureyev, Ya. I. Nesterets, F. de Hoog, G. Schmalz, S. C. Mayo, S. Mohammadi, and G. Tromba, Duality between noise and spatial resolution in linear systems, Opt. Express **22**, 9087 (2014).

[20] 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).

[21] E. J. Candés and M. B. Wakin, An introduction to compressive sampling, IEEE Signal Process. Mag. **25**, 21 (2008).

[22] H. Anton and C. Rorres, *Elementary Linear Algebra with Applications* (John Wiley and Sons, New York, 1987).

[23] M. S.Birman and M. Z. Solomjak, *Spectral Theory of Self-Adjoint Operators in Hilbert Space* (D. Reidel, Dordrecht, 1987).

[24] A. Messiah, *Quantum Mechanics* (North-Holland, Amsterdam, 1961), Vol. 1.

[25] F. Ferri, D. Magatti, L. A. Lugiato, and A. Gatti, Differential Ghost Imaging, Phys. Rev. Lett. **104**, 253603 (2010).

[26] D. Pelliccia, M. P. Olbinado, A. Rack, and D. M. Paganin, Practical x-ray ghost imaging with synchrotron light, arXiv:1710.00727.

[27] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England,1995).

[28] T. E. Gureyev, A. Kozlov, Y. I. Nesterets, D. M. Paganin, and H. M. Quiney, On noise-resolution uncertainty in quantum field theory, Sci. Rep. **7**, 4542 (2017).

[29] A. Gatti, M. Bache, D. Magatti, E. Brambilla, F. Ferri, and L. A. Lugiato, Coherent imaging with pseudo-thermal incoherent light, J. Mod. Opt. **53**, 739 (2006).

[30] J.-E. Oh, Y.-W. Cho, G. Scarcelli, and Y.-H. Kim, Sub-Rayleigh imaging via speckle illumination, Opt. Lett. **38**, 682 (2013).

[31] 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).

[32] X.-H. Chen, F.-H. Kong, Q. Fu, S.-Y. Meng, and L.-A. Wu, Sub-Rayleigh resolution ghost imaging by spatial low-pass filtering, Opt. Lett. **42**, 5290 (2017).

[33] Y. Bromberg, O. Katz, and Y. Silberberg, Ghost imaging with a single detector, Phys. Rev. A **79**, 053840 (2009).

[34] P.-A. Moreau, E. Toninelli, T. Gregory and M. J. Padgett, Ghost imaging using optical correlations, Laser Photon. Rev. **12**, 1700143 (2017).

[35] P. Clemente, V. Duran, E. Tajahuerce, V. Torres-Company, and J. Lancis, Single-pixel digital ghost holography, Phys. Rev. A **86**, 041803(R) (2012).

[36] M. R. Teague, Deterministic phase retrieval: A Green's function solution, J. Opt. Soc. Am. **73**, 1434 (1983).

[37] D. Paganin, A. Barty, P. J. McMahon, and K. A. Nugent, Quantitative phase-amplitude microscopy III: The effects of noise, J. Microsc. **214**, 51 (2004).

[38] 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).

[39] T. E. Gureyev, Ya. 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 **28**, 2251 (2017).

[40] 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).

[41] 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).

[42] Ya. I. Nesterets and T. E. Gureyev, Noise propagation in x-ray phase-contrast imaging and computed tomography, J. Phys. D: Appl. Phys. **47**, 105402 (2014).

[43] T. E. Gureyev, S. C. Mayo, Ya. 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 imaging quality of synchrotron-based phase-contrast mammographic tomography, J. Phys. D: Appl. Phys. **47**, 365401 (2014).

[44] 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).