Information-theoretical resolution limit of a far-field subwavelength diffraction system

Yan-ming Gao [1], Chuan-jin Zu,[1] Xiang-sheng Xie,[2] and Xiang-yang Yu[1,*]

[1]*School of Physics, State Key Laboratory of Optoelectronic Materials and Technologies, Sun Yat-Sen University, Guangzhou 510275, China*
[2]*Department of Physics, College of Science, Shantou University, Shantou, Guangdong 515063, China*

A general analytical model is established to study the change in the signal-to-noise ratio (SNR) of the spatially noisy optical signal after passing through a subwavelength-aperture system. The corresponding spatial resolution limit is redefined based on Shannon's theory of information, and the expression is analytically given derived from nonparaxial vectorial diffraction theory. Contrary to the conventional wisdom, it is demonstrated both analytically and numerically that the SNR of the optical signal will increase with the propagation distance due to the presence of the subwavelength aperture and the corresponding spatial resolution could, thus, exceed the traditional Abbe-Rayleigh diffraction limit even though when the input SNR is low. The theoretical result provides perspectives for optimizing optical resolution and is general, that can be extended to other sciences, such as x-ray imaging and quantum imaging.

I. INTRODUCTION

Classical resolution criteria, such as, most famously, the Rayleigh's are based on the so-called "calculated image" (see Refs. [1,2]), which does not represent a fundamental limit to optical resolution. A more practical and fundamental limit below the quantum limit [3–6] considers the impact of noise on the detected results, and modern resolution theories use statistical and information-theoretical models to redefine optical resolution [2]. Whereas studies associating the optical resolving problem with statistics and information theories have been showing interesting results until very recently [7–15], most of them are initially meant for general large-aperture systems, and the sizes of the apertures are seldom specified. In the theories of light propagation, diffraction results change qualitatively when the scale of the diffraction structure reaches subwavelength level [16–22], and abnormal optical phenomena, including those leading to subdiffraction resolution [16,18], therefore, occur. Although subwavelength apertures are now implemented in actual experimental setups, such as near-field scanning microscopy [23], confocal microscopy [24], and superoscillation microscopy [25,26] to enhance optical performance, the true resolution limit of these systems due to the presence of the subwavelength aperture is still quite unclear, and theoretical study devoted to such an issue is rarely seen.

Despite using different theoretical models to understand optical resolution [2], one could expect that the resolution limit is ultimately governed by the fundamental properties of light propagation and information [7–15]. One of the interpretations of the optical resolution limit is since the amount of information in Shannon's theory of information could be understood as the number of the symbols used in a message [27], then the highest possible resolution is defined as the ratio of the geometric scale of the object L to the mutual information T of the received image under noisy conditions, which directly links to the signal-to-noise ratio (SNR) of the object. Such a definition of optical resolution is recently analytically proposed by Narimanov in Ref. [15], which reveals the essential information-theoretical nature of the optical resolving problem, and offers many unique advantages for analyzing optical resolutions, especially for objects without much *a priori* knowledge.

In this article, we establish a general analytical model based on the strict vectorial diffraction theory [17,19–22] to study the change in SNR of the spatially noisy optical signal after passing through a subwavelength aperture and follow Narimanov's proposal to study the corresponding resolution limit of the general far-field subwavelength diffraction system. Based on the light propagation properties of the large-aperture system, Narimanov points out that the resolution could exceed the traditional Abbe-Rayleigh diffraction limit when the effective input SNR is large, however, the weak logarithmic dependence of the resolution limit on the SNR prevents the system from achieving deep-subwavelength resolution under normal noise conditions. Although such an analytical result is generally in line with the previous works [7–14], we show that it does not describe the resolution limit of a subwavelength-aperture system as the light propagation properties are essentially different. We show that the SNR of the optical signal will increase due to the local spatial coherence imposed by the subwavelength aperture, and the corresponding resolution limit could, thus, exceed the traditional Abbe-Rayleigh limit even when the input SNR is low. We demonstrate the superresolution capability and its physical mechanism and show that by simply using the far-field diffraction result of a scanning subwavelength aperture to reconstruct the amplitude profile of the object, the resolution is multiple times better than the traditional limit and even those defined by the SNR. Our results show avenues for optical far-field

―――――
*cesyxy@mail.sysu.edu.cn
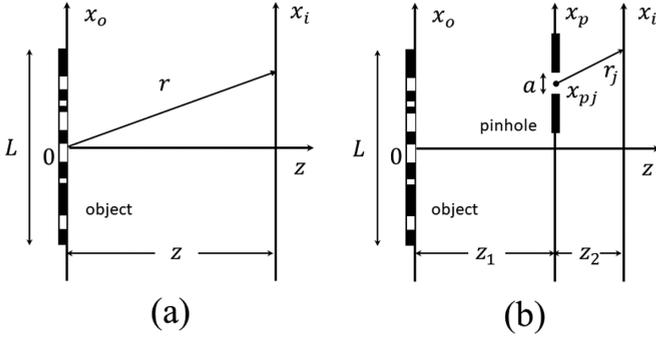
FIG. 1. Configuration for theoretical analysis. (a) General model for the large-aperture imaging system. (b) General model for the large-aperture system with a scanning pinhole.

nonfluorescent superresolution, which can be also extended to other sciences, such as x-ray imaging [28] and quantum imaging that may benefit from quantum metrology [3–6] and nonclassical correlations [29–31].

This article is organized as follows. We start with Narimanov's theory in Sec. II, and give the resolution limit of the general subwavelength-aperture system in Sec. III. We demonstrate the superresolution capability and the physical mechanism behind it in Sec. IV. Conclusions are given in Sec. V. Complete derivations for the light propagation in the subwavelength system are given in Appendix A. We supply more numerical results in Appendix B, and perform a more rigorous approximation in Appendix C.

II. INFORMATION-THEORETICAL LIMIT OF OPTICAL RESOLUTION

We follow the lead of Narimanov [15] and start with a general model to illustrate on the core physics of the problem.

Consider a general imaging model as shown by Fig. 1(a), a one-dimensional object is placed in the object plane under the illumination of a monochrome coherent field with a wavenumber $k = 2\pi/\lambda$. The interaction between the illumination field and the object needs to be considered in the framework of scattering theory [15,19], yet for general consideration, one could regard the light field right behind the object $E(x_o)$ as the input "object function" and the imaging problem as the reconstruction of $E(x_o)$ from its detection results $s(x_i) \equiv E(x_i) + n(x_i)$, where $E(x_i)$ is the calculated image of $E(x_o)$ and $n(x_i)$ is the additional noise. Regardless of any specific noise distribution, we assume $n(x_i)$ to be a spatial stationary random process with a zero mean value and a nonzero mean-square module.

In general, for a linear system (say injective mapping) with a clear analytical relation $E(x_i) = F\{E(x_o)\}$, one can always reconstruct the input object $E(x_o)$ from its ideal image $E(x_i)$ with infinite accuracy mathematically [1], and it is the noise that practically limits the "spatial accuracy" of the reconstruction, which is, in general, the resolution. To analyze the effect of noise on the resolution, mutual information T from Shannon's theory [27] is introduced to define the optical resolution, which is given in a standard way,

$$T = H[\{s\}] - H[\{s\}|E(x_o)], \quad (1)$$

where $H[\{s\}]$ is the information entropy that offers a measure of the expected information received by the detector and $H[\{s\}|E(x_o)]$ is the conditional entropy for a given object function $E(x_o)$ to measure the loss of validity due to the additional noise. By its general logarithmic definition [27], the amount of information can be understood as the number of the symbols to carry information. Consider using such symbols to reconstruct the object, then the resolution limit of an object with a length L under noisy conditions is given by

$$\Delta = \frac{L}{T}, \quad (2)$$

which is the shortest spatial scale of the object whose geometry can be possibly reconstructed, regardless of using any information processing methods.

To calculate the mutual information T, the signal must be analytically expressed based on the imaging model. For a large-aperture system ($L \gg \lambda$), the ideal far-field ($r \gg L$) detected result is naturally the Fourier spectrum of the object [19],

$$s(\boldsymbol{k}) = E(\boldsymbol{k}) + n(\boldsymbol{k}) = \int E(x_o)e^{ik_x x_o}dx_o + n(\boldsymbol{k}), \quad (3)$$

where $\boldsymbol{k} \equiv (k_x, k_z)$ with the magnitude $|\boldsymbol{k}| = k$ is the wave vector, and we have

$$\frac{k}{r} = \frac{k_x}{x_i} = \frac{k_z}{z}. \quad (4)$$

The information entropy is then given by

$$H[\{s\}] = -\int Ds(\boldsymbol{k})\{P[s(\boldsymbol{k})]\log_M P[s(\boldsymbol{k})]\}, \quad (5)$$

where $P[s(\boldsymbol{k})]$ is the distribution function of the detected signal $s(\boldsymbol{k})$ and the functional integral $\int Ds(\boldsymbol{k})$ is defined in the standard way [15]. The logarithmic base M is the number of the possible states of the symbols to reconstruct the object. Similarly, we have

$$H[\{s\}|E(x_o)] = -\int Ds\{P[s|E(x_o)]\log_M P[s|E(x_o)]\}, \quad (6)$$

where $P[s|E(x_o)]$ is the conditional distribution function of the detected signal for the given object $E(x_o)$. For additive noise, $P[s|E(x_o)]$ is simply equal to the noise distribution $P_n[n(\boldsymbol{k})]$. With further consideration of the central limit theorem on the expected signal and the uncorrelated Gaussian noise, both distribution functions $P[s(\boldsymbol{k})]$ and $P_n[n(\boldsymbol{k})]$ can be analytically expressed (for a detailed derivation, please refer to Sec. 9 of Ref. [15]), and finally we have

$$\Delta_M = \frac{\lambda}{2} \frac{1}{\log_M \sqrt{1 + 2\,\text{SNR} + \eta\,\text{SNR}^2}}, \quad (7)$$

which is the resolution limit of the object with M different levels of amplitude, and SNR is the effective ratio,

$$\text{SNR} = \frac{\langle |s(\boldsymbol{k}) - \langle s(\boldsymbol{k})\rangle|^2\rangle}{\langle |n(\boldsymbol{k})|^2\rangle}, \quad (8)$$

and

$$\eta = \frac{\langle[\text{Re}(E(x_o) - \langle E(x_o)\rangle)]^2\rangle\langle[\text{Im}(E(x_o) - \langle E(x_o)\rangle)]^2\rangle}{\langle|E(x_o) - \langle E(x_o)\rangle|^2\rangle}$$
$$- \frac{\langle\text{Re}(E(x_o) - \langle E(x_o)\rangle)\text{Im}(E(x_o) - \langle E(x_o)\rangle)\rangle^2}{\langle|E(x_o) - \langle E(x_o)\rangle|^2\rangle} \quad (9)$$

represents the relative contribution of the absorption of the object; for a transparent object, we have $\eta = 0$.

The analytical result derived based on the large-aperture approximation in Ref. [15] indicates the possibility of a higher resolution than the traditional limit $\lambda/2$, yet, the required SNR grows nearly exponential with the resolution, which prevents the system from achieving deep-subwavelength resolution under the normal noisy condition. Although such a result is generally in line with the prior works, next, we show that it is not sufficient for the case of using a subwavelength aperture in which the far-field diffraction result is qualitatively different from Eq. (3).

III. RESOLUTION LIMIT OF A SUBWAVELENGTH-APERTURE SYSTEM

The case of a subwavelength aperture is now implemented in the actual experimental setups [23–26], yet due to the complicated vectorial boundary conditions, the theoretical analysis for subwavelength diffraction is generally performed numerically, and the analytical study on the problem is rarely seen [17]. Introduction to the problem and detailed derivations for the following discussion are given in Appendix A. A prior study of the following work can be found in Ref. [16].

A simple reconsideration for $Lk < 1$ simply represents an example for a near-field probe [15,23], which is out of the topic. A more appropriate model is considered as shown by Fig. 1(b): A scannable subwavelength pinhole with a width $a \leqslant \lambda$ is added on a "pinhole plane" parallel to both the object and the image plane, and both z_1 and z_2 are much larger than λ. At each scanning position x_{pj}, the detector detects the secondary diffraction field $E_{2j}(x_i)$ to estimate the value of the primary diffraction field of the object at that point $E_1(x_p)$, forming the signal array s_j. The imaging problem is as before the reconstruction of the object function $E_1(x_o)$ from the detection array of s_j.

The imaging process before the pinhole plane is the same as that shown in Fig. 1(a), and the key difference is introduced by the secondary diffraction process of the subwavelength aperture, which changes the returned SNR_r of s_j. For the input signal on the pinhole plane $E_{1n}(x_p) \equiv E_1(x_p) + n_1(x_p)$ with an input SNR_0, the secondary diffraction field at the image plane is given by [see Eqs. (A8)–(A11) in Appendix A],

$$E_{2j}(x_i) = [G_v(r_j)a\langle E_j\rangle][1 + \epsilon_j(x_i)], \quad (10)$$

where $G_v(r_j)$ is the point-spread function (PSF) of the first Rayleigh-Sommerfeld integrals [32] and the subscript v stands for different vector components [see Eqs. (A9) and (A10) in Appendix A],

$$r_j = \sqrt{(x_{pj} - x_i)^2 + z_2^2} \quad (11)$$

is the distance between the jth central scanning position x_{pj} to one point on the image plane x_i; a is the width of the subwavelength pinhole, and

$$\langle E_j\rangle = \frac{1}{a}\int_{-(a/2)+x_{pj}}^{(a/2)+x_{pj}} E_1(x_p)dx_p \quad (12)$$

is the average amplitude of $E_1(x_p)$ at the aperture opening. The complex term ϵ_j in the parentheses is the relative error

between the actual image $E_{2j}(x_i)$ and the estimated result $[G_v(r_j)a\langle E_j\rangle]$, caused by both the noise and the finite size of the aperture. The maximum square modulus of this term is given by [see Inequality (A12) in Appendix A],

$$|\epsilon_j|^2 \leqslant \epsilon_n\epsilon_v, \quad (13)$$

where $\epsilon_n \approx 1/\text{SNR}_0$ is the error caused by the noise distorted boundary [see Eqs. (A13) and (A23) in Appendix A] and ϵ_v is the error caused by the diffraction of the subwavelength aperture itself, which will decrease with the increase in z_2 and eventually tends to 0 [see Eqs. (A14)–(A21) in Appendix A]. For a finite noise power, we finally have [see Eq. (A22) in Appendix A]

$$\lim_{z_2\to\infty}|\epsilon_j(x_j)|^2 = 0. \quad (14)$$

Although Eq. (3) indicates that the far-field diffraction result for the large-aperture system is naturally the Fourier spectrum of the object, the physical meaning of Eq. (10) here is that despite any noise-distorted boundary, the far-field diffraction results of a subwavelength aperture tends to a fixed point-spread-function $G_v(r_j)$ with which the source amplitude only equals the integral amplitude of $E_1(x_p)$ at the opening. Such a result is recently pointed out in a qualitative way in Ref. [16], and here we derived the analytical expression. The critical point follows is that the impact of noise on the intensity from the input of the system decreases with the increase in z_2, and the returned SNR_r, therefore, improved. Consider using the average intensity of $\langle|E_{2j}(x_i)|^2\rangle$ to return the values of the signal array $\langle E_j\rangle \to s_j$ (phase detection can be achieved by multiple optical methods, for example, a self-interference setup [33]) at each scanning position, the returned SNR_r of s_j is then given by

$$\text{SNR}_r = \frac{\langle|1 + \epsilon_j|^2\rangle}{\langle|\epsilon_j|^2\rangle}, \quad (15)$$

which will tend to infinity at the far field as indicated by Eq. (13),

$$\lim_{z_2\to\infty}\text{SNR}_r = \infty, \quad (16)$$

regardless of arbitrary raw input SNR_0.

If not considering the noise of the detector itself, the resolution limit of such a subwavelength diffraction system is given by

$$\Delta_{Mr} = \frac{\lambda}{2}\frac{1}{\log_M\sqrt{1 + 2\,\text{SNR}_r + \eta\,\text{SNR}_r^2}}, \quad (17)$$

which could tend to 0 when $z_2 \to \infty$, regardless of any raw noise input.

IV. DISCUSSIONS: ON THE IMPROVEMENT OF THE RESOLVING POWER

Although Narimanov's theory depending on Eq. (3) indicates that subwavelength details of the object are still hidden in the propagating waves and can be reconstructed at the far field, the result here shows that it can be achieved even for high noise input. Such a result is based on the far-field diffraction result of a subwavelength aperture, which shows
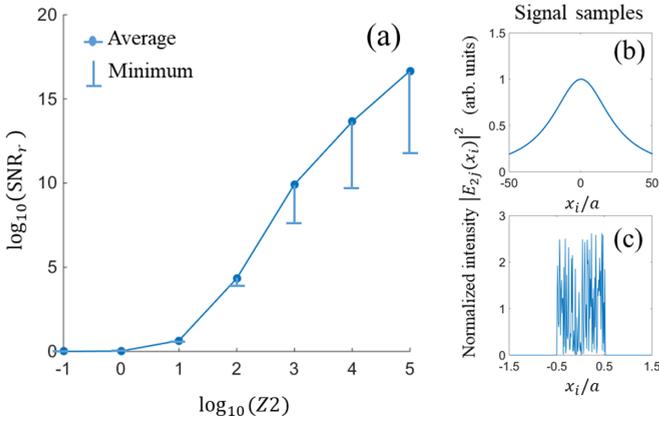
FIG. 2. (a) Statistical results of $\log_{10}(\mathrm{SNR}_r)$ near the optical axis versus a dimensionless parameter $\log_{10}(Z_2)$, where $Z_2 \equiv z_2(4\lambda/a^2)$, calculated by 10 000 realizations of $E_{1n}(x_p)$, $a = 0.1\lambda$, and $x_i \in [-2.5\lambda, 2.5\lambda]$. (b) A sample of the intensity profile of $|E_{2j}(x_i)|^2$ at $Z_2 = 10^3$. (c) A sample of the intensity profile of $|E_{2j}(x_i)|^2$ at $Z_2 = 0.1$.

the intrinsic noise suppression mechanism of subwavelength diffraction in the monochrome coherent field: As indicated by Inequality (13), no matter how much error is caused by the noise on the boundary ϵ_n, it will be offset by the diffraction effect of the aperture itself ϵ_v. Although light diffraction has always been regarded as a "low-pass filter" at which the subwavelength information is lost, the results here give a complete counterexample to such a theory and provide insights for further optimizing optical performance.

A question of interest is how "fast" does the returned SNR_r increase with the propagation distance z_2. For large z_2 and small $\langle |\epsilon_j|^2 \rangle$, the minimum amplification of $\mathrm{SNR}_r/\mathrm{SNR}_0$ near the paraxial region can be analytically given by [see Eqs. (13), (15), and (A20) in Appendix A]

$$\frac{1}{\epsilon_v} \approx \frac{1}{1 - \mathrm{sinc}\left(\frac{ax_i}{\lambda z_2}\right)}. \qquad (18)$$

In general, the smaller aperture a and the closer x_i to the x coordinate of the center of aperture x_p, the faster the magnification of $\mathrm{SNR}_r/\mathrm{SNR}_0$ with the increase in z_2. Note that it is not explicitly related to the specific distribution of the noise, indicating that the results are also valid for the resolution model depending on parameter estimation [3]. The complete numerical results that consider different distributions is given in Appendix B, Fig. 6, and here we give an example for a general demonstration.

Figure 2 shows the calculation results of the returned SNR_r as a function of a dimensionless parameter $Z_2 = z_2(4\lambda/a^2)$ calculated by 10 000 different realizations. As shown in Fig. 2(a), for an initial $\mathrm{SNR}_0 \approx 1$ input at the boundary, the statistical average of the returned SNR_r near the center grows near exponential when $Z_2 > 1$, corresponding to the far-field diffraction region to which our derivation applies. These results suggest that the strong exponential dependence of the resolution increase in the SNR can be effectively offset by the increase in z_2, making it possible to achieve deep-subwavelength resolution within a practical diffraction region. As can be seen the returned SNR_r soon reaches, at least, 10^7
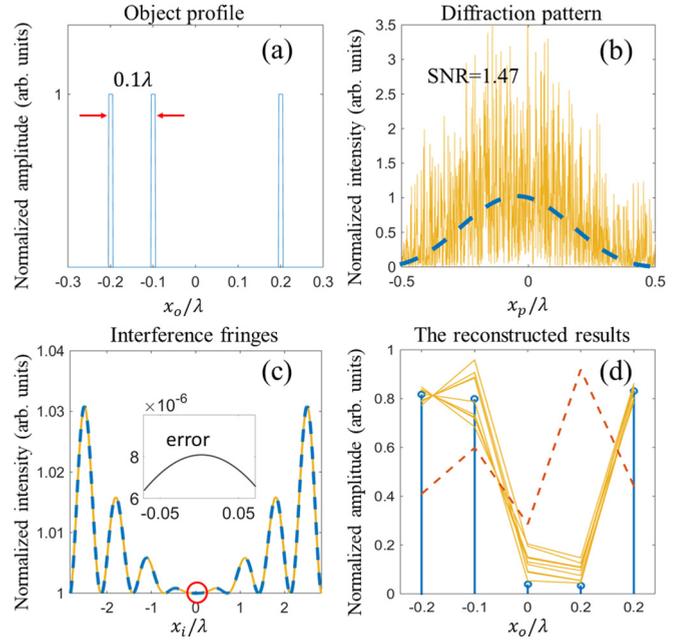
FIG. 3. (a) Amplitude profile of the object. (b) Ideal far-field coherent diffraction pattern of $I_1(x_p) \equiv |E_1(x_p)|^2$ (blue dashed line) and a sample of the noise distorted one $I_{1n}(x_p) \equiv |E_1(x_p) + n_1(x_p)|^2$ (yellow continuous line). (c) Self-interference fringes of the subwavelength pinhole of the ideal boundary $E_1(x_p)$ (blue dashed line) and the distorted one $E_{1n}(x_p)$ (yellow continuous line) at $z_2 = 20\lambda$, and the inset shows the error between the two results near the center. (d) Object profiles recovered from the average amplitude $\langle E_j \rangle$ of the ideal signal $E_1(x_p)$ (blue needles) and ten realizations of the returned signal s_j (yellow lines), and the red dashed line shows an example of the incorrect results reconstructed from the raw noise signal $I_{1n}(x_p)$.

at $Z_2 = 10^3$, corresponding to a resolution of, at least, $\lambda/20$ at $z_2 = 2.5\lambda$ when $M = 2$ for a transparent binary object. For a better visual understanding, two samples of the signals are shown in Figs. 2(b) and 2(c).

To further prove the principle and its practicability, Fig. 3 shows the results of a complete recovery process. Consider a binary transparent object as assumed by Narimanov [15] with an amplitude profile as shown in Fig. 3(a), and the objective resolution here is 0.1λ. For the far-field imaging system, we assume a standard sinc-type PSF with a half-width not exceeding half of the wavelength,

$$\mathrm{PSF}(x) = \mathrm{sinc}\left(\frac{2x}{\lambda}\right). \qquad (19)$$

The vectoral effect and the phase distribution are considered detailed in the Appendices and are not shown here. After passing through a far-field imaging system, the ideal coherent diffraction pattern $I_1 \equiv |E_1(x_p)|^2$ on the pinhole plane is given as shown in Fig. 3(b) (blue dashed line). As can be seen, the PSFs have overlapped with each other and cannot be resolved by any traditional methods. Consider adding a noise with zero-average amplitude and an initial $\mathrm{SNR}_0 \approx 1$ on the pinhole plane, and the distorted pattern $I_{1n} \equiv |E_1(x_p) + n_1(x_p)|^2$ is shown in Fig. 3(b) (yellow continuous line). Consider using the self-interference fringes (can be achieved by a Mach-Zehnder setup [33]) of a subwavelength pinhole with a

width $a = 0.02\lambda$ and a scanning interval $D = 0.1\lambda$ to return the complex amplitude $\langle E_j \rangle$ of each scanning position, and a sample of the interference fringes of one of the scanning positions is shown in Fig. 3(c). As can be seen the error ϵ_j caused by the noise has been reduced to the magnitude of 10^{-6}. According to Eq. (17), this error is small enough for our objective resolution. The error of each scanning position ϵ_j is then added on the return value of $\langle E_j \rangle$ to form the detection array s_j, and the object is then recovered by an inverse matric calculated by the PSF of Eq. (19) [34]. The recovered results are shown in Fig. 3(d), showing that the subwavelength details of the object are successfully recovered.

The recovery program requires no *a priori* constraint on the object, which indicates that the algorithm is far from optimal. By only considering a reverse matric and a general far-field PSF [34], these results clearly show from the fundamental level that the presence of a subwavelength pinhole suppresses the noise and returns results valid for a deep-subwavelength far-field resolution $\leqslant 0.1\lambda$ within a practical scale. If more *a priori* constraints or optimized algorithms are used as they are imposed in standard analog methods, such as that in Refs. [35,36], the requirements on the SNR of the resolution will be further relaxed, which means for an even higher resolution.

Although a subwavelength aperture has been used as a near-field probe in an actual experiment to achieve subwavelength imaging for a long time [23], the results here indicate that it is also possible to use it as a "far-field probe." The technique is general and does not explicitly require any sophisticated optical element, which has great potential to be extended beyond optical imaging to other fields, especially for those in which the conventional optical "lenses" are difficult to be applied, such as x-ray imaging [28].

V. CONCLUSIONS

We established an analytical model to study the change in the SNR of the optical signal after passing through a subwavelength aperture and give the corresponding resolution limit of a subwavelength coherent diffraction system based on Shannon's theory of information. The theoretical results reveal the intrinsic noise suppression mechanism of subwavelength diffraction, which provides perspectives for further optimizing optical resolution.

APPENDIX A: LIGHT PROPAGATION THROUGH A SUBWAVELENGTH APERTURE

To derive Eq. (10), standard assumptions for subwavelength diffraction are applied [17,19–22]. Consider an imaging model as shown in Fig. 4: The monochrome coherent field $\boldsymbol{E}_{1n} = \boldsymbol{E}_1 + \boldsymbol{n}_1$ generated by the object and the noise is incident on a one-dimensional aperture, which consists
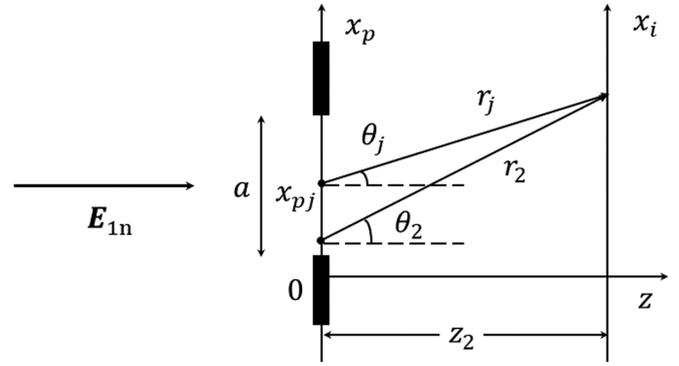
FIG. 4. Configuration for theoretical analysis.

of a perfectly conducting screen of infinite extent and zero thickness and an opening with a width of a centered at x_{pj}. After passing through the aperture, the complete diffraction field is written as $\boldsymbol{E}_{2j}$, and both $\boldsymbol{E}_{1n} = E_{1nx}\hat{x} + E_{1nz}\hat{z}$ and $\boldsymbol{E}_{2j} = E_{2jx}\hat{x} + E_{2jz}\hat{z}$ are assumed to be a polarized electric field. The magnetic field is neglected here, and it can be determined by the Maxwell equation $\boldsymbol{H} = -ic(\nabla \times \boldsymbol{E})/\omega$. The case of the polarized magnetic field and a more complete two-dimensional case can be worked along the same line [19]. For simplification, we set $x_{pj} = 0$, and $\boldsymbol{E}_{2j}$ at any point behind the screen is strictly given by the first Rayleigh-Sommerfeld integral [32],

$$E_{2jx}(x_i) = \frac{-1}{2\pi} \int_{\infty}^{\infty} E_{2jx}(x_p) \frac{\partial G(r_2)}{\partial z_2} dx_p, \quad \text{(A1)}$$

$$E_{2jz}(x_i) = \frac{1}{2\pi} \int_{\infty}^{\infty} E_{2jx}(x_p) \frac{\partial G(r_2)}{\partial x_i} dx_p, \quad \text{(A2)}$$

where $r_2 = \sqrt{(x_p - x_i)^2 + z_2^2}$ is the distance between one point on the pinhole plane and one point on the image plane and $G(r) = e^{ikr}/r$ is the Green's function. $E_{2jx}(x_p)$ is the tangential component of the electric field of $\boldsymbol{E}_{2j}$ at the boundary. For the far-field condition $z_2 \gg \lambda$, it is valid to replace it with $E_{1nx}(x_p)$ at the opening and write [17,37]

$$E_{2jx}(x_i) = \frac{-1}{2\pi} \int_{-a/2}^{a/2} E_{1nx}(x_p) \frac{\partial G(r_2)}{\partial z_2} dx_p, \quad \text{(A3)}$$

$$E_{2jz}(x_i) = \frac{1}{2\pi} \int_{-a/2}^{a/2} E_{1nx}(x_p) \frac{\partial G(r_2)}{\partial x_i} dx_p. \quad \text{(A4)}$$

For $z_2 \gg \lambda > a$, we extract the Green's-function terms in Eqs. (A3) and (A4) from the integrals and write

$$\lim_{z_2 \to \infty} E_{2jx}(x_i) = \frac{-1}{2\pi} \left[\frac{\partial G(r_j)}{\partial z_2} \right] a\langle E_j \rangle, \quad \text{(A5)}$$

$$\lim_{z_2 \to \infty} E_{2jz}(x_i) = \frac{-1}{2\pi} \left[\frac{\partial G(r_j)}{\partial x_i} \right] a\langle E_j \rangle, \quad \text{(A6)}$$

where $r_j = \sqrt{(x_{pj} - x_i)^2 + z_2^2}$ and

$$\langle E_j \rangle \equiv \frac{1}{a} \int_{-a/2}^{a/2} E_{1nx}(x_p) dx_p = \frac{1}{a} \int_{-a/2}^{a/2} E_1(x_p) dx_p \quad \text{(A7)}$$

is the average complex amplitude of E_{1nx} at the aperture opening, and we assume the noise does not affect the mean

of amplitude. Since for the polarized monochrome field the complete electromagnetic field can be determined by one component, we omit subscript x of E_{1nx} and write E_{1n} in the main text and below.

To prove these expressions, the Cauchy-Schwarz inequality is used. Use a uniform expression to rewrite Eqs. (A5) and (A6),

$$\lim_{z_2 \to \infty} \int_{-a/2}^{a/2} E_{1n}(x_p) G_v(r_2) dx_p$$
$$= G_v(r_j) \int_{-a/2}^{a/2} E_{1n}(x_p) dx_p, \quad (A8)$$

where G_v is the point-spread function for the first Rayleigh-Sommerfeld integral and the subscript v stands for x or z. For the x component of the electric field,

$$G_x(r) \equiv \frac{-1}{2\pi} \left[\frac{\partial G(r)}{\partial z} \right] = \frac{-1}{2\pi} \cos(\theta) \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr}, \quad (A9)$$

and for the z component,

$$G_z(r) \equiv \frac{1}{2\pi} \left[\frac{\partial G(r)}{\partial x} \right] = \frac{1}{2\pi} \sin(\theta) \left(\frac{ik}{r} - \frac{1}{r^2} \right) e^{ikr}. \quad (A10)$$

Since r_j is not related to x_p, the relative error between the left and the right sides of Eq. (A8) is as follows:

$$\epsilon_j = \frac{\int_{-a/2}^{a/2} E_{1nx}(x_p)[G_v(r_2) - G_v(r_j)] dx_p}{[a\langle E_j \rangle] G_v(r_j)}. \quad (A11)$$

Let $[G_v(r_2) - G_v(r_j)] = \Delta G_v$. According to the Cauchy-Schwarz inequality we write

$$|\epsilon_j|^2 \leqslant \epsilon_n \epsilon_v. \quad (A12)$$

where

$$\epsilon_n \equiv \frac{\int_{-a/2}^{a/2} |E_{1nx}(x_p)|^2 dx_p}{|a\langle E_j \rangle|^2} \quad (A13)$$

is the error on the intensity caused by the noise distorted boundaries, and

$$\epsilon_v \equiv \frac{\int_{-a/2}^{a/2} |\Delta G_v|^2 dx_p}{|G_v(r_j)|^2} \quad (A14)$$

is the error caused by the diffraction of a finite-size aperture itself. For a signal with finite power and $a\langle E_j \rangle \neq 0$, regardless of any specific distribution, ϵ_n is a finite positive real number, thus, according to the squeeze theorem we only need to prove that the second term brought by the diffraction is small enough,

$$\lim_{z_2 \to \infty} \epsilon_v = 0, \quad (A15)$$

that we prove Eq. (A8).

For the x component of the electric field, substituting Eq. (A9) into Eq. (A14) and expanding ΔG_x, we have

$$\epsilon_x = \int_{-a/2}^{a/2} \left| \frac{\cos(\theta_2) R(r_2) e^{ikr_2}}{\cos(\theta_j) R(r_j) e^{ikr_j}} - 1 \right|^2 dx_p, \quad (A16)$$
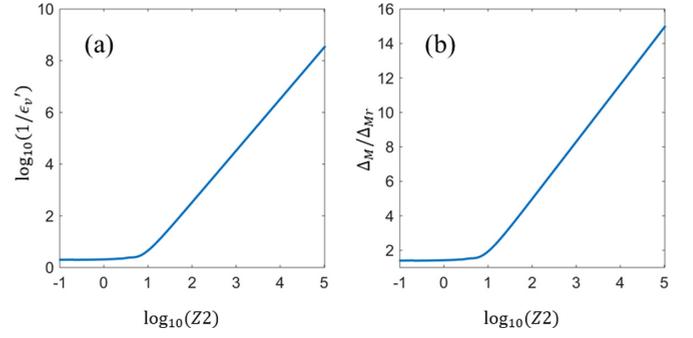
FIG. 5. (a) $\log_{10}(1/\epsilon_v')$ as a function of $\log_{10}(Z_2)$, calculated by Eq. (A24), where $Z_2 \equiv z_2(4\lambda/a^2)$, $P = 5a$, $x_{i0} = 0$, and $a = 0.1\lambda$. (b) Improvement of the resolution limit Δ_M/Δ_{Mr} as a function of $\log_{10}(Z_2)$, calculated by Eqs. (7) and (17), $SNR_0 = 1.5$, $M = 2$, $\eta = 0$, and $\Delta_M = \lambda/2$.

where

$$R(r) = \frac{-1}{2\pi} \left(\frac{ik}{r} - \frac{1}{r^2} \right). \quad (A17)$$

Consider that when $z_2 \gg a$ and $r_j \gg \lambda$, then $\cos(\theta_2) \approx \cos(\theta_j)$ and $R(r_2) \approx R(r_j)$, and we write

$$\epsilon_x = \int_{-a/2}^{a/2} |e^{ik(r_2 - r_j)} - 1|^2 dx_p. \quad (A18)$$

Let $\Delta r = r_2 - r_j$ since $\Delta r \ll r_j$, we have $d\Delta r = \sin(\theta_j) dx_p$, and we write

$$\epsilon_x = \frac{1}{\sin(\theta_j)} \int_{-a \sin(\theta_j)/2}^{a \sin(\theta_j)/2} |e^{ik\Delta r} - 1|^2 d\Delta r. \quad (A19)$$

After integration, when $x_i \ll z_2$ we have

$$\epsilon_x = 1 - \text{sinc}\left(\frac{a}{\lambda} \sin(\theta_j) \right) \approx 1 - \text{sinc}\left(\frac{ax_i}{\lambda z_2} \right). \quad (A20)$$

It is not difficult to write $\epsilon_x = \epsilon_z$. When $a < \lambda$, for any given x_i, we write

$$\lim_{z_2 \to \infty} \epsilon_v(x_i) = 0, \quad (A21)$$

and, finally,

$$\lim_{z_2 \to \infty} \epsilon_j(x_i) = 0. \quad (A22)$$

The physics here as indicated by Inequality (A12) is no matter how much error ϵ_n is caused by the noise on the intensity, it will be offset by the coherent diffraction effect of the subwavelength structure at the far field. If we consider

$$\frac{1}{\epsilon_n} \equiv \frac{|a\langle E_j \rangle|^2}{\int_{-a/2}^{a/2} |E_{1nx}(x_p)|^2 dx_p} \approx \text{SNR}_0 \quad (A23)$$

to be approximately the signal-to-noise ratio of the input signal, then the minimum magnification of $\text{SNR}_r/\text{SNR}_0$ at the far field is analytically given by $1/\epsilon_v$. Consider the smallest pixel size of the detector as P, then the minimum SNR magnification can be adjusted as

$$\frac{1}{\epsilon_v'} \equiv \frac{1}{\int_{x_{i0}-P/2}^{x_{i0}+P/2} \epsilon_v dx_i}, \quad (A24)$$
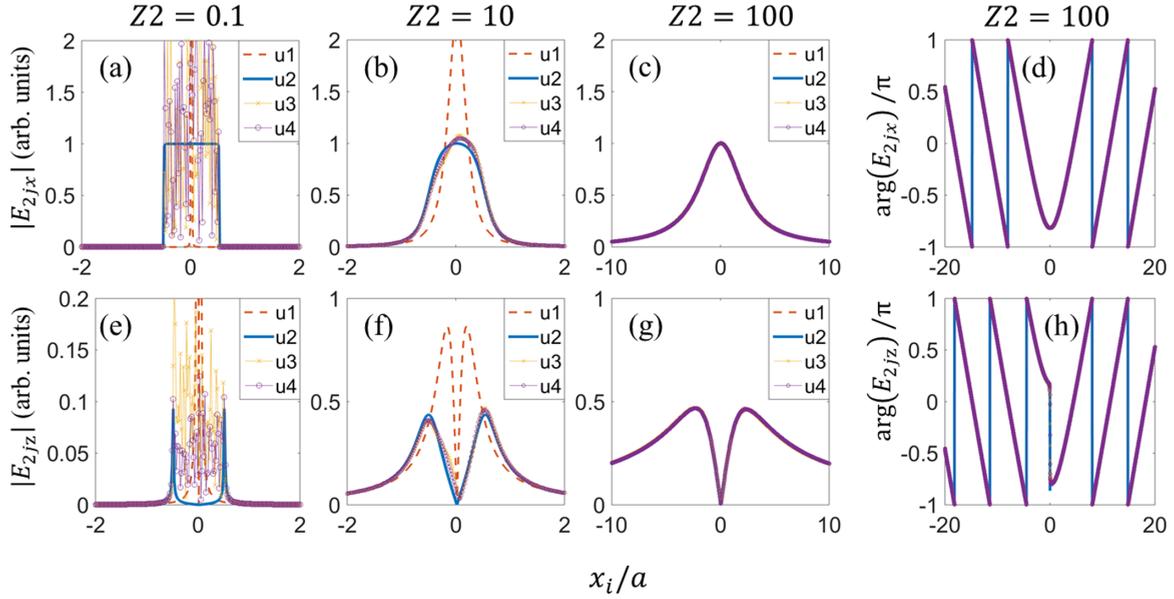
FIG. 6. Vectoral diffraction results $E_{2j}(x_i)$ of different boundary distributions $E_{1n}(x_p)$ at different Z_2's, calculated by the first Rayleigh-Sommerfeld integrals [Eqs. (18) and (19)], $a = 0.1\lambda$. (a)–(c) Normalized amplitude of the x-component $|E_{2jx}(x_i)|$ at different Z_2's, $Z_2 = 0.1, 10, 100$, respectively. (d) Normalized phase distribution of the x-component $\arg[E_{2jx}(x_i)]/\pi$ at $Z_2 = 100$. (e)–(g) Normalized amplitude of the z-component $|E_{2jz}(x_i)|$ at different Z_2's, $Z_2 = 0.1, 10, 100$, respectively. (g) Normalized phase distribution of the z-component $\arg[E_{2jx}(x_i)]/\pi$ at $Z_2 = 100$. u1–u4 stand for different distributions of $E_{1n}(x_p)$. u1: $E_{1n}(x_p) = \delta(x_p)$, where $\delta(x)$ is the Dirac-δ function to represent the ideal point-spread function of G_v. u2: $E_{1n}(x_p) = \mathrm{rect}(x_p/a)$, where $\mathrm{rect}(x_p)$ is the rectangular function of the aperture to represent a uniform light field distribution. u3: $E_{1n}(x_p) = N_{\mathrm{rand}}(x_p)$, where $N_{\mathrm{rand}}(x)$ is a random process consisting of 100 Gaussian random numbers to represent a normally distributed boundary. The Gaussian random numbers are generated by the rand$(\bar{u}, \sigma)$ function in the MATLAB program, where $\bar{u} = 0.5$ is the mean and $\sigma = 0.5$ is the standard deviation. u4: $E_{1n}(x_p) = U_{\mathrm{rand}}(x_p)$, where $U_{\mathrm{rand}}(x_p)$ is a random process consisting of 100 random numbers that are uniformly distributed in the interval [0, 2]. The random numbers are generated by the rand function in the MATLAB program.

where x_{i0} is the central position of the pixel. A plot figure of $\log_{10}(1/\epsilon'_v)$ and the corresponding resolution improvement Δ_M/Δ_{Mr} as a function of Z_2 are as shown in Fig. 5.

APPENDIX B: VECTORAL CALCULATION RESULTS OF DIFFERENT BOUNDARY DISTRIBUTIONS

The complete vectoral diffraction results of $E_{1n}(x_p)$ under different boundary distributions are as shown in Fig. 6. The results show no substantial difference between the vectoral fields calculated by different distributions when $Z_2 \geqslant 100$ even for the nonparaxial region.

APPENDIX C: A MORE RIGOROUS APPROXIMATION

In Eq. (A8), we perform the approximation $G_v(r_2) \approx G_v(r_j)$ and extract it from the integral. Alternatively, we can perform a more rigorous approximation by expanding the phase factor e^{ikr_2} to the first-order $e^{ikr_2} \approx e^{ikr_j}(1 + ik \, \Delta r)$ and write

$$G_v(r_2) \approx G_v(r_j)(1 + ik \, \Delta r), \quad (C1)$$

so that Eq. (A8) is rewritten as

$$\lim_{z_2 \to \infty} \int_{-a/2}^{a/2} E_{1n}(x_p)G_v(r_2)dx_p = G_v(r_j)[a\langle E_j \rangle + f_1], \quad (C2)$$

where

$$f_1 = ik \, \sin(\theta_j) \int_{-a/2}^{a/2} x_p[E_1(x_p) + n(x_p)]dx_p \quad (C3)$$

is the corresponding correction term. Consider that the expected distribution of noise to this term is 0. Also, expand $E_1(x_p)$ to first order,

$$E_{1n}(x_p) \approx E_1(x_{pj}) + (x_p - x_{pj})E'_1(x_{pj}), \quad (C4)$$

and we find

$$f_1 \approx \frac{1}{12}[ik \sin(\theta_j)]a^3 E'_1(x_{pj}), \quad (C5)$$

where $E'_1(x_{pj})$ is the slope of the incident field at x_{pj}. This term represents the contribution of oblique incidence to the diffraction.

[1] V. Ronchi, Resolving power of calculated and detected images, J. Opt. Soc. Am. **51**, 458 (1961).

[2] A. J. den Dekker and A. van den Bos, Resolution: A survey, J. Opt. Soc. Am. A **14**, 547 (1997).

[3] M. Tsang, R. Nair, and X.-M. Lu, Quantum Theory of Super-resolution for Two Incoherent Optical Point Sources, Phys. Rev. X **6**, 031033 (2016).

[4] M. Tsang, Quantum limit to subdiffraction incoherent optical imaging, Phys. Rev. A **99**, 012305 (2019).

[5] C. Lupo and S. Pirandola, Ultimate Precision Bound of Quantum and Subwavelength Imaging, Phys. Rev. Lett. **117**, 190802 (2016).

[6] S. Zhou and L. Jiang, Modern description of Rayleigh's criterion, Phys. Rev. A **99**, 013808 (2019).

[7] S. Van Aert, D. Van Dyck, and A. J. den Dekker, Resolution of coherent and incoherent imaging systems reconsidered - Classical criteria and a statistical alternative, Opt. Express **14**, 3830 (2006).

[8] G. T. di Francia, Resolving power and information, J. Opt. Soc. Am. **45**, 497 (1955).

[9] P. B. Fellgett and E. H. Linfoot, On the assessment of optical images, Philos. Trans. R. Soc. A **247**, 369 (1955).

[10] N. J. Bershad, Resolution, optical-channel capacity and information theory, J. Opt. Soc. Am. **59**, 157 (1969).

[11] E. L. Kosarev, Shannons superresolution limit for signal recovery, Inverse Prob. **6**, 55 (1990).

[12] 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).

[13] T. E. Gureyev, D. M. Paganin, A. Kozlov, Y. I. Nesterets, and H. M. Quiney, Complementary aspects of spatial resolution and signal-to-noise ratio in computational imaging, Phys. Rev. A **97**, 053819 (2018).

[14] T. E. Gureyev, A. Kozlov, Y. I. Nesterets, D. M. Paganin, A. V. Martin, and H. M. Quiney, Signal-to-noise, spatial resolution and information capacity of coherent diffraction imaging, IUCrJ **5**, 716 (2018).

[15] E. Narimanov, Resolution limit of label-free far-field microscopy, Adv. Photon. **1**, 056003 (2019).

[16] Y.-M. Gao, J.-P. Xie, and X.-Y. Yu, Rayleigh-Sommerfeld diffraction on a subwavelength scale: Theories and a resolution criterion, Phys. Rev. A **99**, 023814 (2019).

[17] S. Guha and G. D. Gillen, Description of light propagation through a circular aperture using nonparaxial vector diffraction theory, Opt. Express **13**, 1424 (2005).

[18] X. Luo, D. P. Tsai, M. Gu, and M. Hong, Subwavelength interference of light on structured surfaces, Adv. Opt. Photon. **10**, 757 (2018).

[19] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press, Cambridge, UK, 1999).

[20] H. A. Bethe, Theory of diffraction by small holes, Phys. Rev. **66**, 163 (1944).

[21] C. J. Bouwkamp, On Bethe's theory of diffraction by small holes, Philips Res. Rep. **5**, 321 (1950).

[22] G. Bekefi, Diffraction of electromagnetic waves by an aperture in a large screen, J. Appl. Phys. **24**, 1123 (1953).

[23] E. A. Ash and G. Nicholls, Super-resolution Aperture Scanning Microscope, Nature (London) **237**, 510 (1972).

[24] X. Xie, Y. Chen, K. Yang, and J. Zhou, Harnessing the Point-Spread Function for High-Resolution Far-Field Optical Microscopy, Phys. Rev. Lett. **113**, 263901 (2014).

[25] E. T. F. Rogers, J. Lindberg, T. Roy *et al.*, A super-oscillatory lens optical microscope for subwavelength imaging, Nat. Mater. **11**, 432 (2012).

[26] E. T. F. Rogers and N. I. Zheludev, Optical super-oscillations: Sub-wavelength light focusing and super-resolution imaging, J. Opt. **15**, 094008 (2013).

[27] C. E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379 (1948).

[28] H. Mimura, S. Handa, T. Kimura *et al.*, Breaking the 10 nm barrier in hard-X-ray focusing, Nat. Phys. **6**, 122 (2010).

[29] Z. Chen, Y. Zhou, and J.-T. Shen, Correlation signatures for a coherent three-photon scattering in waveguide quantum electrodynamics, Opt. Lett. **45**, 2559 (2020).

[30] Z. Chen, Y. Zhou, and J.-T. Shen, Photon antibunching and bunching in a ring-resonator waveguide quantum electrodynamics system, Opt. Lett. **41**, 3313 (2016).

[31] Z. Chen, Y. Zhou, and J.-T. Shen, Dissipation-induced photonic-correlation transition in waveguide-QED systems, Phys. Rev. A **96**, 053805 (2017).

[32] R. K. Luneberg, *Mathematical Theory of Optics* (University of California Press, Berkeley, CA., 1964).

[33] G. Pedrini, H. Li, A. Faridian, and W. Osten, Digital holography of self-luminous objects by using a Mach-Zehnder setup, Opt. Lett. **37**, 713 (2012).

[34] *Algorithm*: We use a matric F calculated by the PSF function [Eq. (19)] to calculate the ideal image $y = Fx$, where x is the object array and y is the image array; the reconstructed object x_r is then calculated by the inverse matric of F and the returned signal array $x_r = F^{-1}s_j$.

[35] S. Gazit, A. Szameit, Y. C. Eldar, and M. Segev, Super-resolution and reconstruction of sparse sub-wavelength images, Opt. Express **17**, 23920 (2009).

[36] P. Sidorenko *et al.*, Sparsity-based super-resolved coherent diffraction imaging of one-dimensional objects, Nat. Commun. **6**, 8209 (2015).

[37] Y. Li, Boundary conditions for field distribution in the aperture plane and their effect on diffraction patterns, J. Mod. Opt. **54**, 2723 (2007).