



HAL
open science

Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy

Olivier Lévêque, Caroline Kulcsár, Antony Lee, Hervé Sauer, Artur Aleksanyan, Pierre Bon, Laurent Cognet, François Goudail

► **To cite this version:**

Olivier Lévêque, Caroline Kulcsár, Antony Lee, Hervé Sauer, Artur Aleksanyan, et al.. Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy. *Optics Express*, 2020, 28 (22), pp.32426-32446. 10.1364/OE.402752. hal-03025638

HAL Id: hal-03025638

<https://hal.science/hal-03025638v1>

Submitted on 26 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy

OLIVIER LÉVÊQUE,^{1,*} CAROLINE KULCSÁR,¹ ANTONY LEE,^{2,3}
HERVÉ SAUER,¹ ARTUR ALEKSANYAN,^{2,3} PIERRE BON,^{2,3} LAURENT
COGNET,^{2,3} AND FRANÇOIS GOUDAIL¹

¹Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127, Palaiseau, France.

²Université de Bordeaux, Laboratoire Photonique Numérique et Nanosciences, UMR 5298, 33400, Talence, France.

³Institut d'Optique & CNRS, LP2N UMR 5298, 33400, Talence, France.

*olivier.leveque@institutoptique.fr

Abstract: Single-molecule localization microscopy has become a prominent approach to study structural and dynamic arrangements of nanometric objects well beyond the diffraction limit. To maximize localization precision, high numerical aperture objectives must be used; however, this inherently strongly limits the depth-of-field (DoF) of the microscope images. In this work, we present a framework inspired by “optical co-design” to optimize and benchmark phase masks, which, when placed in the exit pupil of the microscope objective, can extend the DoF in the realistic context of single fluorescent molecule detection. Using the Cramér-Rao bound (CRB) on localization accuracy as a criterion, we optimize annular binary phase masks for various DoF ranges, compare them to Incoherently Partitioned Pupil masks and show that they significantly extend the DoF of single-molecule localization microscopes. In particular we propose different designs including a simple and easy-to-realize two-ring binary mask to extend the DoF. Moreover, we demonstrate that a simple maximum likelihood-based localization algorithm can reach the localization accuracy predicted by the CRB. The framework developed in this paper is based on an explicit and general information theoretic criterion, and can thus be used as an engineering tool to optimize and compare any type of DoF-enhancing phase mask in high resolution microscopy on a quantitative basis.

© 2020 Optical Society of America. Users may use, reuse, and build upon the article, or use the article for text or data mining, so long as such uses are for non-commercial purposes and appropriate attribution is maintained. All other rights are reserved. DOI: [10.1364/OE.402752](https://doi.org/10.1364/OE.402752)

1. Introduction

Over the last decades, a variety of super-resolution fluorescence microscopy techniques have allowed to obtain images with higher resolutions than the diffraction limit [1]. Among these techniques, single-molecule localization microscopy consists in detecting single emitters [2–5], which, when optically isolated, can easily be super-localized, i.e. with nanometric precision [6]. On the one hand, taking advantage of such localization precision, single molecule/particle tracking allows the precise dynamic behavior of molecules to be revealed in complex environments including in live biological cells [7] or in structured materials [8–10]. On the other hand, the ultrastructure of densely labelled entities (e.g. biological specimens [11–14] or nanomaterials [15–17]) can be revealed by controlling the emission properties of the emitters used.

Because molecule diffusion and molecular assemblies are generally not confined within the two dimensions of the imaging microscope plane, several approaches have been designed to extend the super-localization concept to the third dimension [18, 19]. 3D single-molecule localization

proved to be very efficient within the depth-of-field (DoF) of the microscope, even in thick samples [20, 21]. These techniques require sophisticated devices, calibration and processing techniques, and often lead to PSF displaying broadened transverse shapes which might penalize 2D localization accuracy.

Yet, for some applications, extending the DoF without aiming at superlocalizing molecules along the axial direction might be useful, for minimal instrumental and processing complexity or in the case of low photon numbers, for example when imaging at high speed. To this aim, several approaches have been proposed, consisting in extending the DoF by making the PSF of the microscope invariant along the imaging axis to generate volumetric images formed of 2D projections of the 3D imaging volume [22, 23]. This concept was also used in 2-photon excitation fluorescence microscopy for fast volumetric imaging of brain function [24].

In this work, we propose an alternative methodology to optimize phase masks designed to increase the DoF of localization microscopes. Our approach is inspired by “optical co-design” [25–30], which not only takes into account the image formation model and the properties of the optical system to design the phase mask, but also the method of localization extraction to maximize the quality of the final information delivered by the system. Building on this, we propose a rigorous framework to optimize and benchmark phase masks aimed at optimizing 2D location accuracy within a prescribed DoF range. The potential of this framework is illustrated by optimizing and comparing the performance of annular binary phase masks [31] and Incoherently Partitioned Pupil masks [22].

More precisely, this optimization framework is based on the Cramér-Rao Bound (CRB) that represents the fundamental limit of single-molecule localization accuracy [32]. The CRB has already been used in the literature to evaluate the 2D and 3D location capabilities of localization microscopy strategies, and to compare the performance of different strategies [33–35]. However, to the best of our knowledge, it has never been used to design optimized DoF-enhancing phase masks. In order to efficiently localize the particle in practice from the images, we also propose a localization algorithm based on the maximum likelihood (ML) and adapted to the optical characteristics of the optimal masks. We show that contrary to the ML-based methods used in standard focused localization microscopy [33], this algorithm requires segmenting the DoF range in a sufficient number of segments to reach the CRB.

Our approach is in sharp contrast with the works in [31] and [36], which address the problem of DoF extension in classical imaging. In these publications, the optimization criterion is not a location performance expressed by a CRB or a Fisher information. It is the image quality obtained after deconvolution with an averaged Wiener filter. This quality is expressed in terms of mean square error (MSE) between the ideal and deconvolved images. We show in the present paper that since the CRB and MSE-based optimization criteria are different, they lead to optimal masks that are different and that have significantly different localization performance.

This article is organized as follows: Section 2 describes the imaging and noise models, introduces the Fisher information matrix and the CRB, and describes annular binary phase masks aimed at improving the localization accuracy of defocused imaging systems. In Section 3, we present a co-design approach for optimizing DoF-extending phase masks. Using the CRB as a criterion, we optimize annular binary phase masks for various DoF ranges. In Section 4, we design a ML based localization algorithm and demonstrate that it reaches the CRB at the price of a moderate increase of the computational complexity compared to the case when the particle is in focus. In Section 5, we use the developed framework to benchmark the performance of two different types of optimized DoF-extending phase masks, with a particular focus on a simple and easy-to-realize binary mask composed of two rings only. Conclusions and perspectives are drawn in Section 6.

2. Single-molecule localization microscopy and DoF extension

Our goal is to improve the DoF of single-molecule fluorescence microscopes by using phase masks and adapted image processing algorithms. In this section, we first define the signal and noise models considered in this article, then the single-molecule localization accuracy criterion we have chosen. We then describe the type of phase masks we shall use for DoF extension and illustrate their capacity to make localization accuracy nearly invariant to defocus.

2.1. Signal and noise models

Since the emitter is unresolved by the microscope, we observe, in the image plane, the point spread function (PSF) of the microscope objective centered on the geometric image position of the emitter. The irradiance is proportional to

$$f^{\psi, \theta_0}(x, y) = f^{\psi}(x - M\theta_{0x}, y - M\theta_{0y}) \quad (1)$$

where M is the lateral magnification of the imaging setup and $\theta_0 = (\theta_{0x}, \theta_{0y})^t$ is the position of the emitter in the object plane. The superscript t denotes transposition. The function $f^{\psi}(x, y)$ represents the 2D spatial distribution of the PSF and is normalized so that $\iint f^{\psi}(x, y) dx dy = 1$. It is proportional to the squared modulus of the Fourier transform of the normalized complex pupil function defined as $\exp[i\Phi(\xi, \eta)]$, when $\xi^2 + \eta^2 < 1$, and 0 otherwise, with the phase function $\Phi(\xi, \eta)$ expressed, for pure defocus wavefront error, by

$$\Phi(\xi, \eta) = \frac{2\pi\psi}{\lambda} (\xi^2 + \eta^2), \quad (\xi, \eta) \in [-1, 1]^2. \quad (2)$$

In this equation, the defocus parameter ψ is defined as

$$\psi = \frac{NA_o^2 \Delta z_o}{2n_o} \quad (3)$$

where NA_o denotes the object NA of the imaging system, Δz_o the longitudinal distance between the observed particle and the focus point, and n_o the refractive index of the object space. This parameter represents the first non-zero expansion of the optical path difference between the focused and defocused spherical wavefronts at the edge of the exit pupil. This approximation is valid for typical microscope configurations [37, 38]. According to this model, if $\psi = 0$ (the object is in focus), the PSF is the classical Airy pattern with size defined by the object NA of the microscope objective, the absolute value of M , the lateral magnification of the imaging setup and the wavelength λ of the collected light. When ψ increases, the PSF progressively gets wider [39, 40]. It is classically considered that defocus is not critical when $|\psi| < \lambda/4$ (Rayleigh criterion).

The digital image delivered by the sensor is a version of Eq. (1) that has been sampled, filtered by the finite size pixel and corrupted by noise. Let us assume that this image has a width of $2P + 1$ pixels, with $P \in \mathbb{N}^+$. We call s_{ij} the number of photo-electrons observed at pixel $(i, j) \in \mathbb{N}^2$ with $|i|$ and $|j| \leq P$. Assuming that the measurement noise is additive, Gaussian, spatially white, of mean zero and variance σ_n^2 , the number s_{ij} is a Gaussian random variable of probability density function

$$p(s_{ij} | \mu_{ij}^{\psi, \theta_0}) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp \left[-\frac{(s_{ij} - \mu_{ij}^{\psi, \theta_0})^2}{2\sigma_n^2} \right] \quad (4)$$

where

$$\mu_{ij}^{\psi, \theta_0} = N_0 \int_{(i-\frac{1}{2})\Delta_{xy}}^{(i+\frac{1}{2})\Delta_{xy}} \int_{(j-\frac{1}{2})\Delta_{xy}}^{(j+\frac{1}{2})\Delta_{xy}} f^{\psi, \theta_0}(x, y) dx dy \quad (5)$$

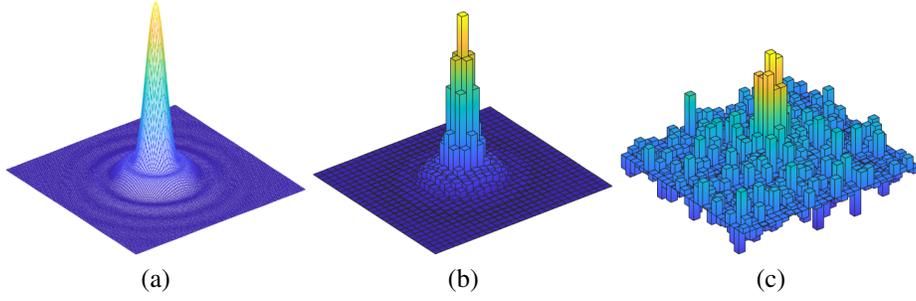


Fig. 1. PSF simulations of a defocused optical system such as $\psi = \pm 0.4\lambda$, $N_0 = 500$ photons, and $\sigma_n^2 = 6$ photons²/pixel (noise variance). The other simulation parameters are defined in Table 1. (a) The 2D spatial distribution of measured photo-electrons $f^{\psi, \theta_0}(x, y)$. (b) The expected number of photo-electrons per pixel $\mu_{ij}^{\psi, \theta_0}$. (c) A realization of the observed data per pixel s_{ij} .

is the mean value of s_{ij} . In Eq. (5), N_0 denotes the total number of photo-electrons expected in the whole image and Δ_{xy} the length of the side of the square pixels. We have represented the photo-electron detection steps in Fig. 1: Fig. 1.a represents the continuous PSF $f^{\psi, \theta_0}(x, y)$ of a defocused optical system with $\psi = \pm 0.4\lambda$; Fig. 1.b represents its sampled/integrated version $\mu_{ij}^{\psi, \theta_0}$; and Fig. 1.c represents a realization of the observed data per pixel s_{ij} according to the statistical distribution in Eq. (4).

For the sake of simplicity, we do not take into account in our model the signal-dependent photon noise, that is Poisson distributed and not additive. Also, we use Fourier optics for modeling image formation of high-aperture microscope objectives and nanometric targets (see [37, 41] for a more accurate electromagnetic-based model). These simplifications ease the proof of the DoF-enhancing potential of binary phase masks. However, the results obtained in this paper can be generalized without difficulty to more thorough imaging and noise models.

2.2. Fundamental limit of localization accuracy

Locating a molecule consists in estimating the particle coordinates θ_0 with high accuracy from the measured data s_{ij} . According to [33], the fundamental limit of localization accuracy can be obtained from the Fisher information matrix, which indicates how the likelihood of the observed data is affected by changes in the values of the parameters of interest. It is defined by

$$\underline{\mathbf{I}}(\theta_0) = \sum_{i=-P}^P \sum_{j=-P}^P \mathbb{E} \left[\left(\left. \frac{\partial \ln p(s_{ij} | \mu_{ij}^{\psi, \theta})}{\partial \theta} \right|_{\theta=\theta_0} \right) \left(\left. \frac{\partial \ln p(s_{ij} | \mu_{ij}^{\psi, \theta})}{\partial \theta} \right|_{\theta=\theta_0} \right)^t \right] \quad (6)$$

where $p(s_{ij} | \mu_{ij}^{\psi, \theta_0})$ is defined in Eq. (4) and the symbol $\mathbb{E}[\cdot]$ denotes the mathematical expectation operation.

The diagonal values of the inverse of the Fisher matrix are the Cramér-Rao bounds (CRB) of the estimates of θ_{0x} and θ_{0y} . They represent a lower bound on the estimation error variance of these parameters that can be obtained with an unbiased estimator. They thus represent the intrinsic difficulty of an estimation problem, independently of the (unbiased) method used to solve it. We consider here equivalently the square root of the CRB, denoted RCRB and which has the dimension of a distance, as the limit of localization accuracy.

Due to the circular symmetry of the PSF $f^{\psi, \theta_0}(x, y)$ for any value of ψ , the off-diagonal terms of the Fisher information matrix are zero and the RCRB is the square root of the inverse of the

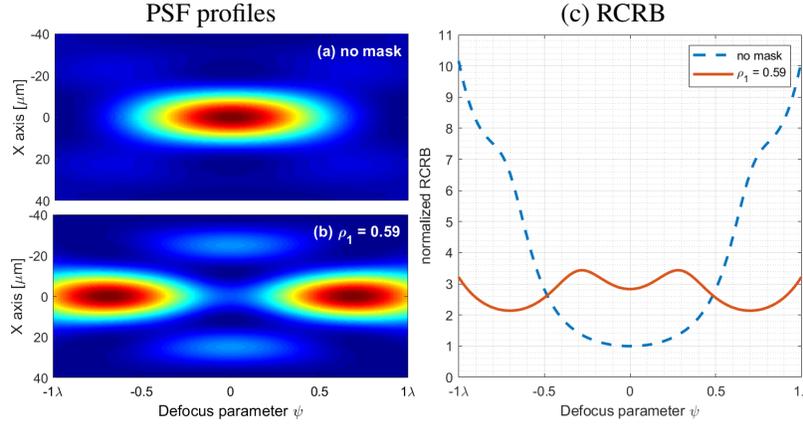


Fig. 2. PSF variability of the x axis profile of (a) an aberration-free diffraction-limited system with a circular aperture, (b) the same optical system using a two-ring binary phase mask with $\rho_1 = 0.59$ and (c) the normalized limit of the localization accuracy, denoted normalized RCRB, along x coordinate axis as a function of defocus. The simulation parameters are defined in Table 1.

diagonal elements of the Fisher information. Thus, the RCRB along the x and y axes have the following expressions:

$$\text{RCRB}_x = \frac{\sigma_n}{N_0} \left[\sum_{i=-P}^P \sum_{j=-P}^P \left(\int_{(j-\frac{1}{2})\Delta_{xy}}^{(j+\frac{1}{2})\Delta_{xy}} \left[f^{\psi, \theta_0} \left(\left(i - \frac{1}{2} \right) \Delta_{xy}, y \right) - f^{\psi, \theta_0} \left(\left(i + \frac{1}{2} \right) \Delta_{xy}, y \right) \right] dy \right)^2 \right]^{-\frac{1}{2}} \quad (7)$$

$$\text{RCRB}_y = \frac{\sigma_n}{N_0} \left[\sum_{i=-P}^P \sum_{j=-P}^P \left(\int_{(i-\frac{1}{2})\Delta_{xy}}^{(i+\frac{1}{2})\Delta_{xy}} \left[f^{\psi, \theta_0} \left(x, \left(j - \frac{1}{2} \right) \Delta_{xy} \right) - f^{\psi, \theta_0} \left(x, \left(j + \frac{1}{2} \right) \Delta_{xy} \right) \right] dx \right)^2 \right]^{-\frac{1}{2}} \quad (8)$$

where σ_n , N_0 , Δ_{xy} , and $f^{\psi, \theta_0}(x, y)$ appear in Eq. (1) and (5). These expressions show this standard result that the fundamental limit on localization is inversely proportional to the signal-to-noise ratio (SNR), defined as $\text{SNR} = N_0/\sigma_n$ and depends on the spatial sampling of the PSF. Moreover, when the PSF is circularly symmetric, which is the case here, one has $\text{RCRB}_x = \text{RCRB}_y$. Thus in the following, we will use the normalized RCRB in order to characterize the influence of defocus for any value of the SNR. The normalized RCRB is defined by dividing RCRB_x by its value when the object is in focus (i.e. $\psi = 0$). It is independent of the value of the SNR.

As an illustration, we have represented in Fig. 2.a the variation of the PSF profile along the x axis as a function of the defocus parameter ψ for an example of microscope configuration defined in Table 1. We express the defocus parameter ψ in units of central wavelength λ of the collected light. We note that the larger the defocus parameter, the more the PSF spreads out and its central lobe gets fainter. We have represented in Fig. 2.c the variation of the normalized RCRB of this optical system as a function of defocus (blue dotted line). It is observed that the RCRB increases very slowly until $\psi \approx \pm\lambda/4$, which corresponds to the Rayleigh criterion, then the increase gets much sharper. At a defocus parameter of $\psi = \pm 1\lambda$, the RCRB is more than ten times larger than for $\psi = 0$ where, for example, the RCRB is equal to 0.04 pixel (or 2.7 nm) when $N_0 = 500$ photons and $\sigma_n^2 = 6$ photons²/pixel. The case with a phase mask, in Fig. 2.b, is discussed in the next section.

Parameters	Symbols	Values
Particle position	θ_0	$(0, 0)^t \mu\text{m}$
Pixel size	Δ_{xy}	$6.45 \mu\text{m}$
Image size	$2P + 1$	27 pixels
Object numerical aperture	NA_o	1.4
Wavelength	λ	520 nm
Lateral magnification	M	100

Table 1. Simulation parameters use in Figs. 1,2,4,5,6,7,8,9,10,12

2.3. DoF extension using annular binary phase masks

Our goal is to minimize the variation of localization accuracy with defocus by placing an optimized phase mask in the exit pupil of the optical system and by adapting the localization algorithm.

There exists many types of phase masks that may improve the DoF [42–47]. We consider in this article annular binary phase masks since they are easy to manufacture and have proven their efficiency in classical imaging applications [36, 48–51]. These masks consist of concentric rings defined by their normalized outer radii. Each ring implements a phase modulation of alternatively 0 or π radians at a nominal wavelength λ . For instance, we have represented in Fig. 3 a four-ring annular binary phase mask defined by the parameter vector $\rho = (\rho_1, \rho_2, \rho_3)^t$, with $0 \leq \rho_1 \leq \rho_2 \leq \rho_3 \leq 1$, where ρ_n is the outer radius of the n -th ring. This mask defines a binary phase function that we denote by $\Phi_{\text{mask}}(\xi, \eta, \rho)$. If such a mask is placed in the exit pupil of a defocused optical system, the phase function in the exit pupil, as defined in Eq. (2), has the following expression:

$$\Phi(\xi, \eta, \rho) = \frac{2\pi\psi}{\lambda} (\xi^2 + \eta^2) + \Phi_{\text{mask}}(\xi, \eta, \rho), \quad (\xi, \eta) \in [-1, 1]^2 \quad (9)$$

where ρ is the mask parameter vector.

We have represented in Fig. 2.b the variation of the PSF profile as a function of the defocus parameter ψ when a two-ring mask defined by $\rho_1 = 0.59$ is placed in the exit pupil (the reason for choosing this value of ρ_1 will become apparent in the next section). The presence of the mask significantly alters the optical properties of the system through defocus when compared with the aberration free PSF in Fig. 2.a. The minimum spread of the PSF is not obtained for perfect focus (i.e. $\psi = 0$), but for $\psi \approx \pm 0.7\lambda$. On the other hand, for $\psi = 0$, the PSF profile is not concentrated but divided in three main lobes. It is also noticed from Fig. 2.a and 2.b that whether or not the mask is present, the PSF is identical for ψ and $-\psi$ (that is, on either side of the focus point). This would not be the case for most other types of DoF-enhancing phase masks, such as the cubic mask [42]. This is an interesting property of binary phase masks with π -phase modulation [31].

We have represented in Fig. 2.c the variation of the normalized RCRB as a function of ψ with this mask placed in the exit pupil (red solid line). The normalized RCRB for a given scenario is equal to the RCRB in this scenario divided by the value of the RCRB without mask and in focus (i.e. $\psi = 0$). By comparison with the curve obtained without mask (blue dotted line), it is seen that the mask allows to get much lower values of the RCRB over the whole defocus range. With the mask, the maximal value of the RCRB over the defocus range is three times smaller than without it. Of course, this value is also three times larger than the value obtained without mask and in focus: the price to pay to extend the DoF is to degrade the localization accuracy in focus.

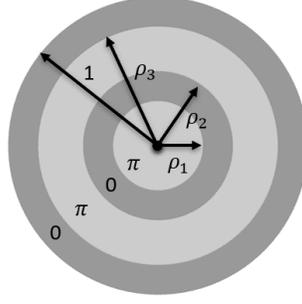


Fig. 3. Example of four-ring annular binary phase mask defined by $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3)^t$. Each ring is defined as an annular region with constant phase modulation. Dark gray areas induce a phase of 0 radians and light gray areas induce a phase of π radians at a nominal wavelength λ .

It is interesting to note that the value of RCRB is good at $\psi = 0$, even if for this defocus parameter, the PSF is spread into three main lobes (see Fig. 2.b). This means that such a PSF, albeit not concentrated in a single lobe, still contains enough information to ensure accurate localization.

Our goal will be to determine the binary phase masks parameters that optimize the RCRB for different numbers of rings and different values of the defocus range, and to determine the algorithms that make it possible to reach a localization accuracy equal to the RCRB in practice.

3. Binary phase mask optimization for localization applications

In this section, we define an optimization criterion for binary phase masks aimed at DoF extension. This criterion is highly non-convex, and we describe a method to perform its optimization. Then, we apply this method to optimize annular binary phase masks composed of two to five rings for various defocus ranges. In each case, we evaluate the obtained trade-off between localization accuracy and DoF extension.

3.1. Optimization method

Using Eq. (7) and (8), we are able to calculate the RCRB of an imaging system equipped with a phase mask for any value of the defocus parameter ψ . Since $\text{RCRB} = \text{RCRB}_x = \text{RCRB}_y$ and the PSF is identical for ψ and $-\psi$ using binary phase masks with π -phase modulation, a reasonable criterion for phase mask optimization is therefore to minimize the value of the RCRB for the worst value of ψ , that is, to define the optimal mask parameters as

$$\boldsymbol{\rho}_{\text{opt}} = \arg \min_{\boldsymbol{\rho}} \left\{ \max_{\psi \in [0, \psi_{\text{max}}]} \text{RCRB}(\boldsymbol{\rho}, \psi) \right\} \quad (10)$$

where $[0, \psi_{\text{max}}]$ is the defocus range on which we want the localization to be accurate. Solving this minimax optimization problem is not simple since the function $\max_{\psi} \text{RCRB}(\boldsymbol{\rho}, \psi)$ is non-convex and presents several local minima. In Fig. 4.a, the value of $\max_{\psi} \text{RCRB}(\rho_1, \psi)$ is represented for a two-ring mask as a function of the radius ρ_1 and for different values of ψ_{max} . We notice that there are several local minima. Similarly, Fig. 4.b represents the value of $\max_{\psi} \text{RCRB}(\rho_1, \rho_2, \psi)$ for a three-ring mask as a function of the radii ρ_1 and ρ_2 for $\psi_{\text{max}} = 2.5\lambda$. We also can see several local minima (the global minimum is marked with a white cross). In these two cases, mask optimization can be solved by an exhaustive search. To achieve the non-convex optimization defined in Eq. (10) with a larger number of rings, we use the particle swarm optimization algorithm defined in [52]. This algorithm, developed by J. Kennedy and R. Eberhart in the late 1980s, relies on the collaboration of individuals. Based on simple displacement rules, a set of

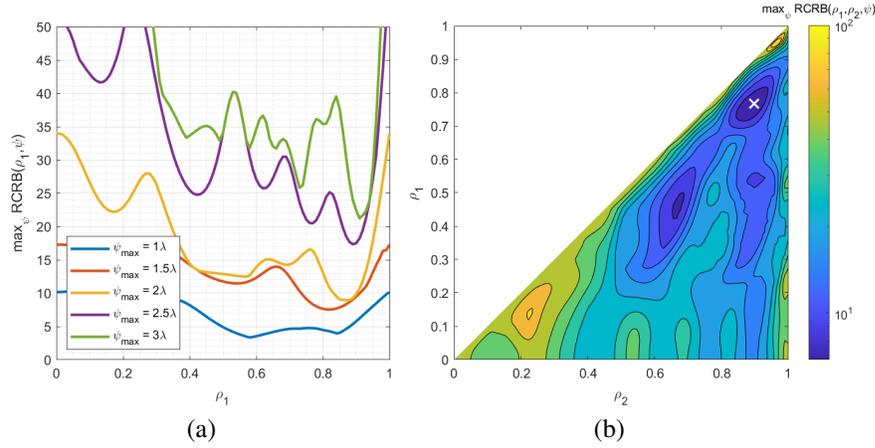


Fig. 4. Evolution of the normalized \max_{ψ} RCRB(ρ, ψ) for (a) two-ring masks computed for $\psi_{\max} = \{1\lambda, 2\lambda, 3\lambda\}$ and (b) three-ring for $\psi_{\max} = 2.5\lambda$. The cross corresponds to the most prominent local minimum. The simulation parameters are defined in Table 1.

particles explores the landscape of the phase mask parameters to be optimized and gradually converges towards a local minimum. Because we have no guarantee to find the global minimum in a single optimization run, we perform a large number of optimization runs with randomly chosen initial values to find a reliable value of the global optimum, similarly to what is done in [31].

3.2. Performance and limits of optimized binary phase masks

By applying the optimization method described in the previous section, we have optimized multi-ring binary phase masks for various defocus ranges. Fig. 5.a represents the maximal value of the normalized RCRB over the defocus range obtained with the optimal mask parameters ρ_{opt} , defined as

$$\text{RCRB}_{\max} = \max_{\psi \in [0, \psi_{\max}]} \text{RCRB}(\rho_{\text{opt}}, \psi) \quad (11)$$

and plotted as a function of the number of rings. Each curve corresponds to a different value of the defocus range ψ_{\max} . The leftmost point of the curves corresponds to an optical system without mask. In this case, RCRB_{\max} naturally increases with ψ_{\max} . When a mask is used, RCRB_{\max} first sharply decreases with the number of rings, then levels off. For example, on the blue curve corresponding to $\psi_{\max} = 1\lambda$, RCRB_{\max} is divided by three by using a two-ring mask, but is not reduced as the number of rings is further increased. This means that to solve the DoF extension problem over this defocus range with an annular binary phase mask, two rings are sufficient. It has to be noted that the optimal two-ring mask is defined by $\rho_1 = 0.59$, which corresponds to the mask used to plot Fig. 3.

If we now consider the orange curve that corresponds to $\psi_{\max} = 1.5\lambda$, we first see that it is above the blue one, since the problem to solve is more difficult. Moreover, for this defocus range, using three rings instead of two brings some improvement in localization accuracy. Globally speaking, it is observed that as the defocus range widens, a larger number of rings is needed to reach the optimal localization accuracy. For the maximal considered value of the defocus range $\psi_{\max} = 3\lambda$, four rings are necessary and sufficient to reach the minimal RCRB. Fig. 5.b represents the profiles of the optimized binary phase masks obtained for this defocus range for two to five rings. We can see that the optimal mask with five rings is quite similar to that with four rings, which is consistent with their similar localization performance.

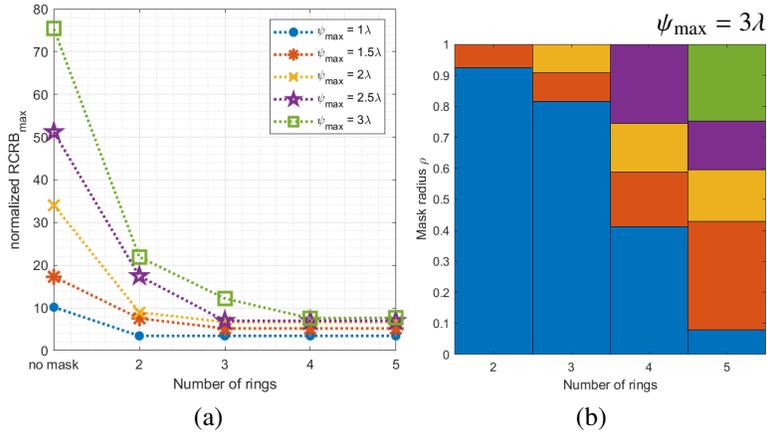


Fig. 5. (a) Evolution of the normalized RCRB_{\max} as a function of the number of rings of binary phase masks for different values of ψ_{\max} . (b) Optimized binary phase masks profiles for $\psi_{\max} = 3\lambda$ and various number of rings. The simulation parameters are defined in Table 1.

The main conclusion of these results is that optimal binary masks significantly improve localization accuracy for all the considered defocus ranges, and that a limited number of rings is sufficient to obtain optimal performance. These conclusions are similar to those obtained in the case of DoF enhancement of classical imaging systems, where the optimization criterion is the quality of the deconvolved image [31]. However, as shown in Appendix A, the optimal masks are different.

In order to gain a deeper insight into the obtained results, let us analyze how the optimal masks modify the PSF of the optical system and how these modifications make it possible to extend the DoF. Fig. 6.a represents the variation of the PSF profiles as a function of the defocus parameter for four different configurations: without mask and with masks optimized, respectively, for $\psi_{\max} = \{1\lambda, 1.5\lambda, 3\lambda\}$. It can be seen that the use of a binary phase mask significantly reduces the fading of the PSF over the defocus range. Fig. 6.b illustrates the variation of the normalized RCRB as a function of ψ for the same four imaging systems. Comparing Fig. 6.a and 6.b is insightful. It is observed that the local minima of the RCRB in Fig. 6.b correspond, for all defocus ranges, to two different types of PSF profiles that are appropriate for localization in Fig. 6.a. The first one is characterized by an important central lobe. It occurs for example at $\psi \approx 0.7\lambda$ for $\psi_{\max} = 1\lambda$ and $\psi = 0$ for $\psi_{\max} = 1.5\lambda$. The second type is characterized by important secondary lobes. It occurs for example at $\psi \approx 0$ for $\psi_{\max} = 1\lambda$ and $\psi \approx 0.7\lambda$ for $\psi_{\max} = 1.5\lambda$. The PSF profiles corresponding to the transition between these two types yield higher values of the RCRB. For example, for $\psi_{\max} = 1.5\lambda$, the worst localization accuracy is obtained for $\psi \approx 0.4\lambda$. In Fig. 6.a, this value of ψ corresponds to a low contrast PSF profile with no distinct central or secondary lobes.

4. Localization algorithm

The phase masks have been optimized using the RCRB criterion defined in Eq. (11), which represents a lower bound on the localization standard deviation of unbiased estimation. It is thus a “potential” performance level, and one has to specify actual estimators able to reach this performance in practice. In the case of well focused images, it has been shown that for sufficient SNR, ML algorithms are able to reach the CRB [33]. However, in our case, the problem is more involved since it depends on another parameter, the defocus parameter ψ , which is unknown and

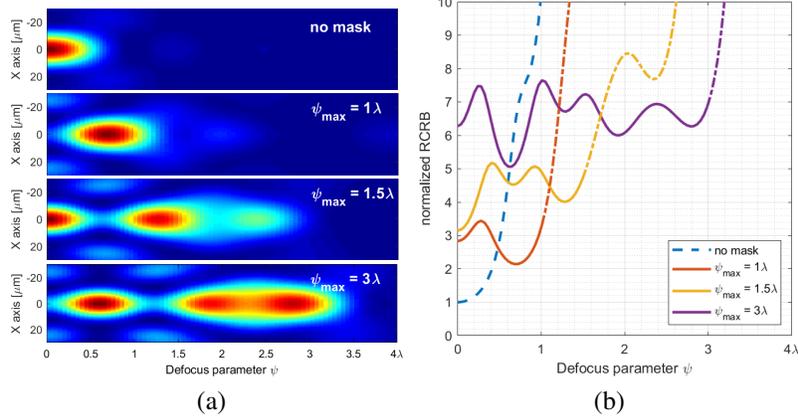


Fig. 6. Evolutions of (a) the PSF profile along x axis and (b) the normalized limit of the localization accuracy, denoted normalized RCRB, as a function of defocus for an optical system using an binary phase mask optimized for $\psi_{\text{max}} = \{1\lambda, 1.5\lambda, 3\lambda\}$. The simulation parameters are defined in Table 1.

will not be retrieved. Thus, the parameter ψ can be considered as a nuisance parameter for our localization problem. If one does not have any *a priori* information on the actual value of ψ , one has to maximize, with respect to θ and ψ , the log-likelihood defined as

$$\ell(\theta, \psi) = \sum_{i=-P}^P \sum_{j=-P}^P \ln p(s_{ij} | \mu_{ij}^{\psi, \theta}) \quad (12)$$

where s_{ij} and $\mu_{ij}^{\psi, \theta}$ are as defined in Eq. (5) and above, and $p(s_{ij} | \mu_{ij}^{\psi, \theta})$ is defined in Eq. (4). It amounts to a joint ML estimation of θ and ψ . The problem is that the defocus parameter ψ is a continuous variable, so that this problem is intractable. In the next section, we propose a first method based on a single averaged kernel, which has the advantage of being as fast as in the standard case of a well-focused particle. After analyzing the performance and drawbacks of this method, we propose another solution based on the division of the defocus range in a limited number of subdomains.

4.1. Estimation based on a global kernel

When the defocus parameter ψ is known, the ML estimate of the position θ can be written as a correlation product:

$$\hat{\theta}_{\text{ml}} = \arg \max_{\theta} \sum_{i=-P}^P \sum_{j=-P}^P \ln p(s_{ij} | \mu_{ij}^{\psi, \theta}) = \arg \min_{\theta} \sum_{i=-P}^P \sum_{j=-P}^P (s_{ij} - \mu_{ij}^{\psi, \theta})^2 \quad (13)$$

$$= \arg \min_{\theta} \left[-2 \sum_{i=-P}^P \sum_{j=-P}^P s_{ij} \mu_{ij}^{\psi, \theta} + \sum_{i=-P}^P \sum_{j=-P}^P s_{ij}^2 + \sum_{i=-P}^P \sum_{j=-P}^P (\mu_{ij}^{\psi, \theta})^2 \right] \quad (14)$$

$$\approx \arg \max_{\theta} \sum_{i=-P}^P \sum_{j=-P}^P s_{ij} \mu_{ij}^{\psi, \theta} \quad \underbrace{\hspace{10em}}_{\text{almost constant}} \quad (15)$$

This estimator requires the knowledge of $\mu_{ij}^{\psi, \theta}$ and thus, of ψ . In practice, we have no information on ψ . A simple way to cope with this problem is to replace $\mu_{ij}^{\psi, \theta}$ in Eq. (15) with a defocus

invariant kernel r_{ij}^θ such that the estimator has the following expression:

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=-P}^P \sum_{j=-P}^P s_{ij} r_{ij}^\theta. \quad (16)$$

The question is how to construct this kernel. First, we show in Appendix B that it is possible, with some approximations, to get a closed-form expression of the variance $\text{var}_{\psi}[\hat{\theta}]$ of the estimator in Eq. (16) for a given value of ψ . We then choose a set of values ψ_k , $k \in [1, K]$, evenly spaced within the defocus range $[0, \psi_{\max}]$ and determine the kernel r_{ij}^θ that maximizes $\sum_k (\text{var}_{\psi_k}[\hat{\theta}])^{-1}$, that is, the sum of inverses of the estimation variance for each value of ψ_k . We show in Appendix C that this kernel has the following expression:

$$r_{ij}^\theta = \sum_{k=1}^K \alpha_k \mu_{ij}^{\psi_k, \theta} \quad (17)$$

where the coefficients α_k are the components of the eigenvector associated with the greatest eigenvalue of the matrix $\underline{\mathbf{W}}$ defined as

$$[\underline{\mathbf{W}}]_{ij} = \iint_{\mathbb{R}^2} v_x v_y \tilde{f}^{\psi_i*}(v_x, v_y) \tilde{f}^{\psi_j}(v_x, v_y) dv_x dv_y \quad (18)$$

where the superscript \sim denotes the Fourier transform, the superscript $*$ denotes the complex conjugate and $f^\psi(x, y)$ is as defined in Eq. (1). It can be noted that this approach consisting in defining a single ‘‘averaged’’ kernel is similar to that used for DoF extension in classical imaging, where an averaged Wiener filter is used to deconvolve the image whatever the value of ψ [28].

Fig. 7.a compares the empirical normalized standard deviation of the estimator defined in Eq. (16) with the normalized RCRB as a function of the defocus parameter ψ . The empirical normalized standard deviation of the estimator for a given scenario is equal to the empirical standard deviation of the estimator in this scenario divided by the value of the RCRB without mask and in focus (i.e. $\psi = 0$). The simulated optical system uses a two-ring binary phase mask optimized for the DoF range $\psi_{\max} = 1\lambda$. The estimator standard deviation is estimated using Monte-Carlo simulations based on 4000 realizations. This curve shows that the empirical normalized standard deviation of the estimator reaches the normalized RCRB only for $\psi > 0.4\lambda$. For values of ψ below this limit, the variance of the estimator is much larger than the RCRB (the values are outside the graph limits in Fig. 7.a).

The reason for this failure is the following. From Fig. 6, it can be seen that the shape of the PSF strongly varies with the defocus. For the defocus range $\psi_{\max} = 1\lambda$, two regimes can be distinguished: one where the shape of the PSF is spread with important secondary lobes (see Fig. 7.c) and the other where the PSF is more concentrated around a major main lobe (see Fig. 7.d). The optimal defocus invariant kernel r_{ij}^θ , represented in Fig. 7.b, is clearly much more similar to the PSF for $\psi \simeq 0.7\lambda$ than to the PSF for $\psi = 0$. This explains why its accuracy is good for larger values of the defocus parameter but dramatically fails for smaller values. In conclusion, it is impossible for a unique kernel - even an optimal one - to encompass the PSF variability over the defocus range $[0, 1\lambda]$. It is therefore necessary to split the defocus range into sub-ranges on which the shape of the PSF is more invariant to apply the estimator defined in Eq. (16). It is important to specify that these sub-ranges do not allow to estimate the defocus parameter, but only to improve the estimation algorithm of the 2D coordinates of the single molecule observed.

4.2. Estimation based on multiple kernels

Suppose that the DoF range $[0, \psi_{\max}]$ is split into M distinct sub-ranges. Using the method described in the previous section, we can define M optimal defocus invariant kernels over each

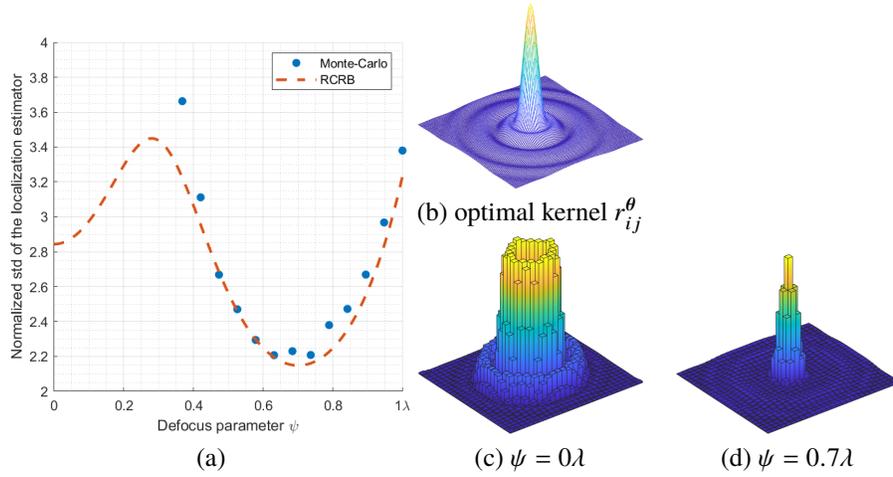


Fig. 7. (a) Comparison between the normalized standard deviation (std) estimator defined in Eq. (16) and the normalized fundamental limit of localization accuracy as a function of defocus. The standard deviation of the estimator is computed empirically with 4000 realizations. (b) Optimal correlation kernel that maximizes $\sum_k (\text{var}_{\psi_k} [\hat{\theta}])^{-1}$ for $\psi_k = \{0, 0.2, 0.4, 0.6, 0.8, 1\lambda\}$. The expected number of photo-electrons per pixel is represented (c) for $\psi = 0\lambda$ and (d) $\psi = 0.7\lambda$. Simulation parameters: $N_0 = 500$ photons, $\sigma_n^2 = 6$ photons²/pixel (noise variance) and $\rho_1 = 0.59$ (mask parameter). The other simulation parameters are defined in Table 1.

defocus sub-ranges, called $r_{ij}^{m,\theta}$ with $m \in \{1, \dots, M\}$. Then, we jointly estimate the particle coordinates $\hat{\theta}$ and the defocus sub-range \hat{m} in the ML sense:

$$(\hat{\theta}, \hat{m}) = \arg \min_{\theta, m} \sum_{i=-P}^P \sum_{j=-P}^P (s_{ij} - \beta r_{ij}^{m,\theta})^2 \quad (19)$$

where β is a scale factor. Indeed, to perform the fit with the data actually present in the image s_{ij} , it is necessary to scale the amplitude of the kernel to the amplitude of s_{ij} with help of the parameter β defined as

$$\beta = \frac{\sum_{i=-P}^P \sum_{j=-P}^P s_{ij} r_{ij}^{m,\theta}}{\sum_{i=-P}^P \sum_{j=-P}^P (r_{ij}^{m,\theta})^2} \quad (20)$$

Let us consider again the imaging system with a two-ring binary phase mask $\rho_1 = 0.59$ studied in Fig. 7. We can define four sub-ranges in which the shape of the PSF is almost invariant: the sub-range $m = 1$ is $\psi \in [0, 0.2\lambda]$, $m = 2$ is $\psi \in [0.2\lambda, 0.35\lambda]$, $m = 3$ is $\psi \in [0.35\lambda, 0.5\lambda]$, and $m = 4$ is $\psi \in [0.5\lambda, 1\lambda]$. For each of these sub-ranges, we compute the optimal defocus invariant kernel $r_{ij}^{m,\theta}$ as defined in Eq. (17). In practice, each of these four kernels is built using six values ψ_k evenly spaced within each sub-range. They are respectively illustrated in Fig. 8.b, 8.c, 8.d, and 8.e. In particular, it is noticed that the kernel corresponding to the first sub-range (see Fig. 8.b) is now similar to the PSF for $\psi = 0$ (see Fig. 7.c). In Fig. 8.a, we have represented with blue dots the empirical normalized standard deviation of the estimator defined in Eq. (19) (estimated with Monte Carlo simulations). If we compare them to the RCRB (red dotted line), we note that in contrast to Fig. 7.a, the actual standard deviation reaches the RCRB for all the defocus

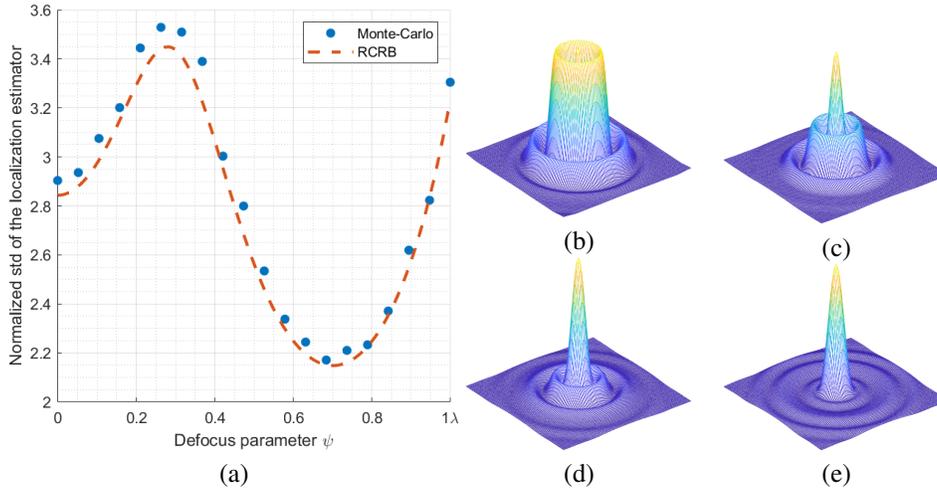


Fig. 8. (a) Comparison between the normalized standard deviation (std) estimator and the normalized fundamental limit of localization accuracy as a function of defocus. The standard deviation of the estimator is computed empirically with 4000 realizations. On the right part, optimal correlation kernels that maximize $\sum_k (\text{var}_{\psi_k} [\hat{\theta}])^{-1}$ are shown for (b) $\psi_k \in [0, 0.2\lambda]$, (c) $\psi_k \in [0.2\lambda, 0.35\lambda]$, (d) $\psi_k \in [0.35\lambda, 0.5\lambda]$, (e) $\psi_k \in [0.5\lambda, 1\lambda]$. Simulation parameters: $N_0 = 500$ photons, $\sigma_n^2 = 6$ photons²/pixel (noise variance) and $\rho_1 = 0.59$ (mask parameter). The other simulation parameters are defined in Table 1.

parameters. Dividing the defocus range $[0, \psi_{\max}]$ into four parts is thus enough to encompass the PSF variability within this range. The price to pay is to perform four correlations instead of a single one in the case of well-focused systems.

Of course, when the defocus range increases, the PSF variability within this range also increases (see Fig. 6.a). As a consequence, one has to use a larger number of sub-ranges to encompass this variability. For example, for a range $\psi_{\max} = 1.5\lambda$, we have used five sub-ranges. We have represented in Fig. 9.a the standard normalized deviation of this estimator, estimated with Monte-Carlo simulations, and we see that it fits the normalized RCRB. Fig. 9.b represents the same values for $\psi_{\max} = 2\lambda$. In this case, we had to use six sub-ranges to fit the normalized RCRB.

As a summary, contrary to well-focused systems, the RCRB of DoF-enhanced systems cannot be reached with an estimator consisting of a single correlation kernel. This is due to the fact that even with the optimal masks, the PSF varies within the defocus range. There is thus a price to pay in terms of computational complexity for DoF extension. We have proposed a method based on the subdivision of the defocus range in a sufficient number of sub-ranges, each one being associated with a single correlation kernel. We have shown that this method makes it possible to reach the RCRB. Its computation complexity is simply proportional to the number M of necessary sub-ranges, which increases with the width of the defocus range. Indeed, for M subranges, localization requires M correlations instead of a single one in the standard case of focused localization.

5. Comparison with other previously proposed DoF-extending mask

The mask optimization approach presented in this article is based on a general and objective localization criterion: the RCRB. This allows the comparison of any type of DoF-extending strategy on a quantitative basis. In order to illustrate this potential, we compare in this section the annular binary masks and the masks introduced in [22]. These masks consist of a series of

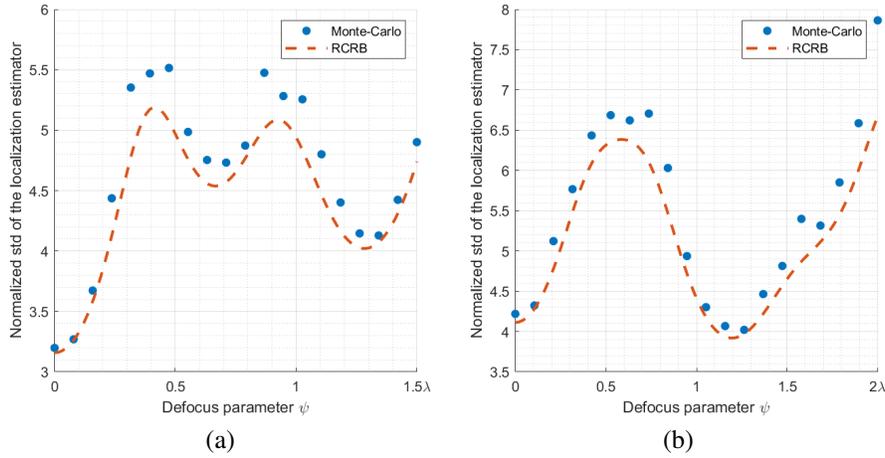


Fig. 9. Comparison between the normalized standard deviation (std) estimator and the normalized fundamental limit of localization accuracy as a function of defocus for a range (a) $\psi_{\max} = 1.5\lambda$ with $\rho = (0.4709, 0.6782)^t$ and (b) $\psi_{\max} = 2\lambda$ with $\rho = (0.7436, 0.8974)^t$. The standard deviation of the estimator is computed empirically with 4000 realizations. Simulation parameters: $N_0 = 500$ photons and $\sigma_n^2 = 6$ photons²/pixel (noise variance). The other simulation parameters are defined in Table 1.

concentric annular sub-apertures introducing phase delays that are much larger than the coherence length of illumination. Hence, the light beams emerging for each sub-aperture are mutually incoherent and the PSF of the mask is simply the incoherent addition of the PSFs produced by each sub-aperture. In the following, these masks will be denoted Incoherently Partitioned Pupil (IPP) masks. Their parameters are the widths of the rings.

To compare these two strategies, we optimize the masks, that is, the ring widths, with the minimax criterion defined in Eq. (10) for different values of the defocus range ψ_{\max} . Fig. 10 represents the normalized value of RCRB_{\max} obtained with different imaging strategies for discrete values of ψ_{\max} ranging from 0 to 1λ with a step of 0.1λ . The black line represents the normalized RCRB_{\max} obtained with a localization microscope with no mask and only limited by diffraction. It will serve as a baseline for comparison. The dotted blue curve is obtained with the optimized annular binary phase mask and the dash-dotted red one with the optimized IPP mask. Note that for each value of ψ_{\max} , the optimal mask may be different. Interestingly, it is first observed that the phase masks do not improve performance for small defocus parameters $\psi_{\max} < 0.4\lambda$, whatever the used strategy, since the three curves are superposed. In this case, DoF is too small for improving localization performance using masks. On the other hand, when the DoF range becomes larger, phase masks yield significant improvement, and this improvement is larger with annular binary phase masks than with IPP masks for any value of ψ_{\max} . For instance, when the defocus range is equal to $\psi_{\max} = 1\lambda$, the optimal annular binary phase mask yields a RCRB_{\max} three times smaller than that obtained without mask, while the optimal IPP mask reduces RCRB_{\max} only by a factor of two.

This result is an illustration of how the framework proposed in the article allows any type of DoF-enhancing strategy to be compared. Of course, localization accuracy may not be the only criterion for the choice of a mask in a given application. For example, manufacturability is also an important criterion. In that regard, annular binary phase masks may be easier to manufacture with photolithographic techniques since they only require one shallow etching level whereas the IPP masks necessitate deep tiers between each ring.

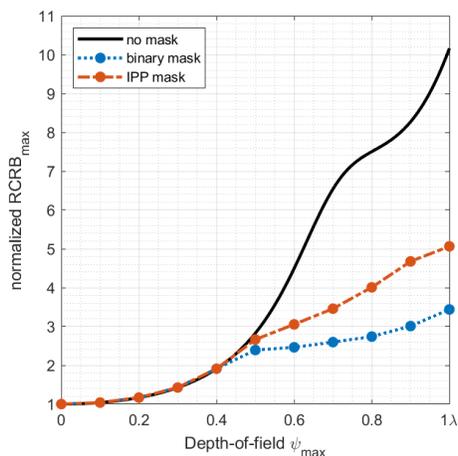


Fig. 10. Evolution of the normalized RCRB_{\max} as a function of the defocus range ψ_{\max} for different types of mask. The simulation parameters are defined in Table 1.

6. Conclusion

We have investigated the problem of DoF extension in the context of single-molecule localization microscopy. We have shown that placing an optimized phase mask in the exit pupil of the microscope and using an adapted processing algorithm allows to significantly increase the localization performance within the required defocus range. We have proposed different binary mask designs to enhance the DoF including a two-ring only solution that is easy to manufacture. Of course, there is a price to pay for DoF extension since the localization accuracy is always lower than for well-focused systems.

A strong asset of the framework developed in this article is to be based on an explicit and general information theoretical criterion. It thus makes it possible to compare any type of DoF extending masks on a realistic and quantitative basis. We have illustrated this potential by comparing optimized annular binary phase masks with another type of optimized phase mask proposed in the literature.

The present work is based on a simple imaging model, and in this sense, the orders of magnitude given in this article can be considered as upper limits on the DoF improvement that can be obtained in practice with phase masks. Thus, this method lays the basis on which more sophisticated and application-dependent strategies can be built. In particular, the imaging model could be improved in several ways, in terms of PSF modeling, optical aberrations or noise model. Another interesting perspective is to apply this framework to more general types of masks, such as continuous pure-phase masks or masks with combined amplitude and phase modulation, and to compare them on a realistic basis.

A. Comparison between CRB and MSE-based optimization criteria

In this Appendix, we compare the binary annular DoF-enhancing phase masks optimal for classical imaging systems and for single-molecule localization microscopy.

We have represented in Fig. 11 the optimal masks for DoF extension of $\psi_{\max} = 1\lambda$ in classical imaging [31] (MSE optimization, Fig. 11.a) and in localization microscopy (CRB optimization, Fig. 11.b). It is seen that these mask are very different. The question is whether they yield different localization performance.

To answer this question, we have plotted in Fig. 12.a and Fig. 12.b the x axis profile of the PSF of a single-molecule microscope as a function of defocus when using these two masks. It

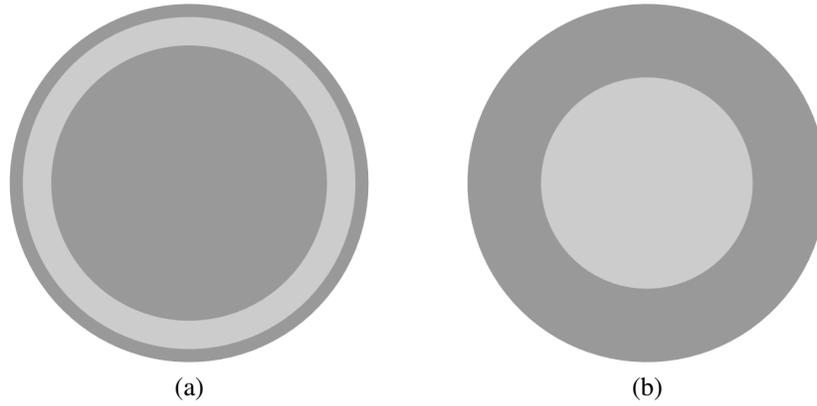


Fig. 11. Annular binary phase masks optimized for DoF extension $\psi_{\max} = 1\lambda$ in (a) classical imaging and in (b) localization microscopy. The mask parameters are respectively equal to $\rho = (0.7684, 0.9272)^t$ and $\rho = 0.59$. Each ring is defined as an annular region with constant phase modulation such as dark gray areas induce a phase of 0 radians and light gray areas induce a phase of π radians at a nominal wavelength λ .

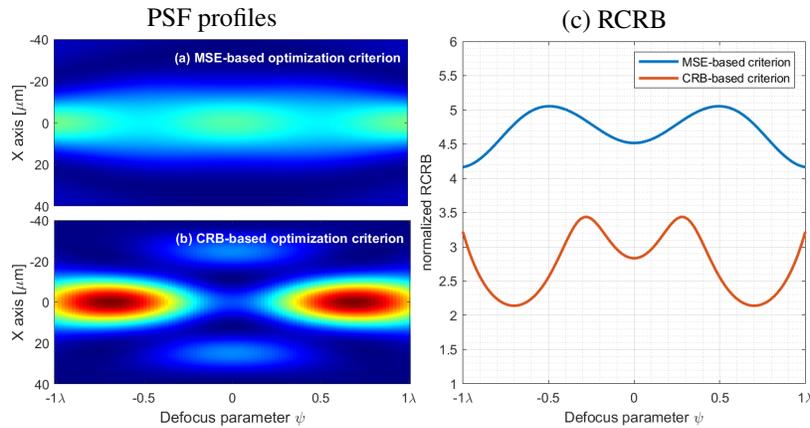


Fig. 12. x axis profile of the PSF for a co-designed optical system using binary phase masks optimized for (a) image quality (MSE criterion) or (b) localization accuracy (CRB criterion). (c) normalized limit of localisation accuracy, denoted normalized RCRB, along the x axis as a function of the defocus parameter ψ . The simulation parameters are defined in Table 1.

is observed that these profiles are quite different. The profile of the mask optimized with the MSE criterion (Fig. 12.a) is smooth since the quality of the deconvolved image has to be constant along the required DoF range. In contrast, the profile of the mask optimized with the CRB criterion (Fig. 12.b) varies more sharply. This is because, as explained in the section 3.2, the three-lobe profile around $\psi = 0$ and the “blob-like” profile around $\psi = 0.7\lambda$ are both appropriate for localization (they yield low values of the RCRB).

Finally, Fig. 12.c represents the localization performance of the two masks represented in Fig. 11, expressed in terms of normalized RCRB, as a function of the defocus ψ . It is seen that this localization performance differ significantly. We can conclude that the DoF enhancing masks optimal for classical imaging [31, 36] are not optimal for localization microscopy.

B. Closed-form expression of the variance $\text{var}_\psi[\hat{\theta}]$

In this Appendix, we show that it is possible, with some approximations, to get a closed-form expression of the variance $\text{var}_\psi[\hat{\theta}]$ of the estimator in Eq. (16) for a given value of ψ . For the sake of simplicity, we model the single-molecule localization problem as an one-dimensional problem. However, the method is also valid for 2D localization. To facilitate the mathematical developments, we assume that the observed signal $s(x)$ is continuous and modeled as

$$s(x) = N_0 f^{\psi, \theta_0}(x) + n(x) \quad (21)$$

where θ_0 is the position of the emitter in the object plane, N_0 is the total number of photo-electrons expected in the whole 1D image, $n(x)$ is a Gaussian zero-mean white noise of power spectral density $S_{nn}(\nu) = q$ and $f^{\psi, \theta_0}(x) = f^\psi(x - \theta_0)$ is proportional to the 1D spatial distribution of irradiance over the sensor for a given defocus parameter ψ .

We want to estimate the coordinate θ_0 of the emitter in the object plane by maximizing

$$\hat{\theta} = \arg \max_{\theta} \int_{\mathbb{R}} s(x) r(x - \theta) dx \quad (22)$$

where $r(x)$ is a defocus invariant kernel.

To quantify the performance of the estimator defined in Eq. (22), we can calculate its bias and variance. Using Eq. (21), the estimator has the following expression:

$$\hat{\theta} = \arg \max_{\theta} [N_0 \Omega(\theta) + n'(\theta)] \quad (23)$$

where $\Omega(\theta) = \int f^{\psi, \theta_0}(x) r(x - \theta) dx$ is a correlation product and $n'(\theta) = \int n(x) r(x - \theta) dx$ is a filtered noise. Let us consider the second order Taylor expansion of $\Omega(\theta)$ when θ is close to the true value θ_0 (i.e. when the variance is low)

$$\Omega(\theta) \simeq \Phi(\theta_0) + (\theta - \theta_0) \left. \frac{\partial \Omega(\theta)}{\partial \theta} \right|_{\theta=\theta_0} + \frac{(\theta - \theta_0)^2}{2} \left. \frac{\partial^2 \Omega(\theta)}{\partial \theta^2} \right|_{\theta=\theta_0}. \quad (24)$$

We suppose that the correlation function $\Omega(\theta)$ reaches its maximum when $\theta = \theta_0$. So using expressions (23) and (24), we can show that

$$\hat{\theta} = \theta_0 - \frac{\left. \frac{\partial n'(\theta)}{\partial \theta} \right|_{\theta=\theta_0}}{N_0 \left. \frac{\partial^2 \Omega(\theta)}{\partial \theta^2} \right|_{\theta=\theta_0}}. \quad (25)$$

The estimated position $\hat{\theta}$ of the emitter is therefore a random variable fluctuating around θ_0 . Using the theorem of derivation under the integral sign, we can write

$$\frac{\partial n'(\theta)}{\partial \theta} = \int_{\mathbb{R}} n(x) \frac{\partial r(x - \theta)}{\partial \theta} dx = - \int_{\mathbb{R}} n(x + \theta) \frac{\partial r(x)}{\partial x} dx. \quad (26)$$

This random process is zero mean, so the estimator $\hat{\theta}$ is unbiased. Moreover, the power spectral density of the filtered noise $n'(\theta)$ is by definition equal to $|2i\pi\nu\tilde{r}(\nu)|^2 S_{nn}(\nu)$ where $S_{nn}(\nu) = q$ and the superscript \sim denotes the Fourier transform, leading to

$$\text{var} \left[\left. \frac{\partial n'(\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right] = 4\pi^2 q \int_{\mathbb{R}} \nu^2 |\tilde{r}(\nu)|^2 d\nu. \quad (27)$$

On the other hand, we know that the Fourier transform of $\frac{\partial^2 \Omega(\theta)}{\partial \theta^2}$ is equal to $(2i\pi\nu)^2 [\tilde{f}^{\psi, \theta_0}(\nu)]^* \tilde{r}(\nu)$ where the superscript $*$ denotes the complex conjugate. By explicitly writing the inverse Fourier transform, we obtain

$$\frac{\partial^2 \Omega(\theta)}{\partial \theta^2} = -4\pi^2 \int_{\mathbb{R}} \nu^2 [\tilde{f}^{\psi, \theta_0}(\nu)]^* \tilde{r}(\nu) e^{2i\pi\nu\theta} d\nu \quad (28)$$

$$= -4\pi^2 \int_{\mathbb{R}} \nu^2 \tilde{f}^{\psi*}(\nu) \tilde{r}(\nu) e^{2i\pi\nu(\theta - \theta_0)} d\nu . \quad (29)$$

So taking into account the previous simplifications, we obtain the following closed-form expression of the variance $\text{var}_{\psi}[\hat{\theta}]$ for a given value of the defocus parameter ψ :

$$\text{var}_{\psi}[\hat{\theta}] = \frac{1}{4\pi^2} \frac{q}{N_0^2} \frac{\int_{\mathbb{R}} \nu^2 |\tilde{r}(\nu)|^2 d\nu}{\left(\int_{\mathbb{R}} \nu^2 \tilde{f}^{\psi*}(\nu) \tilde{r}(\nu) d\nu \right)^2} . \quad (30)$$

C. Invariant kernel based on variance minimization

In this Appendix, we determine the optimal defocus invariant kernel $r(x)$ such as the variance $\text{var}_{\psi}[\hat{\theta}]$ of the estimator defined in Eq. (22) is minimal. Considering a discrete set of values ψ_k , such as $k \in [1, K]$, the optimal defocus invariant kernel can be defined as the one which maximizes

$$\sum_{k=1}^K (\text{var}_{\psi_k}[\hat{\theta}])^{-1} = \frac{4\pi^2 N_0^2}{q} \frac{\sum_{k=1}^K \left(\int_{\mathbb{R}} \nu^2 \tilde{f}^{\psi_k*}(\nu) \tilde{r}(\nu) d\nu \right)^2}{\int_{\mathbb{R}} \nu^2 |\tilde{r}(\nu)|^2 d\nu} . \quad (31)$$

This criterion is reasonable. Of course, other criteria could have been considered, but one of its advantages is that its solution is closed-form. Indeed, optimization of this criterion can be performed by solving the following constrained optimization problem using the method of Lagrange multipliers: maximize function $\sum_k \left(\int \nu^2 \tilde{f}^{\psi_k*}(\nu) \tilde{r}(\nu) d\nu \right)^2$ subject to $\int \nu^2 |\tilde{r}(\nu)|^2 d\nu = A$, where A is a constant. The Lagrangian has the following expression:

$$\mathcal{L}(\tilde{r}) = \sum_{k=1}^K \left(\int_{\mathbb{R}} \nu^2 \tilde{f}^{\psi_k*}(\nu) \tilde{r}(\nu) d\nu \right)^2 - \lambda \left(\int_{\mathbb{R}} \nu^2 |\tilde{r}(\nu)|^2 d\nu - A \right) \quad (32)$$

with λ a Lagrange multiplier.

By annulling its functional derivative, it is easily shown that the Fourier transform of the optimal defocus invariant kernel is a linear combination of the functions $\tilde{f}^{\psi_k}(x)$ such as

$$\tilde{r}(\nu) = \sum_{k=1}^K \alpha_k \tilde{f}^{\psi_k}(\nu) \quad \text{with} \quad \alpha_k = \frac{\int_{\mathbb{R}} \nu^2 \tilde{f}^{\psi_k*}(\nu) \tilde{r}(\nu) d\nu}{\int_{\mathbb{R}} \nu^2 |\tilde{r}(\nu)|^2 d\nu} . \quad (33)$$

By substituting the expression (33) in Eq. (31), we can write $\sum_k (\text{var}_{\psi_k}[\hat{\theta}])^{-1}$ as a generalized Rayleigh quotient such as

$$\sum_{k=1}^K (\text{var}_{\psi_k}[\hat{\theta}])^{-1} = \frac{4\pi^2 N_0^2}{q} \frac{\alpha^t \mathbf{W} \mathbf{W} \alpha}{\alpha^t \mathbf{W} \alpha} \quad (34)$$

where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_K]$ is the vector of linear combination coefficients and \mathbf{W} is a matrix defined by $[\mathbf{W}]_{ij} = \int \nu^2 \tilde{f}^{\psi_i*}(\nu) \tilde{f}^{\psi_j}(\nu) d\nu$.

We can conclude that the optimal defocus invariant kernel that maximizes $\sum_k (\text{var}_{\psi_k}[\hat{\theta}])^{-1}$ is a linear combination of the functions $f^{\psi_k}(x)$ which coefficients are the components of the eigenvector associated with the greatest eigenvalue of the matrix \mathbf{W} .

Acknowledgements

This work was supported by grants from *Agence Nationale de la Recherche* (ANR-15-CE16-0004-03, ANR-18-CE09-0019-02), the *Fondation ARC pour la recherche sur le cancer* (A. L.) and IdEx Bordeaux (ANR-10-IDEX-03-02). This work has also received the support of the French research group GdR ISIS of the CNRS through the *Projet de Recherche Exploratoire* MASK.

Disclosures

The authors declare no conflicts of interest.

References

1. B. Huang, M. Bates, and X. Zhuang, "Super-Resolution Fluorescence Microscopy," *Annu. Rev. Biochem.* **78**, 993–1016 (2009). PMID: 19489737.
2. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science* **313**, 1642–1645 (2006).
3. S. T. Hess, T. P. Girirajan, and M. D. Mason, "Ultra-High Resolution Imaging by Fluorescence Photoactivation Localization Microscopy," *Biophys. J.* **91**, 4258 – 4272 (2006).
4. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**, 793–796 (2006).
5. A. Sharonov and R. M. Hochstrasser, "Wide-field subdiffraction imaging by accumulated binding of diffusing probes," *Proc. Natl. Acad. Sci.* **103**, 18911–18916 (2006).
6. N. Bobroff, "Position measurement with a resolution and noise-limited instrument," *Rev. Sci. Instruments* **57**, 1152–1157 (1986).
7. L. Cognet, C. Leduc, and B. Lounis, "Advances in live-cell single-particle tracking and dynamic super-resolution imaging," *Curr. Opin. Chem. Biol.* **20**, 78 – 85 (2014). Molecular imaging.
8. J. Kirstein, B. Platschek, C. Jung, R. Brown, T. Bein, and C. Bräuchle, "Exploration of nanostructured channel systems with single-molecule probes," *Nat. Mater.* **6**, 303–310 (2007).
9. D. Lasne, A. Maali, Y. Amarouchene, L. Cognet, B. Lounis, and H. Kellay, "Velocity Profiles of Water Flowing Past Solid Glass Surfaces Using Fluorescent Nanoparticles and Molecules as Velocity Probes," *Phys. Rev. Lett.* **100**, 214502 (2008).
10. A. G. Godin, J. A. Varela, Z. Gao, N. Danné, J. P. Dupuis, B. Lounis, L. Groc, and L. Cognet, "Single-nanotube tracking reveals the nanoscale organization of the extracellular space in the live brain," *Nat. Nanotechnol.* **12**, 238–243 (2017).
11. P. Kanchanawong, G. Shtengel, A. M. Pasapera, E. B. Ramko, M. W. Davidson, H. F. Hess, and C. M. Waterman, "Nanoscale architecture of integrin-based cell adhesions," *Nature* **468**, 580–584 (2010).
12. A. Dani, B. Huang, J. Bergan, C. Dulac, and X. Zhuang, "Superresolution Imaging of Chemical Synapses in the Brain," *Neuron* **68**, 843–856 (2010).
13. G. Giannone, E. Hosy, F. Levet, A. Constals, K. Schulze, A. I. Sobolevsky, M. P. Rosconi, E. Gouaux, R. Tampé, D. Choquet, and L. Cognet, "Dynamic Superresolution Imaging of Endogenous Proteins on Living Cells at Ultra-High Density," *Biophys. J.* **99**, 1303–1310 (2010).
14. A. Löschberger, S. van de Linde, M.-C. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, and M. Sauer, "Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution," *J. Cell Sci.* **125**, 570–575 (2012).
15. L. Cognet, D. A. Tsybouski, and R. B. Weisman, "Subdiffraction Far-Field Imaging of Luminescent Single-Walled Carbon Nanotubes," *Nano Lett.* **8**, 749–753 (2008).
16. J. Feng, H. Deschout, S. Caneva, S. Hofmann, I. Lončarić, P. Lazić, and A. Radenovic, "Imaging of Optically Active Defects with Nanometer Resolution," *Nano Lett.* **18**, 1739–1744 (2018).
17. N. Danné, M. Kim, A. G. Godin, H. Kwon, Z. Gao, X. Wu, N. F. Hartmann, S. K. Doorn, B. Lounis, Y. Wang, and L. Cognet, "Ultrashort Carbon Nanotubes That Fluoresce Brightly in the Near-Infrared," *ACS Nano* **12**, 6059–6065 (2018).
18. S. R. P. Pavani and R. Piestun, "Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system," *Opt. Express* **16**, 22048–22057 (2008).
19. B. Hajji, M. El Beheiry, I. Izeddin, X. Darzacq, and M. Dahan, "Accessing the third dimension in localization-based super-resolution microscopy," *Phys. Chem. Chem. Phys.* **16**, 16340–16348 (2014).
20. P. Bon, J. Linares-Loyez, M. Feyeux, K. Alessandri, B. Lounis, P. Nassoy, and L. Cognet, "Self-interference 3D super-resolution microscopy for deep tissue investigations," *Nat. Methods* **15**, 449–454 (2018).
21. F. Xu, D. Ma, K. P. MacPherson, S. Liu, Y. Bu, Y. Wang, Y. Tang, C. Bi, T. Kwok, A. A. Chubykin, P. Yin, S. Calve, G. E. Landreth, and F. Huang, "Three-dimensional nanoscopy of whole cells and tissues with in situ point spread function retrieval," *Nat. Methods* **17**, 531–540 (2020).

22. S. Abrahamsson, S. Usawa, and M. Gustafsson, "A new approach to extended focus for high-speed high-resolution biological microscopy," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIII*, vol. 6090 J.-A. Conchello, C. J. Cogswell, and T. Wilson, eds., International Society for Optics and Photonics (SPIE, 2006), pp. 128 – 135.
23. R. N. Zahreddine and C. J. Cogswell, "Total variation regularized deconvolution for extended depth of field microscopy," *Appl. Opt.* **54**, 2244–2254 (2015).
24. R. Lu, W. Sun, Y. Liang, A. Kerlin, J. Bierfeld, J. D. Seelig, D. E. Wilson, B. Scholl, B. Mohar, M. Tanimoto, M. Koyama, D. Fitzpatrick, M. B. Orger, and N. Ji, "Video-rate volumetric functional imaging of the brain at synaptic resolution," *Nat. Neurosci.* **20**, 620–628 (2017).
25. W. T. Cathey and E. R. Dowski, "New paradigm for imaging systems," *Appl. Opt.* **41**, 6080–6092 (2002).
26. M. D. Robinson and D. G. Stork, "Joint Design of Lens Systems and Digital Image Processing," in *International Optical Design*, (Optical Society of America, 2006), p. WB4.
27. A. R. Harvey, T. Vettenburg, M. Demenikov, B. Lucotte, G. Muyo, A. Wood, N. Bustin, A. Singh, and E. Findlay, "Digital image processing as an integral component of optical design," in *Novel Optical Systems Design and Optimization XI*, vol. 7061 R. J. Koshel, G. G. Gregory, J. D. M. Jr., and D. H. Krevor, eds., International Society for Optics and Photonics (SPIE, 2008), pp. 32 – 42.
28. F. Diaz, F. Goudail, B. Loiseaux, and J.-P. Huignard, "Increase in depth of field taking into account deconvolution by optimization of pupil mask," *Opt. Lett.* **34**, 2970–2972 (2009).
29. V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, "End-to-End Optimization of Optics and Image Processing for Achromatic Extended Depth of Field and Super-Resolution Imaging," *ACM Trans. Graph.* **37** (2018).
30. S. Elmalem, R. Giryas, and E. Marom, "Learned phase coded aperture for the benefit of depth of field extension," *Opt. Express* **26**, 15316–15331 (2018).
31. R. Falcón, F. Goudail, C. Kulcsár, and H. Sauer, "Performance limits of binary annular phase masks codesigned for depth-of-field extension," *Opt. Eng.* **56**, 1 – 12 (2017).
32. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*, Prentice Hall signal processing series (Prentice-Hall PTR, 1993).
33. R. J. Ober, S. Ram, and E. S. Ward, "Localization Accuracy in Single-Molecule Microscopy," *Biophys. J.* **86**, 1185 – 1200 (2004).
34. A. von Diezmann, Y. Shechtman, and W. E. Moerner, "Three-Dimensional Localization of Single Molecules for Super-Resolution Imaging and Single-Particle Tracking," *Chem. Rev.* **117**, 7244–7275 (2017).
35. M. Badieirostami, M. D. Lew, M. A. Thompson, and W. E. Moerner, "Three-dimensional localization precision of the double-helix point spread function versus astigmatism and biplane," *Appl. Phys. Lett.* **97**, 161103 (2010).
36. A. Fontbonne, H. Sauer, C. Kulcsár, A.-L. Coutrot, and F. Goudail, "Experimental validation of hybrid optical-digital imaging system for extended depth-of-field based on co-optimized binary phase masks," *Opt. Eng.* **58**, 1 – 12 (2019).
37. F. Aguet, "Super-Resolution Fluorescence Microscopy Based on Physical Models," Theses, École polytechnique fédérale de Lausanne (2009).
38. M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).
39. H. H. Hopkins and C. R. Burch, "The frequency response of a defocused optical system," *Proc. Royal Soc. London. Ser. A. Math. Phys. Sci.* **231**, 91–103 (1955).
40. L. Tao and C. Nicholson, "The three-dimensional point spread functions of a microscope objective in image and object space," *J. Microsc.* **178**, 267–271 (1995).
41. C. J. R. Sheppard and P. Török, "An electromagnetic theory of imaging in fluorescence microscopy, and imaging in polarization fluorescence microscopy," *Bioimaging* **5**, 205–218 (1997).
42. E. R. Dowski and W. T. Cathey, "Extended depth of field through wave-front coding," *Appl. Opt.* **34**, 1859–1866 (1995).
43. S. S. Sherif, W. T. Cathey, and E. R. Dowski, "Phase plate to extend the depth of field of incoherent hybrid imaging systems," *Appl. Opt.* **43**, 2709–2721 (2004).
44. A. Saucedo and J. O.-C. neda, "High focal depth with fractional-power wave fronts," *Opt. Lett.* **29**, 560–562 (2004).
45. Q. Yang, L. Liu, and J. Sun, "Optimized phase pupil masks for extended depth of field," *Opt. Commun.* **272**, 56 – 66 (2007).
46. N. Caron and Y. Sheng, "Polynomial phase masks for extending the depth of field of a microscope," *Appl. Opt.* **47**, E39–E43 (2008).
47. F. Zhou, G. Li, H. Zhang, and D. Wang, "Rational phase mask to extend the depth of field in optical-digital hybrid imaging systems," *Opt. Lett.* **34**, 380–382 (2009).
48. Z. Zalevsky, A. Shemer, A. Zlotnik, E. B. Eliezer, and E. Marom, "All-optical axial super resolving imaging using a low-frequency binary-phase mask," *Opt. Express* **14**, 2631–2643 (2006).
49. B. Milgrom, N. Konforti, M. A. Golub, and E. Marom, "Pupil coding masks for imaging polychromatic scenes with high resolution and extended depth of field," *Opt. Express* **18**, 15569–15584 (2010).
50. F. Diaz, M.-S. L. Lee, X. Rejeaunier, G. Lehoucq, F. Goudail, B. Loiseaux, S. Bansropun, J. Rollin, E. Debes, and P. Mills, "Real-time increase in depth of field of an uncooled thermal camera using several phase-mask technologies," *Opt. Lett.* **36**, 418–420 (2011).

51. M.-A. Burcklen, F. Diaz, F. Leprêtre, J. Rollin, A. Delboulbé, M.-S. L. Lee, B. Loiseaux, A. Koudoli, S. Denel, P. Millet, F. Duhem, F. Lemonnier, H. Sauer, and F. Goudail, "Experimental demonstration of extended depth-of-field f/1.2 visible High Definition camera with jointly optimized phase mask and real-time digital processing," *Journal of the European Optical Society : Rapid publications* **10**, 15046 (2015).
52. J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4 (1995), pp. 1942–1948 vol.4.