

Weak convexity of Fisher information matrix and superresolved localization of blinking sources of light

Dmitri B. Horoshko^{1,*}, Alexander B. Mikhalychev², Fedor Jelezko¹ and Polina P. Kuzhir³

¹*Institut für Quantenoptik, Universität Ulm, D-89073 Ulm, Germany*

²*Atomicus GmbH, Amalienbadstraße 41C, 76227 Karlsruhe, Germany*

³*University of Eastern Finland, Yliopistokatu 7, 80101 Joensuu, Finland*



(Received 4 February 2025; accepted 11 June 2025; published 17 July 2025)

A group of techniques known by the general name of single-molecule localization microscopy reaches a nanometer-scale spatial resolution of point light emitters, well below the diffraction limit of traditional microscopy. The key feature of these techniques is blinking, alternation of bright and dark states, of each emitter so that no more than one emitter is bright within the width of the point-spread function of the microscope during a time sufficient for its localization. We give a formulation of the optical part of these techniques in terms of quantum metrology, where the limit of precision is determined by the Fisher information on the emitters' positions contained in the measurement data. We show that the advantage in resolution provided by making the emitters blink is a consequence of the fundamental property of Fisher information, its convexity. In particular, we prove the weak matrix convexity and the trace convexity of the Fisher information matrix—two fundamental results in the multiparameter estimation theory. We show also that the advantage in information is inherent to the quantum state of light itself before the measurement and is related to the convexity of the quantum Fisher information matrix.

DOI: [10.1103/t48h-6hcf](https://doi.org/10.1103/t48h-6hcf)

I. INTRODUCTION

The lateral resolution of a conventional far-field microscope is fundamentally restricted by the diffraction limit, equal to about half the wavelength of the illuminating light [1], which is above 200 nm in the visible range. In the last 30 years, numerous methods for surpassing the diffraction limit have been developed and are known by the general name of superresolution microscopy [2–4]. One of the most successful directions in this area of research is related to subdiffraction imaging of pointlike incoherent sources of light represented by fluorescent molecules [5–7], achieving nanometer-scale lateral resolution and distinguished by the Nobel Prize in Chemistry in 2014 "for the development of superresolved fluorescence microscopy." An important subdirection is represented by the techniques of stochastically activated (blinking) fluorophores such as photoactivated localization microscopy (PALM) [8], stochastic optical reconstruction microscopy (STORM) [9], points accumulation for imaging in nanoscale topography (PAINT) [10], and their variants, known by the general name of single-molecule localization microscopy (SMLM) and having found numerous applications in cell biology [6]. The SMLM techniques have also been applied

to color centers in diamond, providing subdiffraction-limit sensing of magnetic or electric fields [11,12].

The general principle of SMLM consists in making the fluorophores blink, by a mechanism specific to each technique, so that not more than one source is present within the width of the point-spread function (PSF) of the microscope during a time sufficient for its precise localization. The latter is achieved by a simple arithmetic mean of the photodetection positions or by taking an analytical model for the PSF and estimating the position of its peak by least squares or maximum-likelihood estimation [13]. The standard deviation of the source position scales as $1/\sqrt{N}$ with the number N of detected photons [14], which allows one to localize isolated fluorophores with an uncertainty of about 1 nm by detecting $N \approx 10^4$ photons for each of them by a technique known as fluorescence imaging with one-nanometer accuracy (FIONA) [15]. Similarly, the positions of multiple sources within the PSF can be found by modeling the intensity distribution on the microscope camera by a sum of PSFs and estimating the positions of its peaks, albeit with a much larger uncertainty [16,17].

The $1/\sqrt{N}$ scaling of the position standard deviation is easily understood in the framework of mathematical statistics, namely, its branch known as point estimation theory [18,19]. The central relation of this theory is the Cramér-Rao bound for the variance of any unbiased estimator of the parameter x :

$$(\Delta x)^2 \geq 1/\underline{F}, \quad (1)$$

where $\underline{F}$ is the Fisher information on this parameter contained in the measurement data. Here and below, underlined letters denote quantities related to the entire measurement

*Contact author: dmitri.horoshko@uni-ulm.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

session typically consisting of many photodetections. In the case of localization of a single emitter, the subsequent photodetections are statistically independent, and the total Fisher information carried by N photons is simply $\underline{F} = NF_{\text{1phot}}$, where F_{1phot} is the Fisher information on the emitter position carried by one photon. The Cramér-Rao bound can be asymptotically reached in the high- N limit when the estimator is given by the maximum-likelihood estimation [18,19]. Thus, the $1/\sqrt{N}$ scaling of the position standard deviation is a mere consequence of the *additivity* of Fisher information with respect to independent observations. The single-photon Fisher information can be calculated for various models of the PSF and the accompanying noise [20,21] providing the Cramér-Rao bound for any given N . The variance of the estimated emitter position was shown to approach the Cramér-Rao bound for $N = 10^2$ – 10^4 photons, allowing thus for superresolution.

The same argument can be applied to determining the separation of two closely spaced emitters: the total Fisher information carried by N photons on the separation is $\underline{F} = NF_{\text{sep,1phot}}$, where $F_{\text{sep,1phot}}$ is the information on the separation carried by one photon. This information tends to zero together with the separation of the emitters [22]; however, for whatever small $F_{\text{sep,1phot}}$, one can make the total information arbitrarily large by collecting a sufficiently large number of photons and thus determine the separation with arbitrary precision. In practice, a fluorophore emits a limited number of photons before photobleaching, and, therefore, individual localization of blinking emitters gives much better results than a simultaneous localization of multiple cofluorescent emitters within the width of PSF. The precision gain attained by making the sources blink is the key feature of SMLM. A question arises: Is it a general law of Nature that blinking emitters send more information on their locations than cofluorescent ones?

In this paper, we give a positive answer to this question, showing that the informational advantage gained by blinking is a consequence of another fundamental property of Fisher information, its *convexity*. We consider the task of localization of an arbitrary number of light sources as a multiparameter estimation problem and prove that, in a scenario where all sources emit the same number of photons and microscopy satisfies the conditions of translation and rotation invariance, the Fisher information matrix possesses the property of *weak matrix convexity*, which results in a higher localization efficiency for blinking sources. We also prove that, in the general case of arbitrary numbers of emitted photons, the Fisher information matrix possesses the property of *trace convexity*, which leads to a higher localization efficiency of blinking sources when it is measured in terms of the new measure, which we call “average eigenparameter precision.”

The paper is structured as follows. In Sec. II we recast the formalism of quantum metrology [23–26] in the context of localization of dim incoherent sources of light [22,27–30], extending a more traditional approach of classical metrology typically used in this context [20,21,31–34], and introduce three measures of localization efficiency. In Sec. III, we prove the fundamental properties of the Fisher information matrix mentioned above. In Sec. IV, we compare two scenarios of cofluorescent and blinking sources and show that the latter always has an informational advantage. We illustrate the

developed theory by an example of two incoherent sources in Sec. V and extend the results to the quantum Fisher information matrix in Sec. VI. The obtained results are summarized in Sec. VII.

II. QUANTUM METROLOGY FORMALISM FOR LOCALIZATION OF POINT SOURCES OF LIGHT

A. Cramér-Rao bound

The objective of quantum metrology is inference of an unknown parameter θ , which may be a vector $\theta = \{\theta_1, \dots, \theta_M\}$, encoded into the state of the object, $\rho_O(\theta)$, which is a density operator in the Hilbert space of the object. For this purpose a generalized quantum measurement of the object is repeated N times, the object being prepared in the same state $\rho_O(\theta)$ for every trial, so that the state of all N object samples is $\underline{\rho}_O(\theta) = \rho_O(\theta)^{\otimes N}$. In the task of localizing point sources, considered here, θ includes the spatial coordinates $\bar{r}_k = (\bar{x}_k, \bar{y}_k)$ of each source on the object plane. In this task, the entire fluorescence session time (≈ 1 ms) is split into time bins of short duration (≈ 1 ns) so that not more than one photon is detected in each time bin. The empty time bins are discarded, while the time bins with detections are numbered with $n = 1, \dots, N$, where N is the total number of photons detected. The point r_n on the image plane where a photon was detected in the n th bin is considered a result of a measurement of the n th sample of the field. That is, it is implied that the sources’ coordinates do not vary with time and the subsequent acts of spontaneous emission are statistically independent. Note that the exact times of photon detections may not be resolved by the camera of the microscope and are not necessary for the statistical model. Even if the probability of photon detection in one time bin is as small as 0.01, the total number of detected photons can be rather high, typically $N \approx 10^4$. Specifically, $\rho_O(\theta)$ can be understood as the density operator of the electronic excitation in a time bin where a photon was detected. Alternatively, it can be understood as the density operator of a spontaneously emitted photon on the object plane.

A generalized quantum measurement consists of measuring a probe system upon its interaction with the object and comprises three main stages: initialization of the probe, interaction between the probe and the object, and measurement of the probe [26].

In the first stage, the probe is prepared in a state $\underline{\rho}_P$, which is, in the simplest case of independent identical systems, just a direct product of the states of each sample of the probe ρ_P : $\underline{\rho}_P = \rho_P^{\otimes N}$. However, more complicated entangled probes can be used for reaching the quantum-enhanced sensitivity [23], which we do not consider here. In optical microscopy, the probe is represented by the part of the optical field captured by the aperture of the microscope (see Fig. 1). For the spontaneous emission of fluorophores, its initial state is vacuum. Thus, its preparation simply means screening the field of view of the microscope from the background light in the frequency band used for imaging.

In the second stage, the probe and object undergo an interaction, which is described in the most general form by a completely positive trace-preserving linear map $\mathcal{E}$, so that the state of all N samples of the object and probe after

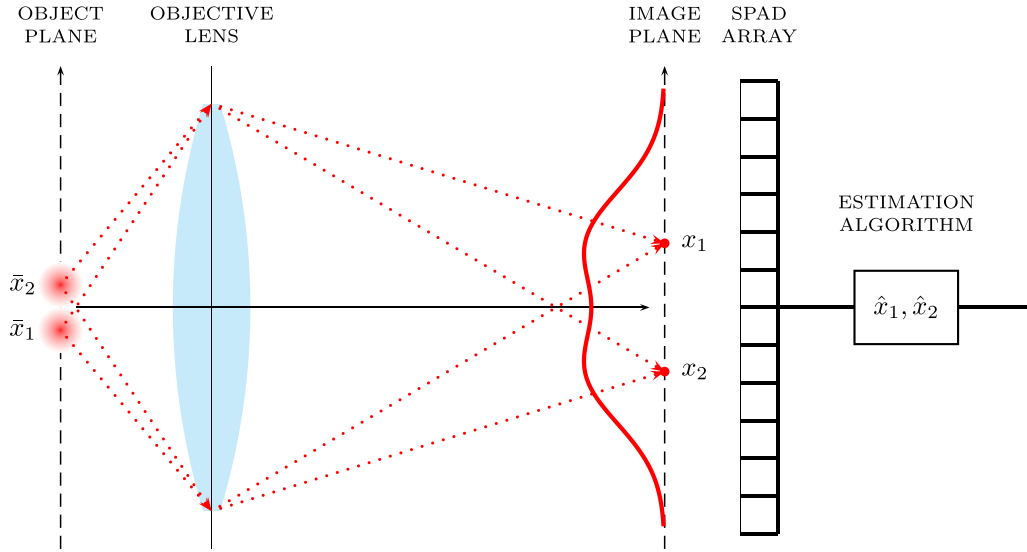


FIG. 1. Scheme for measuring the positions of two incoherent sources by imaging them in a microscope equipped with a single-photon avalanche diode (SPAD) array and estimating the positions from the measurement data. The dotted red lines show the rays that form the geometrical images at points x_1 and x_2 , while the solid red line shows the field intensity distribution in the image plane as a sum of two point-spread functions centered at x_1 and x_2 .

the interaction is $\rho_{OP}(\theta) = \mathcal{E}[\rho_O(\theta) \otimes \rho_P]$. This map can be made unitary on an extended state space by a Stinespring dilation, that is, by adding up an environment system with the initial state ρ_E and tracing out the environment after the interaction: $\rho_{OP}(\theta) = \text{Tr}_E\{U\rho_O(\theta) \otimes \rho_P \otimes \rho_E U^\dagger\}$, where U is the unitary operator describing the interaction of the object, probe, and environment and $\dagger$ stands for Hermitian conjugation. In the simplest case of independent identical interactions, both the state of the environment and the interaction operator are simple direct products of one-sample operators ρ_E and U_θ : $\rho_E = \rho_E^{\otimes N}$ and $U = U^{\otimes N}$. In optical microscopy of fluorophores, the environment is represented by the part of the optical field not captured by the aperture of the microscope and the degrees of freedom of the fluorophore other than the optical transition responsible for the spontaneous emission. The interaction consists in optical excitation of a transition belonging to the environment, from where the excitation is transmitted to the object, which spontaneously emits a photon, belonging to the probe or to the environment.

In the third stage, the probe is measured giving the outcome $\underline{r}$, which is a combination of measurement outcomes r_n for every sample of the probe: $\underline{r} = \{r_1, \dots, r_N\}$. The probability of this outcome is given by the Born rule $p(\underline{r}|\theta) = \text{Tr}_P\{\underline{\Lambda}(\underline{r})\tilde{\rho}_P(\theta)\}$, where $\tilde{\rho}_P(\theta) = \text{Tr}_O\{\rho_{OP}(\theta)\}$ is the state of the probe with the object traced out, and $\underline{\Lambda}(\underline{r})$ is the positive operator-valued measure (POVM) for the outcome $\underline{r}$. As before, for independent identical measurements, the POVM is a direct product of POVMs for every sample: $\underline{\Lambda}(\underline{r}) = \otimes_{n=1}^N \Lambda(r_n)$. Collective measurements of probe samples are studied in some scenarios of quantum metrology [23], but are not considered here. When the magnification of the microscope is sufficiently high, its field of view is covered by a large enough number of pixels, so that the measurement outcome can be represented by the continuous coordinate $r_n = \{x_n, y_n\}$ of the detected photon. The effects of detector pixelation are studied in the literature [20], but are not considered here.

The corresponding single-sample POVM is the projection on the single-photon state at point r_n : $\Lambda(r_n) = |r_n\rangle\langle r_n|$, where $|r_n\rangle = a^\dagger(r_n)|\text{vac}\rangle$, $a^\dagger(r_n)$ being the photon creation operator at point r_n and $|\text{vac}\rangle$ the vacuum state of the field.

Summing up the above relations, we write the conditional probability of the measurement outcome $\underline{r}$ under the condition that the parameter is θ as

$$p(\underline{r}|\theta) = \text{Tr}\{\underline{\Lambda}(\underline{r})U\rho_O(\theta) \otimes \rho_P \otimes \rho_E U^\dagger\}, \quad (2)$$

where the trace is taken over all N samples of the three systems: object, probe, and environment. The value of the parameter θ is inferred from the random outcome $\underline{r}$ by a function $\hat{\theta}(\underline{r})$ known as the estimator, which is again a random variable. The estimator is called unbiased if its mean coincides with the true value of the parameter, $\langle \hat{\theta}(\underline{r}) \rangle = \theta$, where the angular brackets denote the averaging with the conditional probability, Eq. (2). For any unbiased estimator, its covariance matrix $\text{Cov}(\hat{\theta})_{jk} = \langle (\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k) \rangle$ is lower bounded by the Cramér-Rao bound discussed in the Introduction for a single parameter and having a matrix form in the multiparameter case [18,19]:

$$\text{Cov}(\hat{\theta}) \geq \mathbf{F}^{-1}(\theta), \quad (3)$$

where the matrix inequality $\mathbf{A} \geq \mathbf{B}$ for two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$ means that the matrix $\mathbf{A} - \mathbf{B}$ is positive semidefinite, i.e., all its eigenvalues are non-negative [35], and $\mathbf{F}(\theta)$ is the Fisher information matrix, whose elements are defined as

$$F_{ij}(\theta) = \int \frac{1}{p(\underline{r}|\theta)} \frac{\partial p(\underline{r}|\theta)}{\partial \theta_i} \frac{\partial p(\underline{r}|\theta)}{\partial \theta_j} d\underline{r}. \quad (4)$$

Here and in the following, we assume that $p(\underline{r}|\theta)$ is nonzero everywhere, which is true for an imaging system whose point-spread function is modeled by a Gaussian. Note that we use bold font for matrices, italic for their elements, Roman for the

elements of function matrices, and underline, as above, for the quantities related to the entire set of N samples.

In optical microscopy, the typical estimator for the location of a single emitter is the result of fitting the distribution of the field intensity on its image spot by a two-dimensional Gaussian function [9,12]. In the case of multiple emitters, the Cramér-Rao bound is saturated by finding the parameter $\hat{\theta}(\underline{r})$ which maximizes $\underline{p}(\underline{r}|\theta)$ considered as a function of θ and known in this aspect as likelihood, which gives the name of “maximum likelihood” to the estimator [18,19].

B. Localization efficiency

The diagonal elements of the matrix $\mathbf{Cov}(\hat{\theta})$ give variances of the components of θ , $\text{Cov}(\hat{\theta})_{ii} = \text{Var}(\hat{\theta}_i)$, and traditionally the entire diagonal of the covariance matrix is considered the indicator of success in determining the unknown parameters. However, when tens or hundreds of sources are available, one might be interested in one number characterizing the efficiency of mass localization. We prefer to express this “localization efficiency” through the inverse of the parameter variance known as precision [28,36], so that a high efficiency corresponds to a low position variance. This localization efficiency can be introduced in several alternative ways, depending on the order in which the averaging and inversion are performed.

Definition 1. For any unbiased estimator $\hat{\theta}$ of an M -dimensional parameter θ , the average total precision is

$$H_{\text{tot}}(\hat{\theta}) = \frac{M}{\text{Tr}\{\mathbf{Cov}(\hat{\theta})\}}. \quad (5)$$

This quantity is the inverse of the average variance and, according to Eq. (3), satisfies the inequality

$$H_{\text{tot}}(\hat{\theta}) \leq \frac{M}{\text{Tr}\{\underline{\mathbf{F}}^{-1}(\theta)\}}, \quad (6)$$

which is asymptotically saturated by the maximum-likelihood estimator.

However, the average variance can be odd suited to the description of the case where a localization of multiple fluorophores is required. Imagine a situation where one of the 100 fluorophores does not emit light for some reason. In this case, the probability distribution does not depend on the position of this fluorophore, and therefore all rows and columns of the Fisher information matrix corresponding to its coordinates are zeros. The variance of these coordinates is infinite and the average variance is infinite as well, giving $H_{\text{tot}}(\hat{\theta}) = 0$, which means failure of the estimation procedure. On the other hand, absence of one of 100 fluorophores is not important for the task of visualization of a biological object, for which the fluorophores are intended. Thus, we arrive at another measure of the localization efficiency where the individual precisions are averaged rather than the individual variances.

Definition 2. For any unbiased estimator $\hat{\theta}$ of an M -dimensional parameter θ , the average individual precision is

$$H_{\text{ind}}(\hat{\theta}) = \frac{1}{M} \sum_{i=1}^M \frac{1}{\text{Cov}(\hat{\theta})_{ii}}. \quad (7)$$

In the considered case of one dark fluorophore, its individual precision $1/\text{Cov}(\hat{\theta})_{ii}$ is zero and does not contribute to the average individual precision defined by Eq. (7). The upper bound for the average individual precision is set according to Eq. (3), by the inequality $H_{\text{ind}}(\hat{\theta}) \leq \bar{F}_{\text{ind}}(\theta)$, where

$$\bar{F}_{\text{ind}}(\theta) = \frac{1}{M} \sum_{i=1}^M \frac{1}{[\underline{\mathbf{F}}^{-1}(\theta)]_{ii}} \quad (8)$$

is the average individual Fisher information. This bound is asymptotically saturated by the maximum-likelihood estimator.

Both $H_{\text{ind}}(\hat{\theta})$ and $\bar{F}_{\text{ind}}(\theta)$ share the drawback of being non-invariant with respect to rotations in the parameter space. Aiming at an invariant measure, we introduce one more measure of localization efficiency. First, we define the precision matrix [36] $\mathbf{Prec}(\hat{\theta}) = \mathbf{Cov}(\hat{\theta})^{-1}$, whose diagonal entry $\text{Prec}(\hat{\theta})_{ii}$ is larger than or equal to the individual precision $1/\text{Cov}(\hat{\theta})_{ii}$. If we consider an orthogonal rotation in the parameter space, $\varphi = \mathbf{O}^T \theta$, where $\varphi = \{\varphi_1, \dots, \varphi_M\}$, both parameter vectors are understood as column vectors, T denotes a transposition, and $\mathbf{O}$ is an orthogonal matrix such that $\mathbf{Cov}(\hat{\varphi}) = \mathbf{O}^T \mathbf{Cov}(\hat{\theta}) \mathbf{O}$ is diagonal with $\hat{\varphi} = \mathbf{O}^T \hat{\theta}$ an unbiased estimator for φ , then $\text{Prec}(\hat{\varphi})_{ii} = 1/\text{Cov}(\hat{\varphi})_{ii}$. We will call φ the eigenparameter of the localization task. Now, we define the third measure of localization efficiency as follows.

Definition 3. For any unbiased estimator $\hat{\theta}$ of an M -dimensional parameter θ , the average eigenparameter precision is

$$H_{\text{eig}}(\hat{\theta}) = \frac{1}{M} \text{Tr}\{\mathbf{Prec}(\hat{\theta})\}. \quad (9)$$

The upper bound for this quantity can be found from Eq. (3). For any two invertible Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ if and only if $\mathbf{A}^{-1} \leq \mathbf{B}^{-1}$ [35]. Note that the covariance, precision, and Fisher information matrices are real symmetric by definition, and as such are Hermitian. Therefore, Eq. (3) can be rewritten as $\mathbf{Prec}(\hat{\theta}) \leq \underline{\mathbf{F}}(\theta)$. Taking a trace of both sides of this inequality, we obtain $H_{\text{eig}}(\hat{\theta}) \leq \bar{F}(\theta)$, where

$$\bar{F}(\theta) = \frac{1}{M} \text{Tr}\{\underline{\mathbf{F}}(\theta)\} \quad (10)$$

is the average Fisher information. This bound is asymptotically saturated by the maximum-likelihood estimator. The matrix $\mathbf{O}$ diagonalizing the covariance matrix coincides asymptotically with that diagonalizing the Fisher information matrix. However, the matrix $\mathbf{O}$ is itself dependent on the unknown parameters and is typically unknown when the localization procedure starts.

Two commentaries on the introduced measures are necessary. It is typical for multiparameter quantum metrology to introduce a weight matrix assigning different statistical weights to different parameters and to consider the trace of the product of the covariance matrix with the weight matrix as a measure of the multiparameter estimation success [25,26]. We do not see why, in the task of multiple source localization, one source might have preference before another and therefore assign equal weights to the coordinates of all sources, which leads us to a simple summation in all three definitions.

It should also be understood that the measures given by Definitions 2 and 3 are not intended for use in the fitting algorithms. The quantities $H_{\text{ind}}(\hat{\theta})$ and $H_{\text{eig}}(\hat{\theta})$ may be very large in the case where only one source is localized with very high precision but the positions of all other sources are totally unknown, which is not the best localization outcome. We assume that the fitting is realized by the maximum-likelihood estimation and the measure of localization efficiency is applied afterward to the results of such estimation. The evaluation of efficiency here is similar to that in quantum teleportation [37,38]: the verifier knowing the exact positions of the sources gives a sample to the microscopist, who returns the estimated source positions, for which the verifier calculates the measure of efficiency, similar in this respect to the teleportation fidelity.

The three introduced measures are ordered as follows:

$$H_{\text{tot}}(\hat{\theta}) \leq H_{\text{ind}}(\hat{\theta}) \leq H_{\text{eig}}(\hat{\theta}). \quad (11)$$

The first inequality follows from the convexity of the function $f(x) = 1/x$ on the interval $x \in (0, +\infty)$, resulting in

$$f\left(\frac{1}{M} \sum_{i=1}^M x_i\right) \leq \frac{1}{M} \sum_{i=1}^M f(x_i). \quad (12)$$

Substituting $x_i = \mathbf{Cov}(\hat{\theta})_{ii}$, we obtain the first inequality. To prove the second inequality, consider an invertible real symmetric $M \times M$ matrix $\mathbf{A}$ and a unit $M \times 1$ vector v . Let $\mathbf{O}$ be an orthogonal matrix that brings $\mathbf{A}$ to the diagonal form, $\mathbf{A} = \mathbf{O}\mathbf{D}\mathbf{O}^T$, where $\mathbf{D}$ is diagonal. Then $v^T \mathbf{A} v = \tilde{v}^T \mathbf{D} \tilde{v}$, where $\tilde{v} = \mathbf{O}^T v$. Since $\tilde{v}$ is also a unit vector, the numbers $\tilde{v}_i^2$ sum up to unity and can be considered as statistical weights. From the convexity of the function $f(x) = 1/x$ on the interval $x \in (0, +\infty)$, as above, we obtain

$$f\left(\sum_{i=1}^M \tilde{v}_i^2 D_{ii}\right) \leq \sum_{i=1}^M \tilde{v}_i^2 f(D_{ii}), \quad (13)$$

or $(\tilde{v}^T \mathbf{D} \tilde{v})^{-1} \leq \tilde{v}^T \mathbf{D}^{-1} \tilde{v}$ and, as a consequence, $(v^T \mathbf{A} v)^{-1} \leq v^T \mathbf{A}^{-1} v$. Choosing a vector $v_i = \delta_{im}$ for some $1 \leq m \leq M$, we obtain the inequality $1/A_{mm} \leq (\mathbf{A}^{-1})_{mm}$. Substituting $A_{mm} = \text{Cov}(\hat{\theta})_{mm}$ and summing up over m , we obtain the second inequality of Eq. (11).

C. Quantum Cramér-Rao bound

For a given state of the object, the conditional probabilities, Eq. (2), and therefore the parameter variances calculated from them, depend on four quantities: ρ_p , ρ_E , $\underline{U}_\theta$, and $\underline{\Lambda}(r)$. Each of these quantities can be manipulated to optimize the estimation. In particular, optimization of $\underline{\Lambda}(r)$ corresponds to the choice of the best measurement for the given state of the probe, $\tilde{\rho}_p(\theta)$. For this purpose, the quantum Fisher information matrix $\underline{\mathbf{Q}}(\theta)$ is defined as [39–41]

$$\underline{\mathbf{Q}}_{ij}(\theta) = \frac{1}{2} \text{Tr}\{(\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i) \tilde{\rho}_p(\theta)\}, \quad (14)$$

where $\mathcal{L}_i$ is the symmetric logarithmic derivative operator defined by the equation

$$\frac{\partial}{\partial \theta_i} \tilde{\rho}_p(\theta) = \frac{1}{2} [\mathcal{L}_i \tilde{\rho}_p(\theta) + \tilde{\rho}_p(\theta) \mathcal{L}_i]. \quad (15)$$

The Quantum Fisher information matrix satisfies the condition $\underline{\mathbf{F}}^{-1}(\theta) \geq \underline{\mathbf{Q}}^{-1}(\theta)$, where $\underline{\mathbf{F}}^{-1}(\theta)$ is calculated for the same state of the probe with an arbitrary POVM $\underline{\Lambda}(r)$. Moreover, there is often an optimal POVM $\underline{\Lambda}_{\text{opt}}(r)$ (but not always [25]) for which $\underline{\mathbf{F}}_{\text{opt}}^{-1}(\theta) = \underline{\mathbf{Q}}^{-1}(\theta)$. This means that the quantity $\text{Tr}\{\underline{\mathbf{Q}}^{-1}(\theta)\}/M$ gives the lower bound of the average variance for the optimized measurement. As mentioned above, manipulation of ρ_p by entangling the probes is known to further minimize the estimation error in some tasks of quantum metrology [23]. In this paper, we show that the super-resolution obtained by making the sources blink corresponds to manipulation of $\rho_O(\theta)$ without affecting the value of the parameter θ .

III. ADDITIVITY AND WEAK CONVEXITY OF FISHER INFORMATION MATRIX

The Fisher information matrix possesses some remarkable properties, which are summarized below. First, we reformulate its standard definition [18,19] in terms of a functional of the probability density.

Definition 4. Given two positive integers ℓ and M and a random variable $X \in \mathbb{R}^\ell$ depending on parameter $\theta \in \mathbb{R}^M$, so that the probability density for X under condition that the parameter is θ is $q(X|\theta)$, the Fisher information matrix for the information contained in X about θ is the $(M \times M)$ -matrix-valued functional $\mathcal{F}_M^{(\ell)}[q(X|\theta)]$, whose matrix elements are

$$\mathcal{F}_M^{(\ell)}[q(X|\theta)]_{ij} = \int_{\mathbb{R}^\ell} \frac{1}{q(X|\theta)} \frac{\partial q(X|\theta)}{\partial \theta_i} \frac{\partial q(X|\theta)}{\partial \theta_j} d^\ell X. \quad (16)$$

Now, we formulate the property of additivity for this functional.

Lemma 1. The Fisher information matrix possesses the property of additivity:

$$\begin{aligned} \mathcal{F}_M^{(\ell_1+\ell_2)}[q_1(X_1|\theta)q_2(X_2|\theta)] \\ = \mathcal{F}_M^{(\ell_1)}[q_1(X_1|\theta)] + \mathcal{F}_M^{(\ell_2)}[q_2(X_2|\theta)], \end{aligned} \quad (17)$$

where $X_1 \in \mathbb{R}^{\ell_1}$ and $X_2 \in \mathbb{R}^{\ell_2}$.

The proof can be found in the literature [19]. Next, we formulate the property of convexity in the single-parameter case proven by Cohen [42].

Lemma 2 (Cohen). In the single-parameter case, $M = 1$, the Fisher information matrix (reducing to a scalar in this case) possesses the property of strict convexity:

$$\begin{aligned} \mathcal{F}_1^{(\ell)}[\gamma q_1(X|\theta) + (1-\gamma)q_2(X|\theta)] \\ < \gamma \mathcal{F}_1^{(\ell)}[q_1(X|\theta)] + (1-\gamma) \mathcal{F}_1^{(\ell)}[q_2(X|\theta)], \end{aligned} \quad (18)$$

for any $0 < \gamma < 1$ and $q_1(X|\theta) \neq q_2(X|\theta)$.

This property can be extended to the multiple-parameter case by the following theorem.

Theorem 1. Given two positive integers $K \geq 2$ and d , in the general case of M parameters split into K groups of d parameters each, $\theta = \{\bar{\theta}_1, \dots, \bar{\theta}_K\}$, where $\bar{\theta}_k \in \mathbb{R}^d$, $M = Kd$, and given real numbers $\mu_1, \dots, \mu_K$ belonging to $[0, 1)$ and summing up to unity, the Fisher information matrix possesses the following properties.

(A) Strict weak matrix convexity in the case of equal weights $\mu_k = 1/K$:

$$\mathcal{F}_M^{(\ell)} \left[\frac{1}{K} \sum_{k=1}^K q_k(X|\theta) \right] < \frac{1}{K} \sum_{k=1}^K \mathcal{F}_M^{(\ell)} [q_k(X|\theta)], \quad (19)$$

under condition that the probability density $q_k(X|\theta)$ depends on θ_k only and

$$\mathcal{F}_M^{(\ell)} [q_k(X|\theta)]_{ij} = \begin{cases} f_0, & i = j \in [(k-1)d+1, kd] \\ 0, & \text{otherwise,} \end{cases} \quad (20)$$

where f_0 is a positive number independent of k .

(B) Strict trace convexity in the general case of arbitrary weights μ_k :

$$\text{Tr} \left\{ \mathcal{F}_M^{(\ell)} \left[\sum_{k=1}^K \mu_k q_k(X|\theta) \right] \right\} < \sum_{k=1}^K \mu_k \text{Tr} \{ \mathcal{F}_M^{(\ell)} [q_k(X|\theta)] \}. \quad (21)$$

Proof. Let us denote the matrix in the left-hand side of Eq. (21) by $\mathbf{F}$ and that in the right-hand side by $\hat{\mathbf{F}} = \sum_k \mu_k \mathcal{F}_M^{(\ell)} [q_k(X|\theta)]$. $\mathbf{F}$ is a real symmetric matrix and as such can be brought to a diagonal form by an orthogonal rotation in the parameter space: $\mathbf{F} = \mathbf{O} \mathbf{D}_F \mathbf{O}^T$, where $\mathbf{D}_F$ is diagonal and $\mathbf{O}$ is orthogonal. Consider a linear transformation from the (column) vector of parameters $\theta = \{\theta_1, \dots, \theta_M\}$ to the new vector of parameters $\varphi = \{\varphi_1, \dots, \varphi_M\}$, $\varphi = \mathbf{O}^T \theta$. The conditional probability density of X for given φ is $\tilde{q}(X|\varphi) = q(X|\mathbf{O}\varphi)$. Let us denote the Fisher information matrix for the information contained in X about φ by $\hat{\mathbf{F}}$. According to Definition 4, its elements are

$$\begin{aligned} \mathcal{F}_M^{(\ell)} [\tilde{q}(X|\varphi)]_{ij} &= \int_{\mathbb{R}^\ell} \frac{1}{\tilde{q}(X|\varphi)} \frac{\partial \tilde{q}(X|\varphi)}{\partial \varphi_i} \frac{\partial \tilde{q}(X|\varphi)}{\partial \varphi_j} d^\ell X \\ &= \sum_{kl} \mathcal{F}_M^{(\ell)} [q(X|\theta)]_{kl} \frac{\partial \theta_k}{\partial \varphi_i} \frac{\partial \theta_l}{\partial \varphi_j} \\ &= \sum_{kl} \mathcal{F}_M^{(\ell)} [q(X|\theta)]_{kl} O_{ki} O_{lj}, \end{aligned} \quad (22)$$

or $\hat{\mathbf{F}} = \mathbf{O}^T \mathbf{F} \mathbf{O} = \mathbf{D}_F$. This means that φ can be considered as a vector of “eigenparameters” of the matrix $\mathbf{F}$. Let us fix all eigenparameters except for one, φ_i , and write the conditional probability density of X for given φ_i as $\tilde{q}'(X|\varphi_i) = \tilde{q}(X|\varphi)$. The Fisher information contained in X about φ_i is

$$\begin{aligned} \mathcal{F}_1^{(\ell)} [\tilde{q}'(X|\varphi_i)] &= \mathcal{F}_M^{(\ell)} [\tilde{q}(X|\varphi)]_{ii} \\ &= \sum_{mn} \mathcal{F}_M^{(\ell)} [q(X|\theta)]_{mn} O_{mi} O_{ni}. \end{aligned} \quad (23)$$

According to Lemma 2, it satisfies the strict convexity inequality

$$\mathcal{F}_1^{(\ell)} \left[\sum_{k=1}^K \mu_k \tilde{q}'_k(X|\varphi_i) \right] < \sum_{k=1}^K \mu_k \mathcal{F}_1^{(\ell)} [\tilde{q}'_k(X|\varphi_i)]. \quad (24)$$

Applying Eq. (23) to each side of this inequality, we obtain

$$\begin{aligned} \sum_{mn} \mathcal{F}_M^{(\ell)} \left[\sum_{k=1}^K \mu_k q_k(X|\theta) \right]_{mn} O_{mi} O_{ni} \\ < \sum_{mn} O_{mi} O_{ni} \sum_{k=1}^K \mu_k \mathcal{F}_M^{(\ell)} [q_k(X|\theta)]_{mn}, \end{aligned} \quad (25)$$

or, in matrix form,

$$\{\mathbf{O}^T (\hat{\mathbf{F}} - \mathbf{F}) \mathbf{O}\}_{ii} > 0. \quad (26)$$

In the case of part A, we have $\hat{\mathbf{F}} = f_0 \mathbf{I}_M / K$, where $\mathbf{I}_M$ is the unit $M \times M$ matrix, which can be seen when applying the condition of equal weights, $\mu_k = 1/K$, and then the identity $\sum_k \mathcal{F}_M^{(\ell)} [q_k(X|\theta)]_{mn} = f_0 \delta_{mn}$ following from Eq. (20). Since inequality (26) holds for any i and both matrices $\mathbf{O}^T \hat{\mathbf{F}} \mathbf{O} = \hat{\mathbf{F}}$ and $\mathbf{O}^T \mathbf{F} \mathbf{O} = \mathbf{D}_F$ are diagonal, this inequality can be rewritten as $\mathbf{O}^T (\hat{\mathbf{F}} - \mathbf{F}) \mathbf{O} > 0$ or $\mathbf{F} < \hat{\mathbf{F}}$, which gives Eq. (19). Part A of the theorem is proved.

In the general case, we sum Eq. (26) over i and use the invariance of the trace with respect to orthogonal rotations to obtain $\text{Tr} \mathbf{F} < \text{Tr} \hat{\mathbf{F}}$, which coincides with Eq. (21) and proves part B. ■

Note that the area of applicability of the matrix convexity is slightly wider than that stated in part A of Theorem 1. Indeed, what we need for the proof is that $\hat{\mathbf{F}}$ is proportional to the unit matrix, which can be reached by replacing f_0 by f_k in Eq. (20) and demanding, instead of equal weights, that $\mu_k f_k$ are independent of k .

IV. LOCALIZATION OF BLINKING SOURCES

Now, we consider K pointlike mutually incoherent sources of light located at unknown locations $\{\bar{x}_k, \bar{y}_k\}$ on the object plane of the microscope. The sources have brightnesses B_k , summing up to the full brightness $B = \sum_k B_k$. We define relative brightnesses $\mu_k = B_k/B$ determining the part of a given source in the total photon flux. We consider two scenarios for estimating the K unknown locations (see Fig. 2).

In the first scenario, which we call “cofluorescence,” all sources can emit light simultaneously. Every detected photon can come from the k th source with the probability μ_k , so that the state of the source in each time bin is $\rho_O^{\text{coflu}} = \sum_k \mu_k \rho_{Ok}$, where ρ_{Ok} is the state of the object with just the k th source active, and the state of the object samples in all time bins

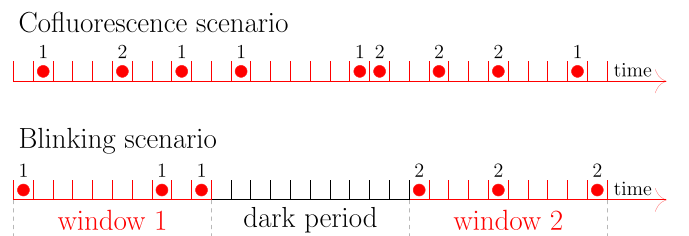


FIG. 2. Photocounts (red balls) appearing in certain time bins in two scenarios of fluorescence. The number above a ball shows from what source the photon comes.

where the photons are detected is

$$\underline{\rho}_O^{\text{coflu}} = \left(\sum_k \mu_k \rho_{Ok} \right)^{\otimes N}. \quad (27)$$

All other quantities in the right-hand side of Eq. (2) correspond to independent identical generalized measurements; that is, they are simple N th powers of the single-sample quantities:

$$\underline{\rho}_P = \rho_P^{\otimes N}, \quad \underline{\rho}_E = \rho_E^{\otimes N}, \quad \underline{U} = U^{\otimes N}, \quad \underline{\Lambda}(r) = \prod_n \Lambda(r_n). \quad (28)$$

Substituting these expressions into Eq. (2), we obtain $\underline{p}(r|\theta) = \prod_n p(r_n|\theta)$, where

$$p(r|\theta) = \text{Tr}\{\Lambda(r)U\rho_O^{\text{coflu}}(\theta) \otimes \rho_P \otimes \rho_E U^\dagger\} \quad (29)$$

is the single-sample conditional probability of measurement outcome $r = \{x, y\}$, which is the location of the detected photon. Substituting this probability into Eq. (4), we find the Fisher information matrix in the form $\underline{\mathbf{F}}_{\text{coflu}}(\theta) = N\mathbf{F}_{\text{coflu}}(\theta)$, where $\mathbf{F}_{\text{coflu}}(\theta)$ is the single-sample Fisher information matrix with the elements

$$F_{\text{coflu},ij}(\theta) = \int \frac{1}{p(r|\theta)} \frac{\partial p(r|\theta)}{\partial \theta_i} \frac{\partial p(r|\theta)}{\partial \theta_j} dr. \quad (30)$$

In the second scenario, which we call “blinking,” the object is affected so that just one source is active in a given time window separated from the neighboring windows by dark periods. Let us denote by $T_k^{(i)}$ the duration of the window of the i th activity of the k th source. This duration is assumed to be long enough for a successful identification of the source number by the localization algorithm. We unite all windows of the same source into the k th superwindow of duration $T_k = \sum_i T_k^{(i)}$. For a sufficiently long experiment, the durations of activity of the sources are approximately equal, $T_k = T$. Therefore, each source emits the total number of $N_k = \mu_k N$ photons (we assume that $\mu_k N$ is an integer, which is a good approximation for a high N). The state of all N_k samples in the k th superwindow is $\rho_{Ok}^{\otimes N_k}$. The joint state of all N samples of the source is

$$\underline{\rho}_O^{\text{blink}} = \rho_{O1}^{\otimes N_1} \otimes \cdots \otimes \rho_{OK}^{\otimes N_K}. \quad (31)$$

All other quantities in the right-hand side of Eq. (2) are simple N th powers of the single-sample quantities, as in the previous scenario. Substituting these expressions into Eq. (2), we obtain $\underline{p}(r|\theta) = \prod_{k=1}^K \underline{p}_k(r_k|\theta)$, where r_k is a part of r with the outcomes in the k th superwindow and $\underline{p}_k(r_k|\theta) = \prod_{l=1}^{N_k} p_k(r_{kl}|\theta)$ with

$$p_k(r|\theta) = \text{Tr}\{\Lambda(r)U\rho_{Ok}(\theta) \otimes \rho_P \otimes \rho_E U^\dagger\} \quad (32)$$

the single-sample conditional probability of measurement outcome when the k th source is only active. Here, the index kl means the l th element in the k th superwindow and equals $kl = l + \sum_{m=1}^{k-1} N_m$. Substituting this probability into Eq. (4), we find the Fisher information matrix in the form $\underline{\mathbf{F}}_{\text{blink}}(\theta) = \sum_k N_k \mathbf{F}^{[k]}(\theta)$, where $\mathbf{F}^{[k]}(\theta)$ is the single-sample Fisher information matrix in the k th window, where just the k th source is

active:

$$F_{ij}^{[k]}(\theta) = \int \frac{1}{p_k(r|\theta)} \frac{\partial p_k(r|\theta)}{\partial \theta_i} \frac{\partial p_k(r|\theta)}{\partial \theta_j} dr. \quad (33)$$

Actually, $\mathbf{F}^{[k]}(\theta)$ depends on the location $\bar{\theta}_k = \bar{r}_k$ of the k th source only, because other sources are inactive in this superwindow.

In order to satisfy the condition of Theorem 1, we need to restrict our consideration to microscopy with certain properties, typical of ordinary microscopes within their field of view [43].

Definition 5. Microscopy with magnification $g \in \mathbb{R}$ is translation and rotation invariant if the conditional distribution of the coordinates $r = \{x, y\}$ of a detected photon on the image plane under condition of a single source at point $\bar{r} = \{\bar{x}, \bar{y}\}$ on the object plane, $p(r|\bar{r})$, satisfies the conditions

$$p(r + g\bar{r}_0|\bar{r} + \bar{r}_0) = p(r|\bar{r}), \quad (34)$$

$$p(\mathbf{R}r|\mathbf{R}\bar{r}) = p(r|\bar{r}), \quad (35)$$

where $\bar{r}_0 = \{\bar{x}_0, \bar{y}_0\}$ is an arbitrary real vector and $\mathbf{R}$ is a rotation matrix on the object plane.

Such a microscopy is characterized by a special form of the Fisher information matrix.

Lemma 3. In translation- and rotation-invariant microscopy, the Fisher information matrix on the position of a single source $\bar{r} = \{\bar{x}, \bar{y}\}$ is proportional to the unit 2×2 matrix $\mathbf{I}$,

$$\mathcal{F}_2^{(2)}[p(r|\bar{r})] = f_0 \mathbf{I}, \quad (36)$$

where f_0 is a positive number independent of $\bar{r}$.

Proof. Rotating the parameters as $\bar{r} \rightarrow \mathbf{R}\bar{r}$, we find from Eq. (16) the corresponding rotation of the Fisher information matrix, $\mathcal{F}_2^{(2)}[p(r|\mathbf{R}\bar{r})] = \mathbf{R}\mathcal{F}_2^{(2)}[p(r|\bar{r})]\mathbf{R}^T$. Since the Jacobian corresponding to a rotation of integration coordinates is $|\det \mathbf{R}| = 1$, we also have $\mathcal{F}_2^{(2)}[p(\mathbf{R}r|\mathbf{R}\bar{r})] = \mathcal{F}_2^{(2)}[p(r|\bar{r})]$. Combining these two equalities with Eq. (35), we obtain invariance of $\mathcal{F}_2^{(2)}[p(r|\bar{r})]$ to rotations, which means that it is proportional to the unit matrix $\mathbf{I}$. Substituting Eq. (34) into Eq. (16), we find $\mathcal{F}_2^{(2)}[p(r|\bar{r} + \bar{r}_0)] = \mathcal{F}_2^{(2)}[p(r|\bar{r})]$, which means that f_0 is independent of $\bar{r}$. ■

Now, we formulate the following theorem.

Theorem 2. For $\underline{\mathbf{F}}_{\text{coflu}}(\theta)$ defined with the object state (27), $\underline{\mathbf{F}}_{\text{blink}}(\theta)$ defined with the object state (31), and all generalized measurements being independent and identical, the following inequalities hold for any ρ_{Ok} , ρ_P , ρ_E , U , and $\Lambda(r)$.

(A) In the case of translation- and rotation-invariant microscopy with $N_k = N/K$,

$$\underline{\mathbf{F}}_{\text{blink}}(\theta) > \underline{\mathbf{F}}_{\text{coflu}}(\theta). \quad (37)$$

(B) In the general case of arbitrary N_k , at least two of which are nonzero,

$$\text{Tr}\{\underline{\mathbf{F}}_{\text{blink}}(\theta)\} > \text{Tr}\{\underline{\mathbf{F}}_{\text{coflu}}(\theta)\}. \quad (38)$$

Proof. Substituting Eqs. (27) and (28) into Eq. (2) and using the property $\text{Tr}\{\mathbf{A} \otimes \mathbf{B}\} = \text{Tr}\mathbf{A}\text{Tr}\mathbf{B}$, we obtain $\underline{p}(r|\theta) = \prod_n p(r_n|\theta)$ with $p(r|\theta)$ defined by Eq. (29), which can be

rewritten as

$$p(r|\theta) = \sum_k \mu_k p_k(r|\theta), \quad (39)$$

where $p_k(r|\theta)$ is defined by Eq. (32). Substituting this form of $p(r|\theta)$ into Eq. (4) and using the additivity of Fisher information by Lemma 1, we obtain

$$\mathbf{F}_{\text{coflu}}(\theta) = N \mathbf{F}_{\text{coflu}}(\theta), \quad (40)$$

where $\mathbf{F}_{\text{coflu}}(\theta)$ is a single-sample Fisher information matrix calculated as a functional of $p(r|\theta)$, $\mathbf{F}_{\text{coflu}}(\theta) = \mathcal{F}_M^{(2)}[p(r|\theta)]$, where the upper index $\ell = 2$ is the dimension of the vector $r = \{x, y\}$.

On the other hand, substituting Eqs. (31) and (28) into Eq. (2), we obtain $\underline{p}(r|\theta) = \prod_{k=1}^K \underline{p}_k(r_k|\theta)$, where $\underline{p}_k(r_k|\theta) = \prod_{l=1}^{N_k} p_k(r_{kl}|\theta)$ and $p_k(r_{kl}|\theta)$ is given by Eq. (32). Substituting this form of $\underline{p}(r|\theta)$ into Eq. (4) and using the additivity of Fisher information by Lemma 1, we obtain

$$\mathbf{F}_{\text{blink}}(\theta) = \sum_k N_k \mathbf{F}^{[k]}(\theta), \quad (41)$$

where $\mathbf{F}^{[k]}(\theta)$ is given by Eq. (33) and calculated as a functional of $p_k(r|\theta)$, $\mathbf{F}^{[k]}(\theta) = \mathcal{F}_M^{(2)}[p_k(r|\theta)]$. Finally, we apply the equality $N_k = N \mu_k$. In the case of $N_k = N/K$, using the convexity of the Fisher information matrix by part A of Theorem 1, we obtain Eq. (37). Condition (20) is satisfied according to Lemma 3. In the case of arbitrary N_k , using the convexity of the trace of the Fisher information matrix by part B of Theorem 1, we obtain Eq. (38). ■

If the matrices $\mathbf{A}$ and $\mathbf{B}$ are invertible and positive, $\mathbf{A} > \mathbf{B}$ is equivalent to $\mathbf{A}^{-1} < \mathbf{B}^{-1}$ [35]. As a corollary to Theorem 2, we have $\mathbf{F}_{\text{blink}}^{-1}(\theta) < \mathbf{F}_{\text{coflu}}^{-1}(\theta)$, and therefore, for a saturated Cramér-Rao bound, in the case of equal brightnesses,

$$H_{\text{blink}}(\hat{\theta}) > H_{\text{coflu}}(\hat{\theta}), \quad (42)$$

where $H(\hat{\theta})$ may be any of the three measures of localization efficiency introduced in Sec. II B, $H_{\text{tot}}(\hat{\theta})$, $H_{\text{ind}}(\hat{\theta})$, or $H_{\text{eig}}(\hat{\theta})$. Moreover, Eq. (42) holds for $H_{\text{eig}}(\hat{\theta})$ also in the case of arbitrary brightnesses.

V. LOCALIZATION OF TWO INCOHERENT SOURCES

In this section, we consider a simple example illustrating the theory developed above and consisting of a one-dimensional localization of two incoherent sources lying at unknown locations $\bar{x}_1$ and $\bar{x}_2$ on the object plane of the microscope and having known relative brightnesses μ_1 and μ_2 . On its image plane, they are represented by two spots centered at $x_1 = g\bar{x}_1$ and $x_2 = g\bar{x}_2$, where g is the microscope magnification. We consider the task of estimating the parameter vector $\theta = \{x_1, x_2\}$, from which $\bar{x}_1$ and $\bar{x}_2$ can easily be found (see Fig. 1).

The localization efficiency is characterized by the three measures introduced in Sec. II B. Operationally, this can be realized by the following protocol: (i) the verifier prepares two sources and passes them to the microscopist, communicating the brightnesses $\mu = \{\mu_1, \mu_2\}$ but not the locations $\theta = \{x_1, x_2\}$; (ii) the microscopist returns the estimates of the locations $\hat{\theta} = \{\hat{x}_1, \hat{x}_2\}$ to the verifier; (iii) such an exchange

is repeated many times, generating a set of data; (iv) the verifier calculates the covariance matrix $\text{Cov}(\hat{\theta})_{jk} = \langle (\hat{x}_j - x_j)(\hat{x}_k - x_k) \rangle$ for the subset of data with given θ and μ ; and (v) from the covariance matrix, the verifier calculates the localization efficiency $H(\theta)$ as a function of θ and μ .

We consider an ordinary wide-field microscope with a circular aperture having the properties of translation and rotation invariance within its field of view. When the number of photons on the image plane within the coherence time is much less than unity (which means that the source is “dim”), the wave function of the photon coming from the k th source is [22,27] $\psi(x - x_k)$, where $\psi(x)$ is the impulse-response function of the optical system of the microscope [43]. For simplicity, we assume that this function is real and model it with a Gaussian function so that its square, the PSF of the microscope, is normalized to unity:

$$\psi(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-x^2/4\sigma^2}, \quad (43)$$

where the standard deviation σ determines the size of each spot.

When just the k th source is active (blinking scenario), the state of the probe (single-photon field) on the image plane is $\rho_{Pk} = |\Psi(x_k)\rangle\langle\Psi(x_k)|$, where

$$|\Psi(x_k)\rangle = \int_{-\infty}^{+\infty} dx \psi(x - x_k) a^\dagger(x) |\text{vac}\rangle, \quad (44)$$

$a^\dagger(x)$ being the photon creation operator at point x . In direct detection, the POVM for detecting a photon at position x is the projector on the one-photon state at position x , i.e., $\Lambda(x) = a^\dagger(x) |\text{vac}\rangle\langle\text{vac}| a(x)$. Then, the probability of measurement outcome x is $p_k(x|x_1, x_2) = \text{Tr}[\rho_{Pk} \Lambda(x)] = \psi^2(x - x_k)$, where we have used the commutation relation $[a(x), a^\dagger(x')] = \delta(x - x')$. The corresponding Fisher information matrix is $F_{ij}^{[k]} = \sigma^{-2} \delta_{ij} \delta_{ik}$. The Fisher information matrix of the entire estimation session with N detected photons, $\mu_1 N$ of which coming from the first source and $\mu_2 N$ coming from the second one, is thus

$$\mathbf{F}_{\text{blink}} = \frac{N}{2\sigma^2} \begin{pmatrix} 1 + \delta & 0 \\ 0 & 1 - \delta \end{pmatrix}, \quad (45)$$

where we express the relative brightnesses, satisfying the relation $\mu_1 + \mu_2 = 1$, through one parameter $0 < \delta < 1$, which we assume to be known, as $\mu_1 = \frac{1}{2}(1 + \delta)$ and $\mu_2 = \frac{1}{2}(1 - \delta)$. The average Fisher information in this scenario is $\bar{F}_{\text{blink}} = N/2\sigma^2$.

When the two sources are active simultaneously (cofluorescence scenario), the state of the probe on the image plane is $\bar{\rho}_P = \mu_1 \bar{\rho}_{P1} + \mu_2 \bar{\rho}_{P2}$ and the probability of the measurement outcome x is

$$p(x|x_1, x_2) = \mu_1 \psi^2(x - x_1) + \mu_2 \psi^2(x - x_2). \quad (46)$$

This function is shown in Fig. 3.

The Fisher information matrix of the entire estimation session with N detected photons distributed according to

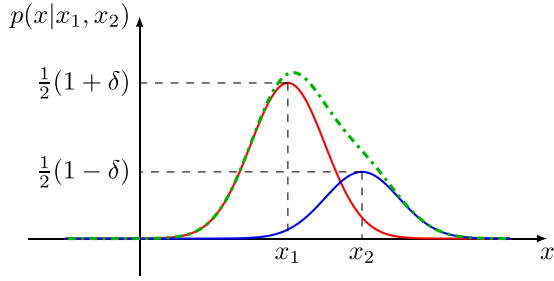


FIG. 3. Statistical model for two cofluorescent incoherent sources. The solid lines show the point-spread functions centered at positions x_1 and x_2 on the image plane of the microscope. The dash-dotted line shows their sum corresponding to one observed spot when the sources are active simultaneously.

Eq. (46) is

$$\underline{F}_{ij}^{\text{coflu}} = \frac{N\mu_i\mu_j}{\sigma^4} \int_{-\infty}^{+\infty} (x - x_i)(x - x_j) \frac{p_i(x|x_i)p_j(x|x_j)}{p(x|x_1, x_2)} dx. \quad (47)$$

The corresponding variance lower bounds $L_{ii} = (\underline{\mathbf{F}}^{-1})_{ii}$ are calculated numerically and shown in Fig. 4(a) for the case of equal brightnesses $\mu_1 = \mu_2 = \frac{1}{2}$. In this case, due to the symmetry with respect to the exchange of parameters x_1 and x_2 , we have $L_{11} = L_{22}$ for both scenarios, and the variance of the cofluorescence scenario is always above that of the blinking scenario, in correspondence with Theorem 2.

In Fig. 4(b), we see a similar situation for the case of unequal brightnesses. The variance lower bounds are different in this case, with the brighter source having a lower position variance. In the blinking scenario, the variance lower bounds for the positions of the first and second sources are $(1 + \delta)^{-1}L_0$ and $(1 - \delta)^{-1}L_0$, respectively, where $L_0 = 1/\bar{F}_{\text{blink}}$. In the cofluorescence scenario, the variances of both positions lie above those of the blinking scenario, though the case of unequal brightnesses is not covered by Theorem 2 for individual bounds.

We also see in Fig. 4(a) that the variances of both positions tend to infinity at a small separation, which means that these positions are difficult to estimate. However, intuitively we understand that the centroid $(x_1 + x_2)/2$ of the overlapping spots can still be estimated with high precision. To show it explicitly, we pass from our parameters $\theta = \{x_1, x_2\}$ to a rotated vector of parameters $\varphi = \{c, s\}$, where $c = (x_1 + x_2)/\sqrt{2}$ is the scaled centroid and $s = (x_2 - x_1)/\sqrt{2}$ is the scaled separation. This rotation in the parameter space can be written as $\varphi = \mathbf{O}_{\pi/4}^T \theta$, where

$$\mathbf{O}_{\xi} = \begin{pmatrix} \cos \xi & -\sin \xi \\ \sin \xi & \cos \xi \end{pmatrix} \quad (48)$$

is the matrix of orthogonal rotation by angle ξ . The Fisher information matrix in the new basis is simply $\tilde{\mathbf{F}} = \mathbf{O}_{\pi/4}^T \underline{\mathbf{F}} \mathbf{O}_{\pi/4}$. Applying this transformation to Eq. (45), we obtain

$$\tilde{\mathbf{F}}_{\text{blink}} = \frac{N}{2\sigma^2} \begin{pmatrix} 1 & -\delta \\ -\delta & 1 \end{pmatrix}, \quad (49)$$

which gives the variance lower bounds $L_{cc} = L_{ss} = L_0/(1 - \delta^2)$. In the cofluorescence scenario, the variance lower

bounds for c and s are calculated numerically and shown in Fig. 4(c) for the case of equal brightnesses. We see that, indeed, the variance of the scaled centroid at zero separation approaches that of blinking sources, but the variance of the scaled separation goes to infinity. Now, we can explain why the positions $x_1 = (c - s)/\sqrt{2}$ and $x_2 = (c + s)/\sqrt{2}$ become completely indeterminate in the $s \rightarrow 0$ limit: they include the scaled separation s whose variance tends to infinity.

The case of unequal brightnesses is illustrated in Fig. 4(d). We see that both c and s become indeterminate at $s \rightarrow 0$ in the cofluorescence scenario. This may be caused by the presence of one variable with a divergent variance in both these parameters, as is the case for x_1 and x_2 analyzed above. Indeed, the limiting value of the Fisher information matrix is

$$\lim_{x_2 \rightarrow x_1} \underline{\mathbf{F}}_{\text{coflu}} = \frac{N}{\sigma^2} \begin{pmatrix} \mu_1^2 & \mu_1\mu_2 \\ \mu_1\mu_2 & \mu_2^2 \end{pmatrix}. \quad (50)$$

This matrix has a zero determinant, but its trace is nonzero. It means that one of its eigenvalues is zero, but another is not: $\lambda_1 = N(\mu_1^2 + \mu_2^2)/\sigma^2$, $\lambda_2 = 0$. As a consequence, at any s , we can introduce two eigenparameters, as discussed in Sec. II B: the best combination of the two positions $b = x_1 \cos \xi + x_2 \sin \xi$ corresponding to the higher eigenvalue of $\underline{\mathbf{F}}_{\text{coflu}}(s)$ and the weighted separation $w = x_2 \cos \xi - x_1 \sin \xi$ corresponding to its lower eigenvalue. The latter parameter has a divergent variance at $s \rightarrow 0$, while the former is well determined.

In Fig. 4(e), we illustrate the behavior of eigenparameter variances and the angle ξ in the case of equal brightnesses. We see that $\xi = \pi/4$ at low separation, which corresponds to $b = c$ and $w = s$, exactly as in Fig. 4(c). At higher separation, the angle becomes $\xi = -\pi/4$, which means that the parameters are relabeled: $b = -s$ and $w = c$. This happens because s has a higher eigenvalue in this region. We also see that both eigenvalues of $\underline{\mathbf{F}}_{\text{coflu}}^{-1}(\theta)$ are larger than the degenerate eigenvalue of $\underline{\mathbf{F}}_{\text{blink}}^{-1}(\theta)$, as prescribed by Eq. (37).

A more interesting case of unequal brightnesses is illustrated in Fig. 4(f). We see that the angle ξ changes continuously, tending to zero at high separation, where the matrix $\underline{\mathbf{F}}_{\text{coflu}}(s)$ becomes diagonal in the (x_1, x_2) basis. The parameter b has always a finite variance, while that of w is divergent at a small separation. Both variances in the cofluorescent scenario lie above their counterparts in the blinking scenario, which means that $\lambda_i(\underline{\mathbf{F}}_{\text{blink}}) \geq \lambda_i(\underline{\mathbf{F}}_{\text{coflu}})$, where $\lambda_i(\mathbf{A})$ is the i th eigenvalue of a Hermitian matrix $\mathbf{A}$, when the eigenvalues are sorted in a nonincreasing order. It is known that for any two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ is a sufficient but not necessary condition for $\lambda_i(\mathbf{A}) \geq \lambda_i(\mathbf{B})$ (see Corollary 7.7.4 in Ref. [35]). This means that we cannot conclude from Fig. 4(f) that inequality (37) holds also in the case of unequal brightnesses, not comprised by Theorem 2. However, an additional numerical study shows that the minimal eigenvalue of the matrix $\underline{\mathbf{F}}_{\text{blink}} - \underline{\mathbf{F}}_{\text{coflu}}$ is positive in the limits of computational error, which is a numerical evidence in favor of validity of inequality (37) in a more general case of unequal brightnesses.

Thus, the variance lower bounds for eigenparameters L_{bb} and L_{ww} appear to be the best characteristics for the overall

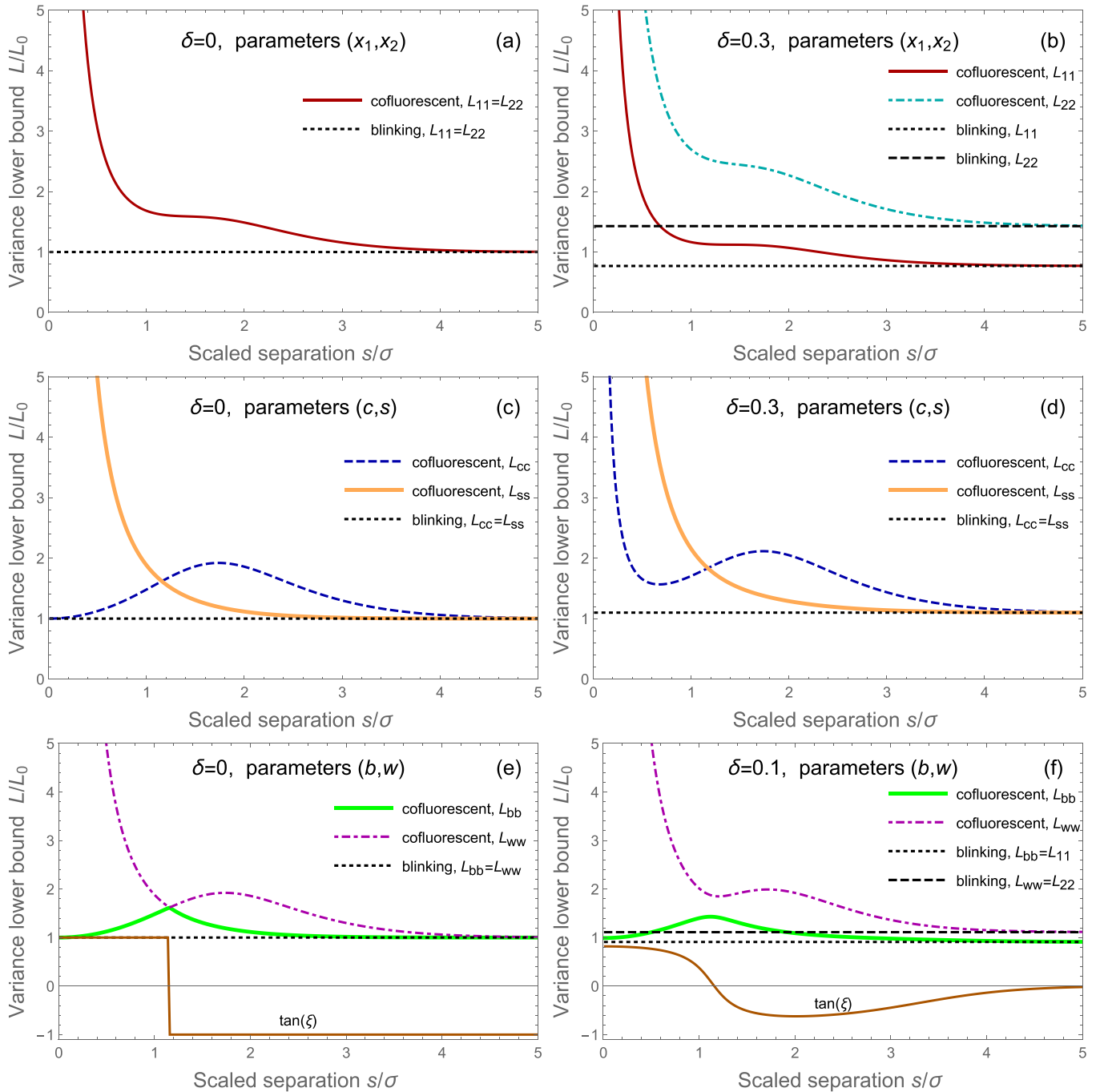


FIG. 4. Variance lower bounds in the units of $L_0 = 2\sigma^2/N$ for estimating the unknown positions (x_1, x_2) of two incoherent sources by detecting N photons in a microscope with the point-spread function of standard deviation σ . δ is the relative brightness difference of the sources. Parameters (c, s) are the scaled centroid and separation, respectively. Parameters (b, w) are the eigenparameters of the Fisher information matrix obtained from (x_1, x_2) via a rotation by angle ξ : $b = x_1 \cos \xi + x_2 \sin \xi$ and $w = x_2 \cos \xi - x_1 \sin \xi$.

localization efficiency. If we wish to combine them in one scalar figure of merit, we should avoid the divergence of L_{ww} at zero separation. The best way to do it is to pass from variances to precisions, as discussed in Sec. II B. The localization efficiency defined in three possible ways is shown in Fig. 5(a) for the case of equal brightnesses and saturated Cramér-Rao bound. This figure confirms what we have already seen in Fig. 4(e): the Fisher information matrix of

blinking sources is larger than that of cofluorescent ones and, therefore, all three measures of localization efficiency are higher for blinking sources.

The case of unequal brightnesses is shown in Fig. 5(b). We see that H_{eig} is higher for blinking sources than for cofluorescent ones, which is a direct consequence of the trace convexity of the Fisher information matrix, proven in Theorem 1. The other two measures, H_{tot} and H_{ind} , are also higher for blinking

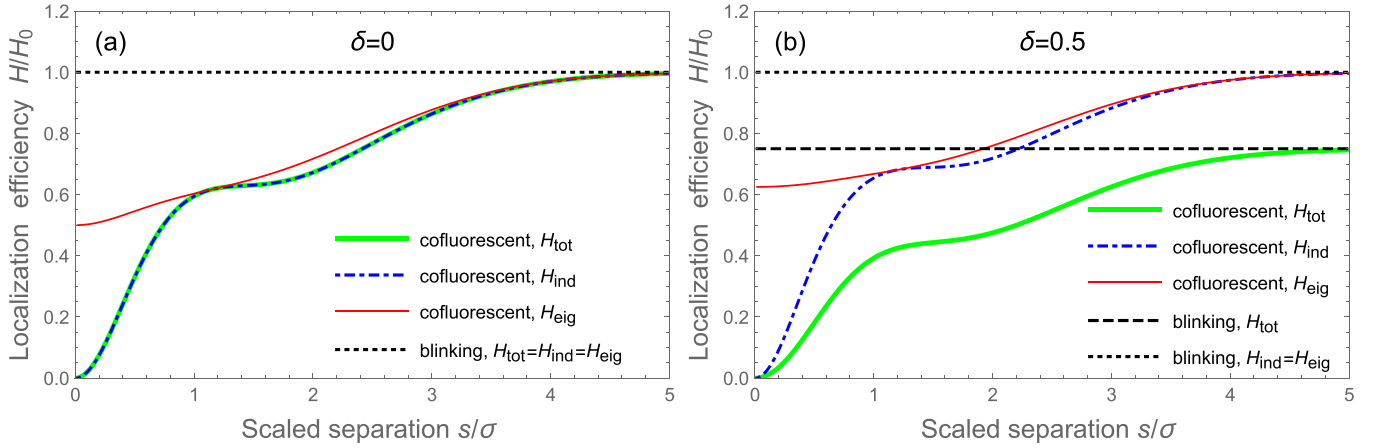


FIG. 5. Localization efficiency in the units of $H_0 = N/2\sigma^2$ for estimating the unknown positions (x_1, x_2) of two incoherent sources by detecting N photons in a microscope with the point-spread function of standard deviation σ . The blinking scenario demonstrates a higher localization efficiency, whatever measure is used for the latter.

sources, though this fact lies outside the scope of the validity of Theorem 2. We also clearly see the ordering of the measures prescribed by Eq. (11).

VI. EXTENSION TO QUANTUM FISHER INFORMATION

The results of the previous sections can be rather straightforwardly extended to the case where the upper bound for localization precision is quantified by quantum Fisher information, which may correspond to a better measurement than the direct detection, as discussed in Sec. II C.

A. Additivity and convexity of quantum Fisher information matrix

We begin by reformulating the definition of the quantum Fisher information matrix [39–41] as a map from a family of density operators in the state space of a quantum system to the space of real symmetric matrices.

Definition 6. Given a positive integer M , a state space of a quantum system $\mathcal{S}$, and a family of density operators $\rho(\theta) \in \mathcal{S}$ depending on parameter $\theta \in \mathbb{R}^M$, the quantum Fisher information matrix for the information contained in $\rho(\theta)$ about θ is the $(M \times M)$ -matrix-valued map $\mathcal{Q}_M^{(S)}[\rho(\theta)]$, whose matrix elements are

$$\mathcal{Q}_M^{(S)}[\rho(\theta)]_{ij} = \frac{1}{2} \text{Tr}\{(\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i) \rho(\theta)\}, \quad (51)$$

where the symmetric logarithmic derivative $\mathcal{L}_i$ is defined by the equation

$$\frac{\partial}{\partial \theta_i} \rho(\theta) = \frac{1}{2} [\mathcal{L}_i \rho(\theta) + \rho(\theta) \mathcal{L}_i]. \quad (52)$$

Now, we formulate the property of additivity for this map, which is well known from the literature [39–41].

Lemma 4. The quantum Fisher information matrix possesses the property of additivity:

$$\begin{aligned} \mathcal{Q}_M^{(S_1 \otimes S_2)}[\rho_1(\theta) \otimes \rho_2(\theta)] \\ = \mathcal{Q}_M^{(S_1)}[\rho_1(\theta)] + \mathcal{Q}_M^{(S_2)}[\rho_2(\theta)], \end{aligned} \quad (53)$$

where $\rho_1(\theta) \in \mathcal{S}_1$ and $\rho_2(\theta) \in \mathcal{S}_2$.

Next, we formulate the property of convexity in the single-parameter case. The property of nonstrict convexity is known from the literature [44,45], but we need the strict convexity property, which is proven below. In the proof, we use the well-known fact that, in the one-parameter case, quantum Fisher information represents the result of maximizing classical Fisher information over all quantum measurements [40,41].

Lemma 5. In the single-parameter case, $M = 1$, the quantum Fisher information matrix (reducing to a scalar in this case) possesses the property of strict convexity:

$$\begin{aligned} \mathcal{Q}_1^{(S)}[\gamma \rho_1(\theta) + (1 - \gamma) \rho_2(\theta)] \\ < \gamma \mathcal{Q}_1^{(S)}[\rho_1(\theta)] + (1 - \gamma) \mathcal{Q}_1^{(S)}[\rho_2(\theta)], \end{aligned} \quad (54)$$

for any $0 < \gamma < 1$ and $\rho_1(\theta)$ and $\rho_2(\theta)$ satisfying the requirement

$$\text{Tr}\{\Lambda_\gamma(X) \rho_1(\theta)\} \neq \text{Tr}\{\Lambda_\gamma(X) \rho_2(\theta)\}, \quad (55)$$

where $\Lambda_\gamma(X)$ is the POVM maximizing the information retrieved from the state $\gamma \rho_1(\theta) + (1 - \gamma) \rho_2(\theta)$ by measuring an ℓ -dimensional variable X .

Proof. The left-hand side of Eq. (47) can be represented as

$$\begin{aligned} \mathcal{Q}_1^{(S)}[\gamma \rho_1(\theta) + (1 - \gamma) \rho_2(\theta)] \\ = \mathcal{F}_1^{(\ell)}[\gamma q_1(X|\theta) + (1 - \gamma) q_2(X|\theta)], \end{aligned} \quad (56)$$

where $\mathcal{F}_1^{(\ell)}$ is the classical Fisher information defined in Definition 4, and the probability distributions $q_i(X|\theta) = \text{Tr}\{\Lambda_\gamma(X) \rho_i(\theta)\}$ correspond to the optimal measurement of $\gamma \rho_1(\theta) + (1 - \gamma) \rho_2(\theta)$.

The strict convexity property of the classical Fisher information (Lemma 2) and Eq. (55) imply that

$$\begin{aligned} \mathcal{F}_1^{(\ell)}[\gamma q_1(X|\theta) + (1 - \gamma) q_2(X|\theta)] \\ < \gamma \mathcal{F}_1^{(\ell)}[q_1(X|\theta)] + (1 - \gamma) \mathcal{F}_1^{(\ell)}[q_2(X|\theta)]. \end{aligned} \quad (57)$$

For each term on the right-hand side of inequality (57), the following relation holds:

$$\mathcal{F}_1^{(\ell)}[q_i(X|\theta)] \leq \mathcal{F}_1^{(\ell)}[q'_i(X|\theta)] = \mathcal{Q}_1^{(S)}[\rho_i(\theta)], \quad (58)$$

where the probability distribution $q'_i(X|\theta) = \text{Tr}\{\Lambda_i(X)\rho_i(\theta)\}$, $i = 1, 2$, corresponds to the measurement $\Lambda_i(X)$ maximizing the information retrieved from the state $\rho_i(\theta)$ and, therefore, defines the quantum Fisher information for that state.

Combining Eqs. (56)–(58), we prove the lemma. ■

This property can be extended to the multiple-parameter case by the following theorem.

Theorem 3. Given two positive integers $K \geq 2$ and d , in the general case of M parameters split into K groups of d parameters each, $\theta = \{\bar{\theta}_1, \dots, \bar{\theta}_K\}$, where $\bar{\theta}_k \in \mathbb{R}^d$, $M = Kd$, and given real numbers $\mu_1, \dots, \mu_K$ belonging to $[0, 1)$ and summing up to unity, the quantum Fisher information matrix possesses the following properties.

(A) Strict weak matrix convexity in the case of equal weights $\mu_k = 1/K$:

$$\mathcal{Q}_M^{(S)} \left[\frac{1}{K} \sum_{k=1}^K \rho_k(\theta) \right] < \frac{1}{K} \sum_{k=1}^K \mathcal{Q}_M^{(S)} [\rho_k(\theta)], \quad (59)$$

under condition that the density operator $\rho_k(\theta)$ depends on $\bar{\theta}_k$ only and

$$\mathcal{Q}_M^{(S)} [\rho_k(\theta)]_{ij} = \begin{cases} q_0, & i = j \in [(k-1)d + 1, kd] \\ 0, & \text{otherwise,} \end{cases} \quad (60)$$

where q_0 is a positive number independent of k .

(B) Strict trace convexity in the general case of arbitrary weights μ_k :

$$\text{Tr} \left\{ \mathcal{Q}_M^{(S)} \left[\sum_{k=1}^K \mu_k \rho_k(\theta) \right] \right\} < \sum_{k=1}^K \mu_k \text{Tr} \{ \mathcal{Q}_M^{(S)} [\rho_k(\theta)] \}. \quad (61)$$

This theorem can be proven exactly the same as Theorem 1, by rotating the parameter vector to the eigenparameters of the quantum Fisher information matrix and applying the single-parameter convexity, established by Lemma 5, to each diagonal entry. In doing this, we assume that condition (55) is satisfied every time convexity is used for a mixture of two density operators. If, by chance, this condition is not satisfied, then the sign “less than” should be replaced by “less than or equal to.” Note that in the task of localization of sources of light, considered here, density operators $\rho_1(\theta)$ and $\rho_2(\theta)$ typically depend on different sets of parameters—the coordinates of the corresponding sources—and the full equivalence of the conditional probability distributions resulting from them is highly unlikely.

B. Quantum Fisher information for localization of blinking sources

The consideration of Sec. IV can be extended to quantum Fisher information. In the cofluorescence scenario, the state of the probe, defined in Sec. II A, is $\tilde{\rho}_P^{\text{coflu}} = (\tilde{\rho}_P^{\text{coflu}})^{\otimes N}$, where

$$\tilde{\rho}_P^{\text{coflu}} = \sum_k \mu_k \tilde{\rho}_{Pk}, \quad (62)$$

and the quantum Fisher information matrix, after Lemma 4, has the form $\underline{\mathbf{Q}}_{\text{coflu}}(\theta) = N \underline{\mathbf{Q}}_{\text{coflu}}(\theta)$, where $\underline{\mathbf{Q}}_{\text{coflu}}(\theta)$ is the

single-sample quantum Fisher information matrix with the elements

$$\mathcal{Q}_{\text{coflu},ij}(\theta) = \frac{1}{2} \text{Tr} \{ (\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i) \tilde{\rho}_P^{\text{coflu}} \}. \quad (63)$$

In the blinking scenario, the state of the probe is

$$\tilde{\rho}_P^{\text{blink}} = \tilde{\rho}_{P1}^{\otimes N_1} \otimes \dots \otimes \tilde{\rho}_{PK}^{\otimes N_K} \quad (64)$$

and the quantum Fisher information matrix is $\underline{\mathbf{Q}}_{\text{blink}}(\theta) = \sum_k N_k \underline{\mathbf{Q}}^{[k]}(\theta)$, where $\underline{\mathbf{Q}}^{[k]}(\theta)$ is the single-sample quantum Fisher information matrix in the k th window, where just the k th source is active:

$$\mathcal{Q}_{ij}^{[k]}(\theta) = \frac{1}{2} \text{Tr} \{ (\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_j \mathcal{L}_i) \tilde{\rho}_{Pk} \}. \quad (65)$$

In both Eqs. (63) and (65) the symmetric logarithmic derivative is defined for the respective density operator according to Eq. (15).

The advantage of a blinking scenario is formulated by the following theorem. For its formulation, we need one more definition.

Definition 7. Microscopy is quantum translation and rotation invariant if the quantum Fisher information matrix of the state $\tilde{\rho}_P(r)$ of a single-photon field on the image plane about the coordinates $r = \{x, y\}$ of the image of a single source having emitted this photon satisfies the condition $\mathcal{Q}_2^{(S)}[\tilde{\rho}_P(r)]_{ij} = q_0 \delta_{ij}$, where q_0 is independent of r .

Theorem 4. For $\underline{\mathbf{Q}}_{\text{coflu}}(\theta)$ defined with the object state (62) and $\underline{\mathbf{Q}}_{\text{blink}}(\theta)$ defined with the object state (64), the following inequalities hold:

(A) In the case of quantum translation- and rotation-invariant microscopy with $N_k = N/K$,

$$\underline{\mathbf{Q}}_{\text{blink}}(\theta) > \underline{\mathbf{Q}}_{\text{coflu}}(\theta). \quad (66)$$

(B) In the general case of arbitrary N_k , at least two of which are nonzero,

$$\text{Tr} \{ \underline{\mathbf{Q}}_{\text{blink}}(\theta) \} > \text{Tr} \{ \underline{\mathbf{Q}}_{\text{coflu}}(\theta) \}. \quad (67)$$

The proof of this theorem follows the lines of that of Theorem 2: applying the additivity and convexity properties of Lemma 4 and Theorem 3.

As a corollary to Theorem 4, we have $\underline{\mathbf{Q}}_{\text{blink}}^{-1}(\theta) < \underline{\mathbf{Q}}_{\text{coflu}}^{-1}(\theta)$ in case A. Does it mean that the localization efficiency is always higher in the blinking scenario for any choice of measurement? In the case of Fisher information discussed in Sec. IV, the answer is positive, because the Cramér-Rao bound is always asymptotically saturable. The quantum Cramér-Rao bound, however, is asymptotically saturable only in the case where the equality

$$\text{Tr} \{ (\mathcal{L}_i \mathcal{L}_j - \mathcal{L}_j \mathcal{L}_i) \tilde{\rho}_P(\theta) \} = 0 \quad (68)$$

holds for any i and j , meaning that the parameters θ_i and θ_j are compatible for a simultaneous quantum estimation [25]. Let us show that this equality holds for the blinking scenario in the case of a real impulse-response function $\psi(x, y)$.

When just the k th source is active, the state of the single-photon field on the image plane is $\tilde{\rho}_{Pk} = |\Psi_k\rangle\langle\Psi_k|$, where, similar to Eq. (44), we write

$$|\Psi_k\rangle = \iint_{-\infty}^{+\infty} dx dy \psi(x - x_k, y - y_k) a^\dagger(x, y) |\text{vac}\rangle, \quad (69)$$

with $a^\dagger(x, y)$ being the photon creation operator at point (x, y) . Substituting this density operator into Eq. (15), we find the symmetric logarithmic derivatives in the form

$$\mathcal{L}_i = |\partial_i \Psi_k\rangle \langle \Psi_k| + |\Psi_k\rangle \langle \partial_i \Psi_k|, \quad (70)$$

where i takes the values x_k, y_k and the states $|\partial_x \Psi_k\rangle$ and $|\partial_y \Psi_k\rangle$ are defined similarly to Eq. (69) but with replacements $\psi(x, y) \rightarrow \partial_x \psi(x, y)$ and $\psi(x, y) \rightarrow \partial_y \psi(x, y)$, respectively. Note that the square of the function $\psi(x, y)$ is normalized to unity, but the square of its derivative is not necessarily that; i.e., the state $|\partial_i \Psi_k\rangle$ is not normalized in general. Differentiating the norm $\langle \Psi_k | \Psi_k \rangle = 1$, we obtain $\text{Re}\{\langle \partial_i \Psi_k | \Psi_k \rangle\} = 0$, and since the impulse-response function is assumed to be real, $|\partial_i \Psi_k\rangle$ is orthogonal to $|\Psi_k\rangle$. Substituting Eq. (70) into Eq. (68), we ascertain that the commutator average is zero. It means that x_k and y_k are compatible for a simultaneous quantum estimation in the window where their source is active. The same argument applies to the locations of all other sources.

In summary, the quantum Cramér-Rao bound is asymptotically saturable in the blinking scenario. We do not discuss here whether it is saturable or not in the cofluorescence scenario, because, in any case, the advantage of the blinking scenario, expressed by inequality (42), holds for the optimal quantum measurement. As in Sec. IV, this inequality holds for any measure of localization efficiency in the case of equal brightnesses and a quantum translation- and rotation-invariant microscopy, and for $H_{\text{eig}}(\hat{\theta})$ in the general case.

It is easy to see that the property of quantum translation and rotation invariance, formulated in Definition 7, is reached for a rotationally invariant impulse-response function, $\psi(\mathbf{R}r) = \psi(r)$, where $\mathbf{R}$ is a matrix of rotation on the (x, y) plane. Indeed, from Eqs. (69) and (70), we obtain $Q_2^{(S)}[\partial_{pk}]_{ij} = J_{ij}$, where

$$J_{ij} = \langle \partial_i \Psi_k | \partial_j \Psi_k \rangle = \iint_{-\infty}^{+\infty} \partial_i \psi(x, y) \partial_j \psi(x, y) dx dy, \quad (71)$$

which reduces to $J_{ij} = q_0 \delta_{ij}$ for a rotationally invariant impulse-response function.

C. Localization in one dimension

We illustrate the developed theory by the same example of two incoherent sources in one spatial dimension as in Sec. V. We find easily from the formalism of the preceding section that

$$\mathbf{Q}_{\text{blink}} = \mathbf{F}_{\text{blink}} = \frac{N}{2\sigma^2} \begin{pmatrix} 1 + \delta & 0 \\ 0 & 1 - \delta \end{pmatrix}. \quad (72)$$

The quantum Fisher information matrix in the case of two cofluorescent incoherent sources with unknown relative brightness was found by Řeháček and co-workers in Ref. [28] for the same Gaussian shape of the impulse-response function. We assume that the relative brightness is known and use from this reference only the 2×2 submatrix related to the centroid $S_0 = (x_1 + x_2)/2$ and the separation $S = x_1 - x_2$, which we denote by $\mathbf{Q}_{\text{coflu}}^R$. We transform this matrix into our parameters x_1 and x_2 as $\mathbf{Q}_{\text{coflu}} = \mathbf{A}^T \mathbf{Q}_{\text{coflu}}^R \mathbf{A}$, where $\mathbf{A}$ is the

transformation matrix for the (column) vectors of the parameters, $\{S_0, S\} = \mathbf{A}\{x_1, x_2\}$, and we obtain

$$\mathbf{Q}_{\text{coflu}} = \frac{N}{2\sigma^2} \begin{pmatrix} 1 + \delta - \beta & -\beta \\ -\beta & 1 - \delta - \beta \end{pmatrix}, \quad (73)$$

where

$$\beta = \frac{1}{8\sigma^2} (1 - \delta^2) (x_1 - x_2)^2 e^{-(x_1 - x_2)^2 / 4\sigma^2}. \quad (74)$$

The variance lower bounds $L_{ii} = (\mathbf{Q}^{-1})_{ii}$ are shown in Fig. 6(a) for the case of equal brightnesses, $\mu_1 = \mu_2 = \frac{1}{2}$, and in Fig. 6(b) for the case of unequal brightnesses, $\mu_1 \neq \mu_2$. We see that, in contrast to Figs. 4(a) and 4(b), the variance lower bound in the cofluorescence scenario does not go to infinity at zero separation. This fact indicates the possibility of an efficient localization of two incoherent sources by means of optimal measurements, such as spatial mode demultiplexing [22, 27, 46–48]. In Fig. 6(a), the lower bound for the cofluorescent scenario lies above that for the blinking scenario, in correspondence with Eq. (66).

A similar behavior is shown in Figs. 6(c) and 6(d) for the scaled centroid and separation. The quantum Fisher information matrix is obtained in this basis as $\tilde{\mathbf{Q}} = \mathbf{O}_{\pi/4}^T \mathbf{Q} \mathbf{O}_{\pi/4}$, which gives the matrix $\tilde{\mathbf{Q}}_{\text{blink}}$ in the form of Eq. (49) and

$$\tilde{\mathbf{Q}}_{\text{coflu}} = \frac{N}{2\sigma^2} \begin{pmatrix} 1 - 2\beta & -\delta \\ -\delta & 1 \end{pmatrix}. \quad (75)$$

The variance lower bounds for the eigenparameters b and w are shown in Figs. 6(e) and 6(f) together with the rotation angle ξ . We see that the latter tends to zero at zero separation in the case of unequal brightnesses, meaning that the best parameter in this limit is the position of the brighter source rather than the centroid as in Fig. 4(f).

The localization efficiency defined in three possible ways for a saturated quantum Cramér-Rao bound is shown in Fig. 7(a) for the case of equal brightnesses. As in Fig. 5(a), we see that all three measures of localization efficiency are higher for blinking sources, which is a consequence of Eq. (66). The case of unequal brightnesses is shown in Fig. 7(b). We see that H_{eig} is higher for blinking sources than for cofluorescent ones, which is a direct consequence of the trace convexity of the quantum Fisher information matrix, proven in Theorem 3.

VII. CONCLUSION

Our comparison of cofluorescent and blinking scenarios for the task of localizing dim incoherent sources of light showed that the blinking scenario always has an informational advantage when the localization efficiency is measured by the average eigenparameter precision, which is a consequence of a fundamental property of the Fisher information matrix, its trace convexity proven in Theorem 1. This advantage is also proven for a more popular measure of efficiency, average position variance, however, only in the case of translation- and rotation-invariant microscopy and equal numbers of photons emitted by each source, which is a consequence of another fundamental property of the Fisher information matrix, its weak convexity proven in the same theorem. Moreover, we

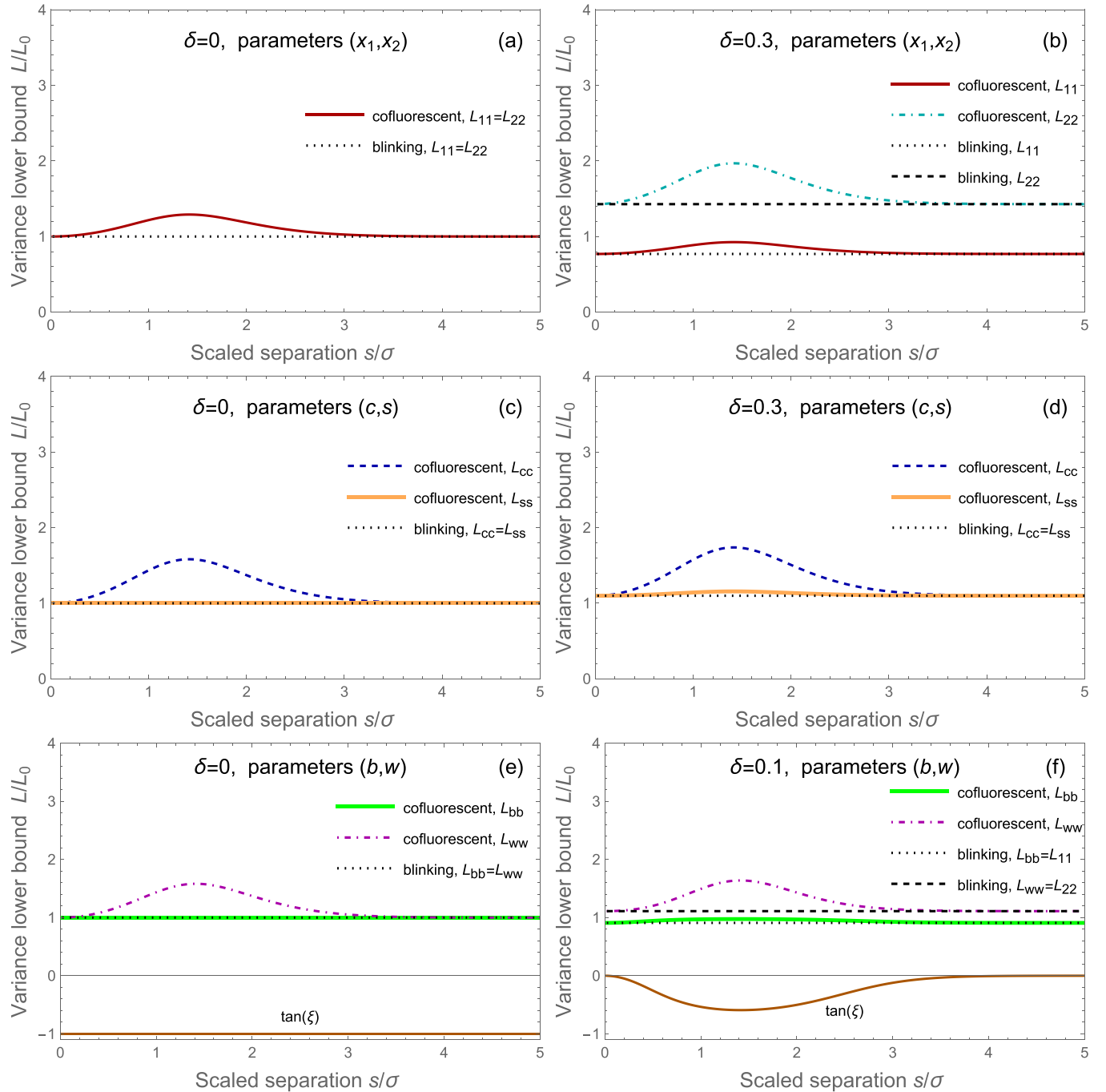


FIG. 6. Variance lower bounds in the units of $L_0 = 2\sigma^2/N$ for estimating the unknown positions (x_1, x_2) of two incoherent sources by measuring optimally N photons in a microscope with the point-spread function of standard deviation σ . δ is the relative brightness difference of the sources. Parameters (c, s) are the scaled centroid and separation, respectively. Parameters (b, w) are the eigenparameters of the Fisher information matrix obtained from (x_1, x_2) via a rotation by angle ξ : $b = x_1 \cos \xi + x_2 \sin \xi$ and $w = x_2 \cos \xi - x_1 \sin \xi$.

have shown that this advantage is a consequence of the corresponding properties—trace and weak matrix convexity—of the quantum Fisher information matrix calculated for light on the image plane of a microscope. The informational advantage of the blinking scenario can be ascribed to a higher *a priori* information, which one has about the number of simultaneously active emitters: 1 in the blinking scenario and $K > 1$ in the cofluorescent one. The advantage in *a priori*

information can be quantified by the advantage in the Fisher information calculated in this work.

We believe that these results are important for both the fields of superresolving optical microscopy and quantum metrology, at the junction of which they are located. On the one hand, they show that the information gain produced by blinking, well known for decades in the superresolving optical microscopy, is a mathematical consequence of such properties

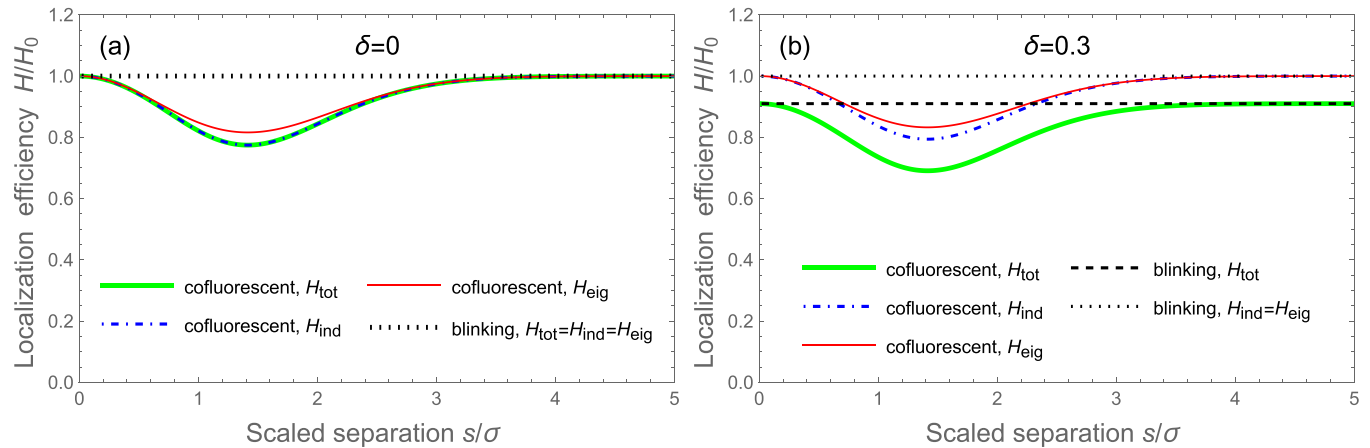


FIG. 7. Localization efficiency in the units of $H_0 = N/2\sigma^2$ for estimating the unknown positions (x_1, x_2) of two incoherent sources by optimally measuring N photons in a microscope with the point-spread function of standard deviation σ under condition of saturated quantum Cramér-Rao bound. The blinking scenario demonstrates a higher localization efficiency, whatever measure is used for the latter.

of the Fisher information matrix as additivity and convexity. On the other hand, it shows that blinking of the sources can increase the Fisher information on their location even to a higher extent than the optimization of the quantum measurement, much studied in the last years in the framework of quantum metrology.

The obtained results are important for a deeper understanding of techniques based on blinking fluorophores [2–4], which are instrumental in advancing superresolution imaging techniques and improving the sensitivity of fluorescence-based biosensors. Both fields are essential for modern biomedical research.

The area of applicability of these results is not reduced to spatial superresolution but also comprises superresolving spectroscopy that allows one to distinguish overlapping spectral lines of optical [49] and magnetic [50,51] fields. Another interesting extension of the developed theory is the analysis of superresolution optical fluctuation imaging (SOFI) [52], where the brightnesses of the sources do not fall to zero, as

in SMLM, but fluctuate in time. A future study will show how the information gain provided by this technique is related to the convexity of the Fisher information.

ACKNOWLEDGMENTS

This work was supported by the QuantERA II Programme (project EXTRASENS) that received funding from Horizon 2020 Framework Programme of the European Union, Grant No. 101017733, the Research Council of Finland, decision 361115, and the Bundesministerium für Bildung und Forschung (Germany), Grant No. 13N16935. It was also funded by the Research Council of Finland (Flagship Programme PREIN, decision 368653; CHARACTER project, decision 357033) and the Horizon Europe Marie Skłodowska-Curie Actions FLORIN Project, Grant No. 101086142.

DATA AVAILABILITY

No data were created or analyzed in this study.

- [1] M. Born and E. Wolf, *Principles of Optics*, 7th ed. (Cambridge University, Cambridge, UK, 1999).
- [2] J. Vangindertael, R. Camacho, W. Sempels, H. Mizuno, P. Dedecker, and K. Janssen, An introduction to optical super-resolution microscopy for the adventurous biologist, *Methods Appl. Fluoresc.* **6**, 022003 (2018).
- [3] L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. Drümgen, Super-resolution microscopy demystified, *Nat. Cell Biol.* **21**, 72 (2019).
- [4] H. Defienne, W. P. Bowen, M. Chekhova, G. B. Lemos, D. Oron, S. Ramelow, N. Treps, and D. Faccio, Advances in quantum imaging, *Nat. Photon.* **18**, 1024 (2024).
- [5] H. Deschout, F. C. Zanacchi, M. Młodzianowski, A. Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, Precisely and accurately localizing single emitters in fluorescence microscopy, *Nat. Methods* **11**, 253 (2014).
- [6] M. Lelek, M. T. Gyparaki, G. Beliu, F. Schueder, J. Griffié, S. Manley, R. Jungmann, M. Sauer, M. Lakadamyali, and C. Zimmer, Single-molecule localization microscopy, *Nat. Rev. Methods Primers* **1**, 39 (2021).
- [7] M. Fazel, K. S. Grussmayer, B. Ferdman, A. Radenovic, Y. Shechtman, J. Enderlein, and S. Pressé, Fluorescence microscopy: A statistics-optics perspective, *Rev. Mod. Phys.* **96**, 025003 (2024).
- [8] 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 (2006).
- [9] M. J. Rust, M. Bates, and X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), *Nat. Methods* **3**, 793 (2006).
- [10] A. Sharonov and R. M. Hochstrasser, Wide-field subdiffraction imaging by accumulated binding of diffusing probes, *Proc. Natl. Acad. Sci. USA* **103**, 18911 (2006).
- [11] M. Gu, Y. Cao, S. Castelletto, B. Kouskousis, and X. Li, Super-resolving single nitrogen vacancy centers within single

- nanodiamonds using a localization microscope, *Opt. Express* **21**, 17639 (2013).
- [12] M. Pfender, N. Aslam, G. Waldherr, P. Neumann, and J. Wrachtrup, Single-spin stochastic optical reconstruction microscopy, *Proc. Nat. Acad. Sci. USA* **111**, 14669 (2014).
- [13] A. Small and S. Stahlheber, Fluorophore localization algorithms for super-resolution microscopy, *Nat. Methods* **11**, 267 (2014).
- [14] R. E. Thompson, D. R. Larson, and W. W. Webb, Precise nanometer localization analysis for individual fluorescent probes, *Biophys. J.* **82**, 2775 (2002).
- [15] A. Yildiz, J. N. Forkey, S. A. McKinney, T. Ha, Y. E. Goldman, and P. R. Selvin, Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization, *Science* **300**, 2061 (2003).
- [16] K. A. Lidke, B. Rieger, T. M. Jovin, and R. Heintzmann, Superresolution by localization of quantum dots using blinking statistics, *Opt. Express* **13**, 7052 (2005).
- [17] F. Huang, S. L. Schwartz, J. M. Byars, and K. A. Lidke, Simultaneous multiple-emitter fitting for single molecule super-resolution imaging, *Biomed. Opt. Express* **2**, 1377 (2011).
- [18] D. R. Cox and D. V. Hinkley, *Theoretical Statistics* (Chapman and Hall, London, 1974).
- [19] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, Berlin, 2006).
- [20] S. Ram, E. S. Ward, and R. J. Ober, Beyond Rayleigh's criterion: A resolution measure with application to single-molecule microscopy, *Proc. Nat. Acad. Sci. USA* **103**, 4457 (2006).
- [21] J. Chao, E. Sally Ward, and R. J. Ober, Fisher information theory for parameter estimation in single molecule microscopy: Tutorial, *J. Opt. Soc. Am. A* **33**, B36 (2016).
- [22] M. Tsang, R. Nair, and X.-M. Lu, Quantum theory of superresolution for two incoherent optical point sources, *Phys. Rev. X* **6**, 031033 (2016).
- [23] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum metrology, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [24] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, *Nat. Photon.* **12**, 724 (2018).
- [25] F. Albarelli, M. Barbieri, M. G. Genoni, and I. Gianani, A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging, *Phys. Lett. A* **384**, 126311 (2020).
- [26] M. Barbieri, Optical quantum metrology, *PRX Quantum* **3**, 010202 (2022).
- [27] M. Tsang, Subdiffraction incoherent optical imaging via spatial-mode demultiplexing, *New J. Phys.* **19**, 023054 (2017).
- [28] J. Řeháček, Z. Hradil, B. Stoklasa, M. Paúr, J. Grover, A. Krzic, and L. Sánchez-Soto, Multiparameter quantum metrology of incoherent point sources: Towards realistic superresolution, *Phys. Rev. A* **96**, 062107 (2017).
- [29] E. Bisketzi, D. Branford, and A. Datta, Quantum limits of localisation microscopy, *New J. Phys.* **21**, 123032 (2019).
- [30] J. R. Hervas, L. L. Sánchez-Soto, A. Z. Goldberg, Z. Hradil, and J. Řeháček, Optimizing measurement tradeoffs in multiparameter spatial superresolution, *Phys. Rev. A* **110**, 033716 (2024).
- [31] A. Mikhalychev, B. Bessire, I. Karuseichyk, A. Sakovich, M. Unternährer, D. Lyakhov, D. L. Michels, A. Stefanov, and D. Mogilevtsev, Efficiently reconstructing compound objects by quantum imaging with higher-order correlation functions, *Commun. Phys.* **2**, 134 (2019).
- [32] S. Vlasenko, A. B. Mikhalychev, I. L. Karuseichyk, D. A. Lyakhov, D. L. Michels, and D. Mogilevtsev, Optimal correlation order in superresolution optical fluctuation microscopy, *Phys. Rev. A* **102**, 063507 (2020).
- [33] S. Kurdzialek and R. Demkowicz-Dobrzański, Super-resolution optical fluctuation imaging—fundamental estimation theory perspective, *J. Opt.* **23**, 075701 (2021).
- [34] A. Mikhalychev, P. Novik, I. Karuseichyk, D. Lyakhov, D. L. Michels, and D. Mogilevtsev, Lost photon enhances superresolution, *npj Quantum Inf.* **7**, 125 (2021).
- [35] R. A. Horn and C. J. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1985).
- [36] J. M. Bernardo and A. F. Smith, *Bayesian Theory* (John Wiley & Sons, New York, 2009).
- [37] S. L. Braunstein and H. J. Kimble, Teleportation of continuous quantum variables, *Phys. Rev. Lett.* **80**, 869 (1998).
- [38] D. B. Horoshko and S. Y. Kilin, Quantum teleportation using quantum nondemolition technique, *Phys. Rev. A* **61**, 032304 (2000).
- [39] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, Cambridge, 1976).
- [40] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
- [41] M. G. A. Paris, Quantum estimation for quantum technology, *Int. J. Quantum Inf.* **07**, 125 (2009).
- [42] M. Cohen, The Fisher information and convexity, *IEEE Trans. Inf. Theory* **14**, 591 (1968).
- [43] J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts & Co., Englewood, CO, 2005).
- [44] G. Tóth and D. Petz, Extremal properties of the variance and the quantum Fisher information, *Phys. Rev. A* **87**, 032324 (2013).
- [45] G. Tóth and F. Fröwis, Uncertainty relations with the variance and the quantum Fisher information based on convex decompositions of density matrices, *Phys. Rev. Res.* **4**, 013075 (2022).
- [46] M. Paúr, B. Stoklasa, Z. Hradil, L. L. Sánchez-Soto, and J. Rehacek, Achieving the ultimate optical resolution, *Optica* **3**, 1144 (2016).
- [47] P. Boucher, C. Fabre, G. Labroille, and N. Treps, Spatial optical mode demultiplexing as a practical tool for optimal transverse distance estimation, *Optica* **7**, 1621 (2020).
- [48] C. Rouvière, D. Barral, A. Grateau, I. Karuseichyk, G. Sorelli, M. Walschaers, and N. Treps, Ultra-sensitive separation estimation of optical sources, *Optica* **11**, 166 (2024).
- [49] L. Motka, B. Stoklasa, M. D'Angelo, P. Facchi, A. Garuccio, Z. Hradil, S. Pascasio, F. V. Pepe, Y. S. Teo, J. Řeháček, and L. L. Sánchez-Soto, Optical resolution from Fisher information, *Eur. Phys. J. Plus* **131**, 130 (2016).
- [50] A. Rotem, T. Gefen, S. Oviedo-Casado, J. Prior, S. Schmitt, Y. Burak, L. McGuinness, F. Jelezko, and A. Retzker, Limits on spectral resolution measurements by quantum probes, *Phys. Rev. Lett.* **122**, 060503 (2019).
- [51] S. Oviedo-Casado, A. Rotem, R. Nigmatullin, J. Prior, and A. Retzker, Correlated noise in Brownian motion allows for super resolution, *Sci. Rep.* **10**, 19691 (2020).
- [52] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI), *Proc. Nat. Acad. Sci. USA* **106**, 22287 (2009).