Discrimination and estimation of incoherent sources under misalignment

De Almeida, J. O.; Koodyski, J.; Hirche, C.; Lewenstein, M.; Skotiniotis, M.

Published in:
Physical Review A

DOI:
10.1103/PhysRevA.103.022406

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
Other

Citation for published version (APA):
Discrimination and estimation of incoherent sources under misalignment

J. O. de Almeida,1,* J. Kołodyński,2 C. Hirche,3 M. Lewenstein,1,4 and M. Skotiniotis5

1ICFO-Institut de Ciencies Fotoniques, Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain
2Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland
3QMATH, Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
4ICREA-Institució Catalana de Recerca i Estudis Avançats, Lluís Companys 23, 08010 Barcelona, Spain
5Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

(Received 6 October 2020; accepted 6 January 2021; published 3 February 2021)

Spatially resolving two incoherent point sources whose separation is well below the diffraction limit dictated by classical optics has recently been shown possible using techniques that decompose the incoming radiation into orthogonal transverse modes. Such a demultiplexing procedure, however, must be perfectly calibrated to the transverse profile of the incoming light as any misalignment of the modes effectively restores the diffraction limit for small source separations. We study by how much can one mitigate such an effect at the level of measurement which, after being imperfectly demultiplexed due to inevitable misalignment, may still be partially corrected by linearly transforming the relevant dominating transverse modes. We consider two complementary tasks: the estimation of the separation between the two sources and the discrimination between one and two incoherent point sources. We show that, although one cannot fully restore super-resolving powers even when the value of the misalignment is perfectly known, its negative impact on the ultimate sensitivity can be significantly reduced. In the case of estimation we analytically determine the exact relation between the minimal resolvable separation as a function of misalignment whereas for discrimination we analytically determine the relation between misalignment and the probability of error, as well as numerically determine how the latter scales in the limit of long interrogation times.

DOI: 10.1103/PhysRevA.103.022406

I. INTRODUCTION

Quantum theory has, over the years, exhibited an innate ability to surpass the limitations in performance set by classical devices in a variety of tasks [1–4], arguably none more so than in the field of statistical inference and decision theory. There the use of distinctive quantum features, such as coherence and entanglement, allows for the existence of ultrasprecise measurements [3] that greatly enhance the performance in a variety of sensing tasks—ghost imaging [5] and quantum illumination [6] to name but a few—that are impossible to achieve by even the best classical means.

One such success of the quantum mechanical formalism concerns the spatial resolution of imaging devices. For over a century it was believed that two sources of incoherent light can barely be resolved if “the maximum of one is over the minimum of the other” [7]; any closer than this and conventional classical imaging techniques cannot resolve the two incoherent sources, even if an asymptotically large number of photons are detected. Despite several efforts [8–11], this limitation of optical imaging systems—known as the diffraction or Rayleigh limit [7]—seemed insurmountable until a proper quantum mechanical treatment of the problem revealed that, just like many other classically derived limitations, it too can be overcome [12]. Rather than imaging directly the incoming radiation it was proven that a simple linear-optical preprocessing of the spatial profile of the electromagnetic field into a predefined set of spatially orthogonal modes, e.g., the Hermite-Gauss modes in the case of Gaussian apertures [13], followed by photon detection over sufficiently long integration time is capable of resolving two incoherent point sources at arbitrary separation. The reason for this drastic improvement is intuitive: spatially orthogonal modes of light provide information about spatial correlations of the incoming photons, whereas direct imaging does not.

The technique of decomposing, or demultiplexing, the optical field into spatially orthogonal modes followed by photon counting has gained increased attention with rapid theoretical and experimental developments (see [14] for a recent review). Its performance has been proven not only in complex estimation tasks, such as resolving multiple sources [15–20], sources of unequal brightness [21,22], sources emitting coherent [23] or nonclassical [24] light, as well as sources localized arbitrarily in space [25], but also for the closely related problem of discrimination beyond the diffraction limit [26].

Moreover, the robustness to imperfections of the proposed schemes has recently been an object of intensive research [27–29], largely motivated by the challenges imposed by up-to-date experimental demonstrations [30–36]. An important
FIG. 1. Super-resolving the separation between incoherent sources under the misalignment of the imaging apparatus. Two incoherent pointlike sources of light are imaged with an optical system exhibiting a Gaussian point spread function of width \( \sigma \) in a way that their separation, \( 2d \), can be most accurately resolved. For this to be possible beyond the diffraction limit, a spatial mode demultiplexing technique is employed, which ideally allows the incoming light to be decomposed into orthogonal transverse modes, whose photon occupation is subsequently measured. In this work, we study the ultimate limits on the resolution in the presence of misalignment of the imaging system, \( \delta = x_c - x_R \ll \sigma \), by applying appropriate linear-optical postprocessing operations \( R(\delta) \) to the two dominant modes of the demultiplexing measurement.

Obstacle pointed out in the original paper of Tsang et al. [12] is the crucial assumption that the centroid \( (x_c) \) —the midpoint between the two light sources whose separation, \( 2d \), is to be resolved—is perfectly aligned \( (x_c = x_R) \) with the detector position \( (x_R) \), where the spatial transverse modes are demultiplexed. In the presence of any misalignment, \( \delta = x_c - x_R \) (see Fig. 1), the Fisher information, \( F \), that quantifies the ultimate resolution no longer approaches a constant with vanishing separation but rather behaves as \( F \sim (d/\delta)^2 \to 0 \) for \( d \to 0 \) [12].

Although the location of the centroid can be estimated via direct imaging, this requires sacrificing an, in principle, large amount of photons in order to do so. A way of leveraging the use of photons between direct imaging and spatial mode demultiplexing, using information of the former to better position the latter, has been recently proposed [37]. Here, in contrast, we take misalignment as experimental fact and investigate the theoretical limits imposed by any misalignment in both the canonical estimation task [12,27–37] as well as the complementary task of hypothesis testing [26,38].

Specifically, we ask the following questions: How do estimation precision and the minimal resolvable distance scale as a function of misalignment in the canonical estimation task, and how does the probability of error—both in single observation, as well as asymptotically—depend on misalignment in a typical hypothesis testing task? We keep our treatment as general as possible by considering the use of passive linear optics after the demultiplexing stage (see Fig. 1). We show that even if the value of misalignment \( \delta \) is perfectly known such linear-optical postprocessing is not capable of restoring the superresolution. We demonstrate that optical postprocessing of the two most dominant modes, tailored to the misalignment \( \delta \), improves estimation precision over the “raw” demultiplexed measurements from \( F \propto 1/\delta^2 \) to \( F \propto 1/\delta^6 \) as \( \delta \approx 0 \) and the minimal resolvable distance from \( \epsilon_{\text{min}} \propto \delta^2 \) to \( \epsilon_{\text{min}} \propto \delta^4 \) for both Gaussian and sinc point-spread functions. Moreover, for single-shot \( (n = 1) \) discrimination the probability of erroneously interpreting a single source for a double one scales as \( P_{\text{err}} \propto \delta^6 \) compared to \( P_{\text{err}} \propto \delta^2 \) for the misaligned demultiplexed measurements for small \( \delta \approx 0 \). Furthermore, we also show that linear postprocessing also helps in the asymptotic \( n \to \infty \) regime, where we observe an enhancement of up to 12% in the exponential decay of the total probability of error.

One may object that if the misalignment is known, then superresolution can easily be restored simply by adjusting the demultiplexing device to the right position. We stress, however, that in our work this assumption is made in order to explore the theoretical limits on superresolution allowed by the most general postprocessing after an imperfect demultiplexing procedure. From this point of view, our results directly link pertinent quantities—the quantum Fisher information, minimal resolvable distance, and probability of error—to the value of the misalignment. From the practical point of view even if the position of the centroid is perfectly known, fabrication of the demultiplexing measurements, or adjustment of their position via mechanical servo mechanisms, entails an inherent uncertainty [39] thus not allowing for perfect alignment. In such instances our theoretical analysis offers insight into additional strategies, which may be implemented efficiently by modulating the phase of integrated waveguides [40,41] that could be used in conjunction with current experimental proposals [37].

The article is structured as follows. First, we review the necessary mathematical background for both classical and quantum mechanical image resolution in Sec. II. In Sec. III we study the effects of misalignment for the problem of estimating the separation between two incoherent point sources, while Sec. IV deals with the effects of misalignment for the problem of discriminating between the one- and two-source hypotheses. Section V summarizes our work and discusses possible future directions of investigation.

II. DIFFRACTION-LIMITED OPTICAL IMAGING

We begin by reviewing the mathematical treatment of optical imaging devices. In Sec. II A we review imaging in classical optics, paying particular attention to how the
diffraction limit comes about in these setups. In Sec. II B we
give a formal quantum mechanical description of the point
spread function (PSF) of an optical imaging system. We re-
strict our attention particularly to one-dimensional Gaussian
and sinc PSFs but the analysis easily extends to other PSFs
and to higher dimensions. We then review a mathematical
approximation to the quantum mechanical state of the PSF—
the qubit model of Chrostowski et al. [42]. The latter will be
used to explore how misalignment of the optical imaging sys-
tem affects its performance, as well as to propose alternative
measurement schemes that compensate for misalignment.

A. Classical theory of diffraction-limited optical imaging

To image light sources that are far away requires specific
lens and aperture systems that allow to process the spatial
distribution of the emitters. Assuming the paraxial approxi-
mation holds, diffraction effects cause variations in radiation
intensity at the image plane—the familiar bright and dark
fringes in imaging stars, or diffraction gratings. Consequently,
the minimum angular distance between two or more emitters
that allows their distinction—the angular resolution of the
imaging device—is fundamentally limited due to diffraction.
Lord Rayleigh was the first to obtain a heuristic rule for the an-
gular resolution of any imaging device [7]: two point sources
can barely be resolved so long as the central maximum in
intensity of one source lies on top of the first minimum in
intensity of the second in the image plane. This rule of thumb
is colloquially known as Rayleigh’s curse or the diffraction
limit in optical imaging.

For the simplest optical imaging device consisting of a
single slit of width \(D\) the diffraction limit can be deduced
by simple geometrical optics and corresponds to the angular
distance, \(\phi\), between the central intensity maximum and first
minimum which is given by

\[
\phi \approx \frac{\lambda}{D},
\]

where \(\lambda\) is the wavelength of the incoming radiation, and
the approximation sign is due to the paraxial approximation.

More formally, the diffraction limit can be obtained by
making use of the Fresnel-Kirchhoff formula which describes
the amplitude of the disturbance in a given direction, \(\phi\), from
the optical axis due to the aperture of the imaging system [43].
For a one-dimensional aperture whose profile is given by \(f(y)\),
the Fresnel-Kirchhoff formula reads

\[
\Psi(\phi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y)e^{i k y \sin \phi} dy,
\]

where \(k = 2\pi/\lambda\) is the wave number. The intensity distribu-
tion, also known as the object’s PSF, at angular separation
\(\phi\) is given by \(|\Psi(\phi)|^2\), and we have implicitly assumed
the intensity is normalized \(\int |\Psi(\phi)|^2 d\phi = 1\).

Equation (2) is the familiar statement that the PSF at the
image plane of an image system is the Fourier transform of
the system’s aperture. The case of the single slit of width \(D\)
corresponds to \(f(y) = \text{rect}(-\frac{D}{2}, \frac{D}{2})\) and gives rise to the sinc
PSF

\[
\Psi(\phi) \propto \text{sinc}\left(\frac{Dk \sin \phi}{2}\right),
\]

and here the sinc function is defined as \(\text{sinc}(x) = \frac{\sin(x)}{x}\). The
first minimum of the sinc function occurs at \(\frac{Dk \sin \phi}{2} = \pi\), i.e.,
\(\phi \approx \frac{\pi}{k}\), which is the result obtained using geometric optics.

For a circular aperture \(f(y) = \sqrt{D^2 - 4y^2}\), where \(D\) is the
diameter of the aperture, the corresponding PSF reads

\[
\Psi(\phi) \propto J_1\left(\frac{Dk}{2} \sin \phi\right),
\]

where \(J_1(z)\) is the Bessel function of the first kind. The first
minimum of the latter occurs when \(\frac{Dk \sin \phi}{2} = 3.8317\), which
sets the angular resolution to \(\phi \approx \frac{\pi}{k}\).

One can, in principle, shape the PSF of an imaging system
to any desired function using apodization which suppresses the
higher-order intensity maxima of the diffraction pattern [44].
Such techniques can be used to turn the Bessel function PSF of
the circular aperture to a Gaussian one. As such techniques do
not alter the shape of the aperture, the diffraction limit above
still holds.

B. Quantum description of two incoherent point sources

Consider two incoherent point sources (e.g., stars or bac-
teria fluorescing) emitting monochromatic light. We assume
that the sources are weak, meaning that the average number
of photons detected by our imaging device is much smaller
than 1. Quantum mechanically we may represent the state of
the incoming radiation by the density operator [12]:

\[
\sigma^{(i)} \approx (1 - \epsilon) |0\rangle\langle 0| + \epsilon \rho^{(i)} + O(\epsilon^2),
\]

where \(\epsilon \ll 1, |0\rangle\langle 0|\) corresponds to the vacuum state, and
\(\rho^{(i)} \in \mathcal{B}(\mathcal{H}_1), i \in \{1, 2\}\), is a one-photon state with the super-
script index labeling the case where the photon is due to one
or two point sources.

As the vacuum offers no information about the nature of
the emitting source our only information comes from the
single-photon events, accumulated over sufficiently long time,
at the image plane of our instrument. Assuming the latter to be
one dimensional we define the image plane position eigenkets
\(|x\rangle = a^\dagger(x)|0\rangle\), where \(a^\dagger(x)\) and \(a(x)\) are the creation and an-
nihilation operators satisfying \([a(x), a^\dagger(y)] = \delta(x - y)\) [45].
The wave function of a single photon can now be expanded in
terms of the position basis of the image plane as

\[
|\Psi(z)\rangle = \int_{-\infty}^{\infty} dx |\Psi(x - z)|x\rangle,
\]

where \(|\Psi(x - z)|^2 = \langle x|\Psi(z)|^2\rangle\) denotes the probability of
detecting a photon at position \(x\) in the image plane—the
object’s PSF. \(|\Psi(x - z)|^2\) is equivalent to \(\Psi(\phi)\) of Eq. (2) in the
far-field regime in which \(\phi \approx 0\) and, hence, \(\sin(\phi) \approx \phi\).

For a Gaussian or square aperture the PSF of a single in-
coherent point source is the corresponding Fourier transform
[13,46],

\[
|\Psi(x - z)|^2 = \frac{1}{\sqrt{2\pi \sigma^2}} \exp\left(-\frac{(x - z)^2}{2\sigma^2}\right),
\]

\[
|\Psi(x - z)|^2 = \frac{1}{\sigma} \text{sinc}^2\left(\frac{\pi}{\sigma}, x - z\right).
\]
FIG. 2. Intensity distribution at the image plane arising from two nearly coinciding incoherent light sources ($\sigma \gg d$) given an imaging system that exhibits a Gaussian point spread function (PSF). Due to the large overlap between the two PSFs, direct imaging does not allow to accurately estimate the positions of each source, and is incapable of discriminating whether the image is the result of one or two sources—hypothesis 1 and hypothesis 2, respectively.

respectively, where $z$ is the mean of the PSF and $\sigma^2$ the corresponding variance (both fully characterized by the imaging system), and we have introduced appropriate normalization factors ensuring that $\int |\Psi(x - z)|^2 dx = 1$. The variance is taken to be $\sigma^2 \approx \frac{D^2}{2\pi}$, where $D$ is the diameter (length) of the Gaussian (square) apertures, respectively Eqs. (7a) and (7b). Note that Eq. (7b) is equivalent to Eq. (3) in the far-field regime, where $x - z = \sin \phi \approx \phi$.

The state of a single photon emanating from a single point source whose PSF is centered around $z = x_0$ is then described by the state

$$\rho^{(1)} = |\Psi(x_0)\rangle \langle \Psi(x_0)|,$$

whereas that for a photon coming from two incoherent point sources with relative intensities $w$ and $1 - w$, whose PSFs are centered around $x_1$ and $x_2$, is described by the density matrix

$$\rho^{(2)} = w|\Psi(x_1)\rangle \langle \Psi(x_1)| + (1 - w)|\Psi(x_2)\rangle \langle \Psi(x_2)|.$$

For the case of two incoherent point sources it is convenient to define the centroid,

$$\bar{x}_c := w x_1 + (1 - w) x_2,$$

and separations,

$$d_i := |x_i - x_0|.$$

For two sources of equal intensity—the case shown in Fig. 1—this reduces to

$$d := \frac{w x_1 + (1 - w) x_2}{w + (1 - w)},$$

and

$$\delta := |x_1 - x_2|.$$

In this work we take misalignment of the demultiplexing device as fact and determine the theoretical limits imposed by such misalignment in both estimation and discrimination tasks (see Fig. 2). To do so we allow ourselves the freedom of performing arbitrary, linear-optical postprocessing of the demultiplexed radiation. Specifically, we analytically determine how such linear-optical postprocessing, based on complete knowledge of the value of the misalignment, affects estimation precision, as well as the minimal resolvable distance. To do so we make use of an approximation of the state of the incoming radiation known as the qubit model [42], which we now review.

C. The qubit model for two incoherent point sources

The qubit model is an approximation of the PSF in the presence of misalignment [42]. The latter can be understood as performing the projective measurement of Eq. (12) about some reference position $x_R$ for $i \in \{0, 1, 2\}$. Assuming that this misalignment is small, i.e., $x_R \approx x_i$, we can Taylor-expand the probability amplitudes of each source,
$\Psi(x - x_i), i \in \{1, 2\}$, about $x_R$ as follows:

\[
\left|\Psi(x)\right| \approx \int_{-\infty}^{\infty} dx \Psi(x - x_R) |x\rangle + (x_i - x_R) \int_{-\infty}^{\infty} dx \frac{d^2\Psi(x - x_i)}{dx_i} |x\rangle,
\]

and identify a qubit subspace with $\langle \psi_1 \rangle$.

An orthonormal basis. Here, $N$ is an appropriate normalization factor which for the Gaussian and sinc PSFs reads

\[
N_G = \frac{1}{4\sigma^2}, \quad N_S = \frac{\pi^2}{3\sigma^2},
\]

respectively.

The state of the incoming radiation can now be described, to a very good approximation, by the following qubit density operators, for one and two sources, respectively:

\[
\rho^{(1)} \approx \frac{1}{1 + (\theta^2)^2N} \left(\begin{array}{cc}
-\theta & -\theta \sqrt{\sigma}N \\
-\theta \sqrt{\sigma}N & (\sigma^2)^2N
\end{array}\right),
\]

\[
\rho^{(2)} \approx \frac{1}{1 + \sigma^2(\sigma^2 + \epsilon^2)N^2} \left(\begin{array}{cc}
-\theta & -\theta \sqrt{\sigma}N \\
-\theta \sqrt{\sigma}N & \sigma^2(\sigma^2 + \epsilon^2)N
\end{array}\right),
\]

where we now introduce dimensionless parameters for misalignment and separation:

\[
\theta := \frac{\delta}{\sigma} = \frac{x_c - x_R}{\sigma} \quad \text{and} \quad \epsilon := \frac{\delta}{\sigma},
\]

respectively. The qubit model allows us to visualize the effects of misalignment on a given PSF in terms of the Bloch representation of qubit density matrices, i.e.,

\[
\rho := \frac{1 + r \cdot \sigma}{2},
\]

where $r \in \mathbb{R}_3$ has elements $r_i = \text{Tr}(\sigma_i\rho)$, and $\sigma := (\sigma_1, \sigma_2, \sigma_3)$ is the vector of Pauli matrices $\sigma$. For the Gaussian and sinc PSFs the corresponding Bloch vectors read

\[
r^{(1)}_G \approx \frac{1}{1 + \frac{\theta^2}{4}} \left(\begin{array}{c}
-\theta \\
0
\end{array}\right), \quad n^{(2)}_G \approx \frac{1}{1 + \frac{\theta^2 + \epsilon^2}{4}} \left(\begin{array}{c}
-\theta \\
0
\end{array}\right),
\]

\[
r^{(1)}_S \approx \frac{1}{1 + \frac{\theta^2}{3}} \left(\begin{array}{c}
-\theta \\
0
\end{array}\right), \quad n^{(2)}_S \approx \frac{1}{1 + \frac{\theta^2 + \epsilon^2}{3}} \left(\begin{array}{c}
-\theta \\
0
\end{array}\right),
\]

respectively. Using the approximations

\[
\frac{1}{1 + x^2} \approx 1 - x^2, \quad 1 - \frac{(\theta^2 + \epsilon^2)}{2} \approx \left(1 - \frac{\theta^2}{2}\right) \left(1 - \frac{\epsilon^2}{2}\right) \approx \cos \theta \left(1 - \frac{\epsilon^2}{2}\right),
\]

and keeping terms up to second order, $O(\theta^2\epsilon^2)$ with $i + j = 2$, the Bloch vectors in Eq. (20) can be further approximated by

\[
r^{(1)}_G \approx \left(-\sin \theta \quad 0 \right)^T \cos \theta, \quad r^{(2)}_G \approx \left(1 - \frac{\epsilon^2}{2}\right) \left(-\sin \theta \quad 0 \right)^T \cos \theta,
\]

\[
r^{(1)}_S \approx \left(-\sin \frac{\theta}{2} \quad 0 \right)^T \cos \frac{\theta}{2}, \quad r^{(2)}_S \approx \left(1 - \frac{\epsilon^2}{2}\right) \left(-\sin \frac{\theta}{2} \quad 0 \right)^T \cos \frac{\theta}{2}.
\]

Consequently the misalignment, $\theta$, can be understood as an infinitesimal rotation about the $y$ axis in the Bloch-sphere picture, whereas the separation, $\epsilon$, between the centers of the two incoherent point sources affects the purity of the state.

Our aim is to use the qubit model to study the effects of misalignment, both in the estimation of the separation between two point sources, as well as in the task of discriminating between the single- and two-source hypotheses. We begin first with estimating the separation between two incoherent point sources.

### III. SEPARATION ESTIMATION UNDER MISALIGNMENT

In this section we review the quantum information tools for multiparameter estimation, after which we use the qubit model to derive the optimal measurement for estimating the separation between two incoherent point sources under misalignment.

#### A. Classical and quantum statistical inference

The task at hand is the estimation of two parameters: the two-source centroid position $x_c$, and their separation $d$ from a finite sample of $n$ measurement outcomes $y := (y_1, \ldots, y_n)^T, y_i \in \mathbb{R}$, in one dimension [48]. For ease of notation let us denote the parameters to be estimated by $\lambda := (\lambda_1, \lambda_2)^T \in \mathbb{R}^2$. Then the data constitute a random variable $y \in Y$ distributed according to $p(y|\lambda)$.

An estimator, $f_i : Y \to \mathbb{R}$, is any function that maps every possible measurement record to an estimate $\hat{\lambda}_i = f_i(y)$ of the parameter $\lambda_i$. An estimator is said to be unbiased if $(\hat{\lambda}_i) = \sum_i p(y|\lambda) \lambda_i = \lambda_i$. Denoting by $\hat{\lambda} \in \mathbb{R}^2$ the two-dimensional vector of estimates of $\lambda$, the Cramér-Rao inequality places a lower bound on the covariance matrix of any unbiased estimator [49]:

\[
(\hat{\lambda} - \lambda) \cdot (\hat{\lambda} - \lambda)^T \geq [nF(p(y|\lambda))]^{-1},
\]

where $F(p(y|\lambda))$ is the Fisher information matrix [50]

\[
F_{ij}(p(y|\lambda)) := \left(\frac{\partial \ln p(y|\lambda)}{\partial \lambda_j}\right)\left(\frac{\partial \ln p(y|\lambda)}{\partial \lambda_i}\right),
\]

quantifying the amount of information the random variable $Y$ carries about the parameters $\lambda$.

An estimator is said to be efficient if it saturates the inequality of Eq. (23). Note that it is possible that no efficient estimator exists if the data sample is finite. However, for an asymptotically large sample size, i.e., $n \to \infty$, it can be shown that the maximum-likelihood estimator always saturates the Cramér-Rao bound [51].
In quantum statistical inference the random variable $Y$ and its corresponding probability distribution $p(y|\lambda)$ arise from performing a quantum measurement on a quantum system. Any set of positive operators, $\{E_{y} \geq 0; y \in Y\}$, satisfying the completeness relation $\sum_{y \in Y} E_{y} = 1$ is an admissible measurement, termed a positive operator-valued measure, or POVM for short. By virtue of positivity $E_{y} = M_{y}^\dagger M_{y}$, where $M_{y}$ constitutes one of the infinitude of square roots of $E_{y}$. If $M_{y}^\dagger = M_{y}^\dagger$ and $M_{y}^2 = M_{y}$ then the POVM consists of projective operators, and there exists a dynamical variable—energy, position, (angular) momentum, etc.—represented by the Hermitian operator $O$, such that $O = \sum_{y} \mu_{y} M_{y}$. Given a POVM the conditional probability of obtaining a given measurement record $y$ is given by

$$p(y|\lambda) = \text{Tr}(E_{y} \rho(\lambda)).$$

Using the natural Riemannian geometry of the space of bounded, positive linear operators one can define the operator analog of the logarithmic derivative in Eq. (24) for each parameter $\lambda_i$—the symmetric logarithmic derivative (SLD), $L_{\lambda_i}$—as the solution to

$$\frac{\partial \rho(\lambda)}{\partial \lambda_i} := \frac{1}{2}(L_{\lambda_i} \rho(\lambda) + \rho(\lambda)L_{\lambda_i}).$$

In the eigendecomposition of $\rho(\lambda)$, $\{\mu_j, |\psi_j\rangle\}$, the SLD operator $L_{\lambda_i}$ is explicitly given by [52,53]

$$L_{\lambda_i} = 2 \sum_{\alpha, \beta} \mu_{\alpha} + \mu_{\beta} \neq 0 \times \langle \psi_{\alpha}(\lambda) | \partial_{\lambda_i} \rho | \psi_{\beta}(\lambda) \rangle \langle \psi_{\beta}(\lambda) | \psi_{\alpha}(\lambda) \rangle,$$

and the quantum Fisher information matrix elements read

$$\mathcal{F}_{ij}(\rho(\lambda)) = \frac{1}{4} \text{Tr}(\rho_{k}[L_{\lambda_i}, L_{\lambda_j}])(28),$$

where $[A, B] = AB + BA$. We thus have the following chain of inequalities for the covariance matrix:

$$\langle (\hat{\lambda} - \lambda) \cdot (\hat{\lambda} - \lambda)^\dagger \rangle \geq \frac{n F(p(y|\lambda))}{n F(p(y|\lambda))}^{-1} \geq \frac{n \mathcal{F}(\lambda)}{n \mathcal{F}(\lambda)}^{-1},$$

(29)

where the latter inequality is commonly referred to as the quantum Cramér-Rao bound.

For each single parameter $\lambda_i$ an asymptotically efficient estimator exists and is given by the maximum-likelihood estimator of the POVM whose elements are the eigenprojectors of the corresponding SLD operator. If all these operators commute, i.e., $[L_{\lambda_i}, L_{\lambda_j}] = 0, \forall i \neq j$, then the quantum Cramér-Rao bound is asymptotically achievable. Note that commutativity is only a sufficient condition; a necessary and sufficient condition—assuming asymptotically many independent and identically distributed copies ($n \gg 1$) of $\rho(\lambda)$—is $\text{Tr}(\rho(\lambda)[L_{\lambda_i}, L_{\lambda_j}]) = 0, \forall i \neq j$ [54]. However, note that the POVM that saturates the quantum Cramér-Rao bound in Eq. (29) may, in general, correspond to a collective measurement on all the $n \gg 1$ copies [55,56].

Hitherto, the application of superresolving measurements in imaging has focused primarily on “beating” the diffraction limit and maximizing the precision in estimating the source separation, $2d$, while assuming full control over all other parameters, particularly the centroid’s position, $x_c$. Of particular importance is the fact that the measurement that attains the quantum Fisher information when estimating only the separation between two incoherent point sources is a projective measurement that does not depend on knowing $d$ in advance [12]. It does, however, require perfect knowledge of the centroid, $x_c$, of the PSF as well as perfect positioning of SPADE so that any misalignment, $\delta \propto \theta$ in Eq. (18), can always be set to zero.

The separation can be estimated without requiring any knowledge about the centroid, if one has access to a quantum memory with a long coherence time so as to store photons collected during several independent experimental rounds ($n > 1$) and be able to implement collective measurements [54]. A proof-of-principle experiment that makes use of a measurement on a doublet of photons ($n = 2$) and allows for simultaneous estimation of both the centroid and the separation of the sources has been reported recently [34]. This has been achieved by encoding the spatial distribution of two incoherent sources into the spatial profile of a single photon generated in the laboratory. Utilizing a pair of such photons and interfering them as in the the Hong-Ou-Mandel experiment [57], the information about both separation and centroid parameters can be harmlessly retrieved, while estimating the former with precision beyond the diffraction limit [34]. On the other hand, a recent theoretical study has proposed the use of direct imaging and SPADE techniques in parallel [37]. Direct imaging is performed repeatedly on part of the incoming radiation, adjusting the exact position of SPADE via a servo feedback mechanism in order to gradually reduce the misalignment, $\theta$, in Eq. (18), with increasing number of experimental repetitions.

In the next section we use the qubit model to obtain the optimum measurement strategy for estimating the separation between two incoherent sources in the presence of misalignment.

**B. Separation estimation under misalignment in the qubit approximation**

Assuming the separation between the incoherent sources to be small—as ensured in the superresolution regime—we use the qubit model in order to construct the optimal measurement for estimating the separation between two point sources under misalignment. We begin by first considering the Gaussian PSF. The eigenvalues and corresponding eigenvectors of $\rho_C$ are

$$\mu_1(\epsilon) = \frac{\epsilon^2}{4}, \quad |\psi_1(\theta)| = \sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle,$$

$$\mu_2(\epsilon) = 1 - \mu_1(\epsilon), \quad |\psi_2(\theta)| = -\cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle.$$

(30)

Using Eq. (27) the corresponding SLD operators are, in the eigenbasis $\{|\psi_1(\theta)\rangle, |\psi_2(\theta)\rangle\}$,

$$L_\alpha = \left(1 - \frac{\epsilon^2}{2}\right) \sigma_x, \quad L_\epsilon = \epsilon \sigma_x^{\epsilon \epsilon - 4}.$$

(31)
Observe that $[\mathcal{L}_\theta, \mathcal{L}_\epsilon] \neq 0$, meaning that the optimal measurements for each of these parameters are incompatible. However, $\text{Tr}(\rho^{(1)}_\theta [\mathcal{L}_\theta, \mathcal{L}_\epsilon]) = 0$, which implies that there exists a possibly joint measurement on all $n$ photons that saturates the quantum Cramér-Rao bound by

$$
\langle (\hat{\theta} - \theta, \hat{\epsilon} - \epsilon)^T (\hat{\theta} - \theta, \hat{\epsilon} - \epsilon) \rangle \geq \frac{1}{n} \frac{1}{\epsilon^2} \left( 0 \ 0 \ \frac{1}{n} \right).
$$

The eigenvectors of the SLD operators [Eq. (31)] are given by

$$
|\theta\rangle = \frac{1}{\sqrt{2}} \left( \left( \sin \frac{\theta}{2} \pm \cos \frac{\theta}{2} \right) |0\rangle + \left( \sin \frac{\theta}{2} \mp \cos \frac{\theta}{2} \right) |1\rangle \right),
$$

$$
|\epsilon\rangle = |\psi_\epsilon(\theta)\rangle \quad \text{with} \quad \alpha \in \{1, 2\},
$$

respectively. As $\mathcal{L}_\epsilon$ is a diagonal operator, the optimal measurement in Eq. (34) for estimating the rescaled separation $\epsilon$ between the two sources according to the qubit model is simply given by a projective measurement in the eigenbasis of Eq. (30). Henceforth, we refer to this measurement as the rotated mode demultiplexer (ROTADE), i.e., the detection scheme depicted schematically in Fig. 1 with the rotation $R(\delta)$ adequately adjusted to $\delta = \theta \sigma$.

In order to compare the quality of the ROTADE measurement, we use Eq. (14) to map the measurement operators into their position-based representation. The latter are shown in Fig. 3. One can then explicitly determine the probability distribution arising from these measurements and hence the corresponding Fisher information using Eq. (24). The results are shown in Fig. 4, where we compare the performance of ROTADE with the quantum Fisher information attained in [12] (dashed lines) and classical Fisher information associated to the ROTADE measurement, $\mathcal{F}_\theta$, for separation $\epsilon$ (red) and misalignment $\theta$ (blue) parameters for perfect alignment $\theta = 0$, as a function of $2\epsilon$. As the POVM Eq. (34) is derived based on the qubit model, it ceases to be optimal with increasing separation of the sources (here for $\epsilon \lesssim 0.1$).

In order to capture the difference between the aforementioned measurements we compare the Taylor expansions of their corresponding Fisher information up to first nontrivial order, for small separation $\epsilon$. These are given by

$$
\mathcal{F}^{(R)}_{\epsilon\epsilon}(\epsilon) \approx \epsilon^2 C^{(R)}(\theta),
$$

$$
\mathcal{F}^{(B)}_{\epsilon\epsilon}(\epsilon) \approx \epsilon^2 C^{(B)}(\theta),
$$

where $C^{(R)}(\theta), C^{(B)}(\theta)$ are coefficients pertaining to the measurements themselves and depend only on the misalignment mode of SPADE are counted, while lumping all other modes to produce a single photon-count outcome. As the probability of detecting the zeroth HG mode occurs regardless of the separation, $d$, B-SPADE is more experimentally friendly but suffers in the same manner as SPADE from the misalignment problem. In Fig. 5 we compare the performance of ROTADE with B-SPADE in estimating the separation under a misalignment $\theta \leq 0.5$.

A simpler measurement that also achieves the quantum bound (Fig. 4) is B-SPADE [12]. This is a coarse-grained version of SPADE where only photons in the fundamental HG
of these coefficients governs the precise minimal resolvable distance for each measurement as we now explain.

The signal-to-noise ratio $\epsilon / \Delta \epsilon$ can be expressed as

$$\epsilon \sqrt{n C^{(\#)}(\theta)} \geq 1,$$  \hspace{1cm} (36)

where $\# \in (R, B)$. The minimal resolvable separation, $\epsilon_{\text{min}}^{(\#)}(\theta)$,
for each measurement is defined as that $\epsilon$ in Eq. (36) for which equality holds. Using the approximations of Eq. (35) one obtains

$$\epsilon_{\text{min}}^{(\#)}(\theta) = \frac{1}{\sqrt{n C^{(\#)}(\theta)}}.$$  \hspace{1cm} (37)

Taylor-expanding the functions $C^{(\#)}(\theta)^{-1}$ to first nontrivial order in $\theta$, one obtains

$$C^{(R)}(\theta)^{-1} \approx \frac{\theta^6}{12},$$

$$C^{(B)}(\theta)^{-1} \approx \theta^2.$$  \hspace{1cm} (38)

It follows that

$$\epsilon_{\text{min}}^{(R)}(\theta) \approx \frac{1}{\sqrt{n \theta^6}},$$

$$\epsilon_{\text{min}}^{(B)}(\theta) \approx \frac{1}{\sqrt{n \theta^2}},$$  \hspace{1cm} (39)

where $\epsilon_{\text{min}}(\theta) \propto n^{-1}$ is a consequence of $F_{\text{ee}} \propto \epsilon^2$ in Eq. (35). In contrast, observe that in the ideal case of no misalignment, for which $F_{\text{ee}} \propto 1$, the minimal resolvable distance scales as $\epsilon_{\text{min}}(0) \propto n^{-1}$. The quadratic increase in the scaling of $\epsilon_{\text{min}}$ for both RO- TADE and B-SPADE due to misalignment mimics closely the behavior of cross-talk between the measurement modes addressed recently by Guerssner et al. [29]. As our qubit approximation puts us in the regime of only monitoring the first two HG modes, and misalignment corresponds to a unitary rotation of the latter, it follows that this unitary rotation can be interpreted as the cross-talk matrix of [29]. As the cross-talk probability between the two modes is proportional to $\sin^2 \theta \approx \theta^2$, $\epsilon_{\text{min}}^{(R)}$ of Eq. (39) follows precisely the analytical model for uniform cross-talk of [29].

Our results show that superresolution is impossible if the initial demultiplexing of the incoming radiation suffers any misalignment, even if the latter is known. Nevertheless, cross-modulation techniques between the two primary HG modes can help in significantly reducing the minimal resolvable distance.

In Appendix A we obtain the optimal measurement under misalignment for the sinc PSF, as well as the minimum resolvable distance. Our results confirm the efficacy of the qubit model; for whatever PSF the first two modes are the most relevant ones in estimating the position of light sources with separation well below the diffraction limit. In the next section, we discuss how the optimal measurement under misalignment derived using the qubit model is also optimal for the task of discriminating whether the incoming radiation is due to two incoherent point sources or one source with twice the power under misalignment.

IV. CLASSICAL AND QUANTUM STATE DISCRIMINATION: ONE OR TWO POINT SOURCES

Hitherto our focus was to estimate the relevant parameters of two incoherent point sources. However, a more pertinent question is whether the incoming radiation is due to two incoherent point sources very close together (the two-source hypothesis $H^{(2)}$), or one point source with twice the power (the one-source hypothesis, $H^{(1)}$). To that end we first review the fundamentals of classical and quantum decision theory and, in particular, simple binary hypothesis testing [52, 53]. We then apply these tools to optimally discriminate between $H^{(1)}$ and $H^{(2)}$ in the presence of misalignment and compare the performance of ROTADE with measurements in the literature, showing that our measurement outperforms all the latter.

A. Classical and quantum hypothesis testing

A fundamental problem in decision theory is to discriminate among several possible hypothesis based on a number, $n$, of observations. The simplest such scenario—known as binary hypothesis testing—occurs when there are two hypotheses, $H^{(1)}$ and $H^{(2)}$, that need to be discriminated. For simplicity, assume that each observation consists of a finite set of possible outcomes $y$ in $Y$ [58]. Under hypothesis $H^{(i)}$, these outcomes are distributed according to $p(y | H^{(i)})$, and thus the problem becomes one of determining from which probability distribution the random variable $Y$ is drawn.

For a single observation ($n = 1$) let $f : Y \rightarrow \{ H^{(1)}, H^{(2)} \}$ be a decision rule. Under such a decision rule the probability of making an error based on a single observation is

$$P_{\text{err}} = \frac{1}{2} [ p(f(y) = H^{(2)} | H^{(1)}) + p(f(y) = H^{(1)} | H^{(2)}) ],$$  \hspace{1cm} (40)

where we have assumed that each hypothesis is equally likely. The conditional probabilities $p(f(y) = H^{(2)} | H^{(1)})$ and $p(f(y) = H^{(1)} | H^{(2)})$ are the type-1 (mistaking one source for two) and type-2 (mistaking two sources for one) errors, respectively. For binary hypothesis testing, the optimal decision rule is to assign the hypothesis with the highest posterior distribution [59, 60], which, for equally likely hypotheses, translates to

$$f(y) = \begin{cases} H^{(1)} & \text{if } p(y | H^{(1)}) > p(y | H^{(2)}) \\ H^{(2)} & \text{if } p(y | H^{(2)}) > p(y | H^{(1)}) \\ \text{any} & \text{if } p(y | H^{(1)}) = p(y | H^{(2)}) \end{cases}$$  \hspace{1cm} (41)

and the corresponding probability of error reads

$$P_{\text{err}} = \sum_{y \in Y} p(y) \min[p(H^{(i)} | y)]$$

$$= \sum_{y \in Y} \min[p(y, H^{(i)})]$$

$$= \frac{1}{2} \left[ 1 - \frac{1}{2} \sum_{y \in Y} | p(y | H^{(1)}) - p(y | H^{(2)}) | \right],$$  \hspace{1cm} (42)

where we have made use of the identity $\min[a, b] = \frac{1}{2} (a + b - |a - b|)$ in order to obtain the last equality.

Quantum hypothesis testing now follows by noting that $p(y | H^{(i)}) = \text{Tr}(E_i \rho^{(i)} )$, where $\{ E_i \}$ constitutes a POVM and the hypotheses, $\rho^{(i)}$, $i \in (1, 2)$, are given by Eqs. (8) and (9).
Doing the appropriate substitutions in Eq. (42), one obtains

\[ P_{\text{err}} = \frac{1}{2} \left( 1 - \frac{1}{2} \text{Tr} \left( \sum_{y \in Y} E_y (\rho^{(1)} - \rho^{(2)}) \right) \right). \]  

(43)

Unlike the classical case, in quantum binary hypothesis testing we are free to choose among all admissible POVMs the one that yields the smallest probability of error. The optimal measurement in this case was derived by Helstrom [52] and corresponds to a two-outcome measurement \( E_0, E_1 \) on the positive and negative eigenspaces of the operator

\[ \Gamma := \frac{1}{2} (\rho^{(2)} - \rho^{(1)}). \]  

(44)

Given \( n \) copies of the initial state, the Helstrom measurement is generally a collective measurement on the positive and negative eigenspaces of \( \Gamma^{\otimes n} = \frac{1}{2} (\rho^{(2)^{\otimes n}} - \rho^{(1)^{\otimes n}}) \). For clarity we call the single-copy optimal measurement the Helstrom measurement, and the overall optimal measurement on \( n \) copies the collective Helstrom measurement.

The probability of error decreases exponentially with the number of copies, \( n \). In order to compare the performance of different measurement strategies one needs to determine the rate at which this error probability decreases. For an asymptotically large (\( n \to \infty \)) number of observations the probability of error saturates Chernoff’s inequality [61]:

\[ P_{\text{err}}(n) \leq e^{-n \xi}, \]  

(45)

where

\[ \xi := -\log_2 \min_{0 \leq s \leq 1} \sum_{y \in Y} p(y) |H^{(1)}_s y\rangle \langle p(y) |H^{(2)}_s 1-y\rangle^{1-s} \]  

(46)

is the Chernoff exponent. In the case of quantum hypothesis testing, the asymptotic error rate is given by the quantum Chernoff exponent [62]:

\[ \xi \leq \xi^{(QM)} := -\log_2 \min_{0 \leq s \leq 1} \text{Tr} ((\rho^{(1)})^s (\rho^{(2)})^{1-s}), \]  

(47)

which is generally larger than its classical counterpart. Note that the quantum Chernoff exponent only depends on the quantum states to be discriminated, and is independent of the measurement performed. Nonetheless, the inequality in Eq. (45) is asymptotically achievable in the limit of infinite copies. In this limit, \( \xi \) reaches the ultimate quantum bound \( \xi^{(QM)} \) of asymptotic (symmetric) hypothesis testing [63]. However, such attainability may require a collective Helstrom measurement to be performed on all the \( n \to \infty \) copies.

Surprisingly, it was already Helstrom [64] who first addressed the problem of discriminating one versus two incoherent point sources of light with tools from hypothesis testing and derived a suboptimal measurement that (i) lacks a physical realization and (ii) requires knowledge of the separation of the two sources. Krovi et al. [38] derived the optimal quantum mechanical measurement that achieves the quantum Chernoff bound for the case where the separation of the two point sources is known and showed how to experimentally implement it. Shortly after, Lu et al. [26] showed that the B-SPADE measurement of [12] achieves the quantum Chernoff bound for one versus two sources of arbitrary separation. However, just like in the estimation case, all these works assumed that the center of the single source, as well as the centroid of the two-source hypothesis, is perfectly aligned with the demultiplexing measurements and neglected any noise at the detectors.

In the next section we analyze the behavior of B-SPADE under misalignment and show that it falls short of the quantum optimal Chernoff bound. Using the qubit model we derive an alternative measurement strategy that is also suboptimal but outperforms the B-SPADE under misalignment by far.

### B. State discrimination in the qubit approximation: The Helstrom measurement

Our aim is to determine whether the PSF observed at the misaligned imaging system is due to two incoherent point sources of equal intensities or a single source with twice the intensity. For the remainder of this section, we work with the Gaussian PSF (results for the sinc PSF can be derived in a similar fashion and are presented in Appendix B). Using the qubit model the matrix \( \Gamma \) of Eq. (44) can be explicitly computed to be

\[ \Gamma = \frac{1}{4} \left( -\cos \theta_0 - \frac{1}{2} \cos \theta_0 (e^2 - 2) \right) \]  

(48)

\[ \sin \theta_0 + \frac{1}{2} \sin \theta_0 (e^2 - 2), \]

\[ \sin \theta_0 + \frac{1}{2} \cos \theta_0 (e^2 - 2) \]

where \( \theta_0 = \frac{\pi \epsilon}{\sigma} \) is the misalignment relative to the center of a single-source PSF, \( \theta_c = \frac{\pi \epsilon}{\sigma} \) is the misalignment relative to the centroid, \( x_c \), of the two-source PSF, and \( \epsilon \) is defined as in Eq. (18). Notice that, in principle, the center of a single source need not coincide with the centroid of two sources, nor with the position of the demultiplexing measurement, \( x_0 \neq x_c \neq x_g \) (\( \theta_0 \neq \theta_c \)). Nonetheless, hereafter we restrict our analysis to the case where only the demultiplexing measurements are misaligned; hence we define

\[ \theta := \theta_0 = \theta_c. \]  

(49)

In this regime, the Helstrom measurement is independent of separation and is equivalent to ROTADE.

In case the detector and centroid are perfectly aligned, \( \theta = 0 \), ROTADE is only the projection onto the zeroth and first HG modes. We refer to this measurement as SPADE01, in order to distinguish it from B-SPADE, which projects only on the zeroth mode. We remark that all measurement strategies reach the quantum bound for zero misalignment. The main advantages of SPADE01 for aligned measurement devices are that it is independent of the two-source separation, it is only necessary to count photons in the first two HG modes (photons coupling to higher modes correspond to no-clicks and are
TABLE I. Taylor expansion to the first nontrivial order in $\theta$ for the type-1 (second column) and type-2 (third column) error probabilities for ROTADE, SPADE01, and B-SPADE.

| Measurement | $p(f(y) = H(2)|H(1))$ | $p(f(y) = H(1)|H(2))$ |
|-------------|-------------------------|-------------------------|
| ROTADE      | $\frac{\sigma^2}{4\pi\sigma^2}$ | $\exp\left(-\frac{\sigma^2}{4\pi}\right)(1 - \frac{\sigma^2}{4\pi})$ |
| SPADE01     | $\frac{\sigma^2}{4\pi\sigma^2}$ | $\exp\left(-\frac{\sigma^2}{4\pi}\right)(1 + \frac{\sigma^2}{4\pi})$ |
| B-SPADE     | $\frac{d^2\rho_{\text{coh}}}{d\theta^2}$ | $\exp\left(-\frac{\sigma^2}{4\pi}\right)(1 + \frac{\sigma^2}{4\pi})$ |

insignificant to the measurement statistics, and there is an unambiguous two-source discrimination whenever a photon is detected in the first HG mode. These results are shown in Appendix C.

Table I shows how the one-shot error probability scales as a function of the misalignment for the first nontrivial order of the Taylor expansion around $\theta = 0$. Notice that for ROTADE, the type-1 error, responsible for the unambiguous determination of the two-source hypothesis, is four orders of magnitude smaller compared to that of SPADE01 and B-SPADE. Hence in the single-shot scenario ROTADE significantly outperforms both these measurements.

The Chernoff exponent of the SPADE01 measurement under misalignment behaves similarly to that of B-SPADE; the asymptotic results of all measurement strategies under misalignment as a function of separation are represented in Fig. 6. However, in contrast with the aligned scenario, for $\theta \neq 0$ the probability of detecting photons into higher HG modes is non-negligible, and corresponds to the no-click probability. This probability represents the intrinsic error of the qubit model and it increases with misalignment (for details see Appendix C).

Unfortunately, we are unable to obtain an analytic expression for the Chernoff exponent under misalignment for any of the three strategies. This is because the $s$ that minimizes the Chernoff exponent in Eq. (46) explicitly depends on $\theta$. Figure 7 presents a numerical optimization for the Chernoff exponent as a function of the misalignment. We observe that for all $\theta > 0$ ROTADE outperforms both SPADE01 and B-SPADE, which is to be expected as ROTADE includes the knowledge on the amount of misalignment. Nonetheless, for exactly $\theta = 0$ all the corresponding Chernoff exponents coincide with the quantum bound, which manifests their discontinuity as $\theta \to 0$.**

V. CONCLUSIONS

In this work we have analyzed the impact of a de facto misalignment in the demultiplexing measurements of an optical imaging system in both estimating the separation of incoherent light sources, as well as discriminating between one and two incoherent point sources. By allowing for linear-optical postprocessing of the two dominant demultiplexed modes, we have shown that superresolution cannot be perfectly restored, even if the value of the misalignment is a priori known. Using quantum information methods, we have constructed misalignment-dependent strategies that, in the case of estimation, allow for subdiffraction-limited estimation of the separation of two incoherent light sources and analytically determined the dependence of both the estimation precision as well as the minimal resolvable distance as a function of the misalignment. Remarkably, the same measurement exhibits improved performance also in the task of discriminating between the one- and two-source hypotheses, showing significant improvement in both the single-shot as well as asymptotic probability of error.

Several interesting questions still remain. How does misalignment affect estimation precision when both the separation as well as the relative intensities of the two incoherent point sources need to be estimated? In the case of discrimination an interesting question occurs when the centers of the two hypotheses do not coincide, i.e., $x_0 - x_c \ll \sigma^2$ but neither $x_0$ nor $x_c$ coincides with $x_R$. Then, the optimal Helstrom measurement does depend on knowing the separation between the two sources and it remains an open question if there exists a classical measurement with superresolving power. We hope to answer these questions in the future.

ACKNOWLEDGMENTS

We thank Wojciech Wasilewski and Michal Parniak for helpful comments. J.O.d.A. and M.L. acknowledge support from ERC AdG NOQIA, Spanish Ministry of Economy and Competitiveness ("Severo
In this Appendix, we present the results of estimating the separation between two incoherent point sources imaged by a system with a rectangular aperture. The PSF of such a system is given by the sinc function [see Eq. (3)].

Repeating the calculation in Sec. III A the eigenvalues and corresponding eigenvectors of $\rho^{(2)}_S$ are

$$\mu_1(\epsilon) = \frac{\epsilon^2}{3}, \quad |\psi_1(\theta)| = \sin \frac{\theta}{\sqrt{3}} |0\rangle + \cos \frac{\theta}{\sqrt{3}} |1\rangle,$$

$$\mu_2(\epsilon) = 1 - \mu_1(\epsilon), \quad |\psi_2(\theta)| = -\cos \frac{\theta}{\sqrt{3}} |0\rangle + \sin \frac{\theta}{\sqrt{3}} |1\rangle,$$

and using Eq. (27) the corresponding SLD operators in the eigenbasis $\{|\psi_1(\theta)\rangle, |\psi_2(\theta)\rangle\}$ are given by

$$\mathcal{L}_\theta = \left( \frac{6 - 4\epsilon^2}{3\sqrt{3}} \right) \sigma_x,$$

$$\mathcal{L}_\epsilon = \frac{2}{\epsilon} \left( \frac{1}{\epsilon^2} \frac{1}{\epsilon^2} \right).$$

The eigenvectors of the SLD operators can now easily be computed to be

$$|\theta_{\epsilon}\rangle = \frac{1}{\sqrt{2}} \left( \left( \sec \frac{2\theta}{\sqrt{3}} \pm \tan \frac{2\theta}{\sqrt{3}} \right) \sqrt{1 \mp \sin \frac{2\theta}{\sqrt{3}}} |0\rangle \right.
\left. + \sqrt{1 \pm \sin \frac{2\theta}{\sqrt{3}}} |1\rangle \right)$$

$$|\epsilon_{\alpha}\rangle = |\psi_\alpha(\theta)\rangle,$$

and using Eq. (27) the corresponding SLD operators in the eigenbasis $\{|\psi_1(\theta)\rangle, |\psi_2(\theta)\rangle\}$ are given by

$$\mathcal{L}_\theta = \left( \frac{6 - 4\epsilon^2}{3\sqrt{3}} \right) \sigma_x,$$

$$\mathcal{L}_\epsilon = \frac{2}{\epsilon} \left( \frac{1}{\epsilon^2} \frac{1}{\epsilon^2} \right).$$

The eigenvectors of the SLD operators can now easily be computed to be

$$|\theta_{\epsilon}\rangle = \frac{1}{\sqrt{2}} \left( \left( \sec \frac{2\theta}{\sqrt{3}} \pm \tan \frac{2\theta}{\sqrt{3}} \right) \sqrt{1 \mp \sin \frac{2\theta}{\sqrt{3}}} |0\rangle \right.
\left. + \sqrt{1 \pm \sin \frac{2\theta}{\sqrt{3}}} |1\rangle \right)$$

$$|\epsilon_{\alpha}\rangle = |\psi_\alpha(\theta)\rangle,$$
Similarly to the results in the main text, we verify in the asymptotic limit that ROTADE performs better than SPADE01.

APPENDIX C: PERFORMANCE OF ROTADE IN DISCRIMINATION

Here we analyze the performance of ROTADE for the task of discriminating one and two light sources. As ROTADE involves only three two-dimensional subspaces spanned by the zeroth and first HG modes, an intrinsic error probability arises when the incoming radiation couples into higher HG modes. This probability is useful for defining the regime of validity of the qubit model.

For example, Fig. 11 presents the error and success probabilities in the regime where the center of each distribution is aligned: $\theta = 0$. We observe that ROTADE has constant value (less than 4% variation); e.g., at $\theta = 0$ the error probability has value $P_{\text{err}} = \frac{1}{2}(P_{\text{err}1} + P_{\text{err}2}) = \frac{1}{2}(0 + e^{\frac{-d^2}{4\sigma^2}})$, and the success probability $P_{\text{suc}} = \frac{1}{2}(P_{\text{suc}1} + P_{\text{suc}2}) = \frac{1}{2}(1 + e^{\frac{-d^2}{4\sigma^2}})$. As $d$ increases, the likelihood that photons couple to higher HG modes increases and hence the error (success) probability moves further away from the priors, 0.5. This is a consequence of the intrinsic error of the qubit model.

The intrinsic error is the distance between the sum of the error and success probabilities from unity. It dictates until which separation and reference position the qubit model—and consequently ROTADE—are adequate. For $|\theta| < \frac{1}{2}$, or when the separation between the sources is comparable to $\sigma$, $\epsilon < \frac{1}{2}$, this error is negligible. These features are presented in Figs. 12 and 13, respectively.

In Fig. 12 we present the intrinsic error in function of the misalignment $\theta$. For a range of misalignments, $|\theta| < \frac{1}{2}$, ROTADE has negligible intrinsic error. Figure 13 shows the intrinsic error in function of the two-source separation $\epsilon$, for misaligned source distributions; i.e., the centroid of the two sources is different from the center of one source ($x_c \neq x_0$). We observe that the qubit model is adequate when placing the measurement in between the distribution centroids $x_c \leq x_R \leq x_0$ (in between red and orange lines) and the intrinsic error of the model is minimum when $\theta_0 = \theta_c$, i.e., when the centers of the two distributions coincide. Notice that when the centroids of the two distributions do not coincide the ROTADE measurement will, in general, depend on the separation of the two-source hypothesis.