ORIGINAL ARTICLE

A stochastic analysis of performance limits
for optical microscopes

Sripad Ram · E. Sally Ward · Raimund J. Ober

Abstract   The optical microscope is a powerful instrument for observing cellular events. Recently, the increased use of microscopy in quantitative biological research, including single molecule microscopy, has generated significant interest in determining the performance limits of an optical microscope. Here, we formulate this problem in the context of a parameter estimation approach in which the acquired imaging data is modeled as a spatio-temporal stochastic process. We derive formulations of the Fisher information matrix for models that allow both stationary and moving objects. The effects of background signal, detector size, pixelation and noise sources are also considered. Further, formulations are given that allow the study of defocused objects. Applications are discussed for the special case of the estimation of the location of objects, especially single molecules. Specific emphasis is placed on the derivation of conditions that guarantee block diagonal or diagonal Fisher information matrices.

Keywords   Spatio-temporal stochastic processes · Fisher information matrix · Cramer-Rao lower bound · Parameter estimation · Fluorescence microscopy · Optical imaging · Single Molecule Microscopy · Localization accuracy

S. Ram · E. S. Ward · R. J. Ober (✉)
Center for Immunology, NB9.106
University of Texas Southwestern Medical Center at Dallas
6000 Harry Hines Boulevard,
Dallas, TX 75390-8576, USA
e-mail (Sally Ward): sally.ward@utsouthwestern.edu

S. Ram
Joint Biomedical Engineering Graduate Program,
University of Texas at Arlington and
University of Texas Southwestern Medical Center at Dallas, Texas
e-mail: sripad.ram@utsouthwestern.edu

R. J. Ober
Department of Electrical Engineering,
University of Texas at Dallas Richardson,
TX 75083-0688, USA
e-mail: ober@utdallas.edu

Introduction

The optical microscope has been an invaluable tool in cell biological research (Michalet et al., 2003). In the more recent past, advances in photodetector technology, high speed data acquisition and fluorescent labeling techniques have significantly enhanced the capabilities of this instrument. Presently, optical microscopes perform tasks ranging from long term (hours time scale) three dimensional imaging of live cells to fast imaging (milliseconds time scale) of molecular interactions within a cellular environment even at the single molecule level (see e.g., Weiss, 1999; Moerner & Fromm, 2003; Ober, Martinez, Lai, Zhou & Ward, 2004).

In several biological applications such as single molecule studies, the data acquired through an optical microscope requires extensive quantitative analysis (e.g., see Saxton & Jacobson, 1997). In order to carry out such studies, it is important for an experimenter to know the capabilities of the instrument. This not only provides insight into determining the feasibility of a particular experiment, but it also helps in designing an optimal experimental setup.

In this paper we present results to calculate performance limits that quantify the capabilities of an optical microscope. Due to the random nature of the acquired data, we use the tools of statistical estimation theory (see e.g., Kay, 1993) to determine the performance limits. We also show how experimental factors such as background noise, detector size, detector shape etc., affect the performance of an optical microscope. The present results are applicable to several microscopic techniques such as fluorescence microscopy, brightfield/transmitted-light microscopy, etc.

Parameter estimation problems play an important role in single molecule microscopy. Examples relate to determining the location of single molecules (Ober, Ram, & Ward, 2004), the determination of the photon detection rate, the estimation of the level of defocus etc (Ram, Ward, & Ober, 2005). Common to all these problems is that it is helpful for the experimenter to have an analytical method to assess with which accuracy the various parameters can be estimated. The approach that we take is based on calculating the Fisher information matrix/Cramer Rao lower bound for the corresponding estimation problems. In Ober, Ram et al., (2004) we have analyzed the localization problem for in focus stationary single molecules using the methodology that is used here. In the current paper we significantly expand on the scope of the approach and address more general parameter estimation problems that include the localization problem but are not limited to it. Importantly from the point of view of applications we also consider time-varying problems, such as when the imaged object is not stationary. We also investigate under which conditions the Fisher information matrix is diagonal and determine in detail the influence of the detector size on the accuracy with which a parameter can be estimated.

It should be pointed out that there have been several reports that addressed specific problems such as the localization accuracy of point sources for a particular estimation technique (see e.g., Bobroff, 1986). In Winick (1986) the Cramer Rao lower bound was calculated for a position estimation problem in the special case where the image of the object is given by a Gaussian function and the detector is pixelated.

The organization of the paper is as follows. In the Section General stochastic framework we present the statistical description of the acquired data. In the Section Performance limits and Fisher information matrix we derive general expressions for the Fisher information matrix relating to the parameter estimation problem. In the Section Image function we derive expressions for the Fisher information matrix that are of relevance when the optical system is spatially shift invariant and an image function can be assumed to exist. In the Section Effects

of reduced detector size we show how the detector size affects the performance of an optical microscope. Throughout the paper we provide examples to illustrate our results with specific profiles that describe the image of a point source.

General stochastic framework

A basic optical microscope setup consists of an object located in the object space, a lens system and a detector in the image space that captures the image of the object (see Fig. 1). For example, the object could be an individual point source (e.g., a fluorescent single molecule or a fluorescent nano-particle), a collection of two or more point sources, or a fluorescently labelled cellular organelle. Here, we are primarily interested in experiments in which the detector detects photons from the object of interest for a fixed acquisition time. Since the photon detection process is inherently a random phenomenon (see e.g., Saleh, 1978), the recorded image of the object is stochastic in nature.

We assume that the acquired data consists of the spatial coordinates of the arrival location of the detected photons on the detector and the time points at which the photons are detected. In a typical quantitative experiment, some attributes of the object such as the location, distance of separation from other objects, orientation, size etc., are determined from the acquired data by using a specific estimation procedure. The accuracy of the estimates can be determined by calculating the standard deviation of the estimates of this attribute assuming repeated experiments.

In any estimation problem, it is important to know whether the specific estimation technique used to estimate the desired attribute indeed comes close to the best possible accuracy. This can be determined by calculating the Fisher information matrix (Zacks, 1971;
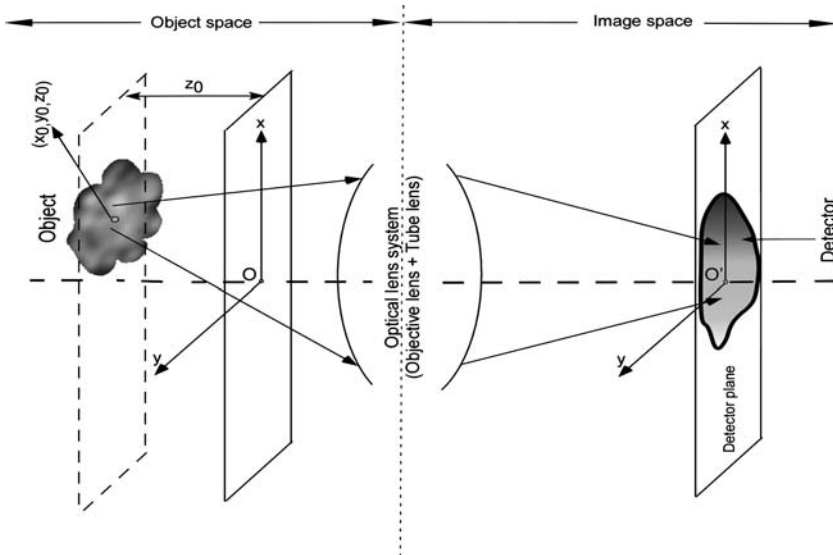

Fig. 1 The schematic shows the main components of an optical microscope based imaging setup. Here, an object located in the object space is imaged by an optical lens system and the image of the object is captured by the detector that is located in the image space. The location of the object in the object space is denoted by (x_0, y_0, z_0)

Kay, 1993) for the underlying random process that characterizes the acquired data. According to the Cramer-Rao inequality (Zacks, 1971; Kay, 1993), the (co)variance (matrix) of any unbiased estimator $\hat{\theta}$ of an unknown vector parameter θ is bounded from below by the inverse of the Fisher information matrix $\mathbf{I}(\theta)$, i.e.,

$$\text{Cov}(\hat{\theta}) \geq \mathbf{I}^{-1}(\theta).$$

Since we have defined the accuracy of an estimator in terms of its standard deviation, the square root of the inverse Fisher information matrix provides a lower bound to the best possible accuracy. It is important to note that the Fisher information matrix is independent of the estimation procedure used to estimate the parameter θ and only depends on the statistical nature of the acquired data. For instance, if the desired attribute is the location of an object, then the above equation implies that for any (asymptotically) unbiased estimator of the location, the accuracy of its location estimates can never be smaller than the square root of the inverse Fisher information matrix. Therefore, the square root of the inverse Fisher information matrix provides a *limit* to the accuracy with which the location of the object can be determined. Generalizing this, in an optical microscope the performance limit in determining a specific attribute of an object is defined as the square root of the inverse Fisher information matrix calculated for that attribute.

Due to its stochastic nature, the acquired data is modeled as a space-time random process (see e.g., Snyder & Miller, 1991) which we refer to as the image detection process $\mathcal{G}$. The temporal part of $\mathcal{G}$ describes the time points of the detected photons and is modeled as a temporal Poisson process with intensity function Λ_θ. The spatial part of $\mathcal{G}$ describes the spatial coordinates of the arrival location of the detected photons and is modeled as a family of mutually independent random variables $\{U_\tau\}_{\tau \geq t_0}$ with probability densities $\{f_{\theta,\tau}\}_{\tau \geq t_0}$ defined on the detector $\mathcal{C}$, where τ denotes the time point of a detected photon. The time dependence of the random variables $\{U_\tau\}_{\tau \geq t_0}$ denotes the fact that the spatial distribution of the detected photons can change with time. For example, this is the case when photons from a moving object are detected. In some applications the spatial part of $\mathcal{G}$ is independent of τ and in that case the random variables are independent and identically distributed. In all cases, we assume that the spatial and temporal parts of $\mathcal{G}$ are mutually independent of each other. We note that the probability density $f_{\theta,\tau}$ satisfies certain regularity conditions that are necessary for the calculation of the Fisher information matrix (see Zacks, 1971).

Definition 1 Let $\mathcal{C}$ denote a *detector*, i.e., an open subset of $\mathbb{R}^2$ with non-zero Lebesgue measure. Let Θ denote the *parameter space* that is an open subset of $\mathbb{R}^n$ and let $t_0 \in \mathbb{R}$. For $\theta \in \Theta$, an *image detection process* $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ is defined as a spatio-temporal process whose temporal part describes the time points of the photons detected on the detector $\mathcal{C}$ and the spatial part describes the spatial coordinates of the arrival location of the photons detected on the detector $\mathcal{C}$.

The temporal part is modeled as a Poisson process $\{Z(\tau); \tau \geq t_0\}$ with intensity Λ_θ, called the *photon detection rate*, such that

(C1) $\Lambda_\theta(\tau)$ is piecewise continuously differentiable with respect to θ for each $\tau \geq t_0$.

(C2) $\Lambda_\theta(\tau)$ is piecewise continuous with respect to τ for each $\theta \in \Theta$.
Let $\mathcal{F}_\Theta$ be the set of probability densities f_θ on $\mathcal{C}$ parameterized by θ that satisfy the following regularity conditions

(C3) $\partial f_\theta(r)/\partial \theta_i$ exists for $r \in \mathcal{C}$, $i = 1, \ldots, n$ and $\theta \in \Theta$.

(C4) $\int_{\mathcal{C}} |\partial f_\theta(r)/\partial\theta_i|\,dr < \infty$ for $i = 1, \ldots, n$ and $\theta \in \Theta$.

(C5) The integral $\displaystyle\int_{\mathcal{C}} \frac{1}{f_\theta(r)} \frac{\partial f_\theta(r)}{\partial\theta_i} \frac{\partial f_\theta(r)}{\partial\theta_j}\,dr$ exists and is finite for $i, j = 1, \ldots, n$ and $\theta \in \Theta$.

The spatial part of the image detection process is modeled as a family of mutually independent random variables $\{U_\tau\}_{\tau \geq t_0}$ that is assumed to be independent of $\{Z(\tau); \tau \geq t_0\}$. The corresponding family of probability densities $\{f_{\theta,\tau}\}_{\tau \geq t_0} \subseteq \mathcal{F}_\Theta$ is called the *photon distribution profile*, if U_τ with probability density $f_{\theta,\tau}$ describes the spatial distribution of the location of the point of detection of a photon on the detector $\mathcal{C}$ that is detected at time τ, $\tau \geq t_0$.

We next derive an expression for the Fisher information matrix of the image detection process $\mathcal{G}$.

Theorem 2 *Let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ be an image detection process. Then for $\theta \in \Theta$ the Fisher information matrix $\mathbf{I}(\theta)$ of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by*

$$
\begin{aligned}
\mathbf{I}(\theta) &= \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial\theta}\right)^T \left(\frac{\partial \Lambda_\theta(\tau)}{\partial\theta}\right) d\tau \\
&\quad + \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda_\theta(\tau)}{f_{\theta,\tau}(r)} \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta}\right)^T \left(\frac{\partial f_{\theta,\tau}(r)}{\partial\theta}\right) dr\,d\tau \\
&= \int_{\mathcal{C}} \int_{t_0}^{t} \frac{1}{\Lambda_\theta(\tau)f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau)f_{\theta,\tau}(r)]}{\partial\theta}\right)^T \left(\frac{\partial[\Lambda_\theta(\tau)f_{\theta,\tau}(r)]}{\partial\theta}\right) dr\,d\tau.
\end{aligned}
$$

Proof This result is a generalization of the Fisher information matrix for a spatio-temporal random process whose temporal component is a Poisson process and the spatial component is independent of the time points (see Snyder & Miller, 1991, p. 213). Due to space limitations the proof of this generalization is omitted, but can be found in Ram (in preparation). □

In deriving the above result we made no specific assumptions about the geometry of the imaging setup or the analytical expression for the photon distribution profile f_θ. Hence the above theorem provides a general result to calculate the Fisher information matrix for a wide range of situations. Note that the two-term expression of $\mathbf{I}(\theta)$ shows explicitly the dependence of $\mathbf{I}(\theta)$ on the temporal and spatial components of the image detection process.

We next consider the superposition of two image detection processes. In many concrete situations the detected photons originate from different sources. For example, the detected photons can result from a background component in addition to those detected from the object of interest. In an incoherent imaging setup, such as in fluorescence microscopy, the photon detection process that describes the collection of all the detected photons is then the superposition of the object and the background image detection process.

Theorem 3 *Let $\mathcal{G}^1(\Lambda_\theta^1, \{f_{\theta,\tau}^1\}_{\tau \geq t_0}, \mathcal{C})$ and $\mathcal{G}^2(\Lambda_\theta^2, \{f_{\theta,\tau}^2\}_{\tau \geq t_0}, \mathcal{C})$ be two independent image detection processes. Then the superposition process is an image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ whose photon detection rate Λ_θ is given by*

$$
\Lambda_\theta(\tau) = \Lambda_\theta^1(\tau) + \Lambda_\theta^2(\tau), \quad \tau \geq t_0, \quad \theta \in \Theta, \tag{1}
$$

and the photon distribution profile $\{f_{\theta,\tau}\}_{\tau \geq t_0}$ is given by

$$
f_{\theta,\tau}(r) = \epsilon_\theta^1(\tau) f_{\theta,\tau}^1(r) + \epsilon_\theta^2(\tau) f_{\theta,\tau}^2(r), \quad r \in \mathcal{C}, \quad \theta \in \Theta, \quad \tau \geq t_0, \tag{2}
$$

where $\epsilon_\theta^1(\tau) = \Lambda_\theta^1(\tau)/\Lambda_\theta(\tau)$, $\epsilon_\theta^2(\tau) = \Lambda_\theta^2(\tau)/\Lambda_\theta(\tau)$, $\theta \in \Theta, \tau \geq t_0$.

Proof This result is analogous to the result of the superposition of Poisson processes (see e.g., Snyder & Miller, 1991). For a detailed proof please see Ram (in preparation). □

Performance limits and Fisher information matrix

Using the results derived in the previous section, we next obtain expressions of the Fisher information matrix for a parameter estimation problem in which the parameter vector $\theta :=$ $(\theta_l, \theta_a, \theta_\Lambda)$ is decomposed into three components, where θ_l denotes the location component, θ_a denotes an auxiliary component and θ_Λ denotes the rate component. The location component θ_l typically consists of the x and y coordinates of the object location. The auxiliary component θ_a, if present, may consist of other relevant parameters such as the z coordinate of the object location. In single molecule microscopy it is often reasonable to assume that the photon detection rate is a constant, i.e., $\Lambda_\theta(\tau) = \Lambda_0$, $\tau \geq t_0$ and that the intensity level Λ_0 needs to be estimated. In this case θ_Λ would consist of the parameter Λ_0. For example, in tracking problems for vesicles, due to photobleaching effects the photon detection rate can often be modeled as $\Lambda_\theta(\tau) = \Lambda_0 \exp(-(\tau - t_0)k_b)$, $\tau \geq t_0$. In this case θ_Λ would consist of the parameters Λ_0 and k_b.

In the following theorem, we consider two independent image detection processes $\mathcal{G}^1$ and $\mathcal{G}^2$. The image detection process $\mathcal{G}^1$ is such that its photon detection rate Λ_θ^1 only depends on the rate component θ_Λ, and its photon distribution profile $f_{\theta,\tau}^1$ only depends on the location component θ_l and the auxiliary component θ_a. The image detection process $\mathcal{G}^2$ is such that its photon detection rate Λ^2 and photon distribution profile f_τ^2 are independent of θ. For example, $\mathcal{G}^1$ can model the detected photons from the object of interest, whereas $\mathcal{G}^2$ might model a background component.

Theorem 4 *Let $\mathcal{G}^1(\Lambda_\theta^1, \{f_{\theta,\tau}^1\}_{\tau \geq t_0}, \mathcal{C})$ and $\mathcal{G}^2(\Lambda^2, \{f_\tau^2\}_{\tau \geq t_0}, \mathcal{C})$ be two independent image detection processes such that $\mathcal{G}^2$ is independent of θ, $\theta \in \Theta$. Let $\mathcal{G}$ be the superposition of $\mathcal{G}^1$ and $\mathcal{G}^2$. Assume that*

(A1) *for $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$ and $\tau \geq t_0$, $\partial f_{\theta,\tau}^1(r)/\partial \theta_\Lambda = 0, r \in \mathcal{C}$, $\partial \Lambda_\theta^1(\tau)/\partial \theta_l = 0$, $\partial \Lambda_\theta^1(\tau)/\partial \theta_a = 0$.*

Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & \mathbf{I}_{l,a}(\theta) & \mathbf{I}_{l,\Lambda}(\theta) \\ \mathbf{I}_{l,a}^T(\theta) & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ \mathbf{I}_{l,\Lambda}^T(\theta) & \mathbf{I}_{a,\Lambda}^T(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix}, \tag{3}$$

where, for $\theta \in \Theta$ and $\alpha, \beta \in \{\ell, a\}$,

$$\mathbf{I}_{\alpha,\beta}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{[\Lambda_\theta^1(\tau)]^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_\alpha} \right)^T \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_\beta} \right) dr \, d\tau, \tag{4}$$

$$\mathbf{I}_{\alpha,\Lambda}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r)}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_\alpha} \right)^T \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right) dr \, d\tau, \tag{5}$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) := \int_{t_0}^t \int_{\mathcal{C}} \frac{(f_{\theta,\tau}^1(r))^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)} \left(\frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} \right)^T \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right) dr \, d\tau. \tag{6}$$

Proof By Theorem 3, for the image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$, $\Lambda_\theta(\tau) := \Lambda_\theta^1(\tau) + \Lambda^2(\tau)$ and $f_{\theta,\tau}(r) := \epsilon_\theta^1(\tau) f_{\theta,\tau}^1(r) + \epsilon_\theta^2(\tau) f_\tau^2(r)$ for $r \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, where $\epsilon_\theta^1(\tau) := \Lambda_\theta^1(\tau)/(\Lambda_\theta^1(\tau) + \Lambda^2(\tau))$ and $\epsilon_\theta^2(\tau) := \Lambda^2(\tau)/(\Lambda_\theta^1(\tau) + \Lambda^2(\tau))$, $\tau \geq t_0$, $\theta \in \Theta$. Thus we have $\Lambda_\theta(\tau) f_{\theta,\tau}(r) = (\Lambda_\theta^1(\tau) + \Lambda^2(\tau))(\epsilon_\theta^1(\tau) f_{\theta,\tau}^1(r) + \epsilon_\theta^2(\tau) f_\tau^2(r)) = \Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)$, $r \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$. Substituting this in Theorem 2, using condition (A1) and the fact that Λ^2 and f_τ^2 are independent of θ, we get

$$\mathbf{I}(\theta) = \int_\mathcal{C} \int_{t_0}^t \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right) dr \, d\tau$$

$$= \int_{t_0}^t \int_\mathcal{C} \frac{1}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)} \begin{bmatrix} \Lambda_\theta^1(\tau) \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_l} \right)^T \\ \Lambda_\theta^1(\tau) \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} \right)^T \\ f_{\theta,\tau}^1(r) \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right)^T \end{bmatrix}$$

$$\begin{bmatrix} \Lambda_\theta^1(\tau) \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_l} & \Lambda_\theta^1(\tau) \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} & f_{\theta,\tau}^1(r) \frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \end{bmatrix} dr \, d\tau, \quad \theta \in \Theta.$$

From the above equation the result immediately follows.   □

In many practical situations it is important to know whether the Fisher information matrix $\mathbf{I}(\theta)$ is diagonal, as the diagonality of $\mathbf{I}(\theta)$ has several implications. For example, it is well known that under certain conditions the maximum likelihood estimator of a vector parameter θ is asymptotically Gaussian distributed with asymptotic mean θ and covariance $\mathbf{I}^{-1}(\theta)$ (see e.g., Zacks, 1971). Here, if $\mathbf{I}(\theta)$ is diagonal, this implies that the components of the maximum likelihood estimate of θ are asymptotically independent. Note that if an efficient estimator of θ exists (i.e., an unbiased estimator whose covariance matrix is equal to $\mathbf{I}^{-1}(\theta)$, $\theta \in \Theta$), then a diagonal $\mathbf{I}(\theta)$ ensures that the estimates of θ are uncorrelated.

In general, for $\theta = (\theta_1, \ldots, \theta_n) \in \Theta$, if $\mathbf{I}(\theta)$ is diagonal, then this implies that the *limit* of the accuracy of the unbiased estimates of θ_i, $i = 1, \ldots, n$, does not depend on the other unknown parameters in θ. For an unbiased estimator of the object location, this means that the limit of the localization accuracy of the x coordinate of the location is the same whether or not the y and z coordinates of the location are known.

We next investigate the conditions under which the Fisher information matrix given in Theorem 4 is block diagonal. As will be shown, for the parameter vector $\theta = (\theta_l, \theta_a, \theta_\Lambda)$, it turns out that $\mathbf{I}(\theta)$ is block diagonal when the detector $\mathcal{C}$ and the photon distribution profiles $f_{\theta,\tau}^1$ and f_τ^2 satisfy certain symmetry conditions. Furthermore, for some special cases of θ, $\mathbf{I}(\theta)$ becomes fully diagonal (see Corollary 9). We next define a symmetric detector and a symmetric function.

Definition 5
1. A detector $\mathcal{C}$ is said to be *symmetric* if there exists a point $(c_x, c_y) \in \mathbb{R}^2$, known as the *center* of $\mathcal{C}$, such that for every $(x, y) \in \mathcal{C}$, $(2c_x - x, y) \in \mathcal{C}$, $(x, 2c_y - y) \in \mathcal{C}$ and $(2c_x - x, 2c_y - y) \in \mathcal{C}$.
2. Let $\mathcal{C}$ be a symmetric detector with center (c_x, c_y). A function $f: \mathcal{C} \to \mathbb{R}$ is said to be *symmetric (antisymmetric) along the x axis with respect to* c_x if for every $(x, y) \in \mathcal{C}$, $f(x, y) = f(2c_x - x, y)$ ($f(x, y) = -f(2c_x - x, y)$). If f is symmetric along both the x and y axes with respect to c_x and c_y respectively, then f is said to be *symmetric with respect to the center of* $\mathcal{C}$.

In the following theorem we assume the location component θ_l to be $\theta_l = (\theta_1, \theta_2)$. Further, we assume that $(\partial f^1_{\theta,\tau}(r)/\partial\theta_1) = -\vartheta_x(\partial f^1_{\theta,\tau}(r)/\partial x)$ and $(\partial f^1_{\theta,\tau}(r)/\partial\theta_2) = -\vartheta_y$ $(\partial f^1_{\theta,\tau}(r)/\partial y)$, $r = (x, y) \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, where ϑ_x and ϑ_y are constants that are independent of $(x, y) \in \mathcal{C}$. This assumption is satisfied if the photon distribution profile $f^1_{\theta,\tau}$ is a function of $(x/\vartheta_x - \theta_1, y/\vartheta_y - \theta_2)$, where $(x, y) \in \mathcal{C}$ (see Section Image function).

Theorem 6 *Let* $\mathcal{G}^1(\Lambda^1_\theta, \{f^1_{\theta,\tau}\}_{\tau\geq t_0}, \mathcal{C})$, $\mathcal{G}^2(\Lambda^2, \{f^2_\tau\}_{\tau\geq t_0}, \mathcal{C})$ *and* $\mathcal{G}$ *be image detection processes as given in Theorem 4. Assume that*

(A1) *for* $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, $\theta_l = (\theta_1, \theta_2)$, $\theta_a = (\theta_3, \ldots, \theta_k)$ *and* $\theta_\Lambda = (\theta_{k+1}, \ldots, \theta_n)$, $\partial f^1_{\theta,\tau}(r)/\partial\theta_\Lambda = 0$, $r \in \mathcal{C}$, $\tau \geq t_0$, $\partial\Lambda^1_\theta(\tau)/\partial\theta_l = 0$, $\tau \geq t_0$ *and* $\partial\Lambda^1_\theta(\tau)/\partial\theta_a = 0$, $\tau \geq t_0$,

(A2) $\partial f^1_{\theta,\tau}(r)/\partial\theta_1 = -\vartheta_x(\partial f^1_{\theta,\tau}(r)/\partial x)$, $\partial f^1_{\theta,\tau}(r)/\partial\theta_2 = -\vartheta_y(\partial f^1_{\theta,\tau}(r)/\partial y)$, $r = (x, y) \in \mathcal{C}$, $\theta \in \Theta$, $\tau \geq t_0$, *where* ϑ_x *and* ϑ_y *are constants that are independent of x and y,*

(A3) *the detector* $\mathcal{C}$ *is symmetric, and*

(A4) $f^1_{\theta,\tau}$ *and* f^2_τ *are symmetric with respect to the center of the detector* $\mathcal{C}$ *for* $\theta \in \Theta$ *and* $\tau \geq t_0$.

1. *Then for* $\theta \in \Theta$ *the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$ *is given by*

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & \mathbf{I}_{l,a}(\theta) & 0 \\ \mathbf{I}^T_{l,a}(\theta) & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ 0 & \mathbf{I}^T_{a,\Lambda}(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where for $\theta \in \Theta$, $\mathbf{I}_{l,a}(\theta)$ *and* $\mathbf{I}_{a,a}(\theta)$ *are given by Eq. 4,* $\mathbf{I}_{a,\Lambda}(\theta)$ *is given by Eq. 5,* $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ *is given by Eq. 6 and*

$$\mathbf{I}_{l,l}(\theta) := \mathrm{Diag}\left[\vartheta^2_x \int_{t_0}^t \int_\mathcal{C} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau)f^1_{\theta,\tau}(r) + \Lambda^2(\tau)f^2_\tau(r)}\left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial x}\right)^2, \right.$$

$$\left. \vartheta^2_y \int_{t_0}^t \int_\mathcal{C} \frac{[\Lambda^1_\theta(\tau)]^2}{\Lambda^1_\theta(\tau)f^1_{\theta,\tau}(r) + \Lambda^2(\tau)f^2_\tau(r)}\left(\frac{\partial f^1_{\theta,\tau}(r)}{\partial y}\right)^2 \, dr d\tau\right], \qquad (7)$$

where Diag denotes the diagonal matrix.

2. *In addition to conditions (A1–A4), assume that*

(A5) *all the elements of the vector* $\partial f^1_{\theta,\tau}(r)/\partial\theta_a$, $r \in \mathcal{C}$, *are symmetric with respect to the center of the detector* $\mathcal{C}$ *for* $\theta \in \Theta$ *and* $\tau \geq t_0$.

Then for $\theta \in \Theta$ *the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$ *is given by*

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & \mathbf{I}_{a,\Lambda}(\theta) \\ 0 & \mathbf{I}^T_{a,\Lambda}(\theta) & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where all the non-zero entries of $\mathbf{I}(\theta)$ *are given in result 1.*

3. *In addition to conditions* (A1–A5), *assume that*

(A6) *the photon detection rate of* $\mathcal{G}^2$ *is zero, i.e.,* $\Lambda^2(\tau) = 0$, $\tau \geq t_0$.

Then for $\theta \in \Theta$ *the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$ *is given by*

$$
\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & 0 \\ 0 & 0 & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},
\tag{8}
$$

where for $\theta \in \Theta$,

$$
\mathbf{I}_{l,l}(\theta) := \begin{bmatrix} \vartheta_x^2 \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial x} \right)^2 dr \, d\tau & 0 \\ 0 & \vartheta_y^2 \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial y} \right)^2 dr \, d\tau \end{bmatrix},
\tag{9}
$$

$$
\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \int_{\mathcal{C}} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} \right)^T \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} dr \, d\tau,
\tag{10}
$$

$$
\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} \right)^T \frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} d\tau.
\tag{11}
$$

Proof
1. Using condition (A1) and the fact that Λ^2 and f_τ^2 are independent of θ, we can show that the general expression for the Fisher information matrix is given by Theorem 4 (see Eq. 3). Consider the matrix $\mathbf{I}_{l,l}(\theta)$ that is given by Eq. 4. Using condition (A2) we immediately obtain the integral expressions of $[\mathbf{I}_{l,l}(\theta)]_{11}$ and $[\mathbf{I}_{l,l}(\theta)]_{22}$ that are given in Eq. 7. To obtain the desired result, we need to show that the off-diagonal terms of $\mathbf{I}_{l,l}(\theta)$ (i.e., $[\mathbf{I}_{l,l}(\theta)]_{12}$ and $[\mathbf{I}_{l,l}(\theta)]_{21}$) and all the terms of $\mathbf{I}_{l,\Lambda}(\theta)$ (i.e., $[\mathbf{I}_{l,\Lambda}(\theta)]_{ij}, i = 1, 2, j = k+1, \ldots, n$) are zero.

Let (c_x, c_y) denote the center of the detector $\mathcal{C}$ and define $T_X : \mathcal{C} \to \mathbb{R}^2$, $(x, y) \mapsto (2c_x - x, y)$. By condition (A4), $f_{\theta,\tau}^1(x, y) = (f_{\theta,\tau}^1 \circ T_X)(x, y)$, $(x, y) \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$. Using the chain rule of differentiation we get

$$
\frac{\partial f_{\theta,\tau}^1(x, y)}{\partial x} = \frac{\partial [(f_{\theta,\tau}^1 \circ T_X)(x, y)]}{\partial x} = \frac{\partial f_{\theta,\tau}^1(2c_x - x, y)}{\partial x} \frac{\partial (2c_x - x)}{\partial x}
$$
$$
= -\left(\frac{\partial f_{\theta,\tau}^1}{\partial x} \circ T_X \right)(x, y),
\tag{12}
$$

where $(x, y) \in \mathcal{C}$, $\theta \in \Theta$ and $\tau \geq t_0$. Similarly, by using condition (A4) we can show that for $\theta \in \Theta$ and $\tau \geq t_0$,

$$
\frac{\partial f_{\theta,\tau}^1(x, y)}{\partial y} = \frac{\partial [(f_{\theta,\tau}^1 \circ T_X)(x, y)]}{\partial y} = \left(\frac{\partial f_{\theta,\tau}^1}{\partial y} \circ T_X \right)(x, y), \quad (x, y) \in \mathcal{C},
\tag{13}
$$

$$
\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(x, y) + \Lambda^2(\tau) f_\tau^2(x, y) = \left((\Lambda_\theta^1(\tau) f_{\theta,\tau}^1 + \Lambda^2(\tau) f_\tau^2) \circ T_X \right)(x, y), \quad (x, y) \in \mathcal{C}.
\tag{14}
$$

Consider the term $[\mathbf{I}(\theta)]_{12}$ that is given by Eq. 4. Hence using this result, condition (A2) and Eqs. 12–14, we get for $\theta \in \Theta$,

$$[\mathbf{I}_{l,l}(\theta)]_{12} = \int_{t_0}^{t} \int_{\mathcal{C}} \frac{[\Lambda_\theta^1(\tau)]^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r) + \Lambda^2(\tau) f_\tau^2(r)} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_1} \frac{\partial f_{\theta,\tau}(r)}{\partial \theta_2} \, dr \, d\tau$$

$$= \vartheta_x \vartheta_y \int_{t_0}^{t} \int_{\mathcal{C}} \frac{[\Lambda_\theta^1(\tau)]^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(x,y) + \Lambda^2(\tau) f_\tau^2(x,y)} \frac{\partial f_{\theta,\tau}^1(x,y)}{\partial x} \frac{\partial f_{\theta,\tau}^1(x,y)}{\partial y} dx dy d\tau$$

$$= -\vartheta_x \vartheta_y \int_{t_0}^{t} \int_{\mathcal{C}} \left(\left(\frac{[\Lambda_\theta^1(\tau)]^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1 + \Lambda^2(\tau) f_\tau^2} \frac{\partial f_{\theta,\tau}^1}{\partial x} \frac{\partial f_{\theta,\tau}^1}{\partial y} \right) \circ T_X \right)(x,y) \, dx \, dy \, d\tau$$

$$= -\vartheta_x \vartheta_y \int_{t_0}^{t} \int_{\mathcal{C}} \frac{[\Lambda_\theta^1(\tau)]^2}{\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(x,y) + \Lambda^2(\tau) f_\tau^2(x,y)} \times$$

$$\frac{\partial f_{\theta,\tau}^1(x,y)}{\partial x} \frac{\partial f_{\theta,\tau}^1(x,y)}{\partial y} dx \, dy \, d\tau = -[\mathbf{I}_{l,l}(\theta)]_{12}, \tag{15}$$

where we have used the Theorem on change of variables (see Rudin, 1987, pp. 153–155) in the final step. Similarly we can show that $[\mathbf{I}_{l,l}(\theta)]_{21} = -[\mathbf{I}_{l,l}(\theta)]_{21}, \theta \in \Theta$. Thus we have $[\mathbf{I}_{l,l}(\theta)]_{12} = [\mathbf{I}_{l,l}(\theta)]_{21} = 0, \theta \in \Theta$.

If $T_Y: \mathcal{C} \to \mathbb{R}^2, (x,y) \mapsto (x, 2c_y - y)$, then similar to Eqs. 12 and 14, we can show that for $\theta \in \Theta$ and $\tau \geq t_0$,

$$\frac{\partial f_{\theta,\tau}^1(x,y)}{\partial y} = \frac{\partial[(f_{\theta,\tau}^1 \circ T_Y)(x,y)]}{\partial y}$$

$$= -\left(\frac{\partial f_{\theta,\tau}^1}{\partial y} \circ T_Y \right)(x,y), \quad (x,y) \in \mathcal{C}, \quad \tau \geq t_0, \quad \theta \in \Theta, \tag{16}$$

$$\Lambda_\theta^1(\tau) f_{\theta,\tau}^1(x,y) + \Lambda^2(\tau) f_\tau^2(x,y) = \left((\Lambda_\theta^1(\tau) f_{\theta,\tau}^1 + \Lambda^2(\tau) f_\tau^2) \circ T_Y \right)(x,y), \quad (x,y) \in \mathcal{C}. \tag{17}$$

Hence by using Eqs. 12, 14, 16 and 17 and condition (A2), we can show that $[\mathbf{I}_{l,\Lambda}(\theta)]_{ij} = -[\mathbf{I}_{l,\Lambda}(\theta)]_{ij}$, for $i = 1, 2, j = k+1, \ldots, n$ and $\theta \in \Theta$. Hence $\mathbf{I}_{l,\Lambda}(\theta) = 0, \theta \in \Theta$. From this the result follows.

2. By condition (A5), for $\zeta \in \{X, Y\}$, $\partial f_{\theta,\tau}^1(x,y)/\partial \theta_a = \partial[(f_{\theta,\tau}^1 \circ T_\zeta)(x,y)]/\partial \theta_a = \left((\partial f_{\theta,\tau}^1/\partial \theta_a) \circ T_\zeta \right)(x,y), (x,y) \in \mathcal{C}, \theta \in \Theta$ and $\tau \geq t_0$, where T_X and T_Y are defined above. Hence using this result, Eqs. 12–14, Eqs. 16–17, condition (A2) and by taking an approach similar to that of Eq. 15, we can show that $[\mathbf{I}_{l,a}(\theta)]_{ij} = -[\mathbf{I}_{l,a}(\theta)]_{ij}, i = 1, 2, j = 3, \ldots, k, \theta \in \Theta$. From this it follows that $\mathbf{I}_{l,a}(\theta) = 0, \theta \in \Theta$. Substituting this in result 1 of this Theorem the result follows immediately.

3. Substituting for Λ^2 in Eq. 7 we immediately obtain the expression for $\mathbf{I}_{l,l}(\theta)$ that is given by Eq. 9. Consider the term $\mathbf{I}_{a,\Lambda}(\theta)$ that is given by Eq. 5 (see Theorem 4). Since $f_{\theta,\tau}^1$ is a density function that satisfies conditions (C3–C5) of Definition 1, $\int_{\mathcal{C}} (\partial f_{\theta,\tau}^1(r)/\partial \theta) = 0$, $\theta \in \Theta$ and $\tau \geq t_0$ (Zacks, 1971 pp. 182–183). Using this result and substituting for Λ^2, we get

$$\mathbf{I}_{a,\Lambda}(\theta) = \int_{t_0}^{t} \left(\int_{\mathcal{C}} \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} \right)^T dr \frac{\partial \Lambda_\theta(\tau)}{\partial \theta_\Lambda} \, d\tau = 0, \quad \theta \in \Theta.$$

Finally, consider the term $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ that is given by Eq. 6. Substituting for Λ^2 and using the fact that $\int_{\mathcal{C}} f^1_{\theta,\tau}(r)\,\mathrm{d}r = 1$, $\tau \geq t_0$, $\theta \in \Theta$, we obtain the desired expression that is given in Eq. 11. $\square$

Remark 7 *Note that when condition* (A6) *is satisfied, the photon detection rate and the photon distribution profile of $\mathcal{G}$ are equal to that of $\mathcal{G}^1$. In this case, the condition that* (*see condition* (A4)) *f^2_τ is symmetric with respect to the center of the detector is no longer required.*

From the above theorem we see that the number of symmetry conditions imposed on the partial derivatives of $f^1_{\theta,\tau}$ (i.e., $\partial f^1_{\theta,\tau}(r)/\partial\theta_l$ and $\partial f^1_{\theta,\tau}(r)/\partial\theta_a$) determines the number of off-diagonal terms that are zero in the Fisher information matrix $\mathbf{I}(\theta)$. Note that the photon detection rate Λ^2 also plays a crucial role in making $\mathbf{I}(\theta)$ block diagonal. Consider a special case, where the parameter vector $\theta = (\theta_l, \theta_\Lambda)$ only consists of the location θ_l and the rate θ_Λ components. In this case the Fisher information matrix given in result 1 of the above theorem is block diagonal, even when the photon detection rate $\Lambda^2(\tau) \neq 0$, $\tau \geq t_0$. Finally, we note that in result 3 of the above theorem if the auxiliary component θ_a and the rate component θ_Λ are scalars, then $\mathbf{I}(\theta)$ is diagonal.

Image function

In the previous sections we made no assumptions about the specific functional form of the photon distribution profile. In an optical microscope, the image of an object can often be considered to be invariant with respect to shifts in the object location (Goodman, 1996). Hence the photon distribution profile $f^1_{\theta,\tau}$ can be expressed as a scaled and shifted version of the image of the object. For example, in the case of a moving object, $f^1_{\theta,\tau}$ can be written as $f^1_{\theta,\tau}(x, y) = \frac{1}{M^2} q_{\theta_e}\left(\frac{x}{M} - x_{0,\tau}, \frac{y}{M} - y_{0,\tau}\right)$, $(x, y) \in \mathbb{R}^2$, $\theta \in \Theta$, $\tau \geq t_0$, where q_{θ_e} denotes an image function, $M > 0$ denotes the lateral magnification and $(x_{0,\tau}, y_{0,\tau})$ denotes the time dependent $x - y$ location of the object. An image function q_{θ_e} describes the image of a fixed object on the detector plane at unit lateral magnification when the object is located along the z axis in the object space (see Fig. 1). Here, θ_e is a vector that parameterizes the image function. For example, θ_e could be the z position of the object and/or the angles that specify the 3D orientation of the object. In some applications the θ_e parameterization is not required and in such cases the image function is denoted as q. Since $f^1_{\theta,\tau}$ is a probability density function that satisfies conditions (C3–C5) of Definition 1, to express $f^1_{\theta,\tau}$ in terms of q_{θ_e} we impose appropriate conditions on the image function that are given below.

Definition 8 Let $\Theta_e \subseteq \mathbb{R}^m$ be a parameter space. For $\theta_e = (\theta_{e,1}, \ldots, \theta_{e,m}) \in \Theta_e$, we define $q_{\theta_e} : \mathbb{R}^2 \to [0, \infty)$ to be an *image function* if the following properties are satisfied.

1. $\int_{\mathbb{R}^2} q_{\theta_e}(x, y)\,\mathrm{d}x\,\mathrm{d}y = 1$,
2. $\frac{\partial q_{\theta_e}(x,y)}{\partial x}$, $\frac{\partial q_{\theta_e}(x,y)}{\partial y}$ and $\frac{\partial q_{\theta_e}(x,y)}{\partial\theta_{e,i}}$ exist for every $(x, y) \in \mathbb{R}^2$,
3. $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial x}\right|\,\mathrm{d}x\,\mathrm{d}y < \infty$, $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial y}\right|\,\mathrm{d}x\,\mathrm{d}y < \infty$ and $\int_{\mathbb{R}^2} \left|\frac{\partial q_{\theta_e}(x,y)}{\partial\theta_{e,i}}\right|\,\mathrm{d}x\,\mathrm{d}y < \infty$, and
4. $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)} \frac{\partial q_{\theta_e}(x,y)}{\partial\zeta_k} \frac{\partial q_{\theta_e}(x,y)}{\partial\zeta_l}\,\mathrm{d}x\,\mathrm{d}y$, $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)} \frac{\partial q_{\theta_e}(x,y)}{\partial\zeta_k} \frac{\partial q_{\theta_e}(x,y)}{\partial\theta_{e,i}}\,\mathrm{d}x\,\mathrm{d}y$ and $\int_{\mathbb{R}^2} \frac{1}{q_{\theta_e}(x,y)}$ $\frac{\partial q_{\theta_e}(x,y)}{\partial\theta_{e,i}} \frac{\partial q_{\theta_e}(x,y)}{\partial\theta_{e,j}}\,\mathrm{d}x\,\mathrm{d}y$ exist and are finite, where $\zeta_1 = x$, $\zeta_2 = y$, $k, l = 1, 2$ and $i, j = 1, \ldots, m$.

The image function q_{θ_e} and its derivative $(\partial q_{\theta_e}/\partial \theta_{e,i})$ are said to be symmetric if, for $\theta_e \in \Theta_e$, $q_{\theta_e}(x, y) = q_{\theta_e}(-x, y) = q_{\theta_e}(x, -y)$ and $\frac{\partial q_{\theta_e}(x,y)}{\partial \theta_{e,i}} = \frac{\partial q_{\theta_e}(-x,y)}{\partial \theta_{e,i}} = \frac{\partial q_{\theta_e}(x,-y)}{\partial \theta_{e,i}}$, respectively, for $(x, y) \in \mathbb{R}^2$.

In several applications the image of the object can be considered to be invariant with respect to time, for example, when the object is not moving during the acquisition of its image. In such cases, the expression for the photon distribution profile will be independent of time. In the following Corollary we derive a general expression for the Fisher information matrix for such applications. Here, the parameter vector is set to be $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, where $\theta_l = (x_0, y_0)$ denotes the x-y location of the object and $\theta_\Lambda = \Lambda_0$ is a scalar parameter that characterizes the photon detection rate Λ_θ^1. We assume that the photon distribution profile $f_{\theta,\tau}^1$ is given in terms of a symmetric image function. We also assume the detector to be infinite, i.e., $\mathcal{C} = \mathbb{R}^2$. An infinite detector provides the best case scenario, where all the photons that reach the detector plane are detected by the detector (see Fig. 1). Hence the square root of the inverse Fisher information matrix for an infinite detector provides the *fundamental limit* to the accuracy with which the components of the parameter vector θ can be determined. We then consider a special case, where θ_l and θ_Λ are as given above and the auxiliary component $\theta_a = \epsilon_0$ is a scalar. For this special case we show that the Fisher information matrix is diagonal. Finally, we assume the photon distribution profile $f_{\theta,\tau}^1$ to be independent of θ_a and the parameter vector to be $\theta = (\theta_l, \theta_\Lambda)$ with θ_l and θ_Λ as given above.

Corollary 9 Let $\Theta \subseteq \mathbb{R}^n$ be a parameter space. Let $\mathcal{G}^1(\Lambda_\theta^1, \{f_{\theta,\tau}^1\}_{\tau \geq t_0}, \mathbb{R}^2)$ and $\mathcal{G}^2(\Lambda^2, \{f_\tau^2\}_{\tau \geq t_0}, \mathbb{R}^2)$ be two independent image detection processes such that $\mathcal{G}^2$ is independent of θ. Let $\mathcal{G}$ be the superposition of $\mathcal{G}^1$ and $\mathcal{G}^2$. For $\theta = (\theta_l, \theta_a, \theta_\Lambda) \in \Theta$, let $\theta_l = (x_0, y_0)$ and $\theta_\Lambda = \Lambda_0$, where x_0, y_0 and Λ_0 are scalar parameters. Assume that for $\theta \in \Theta$,

(A1) $\partial \Lambda_\theta^1(\tau)/\partial \theta_l = 0$, $\partial \Lambda_\theta^1(\tau)/\partial \theta_a = 0$ and $\Lambda^2(\tau) = 0$ for $\tau \geq t_0$,
(A2) there exists a symmetric image function q_{θ_a} such that for $M > 0$, the photon distribution profile $f_{\theta,\tau}^1$ is given by

$$f_{\theta,\tau}^1(x, y) = \frac{1}{M^2} q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right), \quad (x, y) \in \mathbb{R}^2, \quad \tau \geq t_0,$$

(A3) all the elements of the vector $(\partial q_{\theta_a}(x, y)/\partial \theta_a)$ are symmetric, $(x, y) \in \mathbb{R}^2$.

1. Then for $\theta \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \mathbf{I}_{l,l}(\theta) & 0 & 0 \\ 0 & \mathbf{I}_{a,a}(\theta) & 0 \\ 0 & 0 & \mathbf{I}_{\Lambda,\Lambda}(\theta) \end{bmatrix},$$

where

$$\mathbf{I}_{l,l}(\theta) = Diag\left[\int_{t_0}^t \Lambda_\theta^1(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x, y)}\left(\frac{\partial q_{\theta_a}(x, y)}{\partial x}\right)^2 dx\, dy, \right.$$

$$\left.\int_{t_0}^t \Lambda_\theta^1(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x, y)}\left(\frac{\partial q_{\theta_a}(x, y)}{\partial y}\right)^2 dx\, dy\right], \quad \theta \in \Theta,$$

where Diag denotes the diagonal matrix.

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a}\right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dx\, dy, \quad \theta \in \Theta,$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \Lambda_0}\right)^2 d\tau, \quad \theta \in \Theta.$$

2. If $\theta_a = \epsilon_0$ is scalar, then for $\theta = (x_0, y_0, \epsilon_0, \Lambda_0) \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \mathrm{Diag}\left[\int_{t_0}^{t} \Lambda_\theta^1(\tau)\, d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial x}\right)^2 dx dy, \right.$$

$$\int_{t_0}^{t} \Lambda_\theta^1(\tau)\, d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial y}\right)^2 dx dy,$$

$$\int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial \epsilon_0}\right)^2 dx dy,$$

$$\left. \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \Lambda_0}\right)^2 d\tau\right], \quad \theta \in \Theta, \tag{18}$$

where Diag denotes the diagonal matrix.

3. If $\theta = (x_0, y_0, \Lambda_0) \in \Theta$, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \int_{t_0}^{t} \Lambda_\theta^1(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q(x,y)}\left(\frac{\partial q(x,y)}{\partial x}\right)^2 dx dy & 0 & 0 \\ 0 & \int_{t_0}^{t} \Lambda_\theta^1(\tau)d\tau \int_{\mathbb{R}^2} \frac{1}{q(x,y)}\left(\frac{\partial q(x,y)}{\partial y}\right)^2 dx dy & 0 \\ 0 & 0 & \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)}\left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \Lambda_0}\right)^2 d\tau \end{bmatrix}. \tag{19}$$

Proof

1. Let $M > 0$. By conditions (A1) and (A2), $\partial f_{\theta,\tau}^1(r)/\partial \theta_\Lambda = 0$, $\partial \Lambda_\theta^1(\tau)/\partial \theta_l = 0$ and $\partial \Lambda_\theta^1(\tau)/\partial \theta_a = 0$ for $r \in \mathbb{R}^2$, $\theta \in \Theta$ and $\tau \geq t_0$. Further, from condition (A2) we can verify that $\partial f_{\theta,\tau}^1(r)/\partial x_0 = -M(\partial f_{\theta,\tau}^1(r)/\partial x)$ and $\partial f_{\theta,\tau}^1(r)/\partial y_0 = -M(\partial f_{\theta,\tau}^1(r)/\partial y)$ for $r = (x,y) \in \mathbb{R}^2, \theta \in \Theta$ and $\tau \geq t_0$. Since $(Mx_0, My_0) \in \mathbb{R}^2$, for every $(x,y) \in \mathbb{R}^2$, $(2Mx_0 - x, y) \in \mathbb{R}^2$, $(x, 2My_0 - y) \in \mathbb{R}^2$ and $(2Mx_0 - x, 2My_0 - y) \in \mathbb{R}^2$. Hence $\mathbb{R}^2$ is symmetric with respect to the point (Mx_0, My_0) (see Definition 5). From conditions (A2 − A3) we can easily verify that $f_{\theta,\tau}^1(x,y)$ and all the elements of the vector $\partial f_{\theta,\tau}^1(x,y)/\partial \theta_a = (1/M^2)(\partial q_{\theta_a}(x/M - x_0, y/M - y_0)/\partial \theta_a)$ are symmetric with respect to the point (Mx_0, My_0), where $(x,y) \in \mathbb{R}^2$, $\theta \in \Theta$ and $\tau \geq t_0$. Finally we note that by condition (A1), $\Lambda^2(\tau) = 0$, $\tau \geq t_0$.

Thus from the above and from Remark 7, we see that the photon distribution profile and the photon detection rate of $\mathcal{G}^1$ and $\mathcal{G}^2$ satisfy all of the conditions of result 3 of Theorem 6. Hence for the present case, the Fisher information matrix $\mathbf{I}(\theta)$ is block diagonal (see Eq. 8). From Eq. 9 we see that $\mathbf{I}_{l,l}(\theta)$ is diagonal and we evaluate its diagonal terms $[\mathbf{I}_{l,l}(\theta)]_{11}$ and $[\mathbf{I}_{l,l}(\theta)]_{22}$. Substituting for Λ_θ^1 and $f_{\theta,\tau}^1$ in the integral expression of $[\mathbf{I}_{l,l}(\theta)]_{11}$, we get

$$[\mathbf{I}_{l,l}(\theta)]_{11} = M^2 \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{\Lambda_\theta^1(\tau)}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial x}\right)^2 dr d\tau$$

$$= M^2 \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{\frac{1}{M^2} q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}$$

$$\times \left(\frac{1}{M^2} \frac{\partial q_{\theta_a}\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x}\right)^2 dx dy$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{\Lambda_\theta^1(\tau)}{q_{\theta_a}(u,v)} \left(\frac{\partial q_{\theta_a}(u,v)}{\partial u} \frac{1}{M}\right)^2 (M du)(M dv)$$

$$\left(u := \frac{x}{M} - x_0, \ v := \frac{y}{M} - y_0\right)$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial x}\right)^2 dx dy, \quad \theta \in \Theta.$$

Similarly, we can show that

$$[\mathbf{I}_{l,l}(\theta)]_{22} = \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial y}\right)^2 dx dy,$$

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{f_{\theta,\tau}^1(r)} \left(\frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a}\right)^T \frac{\partial f_{\theta,\tau}^1(r)}{\partial \theta_a} dr$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a}\right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dx dy,$$

$$\mathbf{I}_{\Lambda,\Lambda}(\theta) = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda}\right)^T \frac{\partial \Lambda_\theta^1(\tau)}{\partial \theta_\Lambda} d\tau = \int_{t_0}^{t} \frac{1}{\Lambda_\theta^1(\tau)} \left(\frac{\partial \Lambda_\theta^1(\tau)}{\partial \Lambda_0}\right)^2 d\tau,$$

where $\theta \in \Theta$, and $\mathbf{I}_{a,a}(\theta)$ and $\mathbf{I}_{\Lambda,\Lambda}(\theta)$ are given by Eqs. 10 and 11 respectively. From this the result follows.

2. If $\theta_a = \epsilon_0$ is scalar, then for $\theta \in \Theta$,

$$\mathbf{I}_{a,a}(\theta) = \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\theta_a}(x,y)} \left(\frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a}\right)^T \frac{\partial q_{\theta_a}(x,y)}{\partial \theta_a} dx dy$$

$$= \int_{t_0}^{t} \Lambda_\theta^1(\tau) d\tau \int_{\mathbb{R}^2} \frac{1}{q_{\epsilon_0}(x,y)} \left(\frac{\partial q_{\epsilon_0}(x,y)}{\partial \epsilon_0}\right)^2 dx dy.$$

Substituting this in result 1 of this Corollary we obtain the desired result.
3. The result immediately follows from result 2 of this Corollary.   □

From the above Corollary we see that the Fisher information matrix $\mathbf{I}(\theta)$ is independent of (x_0, y_0) and only depends on the image function and its partial derivatives. Moreover, $\mathbf{I}(\theta)$ is diagonal when $\theta = (x_0, y_0, \epsilon_0, \Lambda_0)$ and $\theta = (x_0, y_0, \Lambda_0)$. Note that if $\epsilon_0 = z_0$ denotes the z coordinate of the object location, then $\mathbf{I}(\theta)$ that is given in result 2 of the above Corollary can be used to calculate the three dimensional fundamental limit of the localization accuracy of the object. In Ober, Ram et al., (2004), we recently reported integral expressions for $\mathbf{I}(\theta)$ that are analogous to Eqs. 18 and 19, where the parameter vector was set to be $\theta = (x_0, y_0)$,

Λ_0 was assumed to be known and $\mathbf{I}(\theta)$ was a 2×2 diagonal matrix. We note that Eqs. 18 and 19 are generalizations of our earlier result and show that the diagonality of $\mathbf{I}(\theta)$ is preserved even when additional parameters such as ϵ_0 and Λ_0 are assumed to be unknown.

Example-1

We now illustrate the results derived in this section by considering specific image functions that describe the image of a fixed point source. According to optical diffraction theory, when a point source is in focus with respect to the detector, the intensity distribution of the image of the point source is described by the Airy profile (Born & Wolf, 1999, pp. 440). The 2D Gaussian profile, on the other hand, has been widely used to approximate the Airy profile, for example, in the analysis of data from single molecule fluorescence experiments (Kubitscheck, Kückmann, Kues, & Peter, 2000; Santos & Young, 2000; Thompson, Larson, & Webb, 2002). In the following Corollary, the parameter vector is set to be $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and the photon distribution profile $f_{\theta,\tau}^1$ is assumed to be given in terms of an image function q. The photon detection rate is assumed to be a constant, i.e., $\Lambda_\theta^1(\tau) = \Lambda_0, \tau \geq t_0$. For each image function, we derive a simple formula for the fundamental limit of the localization accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{11}}$ ($\sqrt{[\mathbf{I}^{-1}(\theta)]_{22}}$) of x_0 (y_0) and for the fundamental limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$ of Λ_0. We note that the following results are extensions of the results reported in Ober, Ram et al., (2004).

Corollary 10 Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space. Let $\mathcal{G}^1$, $\mathcal{G}^2$ and $\mathcal{G}$ be image detection processes that are given in Corollary 9. For $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$, let $\Lambda^2(\tau) = 0$, $\Lambda_\theta^1(\tau) = \Lambda_0$ and for $M > 0$, assume that there exist a symmetric image function q such that $f_{\theta,\tau}^1(x, y) = (1/M^2)q(x/M - x_0, y/M - y_0), (x, y) \in \mathbb{R}^2$.

1. *Airy profile:* If, for $n_a, \lambda > 0$, q is given by

$$q(x, y) = \frac{J_1^2(\frac{2\pi n_a}{\lambda}\sqrt{x^2 + y^2})}{\pi(x^2 + y^2)}, \qquad (x, y) \in \mathbb{R}^2, \tag{20}$$

then for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \frac{(2\pi n_a)^2 \Lambda_0(t-t_0)}{\lambda^2} & 0 & 0 \\ 0 & \frac{(2\pi n_a)^2 \Lambda_0(t-t_0)}{\lambda^2} & 0 \\ 0 & 0 & \frac{t-t_0}{\Lambda_0} \end{bmatrix}.$$

Further, the fundamental limit of the localization accuracy $\delta_{x_0}^{2d}$ ($\delta_{y_0}^{2d}$) of x_0 (y_0) and the fundamental limit of the accuracy $\delta_{\Lambda_0}^{2d}$ of Λ_0 are given by

$$\delta_{x_0}^{2d} = \delta_{y_0}^{2d} = \frac{\lambda}{2\pi n_a \sqrt{\Lambda_0(t - t_0)}}, \qquad \delta_{\Lambda_0}^{2d} = \sqrt{\frac{\Lambda_0}{(t - t_0)}}. \tag{21}$$

2. *2D Gaussian profile:* If, for $\sigma > 0$, q is given by

$$q(x, y) := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \qquad (x, y) \in \mathbb{R}^2, \tag{22}$$

then for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \begin{bmatrix} \frac{\Lambda_0(t-t_0)}{\sigma^2} & 0 & 0 \\ 0 & \frac{\Lambda_0(t-t_0)}{\sigma^2} & 0 \\ 0 & 0 & \frac{t-t_0}{\Lambda_0} \end{bmatrix}.$$

Further, the fundamental limit of the localization accuracy $\delta_{x_0}^{gau}$ ($\delta_{y_0}^{gau}$) of x_0 (y_0) and the fundamental limit of the accuracy $\delta_{\Lambda_0}^{gau}$ of Λ_0 are given by

$$\delta_{x_0}^{gau} = \delta_{y_0}^{gau} = \frac{\sigma}{\sqrt{\Lambda_0(t-t_0)}}, \qquad \delta_{\Lambda_0}^{gau} = \sqrt{\frac{\Lambda_0}{(t-t_0)}}. \qquad (23)$$

Proof

1. It can be verified that the Airy profile is a symmetric image function. By definition, for $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$, $\Lambda^2(\tau) = 0$, $\partial \Lambda_\theta^1(\tau)/\partial x_0 = \partial \Lambda_\theta^1(\tau)/\partial y_0 = 0$, $\partial f_{\theta,\tau}^1(r)/\partial \Lambda_0 = 0$, $r \in \mathbb{R}^2$, and $f_{\theta,\tau}^1$ is expressed as a shifted and scaled version of q. Hence for the present case result 3 of Corollary 9 holds and the Fisher information matrix $\mathbf{I}(\theta)$ is diagonal (see Eq. 19). Using Eq. 19 we can easily show that $[\mathbf{I}(\theta)]_{33} = (t - t_0)/\Lambda_0$. Let $\alpha := 2\pi n_a/\lambda$. Using the identity $(\partial/\partial x)\left[x^{-n}J_n(x)\right] = -x^{-n}J_{n+1}(x)$, $x \in \mathbb{R}$ (see Watson, 1958, pp. 18) with $n = 1$, we can show that

$$\frac{\partial}{\partial \zeta} \frac{J_1^2(\alpha\sqrt{x^2+y^2})}{x^2+y^2} = -2\zeta\alpha \frac{J_1(\alpha\sqrt{x^2+y^2})}{\sqrt{x^2+y^2}} \frac{J_2(\alpha\sqrt{x^2+y^2})}{x^2+y^2}, \quad (x,y) \in \mathbb{R}^2, \ \zeta \in \{x, y\}. \tag{24}$$

Hence by using this result and the integral identity $\int_0^\infty (J_n^2(t)/t)dt = 1/(2n)$ (Watson, 1958, pp. 405) with $n = 2$, we can show that $[\mathbf{I}(\theta)]_{11} = [\mathbf{I}(\theta)]_{22} = (2\pi n_a)^2 \Lambda_0(t - t_0)/\lambda^2$ (see also Ober, Ram et al., 2004).

2. We can easily verify that the 2D Gaussian profile is a symmetric image function. Further, we can show that for the present situation the Fisher information matrix is diagonal and is given by Eq. 19 (see proof of result 1). Substituting for Λ_θ^1 and q in Eq. 19 we get $[\mathbf{I}(\theta)]_{33} = (t - t_0)/\Lambda_0$, and it can be shown that $[\mathbf{I}(\theta)]_{11} = [\mathbf{I}(\theta)]_{22} = \Lambda_0(t - t_0)/\sigma^2$ (see Ober, Ram et al., 2004).

In both cases, the fundamental limit of the localization accuracy of x_0 (y_0) and the fundamental limit of the accuracy of Λ_0 are obtained by inverting the Fisher information matrix and taking the square root of the corresponding leading diagonal elements. □

The Airy profile depends on the term α that is given by $\alpha = 2\pi n_a/\lambda$, where n_a denotes the numerical aperture of the objective lens and λ denotes the wavelength of the detected photons. For a given experimental configuration, the numerical values of n_a and λ are known and hence α is known. On the other hand, the 2D Gaussian profile depends on the term σ that needs to be empirically determined from calibration experiments (see Kubitscheck et al., 2000; Thompson et al., 2002).

It can be shown that the maximum likelihood estimator of the photon detection rate Λ_0 is given by $\hat{\Lambda}_0 := N_{tot}/(t - t_0)$, where N_{tot} denotes the total number of detected photons and $t - t_0$ denotes the acquisition time (see e.g., Snyder & Miller, 1991 pp. 74–75), and the standard deviation of $\hat{\Lambda}_0$ is given by $\sqrt{\Lambda_0/(t - t_0)}$. From the above results, we see that for both the Airy profile and the 2D Gaussian profile, the performance limit to determining the parameter Λ_0 is given by $\sqrt{\Lambda_0/(t - t_0)}$, which is equal to the standard deviation of $\hat{\Lambda}_0$. Thus for the above scenario, the maximum likelihood estimator of the photon detection rate is an efficient estimator.

Effects of reduced detector size

In the previous section we derived the Fisher information matrix for the 2D Gaussian profile and the Airy profile, where it is assumed that the detector occupies the full detector plane, i.e., $\mathcal{C} = \mathbb{R}^2$. These models imply that photons impact the detector plane not only close to the center of the image profile, but also far away from the center of the image. This is, however, not a practical assumption since detectors have a finite size. In addition, when analyzing microscopy images typically only small regions of interest are used. This raises the question of how the detector size or the region of interest influences the performance limits.

In the following proposition we show how an image detection process $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ has to be adjusted when instead of the detector $\mathcal{C}$, the photons are detected on a reduced part $\mathcal{C}^{rd}$ of $\mathcal{C}$, i.e., on an open subset $\mathcal{C}^{rd}$ of $\mathcal{C}$.

Proposition 11 *Let $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ be an image detection process and let $\mathcal{C}^{rd} \subseteq \mathcal{C}$ be open. For $\theta \in \Theta$ and $\tau \geq t_0$, let $\alpha_{\theta,\tau} := \int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r) \mathrm{d}r$. The time points and the spatial coordinates of the arrival location of the photons detected on the reduced detector $\mathcal{C}^{rd}$ are described by an image detection process $\mathcal{G}^{rd}$ whose photon detection rate Λ_θ^{rd} and photon distribution profile $f_{\theta,\tau}^{rd}$ are given by*

$$\Lambda_\theta^{rd}(\tau) = \alpha_{\theta,\tau} \Lambda_\theta(\tau), \quad \tau \geq t_0, \quad \theta \in \Theta,$$

$$f_{\theta,\tau}^{rd}(r) = \frac{1}{\alpha_{\theta,\tau}} f_{\theta,\tau}(r), \quad r \in \mathcal{C}^{rd}, \quad \theta \in \Theta, \quad \tau \geq t_0.$$

Proof By definition of the image detection process $\mathcal{G}$, the time points of the detected photons on the detector $\mathcal{C}$ are modeled as a Poisson process with intensity function Λ_θ. It then follows that the time points of the detected photons on the detector $\mathcal{C}^{rd}$ form a Poisson process with intensity function $\alpha_{\theta,\tau} \Lambda_\theta$, $\tau \geq t_0$, $\theta \in \Theta$ (Papoulis & Pillai, 2002 pp. 381).

Let $\mathcal{A} \subseteq \mathcal{C}^{rd}$. Let U_τ denote the random variable that describes the arrival location of a photon that is detected on the detector $\mathcal{C}^{rd}$ at time τ, $\tau \geq t_0$. Then the probability that the arrival location of the detected photon is in the set $\mathcal{A}$ given that the arrival location is in the detector $\mathcal{C}^{rd}$ is given by

$$P[U_\tau \in \mathcal{A} \mid U_\tau \in \mathcal{C}^{rd}] = \frac{P[(U_\tau \in \mathcal{A}) \cap (U_\tau \in \mathcal{C}^{rd})]}{P[U_\tau \in \mathcal{C}^{rd}]} = \frac{P[U_\tau \in \mathcal{A}]}{P[U_\tau \in \mathcal{C}^{rd}]}$$
$$= \frac{\int_{\mathcal{A}} f_{\theta,\tau}(r)m(\mathrm{d}r)}{\int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r)m(\mathrm{d}r)} = \frac{\int_{\mathcal{A}} f_{\theta,\tau}(r)m(\mathrm{d}r)}{\alpha_{\theta,\tau}},$$

where m denotes the Lebesgue measure in $\mathbb{R}^2$. Since the above equation holds for every $\mathcal{A} \subseteq \mathcal{C}^{rd}$ and $\tau \geq t_0$, the term $P[U_\tau \in \mathcal{A} \mid U_\tau \in \mathcal{C}^{rd}]$ is absolutely continuous with respect to m. Hence there exists a probability density function $f_{\theta,\tau}^{rd}$ such that

$$f_{\theta,\tau}^{rd}(r) = \frac{1}{\alpha_{\theta,\tau}} f_{\theta,\tau}(r), \quad r \in \mathcal{C}^{rd}, \quad \theta \in \Theta, \quad \tau \geq t_0.$$

Since, by definition $f_{\theta,\tau}$ satisfies conditions (C3–C5) of Definition 1, it can be verified that $f_{\theta,\tau}^{rd}$ also satisfies these conditions. $\qquad \square$

We refer to the image detection process $\mathcal{G}^{rd}$ as the *reduced version* of $\mathcal{G}$ corresponding to the detector $\mathcal{C}^{rd}$. We next derive a general expression for the Fisher information matrix of $\mathcal{G}^{rd}$. We also derive a formula to calculate the loss of information when a detector of reduced size is used.

Theorem 12 *Let* $\mathcal{G}(\Lambda_\theta, \{f_{\theta,\tau}\}_{\tau \geq t_0}, \mathcal{C})$ *be an image detection process and* $\mathcal{G}^{rd}$ *be the reduced version of* $\mathcal{G}$ *corresponding to the detector* $\mathcal{C}^{rd}$, *where* $\mathcal{C}^{rd} \subseteq \mathcal{C}$. *For* $\theta \in \Theta$, *let* $\mathbf{I}(\theta)$ *denote the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$. *Then for* $\theta \in \Theta$,

1. *the Fisher information matrix of* $\mathcal{G}^{rd}$ *corresponding to the time interval* $[t_0, t]$ *is given by*

$$\mathbf{I}_{rd}(\theta) = \int_{t_0}^t \int_{\mathcal{C}^{rd}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right) \mathrm{d}r \mathrm{d}\tau,$$

2.
$$\Delta \mathbf{I}(\theta) := \mathbf{I}(\theta) - \mathbf{I}_{rd}(\theta) = \int_{t_0}^t \int_{\mathcal{C} \setminus \mathcal{C}^{rd}} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right)^T \times$$
$$\left(\frac{\partial [\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial \theta} \right) \mathrm{d}r \mathrm{d}\tau,$$

3. $\mathbf{I}(\theta) \geq \mathbf{I}_{rd}(\theta)$.

Proof
1. For $\theta \in \Theta$ and $\tau \geq t_0$, let $\alpha_{\theta,\tau} = \int_{\mathcal{C}^{rd}} f_{\theta,\tau}(r) \, \mathrm{d}r$. For the image detection process $\mathcal{G}^{rd}$, by Proposition 11 the photon detection rate $\Lambda_\theta^{rd}(\tau) = \alpha_{\theta,\tau} \Lambda_\theta(\tau)$ and the photon distribution profile $f_{\theta,\tau}^{rd}(r) = (1/\alpha_{\theta,\tau}) f_{\theta,\tau}(r)$ for $r \in \mathcal{C}^{rd}, \theta \in \Theta, \tau \geq t_0$. Substituting for Λ_θ^{rd} and $f_{\theta,\tau}^{rd}$ in Theorem 2 the result immediately follows.
2. The result immediately follows by using the expressions for the Fisher information matrix of $\mathcal{G}$ and $\mathcal{G}^{rd}$ that are given in Theorem 2 and in part 1 of this Theorem, respectively.
3. The integrand in the integral expression of $\Delta \mathbf{I}(\theta)$ given in result 2 of this theorem is non-negative. This implies that $\Delta \mathbf{I}(\theta)$ is positive semidefinite for $\theta \in \Theta$ and from this the result follows.   □

From result 1 of the above theorem we see that the expression for the Fisher information matrix of $\mathcal{G}^{rd}$ is analogous to that of $\mathcal{G}$ (see Theorem 2) with the only difference being that the region of integration of the spatial integral is now the reduced detector $\mathcal{C}^{rd}$.

Upper and lower bounds to the performance limits

In Example-1 the integral expressions of the Fisher information matrix $\mathbf{I}(\theta)$ for an infinite detector reduced to simple formulae. However, in a practical situation the calculation of $\mathbf{I}(\theta)$ can become cumbersome, for example, due to the shape of the finite sized detector $\mathcal{C}$. Hence determining the limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$ for the components of θ can become difficult. We next address this concern by deriving integral expressions for matrices $\mathbf{I}_u(\theta)$ and $\mathbf{I}_l(\theta)$ that provide an upper and lower bound to the Fisher information matrix $\mathbf{I}(\theta)$, respectively, i.e., $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta)$. Note that if $\mathbf{I}_l(\theta)$ is invertible, then it can be shown that $\sqrt{[\mathbf{I}_l^{-1}(\theta)]_{ii}}$ and $\sqrt{[\mathbf{I}_u^{-1}(\theta)]_{ii}}$ provide an upper and lower bound to $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$, respectively, i.e., $\sqrt{[\mathbf{I}_u^{-1}(\theta)]_{ii}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}} \leq \sqrt{[\mathbf{I}_l^{-1}(\theta)]_{ii}}, i = 1, \dots, n$. This is of particular relevance since in a number of situations the upper and lower bounds are diagonal matrices whose diagonal entries can be analytically calculated. We will show that this is the case for the 2D Gaussian profile and the Airy profile if the 'bounding detectors' are circular with center at the center of the image profile. The integral expression for $\mathbf{I}_u(\theta)$ ($\mathbf{I}_l(\theta)$) is derived in such a way that its integrand is identical to that of $\mathbf{I}(\theta)$ and its spatial integral is evaluated over a circular

region $\mathcal{B}_{r_c}(u)$ $(\mathcal{B}_{r_c}(l))$ known as the upper (lower) circular bounding detector that is centered at a point $r_c \in \mathcal{C}$ with radius u (l). The circular bounding detectors are defined below.

Definition 13 Let $\mathcal{C}$ be a detector and $r_c = (r_{c,x}, r_{c,y}) \in \mathcal{C}$. For $\rho > 0$, let $\mathcal{B}_{r_c}(\rho) := \{(x, y) \mid (x - r_{c,x})^2 + (y - r_{c,y})^2 < \rho^2, (x, y) \in \mathbb{R}^2\}$ denote a circular region centered at r_c with radius ρ and let $\mathcal{B}_{r_c}(\infty) := \mathbb{R}^2$. We define $\mathcal{B}_{r_c}(u)$ and $\mathcal{B}_{r_c}(l)$ to be the *upper and lower circular bounding detectors of* $\mathcal{C}$, respectively, if $l = \sup\{\rho \mid \mathcal{B}_{r_c}(\rho) \subseteq \mathcal{C}\}$ and $u = \inf\{\rho \mid \mathcal{C} \subseteq \mathcal{B}_{r_c}(\rho)\}$.

Earlier in this section we discussed the relationship between an image detection process described on a large, possibly infinite, detector and a more realistic smaller detector (see Proposition 11). In the derivation of the upper and lower bounds for the Fisher information matrix it will be useful to have the notion of an 'extended' version of an image detection process. We refer to any image detection process $\mathcal{G}^e$ as an extended version of $\mathcal{G}$ if $\mathcal{G}$ is the reduced version of $\mathcal{G}^e$.

Theorem 14 *Let* $\Theta \subseteq \mathbb{R}^n$ *be a parameter space and let* $\mathcal{G}$ *be an image detection process that is defined over the detector* $\mathcal{C}$. *Assume that* $\mathcal{G}^e(\Lambda_\theta^e, \{f_{\theta,\tau}^e\}_{\tau \geq t_0}, \mathcal{C}^e)$ *is an extended version of* $\mathcal{G}$. *For* $r_c \in \mathcal{C}$, *let* $\mathcal{B}_{r_c}(l)$ *and* $\mathcal{B}_{r_c}(u)$ *denote the circular bounding detectors of* $\mathcal{C}$. *Let* $\mathbf{I}(\theta)$ *be the Fisher information matrix of* $\mathcal{G}$ *corresponding to the time interval* $[t_0, t]$. *If* $\mathcal{C} \subseteq \mathcal{B}_{r_c}(u) \subseteq \mathcal{C}^e$, *then*

1. $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta), \quad \theta \in \Theta,$ $\qquad(25)$

where for $\beta \in \{u, \ell\}$

$$\mathbf{I}_\beta(\theta) := \int_{t_0}^{t} \int_{\mathcal{B}_{r_c}(\beta)} \frac{1}{\Lambda_\theta^e(\tau) f_{\theta,\tau}^e(r)} \left(\frac{\partial[\Lambda_\theta^e(\tau) f_{\theta,\tau}^e(r)]}{\partial \theta}\right)^T \left(\frac{\partial[\Lambda_\theta^e(\tau) f_{\theta,\tau}^e(r)]}{\partial \theta}\right) dr d\tau, \quad(26)$$

2. *Further, if* $\mathbf{I}_l(\theta)$ *is invertible, then*

$$[\mathbf{I}_u^{-1}(\theta)]_{ii} \leq [\mathbf{I}^{-1}(\theta)]_{ii} \leq [\mathbf{I}_l^{-1}(\theta)]_{ii}, \quad i = 1, \ldots, n, \quad \theta \in \Theta. \quad(27)$$

Proof
1. Since $\mathcal{G}^e$ is an extended extension of $\mathcal{G}$, it follows that $\mathcal{G}$ is the reduced version of $\mathcal{G}^e$ and the expression for the Fisher information matrix $\mathbf{I}(\theta)$ of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by result 1 of Theorem 12. Note that $\mathcal{B}_{r_c}(\beta)$ is open and $\mathcal{B}_{r_c}(\beta) \subseteq \mathcal{C}^e$, $\beta \in \{u, l\}$. Hence from Proposition 11 it follows that the detected photons on the circular bounding detector $\mathcal{B}_{r_c}(u)$ $(\mathcal{B}_{r_c}(l))$ can be modeled as an image detection process $\mathcal{G}_u^{rd}$ $(\mathcal{G}_l^{rd})$, which is the reduced version of $\mathcal{G}^e$ corresponding to $\mathcal{B}_{r_c}(u)$ $(\mathcal{B}_{r_c}(l))$. If the Fisher information matrix of $\mathcal{G}_u^{rd}$ $(\mathcal{G}_l^{rd})$ corresponding to the time interval $[t_0, t]$ is denoted as $\mathbf{I}_u(\theta)$ $(\mathbf{I}_l(\theta))$, then from result 1 of Theorem 12 we obtain the desired integral expression that is given by Eq. 26. Further, since $\mathcal{B}_{r_c}(l) \subseteq \mathcal{C} \subseteq \mathcal{B}_{r_c}(u)$, from result 3 of Theorem 12 it can be deduced that $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta), \theta \in \Theta$.
2. If $\mathbf{I}_l(\theta)$ is invertible for $\theta \in \Theta$, then by result 1 of this Theorem $\mathbf{I}(\theta)$ and $\mathbf{I}_u(\theta)$ are also invertible. It then follows that $\mathbf{I}_u^{-1}(\theta) \leq \mathbf{I}^{-1}(\theta) \leq \mathbf{I}_l^{-1}(\theta)$ (Zhang, 1999, pp. 169) and from this the result follows. $\qquad \square$

The above result provides a general formula to calculate the upper and lower bound to the (inverse) Fisher information matrix $\mathbf{I}(\theta)$. We next consider a special case in which the extended version $\mathcal{G}^e$ is defined over $\mathbb{R}^2$. Here, the parameter vector is given by $\theta = (x_0, y_0, \Lambda_0) \in \Theta$, the photon distribution profile $f_{\theta,\tau}^e$ of $\mathcal{G}^e$ is assumed to be given in terms

of a symmetric image function q and the photon detection rate Λ_θ^e of $\mathcal{G}^e$ is independent of x_0 and y_0 for $\theta \in \Theta$. We assume that the circular bounding detectors of $\mathcal{C}$ are centered at the point $r_c := (Mx_0, My_0) \in \mathcal{C}$, where $M > 0$ denotes the lateral magnification. For this case, we show that the matrices $\mathbf{I}_l(\theta)$ and $\mathbf{I}_u(\theta)$ are diagonal.

Corollary 15 Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space and let $\mathcal{G}$ be an image detection process that is defined over the detector $\mathcal{C}$. Let $M > 0$. Assume that $\mathcal{G}^e(\Lambda_\theta^e, \{f_{\theta,\tau}^e\}_{\tau \geq t_0}, \mathbb{R}^2)$ is an extended version of $\mathcal{G}$ and that there exist a symmetric image function q such that $f_{\theta,\tau}^e(r) = \frac{1}{M^2} q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)$ for $r = (x, y) \in \mathbb{R}^2$, $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$. Assume that $\partial \Lambda_\theta^e(\tau)/\partial x_0 = \partial \Lambda_\theta^e(\tau)/\partial y_0 = 0$ for $\tau \geq t_0$ and $\theta \in \Theta$. Let $\theta \in \Theta$, assume that $r_c := (Mx_0, My_0) \in \mathcal{C}$, and let $\mathcal{B}_{r_c}(l)$ and $\mathcal{B}_{r_c}(u)$ denote the circular bounding detectors of $\mathcal{C}$. Then

1. $\mathbf{I}_u(\theta) \geq \mathbf{I}(\theta) \geq \mathbf{I}_l(\theta).$ (28)

where

$$\mathbf{I}(\theta) = \int_{t_0}^t \int_{\mathcal{C}} \frac{\Lambda_\theta^e(\tau)}{f_{\theta,\tau}^e(r)} \begin{bmatrix} M\frac{\partial f_{\theta,\tau}^e(r)}{\partial x} \\ M\frac{\partial f_{\theta,\tau}^e(r)}{\partial y} \\ -\frac{f_{\theta,\tau}^e(r)}{\Lambda_\theta(\tau)}\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} \begin{bmatrix} M\frac{\partial f_{\theta,\tau}^e(r)}{\partial x} & M\frac{\partial f_{\theta,\tau}^e(r)}{\partial y} & -\frac{f_{\theta,\tau}^e(r)}{\Lambda_\theta(\tau)}\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} dr d\tau,$$

(29)

$$\mathbf{I}_\beta(\theta) = \begin{bmatrix} \int_{t_0}^t \Lambda_\theta^e(\tau) d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x,y)}\left(\frac{\partial q(x,y)}{\partial x}\right)^2 dxdy & 0 & 0 \\ 0 & \int_{t_0}^t \Lambda_\theta^e(\tau) d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x,y)}\left(\frac{\partial q(x,y)}{\partial y}\right)^2 dxdy & 0 \\ 0 & 0 & \int_{t_0}^t \frac{1}{\Lambda_\theta^e(\tau)}\left(\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}\right)^2 d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} q(x,y) dxdy \end{bmatrix},$$

(30)

with $\mathcal{B}_0(\beta/M) = \{(x, y) \mid \sqrt{x^2 + y^2} < \frac{\beta}{M}\}$ and $\beta \in \{u, l\}$.

2. Further, if $\mathbf{I}_l(\theta)$ is invertible, then

$$[\mathbf{I}_u^{-1}(\theta)]_{ii} \leq [\mathbf{I}^{-1}(\theta)]_{ii} \leq [\mathbf{I}_l^{-1}(\theta)]_{ii}, \quad i = 1, 2, 3. \tag{31}$$

Proof Let $\theta \in \Theta$. Equations 28 and 31 immediately follow by noting that $\mathcal{B}_{r_c}(l) \subseteq \mathcal{C} \subseteq \mathcal{B}_{r_c}(u) \subseteq \mathbb{R}^2$ and that the results of Theorem 14 hold for the present case. In rest of this proof we derive the integral expressions for $\mathbf{I}(\theta)$ and $\mathbf{I}_\beta(\theta)$ that are given in Eqs. 29 and 30, respectively.

By assumption, $\partial \Lambda_\theta^e(\tau)/\partial x_0 = \partial \Lambda_\theta^e(\tau)/\partial y_0 = \partial f_{\theta,\tau}^e(r)/\partial \Lambda_0 = 0$ and it can be shown that $\partial f_{\theta,\tau}^e(r)/\partial x_0 = -M(\partial f_{\theta,\tau}^e(r)/\partial x)$ and $\partial f_{\theta,\tau}^e(r)/\partial y_0 = -M(\partial f_{\theta,\tau}^e(r)/\partial y)$, for $r = (x, y) \in \mathbb{R}^2$ and $\tau \geq t_0$. Using these results and substituting for Λ_θ^e and $f_{\theta,\tau}^e$ in result 1 of Theorem 12, we obtain the expression for $\mathbf{I}(\theta)$ that is given in Eq. 29. By definition, for $\beta \in \{u, l\}$, $\mathcal{B}_{r_c}(\beta) = \{(x, y) \mid \sqrt{(x/M - x_0)^2 + (y/M - y_0)^2} < \beta/M\}, (x, y) \in \mathbb{R}^2\}$. Substituting for $f_{\theta,\tau}^e$ and Λ_θ^e in Eq. 26, we have for $\beta \in \{u, l\}$,

$$\mathbf{I}_\beta(\theta) = \int_{t_0}^t \int_{\mathcal{B}_{r_c}(\beta)} \frac{1}{\Lambda_\theta^e(\tau) f_{\theta,\tau}^e(r)} \begin{bmatrix} \Lambda_\theta^e(\tau)\frac{\partial f_{\theta,\tau}^e(r)}{\partial x_0} \\ \Lambda_\theta^e(\tau)\frac{\partial f_{\theta,\tau}^e(r)}{\partial y_0} \\ f_{\theta,\tau}^e(r)\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix}$$

$$\times \begin{bmatrix} \Lambda_\theta^e(\tau)\frac{\partial f_{\theta,\tau}^e(r)}{\partial x_0} & \Lambda_\theta^e(\tau)\frac{\partial f_{\theta,\tau}^e(r)}{\partial y_0} & f_{\theta,\tau}^e(r)\frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0} \end{bmatrix} dr d\tau$$

$$
= \frac{1}{M^2} \int_{t_0}^{t} \Lambda_\theta^e(\tau) \int_{\mathcal{B}_{rc}(\beta)} \frac{1}{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}
$$

$$
\times
\begin{bmatrix}
-M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x} \\
-M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial y} \\
\frac{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}
\begin{bmatrix}
-M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial x} \\
-M \frac{\partial q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\partial y} \\
\frac{q\left(\frac{x}{M} - x_0, \frac{y}{M} - y_0\right)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}^T
\mathrm{d}x\mathrm{d}y\mathrm{d}\tau
$$

$$
= \frac{1}{M^2} \int_{t_0}^{t} \Lambda_\theta^e(\tau) \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(u, v)}
\begin{bmatrix}
M \frac{\partial q(u, v)}{\partial u} \frac{\partial u}{\partial x} \\
M \frac{\partial q(u, v)}{\partial v} \frac{\partial v}{\partial y} \\
-\frac{q(u,v)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}
$$

$$
\times
\begin{bmatrix}
M \frac{\partial q(u, v)}{\partial u} \frac{\partial u}{\partial x} \\
M \frac{\partial q(u, v)}{\partial v} \frac{\partial v}{\partial y} \\
-\frac{q(u,v)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}^T
(M\mathrm{d}u)(M\mathrm{d}v)\mathrm{d}\tau
\qquad
\left(
\begin{aligned}
u &:= \frac{x}{M} - x_0 \\
v &:= \frac{y}{M} - y_0
\end{aligned}
\right)
$$

$$
= \int_{t_0}^{t} \Lambda_\theta^e(\tau) \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x, y)}
\begin{bmatrix}
\frac{\partial q(x,y)}{\partial x} \\
\frac{\partial q(x,y)}{\partial y} \\
-\frac{q(x,y)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}
$$

$$
\times
\begin{bmatrix}
\frac{\partial q(x,y)}{\partial x} & \frac{\partial q(x,y)}{\partial y} & -\frac{q(x,y)}{\Lambda_\theta^e(\tau)} \frac{\partial \Lambda_\theta^e(\tau)}{\partial \Lambda_0}
\end{bmatrix}
\mathrm{d}x\mathrm{d}y\mathrm{d}\tau.
\tag{32}
$$

Since the image function q is symmetric, it can be shown that $\partial q(x, y)/\partial x = -\partial q(-x, y)/\partial x$ and $\partial q(x, y)/\partial y = \partial q(-x, y)/\partial y$ for $(x, y) \in \mathbb{R}^2$ (see Eqs. 12 – 13). Thus we have

$$
[\mathbf{I}_\beta(\theta)]_{12} = [\mathbf{I}_\beta(\theta)]_{21} = \int_{t_0}^{t} \Lambda_\theta^e(\tau)\mathrm{d}\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(x, y)} \frac{\partial q(x, y)}{\partial x} \frac{\partial q(x, y)}{\partial y} \, \mathrm{d}x\mathrm{d}y
$$

$$
= -\int_{t_0}^{t} \Lambda_\theta^e(\tau) \, \mathrm{d}\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(-x, y)} \frac{\partial q(-x, y)}{\partial x} \frac{\partial q(-x, y)}{\partial y}\mathrm{d}x\mathrm{d}y
$$

$$
= -\int_{t_0}^{t} \Lambda_\theta^e(\tau) \, \mathrm{d}\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(u, y)} \frac{\partial q(u, y)}{\partial u} \frac{\partial u}{\partial x} \frac{\partial q(u, y)}{\partial y} (-\mathrm{d}u)\mathrm{d}y \quad (u := -x)
$$

$$
= -\int_{t_0}^{t} \Lambda_\theta^e(\tau) \, \mathrm{d}\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{q(u, y)} \frac{\partial q(u, y)}{\partial u} \frac{\partial q(u, y)}{\partial y} \mathrm{d}u\mathrm{d}y
$$

$$
= -[\mathbf{I}_l(\theta)]_{12} = -[\mathbf{I}_l(\theta)]_{21}, \quad \beta \in \{u, l\}.
$$

Hence $[\mathbf{I}_\beta(\theta)]_{12} = [\mathbf{I}_\beta(\theta)]_{21} = 0$, $\beta \in \{u, l\}$. Similarly, we can show that $[\mathbf{I}_\beta(\theta)]_{13} = [\mathbf{I}_\beta(\theta)]_{31} = 0$, $\beta \in \{u, l\}$. Further, by using the symmetry property of q we can also show that $\partial q(x, y)/\partial y = -\partial q(-x, y)/\partial y$ for $(x, y) \in \mathbb{R}^2$ (see Eq. 16). From this it follows that $[\mathbf{I}_\beta(\theta)]_{23} = [\mathbf{I}_\beta(\theta)]_{32} = 0$, $\beta \in \{u, l\}$. Substituting these in Eq. 32 the result follows. □

For the localization accuracy problem the upper and lower bounds of the limit of the localization accuracy of (x_0, y_0) are referred to as the *localization accuracy bounds*.

Example-2

We now illustrate the results derived in this section by considering specific image functions that describe the image of a point source. Here, the parameter vector is set to be $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and the photon distribution profile $f_{\theta,\tau}^e$ is assumed to be given in terms of a symmetric image function q. Further, the photon detection rate is assumed to be a constant, i.e., $\Lambda_\theta^e(\tau) = \Lambda_0, \tau \geq t_0$. For each image function, we derive the expression for the Fisher information matrix $\mathbf{I}(\theta)$ corresponding to the detector $\mathcal{C}$ and also derive the upper and lower bound for $\sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$, $i = 1, 2, 3$, which denotes the limit of the accuracy of the components of θ.

Corollary 16 Let $\Theta \subseteq \mathbb{R}^3$ be a parameter space and let $\mathcal{G}$ be an image detection process that is defined over the detector $\mathcal{C}$. Let $M > 0$. Assume that $\mathcal{G}^e(\Lambda_\theta^e, \{f_{\theta,\tau}^e\}_{\tau \geq t_0}, \mathbb{R}^2)$ is an extended version of $\mathcal{G}$ and that there exists a symmetric image function q such that $f_{\theta,\tau}^e(r) = \frac{1}{M^2} q \left(\frac{x}{M} - x_0, \frac{y}{M} - y_0 \right)$ for $r = (x, y) \in \mathbb{R}^2$, $\theta = (x_0, y_0, \Lambda_0) \in \Theta$ and $\tau \geq t_0$. Assume that $\Lambda_\theta^e(\tau) = \Lambda_0$ for $\tau \geq t_0$ and $\theta \in \Theta$. Let $\theta \in \Theta$, assume that $r_c := (Mx_0, My_0) \in \mathcal{C}$, and let $\mathcal{B}_{r_c}(l)$ and $\mathcal{B}_{r_c}(u)$ denote the circular bounding detectors of $\mathcal{C}$.

1. *Airy profile:* If q is given by Eq. 20, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \Lambda_0(t - t_0) \int_{\mathcal{C}} \frac{1}{\frac{J_1^2(a||r - r_c||)}{\pi ||r - r_c||^2}} \mathcal{Q}_\theta^T(r) \mathcal{Q}_\theta(r) \, dr,$$

where $a = 2\pi n_a/(\lambda M)$, $||r - r_c|| := \sqrt{(x - Mx_0)^2 + (y - My_0)^2}$ and

$$\mathcal{Q}_\theta(r) := \frac{J_1(a||r - r_c||)}{\pi ||r - r_c||} \left[\frac{2aM(x - Mx_0)}{||r - r_c||^2} J_2(a||r - r_c||) \quad \frac{2aM(y - My_0)}{||r - r_c||^2} J_2(a||r - r_c||) \quad -\frac{1}{\Lambda_0} \frac{J_1(a||r - r_c||)}{||r - r_c||} \right].$$

Moreover, if u and l are as defined above, then

$$\frac{\lambda / \left(2\pi n_a \sqrt{\Lambda_0(t - t_0)} \right)}{\sqrt{1 - \left(J_0^2(au) + 2J_1^2(au) + J_2^2(au) \right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$$

$$\leq \frac{\lambda / \left(2\pi n_a \sqrt{\Lambda_0(t - t_0)} \right)}{\sqrt{1 - \left(J_0^2(al) + 2J_1^2(al) + J_2^2(al) \right)}}, \quad i = 1, 2, \tag{33}$$

$$\frac{\sqrt{\Lambda_0/(t - t_0)}}{\sqrt{1 - \left(J_0^2(au) + J_1^2(au) \right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{33}} \leq \frac{\sqrt{\Lambda_0/(t - t_0)}}{\sqrt{1 - \left(J_0^2(al) + J_1^2(al) \right)}}, \tag{34}$$

where J_n denotes the nth order Bessel function of the first kind, $n = 0, 1, 2$.

2. *2D Gaussian profile:* If q is given by Eq. 22, then the Fisher information matrix of $\mathcal{G}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \Lambda_0(t - t_0) \int_{\mathcal{C}} \frac{2\pi(M\sigma)^2}{e^{-\frac{\|r - r_c\|^2}{2(M\sigma)^2}}} \mathcal{Q}_\theta^T(r) \mathcal{Q}_\theta(r) \, dr,$$

where $\sigma > 0$, $\|r - r_c\| := \sqrt{(x - Mx_0)^2 + (y - My_0)^2}$ and

$$\mathcal{Q}_\theta(r) := \frac{1}{2\pi(M\sigma)^2} e^{-\frac{\|r - r_c\|^2}{2(M\sigma)^2}} \left[\frac{(x - Mx_0)}{M\sigma^2} \quad \frac{(y - My_0)}{M\sigma^2} \quad -\frac{1}{\Lambda_0} \right].$$

If u and l are as defined above, then

$$\frac{\sigma/\sqrt{\Lambda_0(t - t_0)}}{\sqrt{1 - e^{-\frac{1}{2}\left(\frac{u}{M\sigma}\right)^2}\left(1 + \frac{1}{2}\left(\frac{u}{M\sigma}\right)^2\right)}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{ii}}$$

$$\leq \frac{\sigma/\sqrt{\Lambda_0(t - t_0)}}{\sqrt{1 - e^{-\frac{1}{2}\left(\frac{l}{M\sigma}\right)^2}\left(1 + \frac{1}{2}\left(\frac{l}{M\sigma}\right)^2\right)}}, \quad i = 1, 2, \tag{35}$$

$$\frac{\sqrt{\Lambda_0/(t - t_0)}}{\sqrt{1 - e^{-\frac{1}{2}\left(\frac{u}{M\sigma}\right)^2}}} \leq \sqrt{[\mathbf{I}^{-1}(\theta)]_{33}} \leq \frac{\sqrt{\Lambda_0/(t - t_0)}}{\sqrt{1 - e^{-\frac{1}{2}\left(\frac{l}{M\sigma}\right)^2}}}. \tag{36}$$

Proof We can show that the Airy profile and the 2D Gaussian profile satisfy the properties of a symmetric image function. Let $\theta \in \Theta$. It can be verified that for the present case the results of Corollary 15 hold. Then the expressions for the Fisher information matrix $\mathbf{I}(\theta)$ immediately follow by substituting the corresponding image function in Eq. 29. Further, we can also verify that $\mathbf{I}^{-1}(\theta)$ exists for each image function. In rest of this proof we derive the expressions for the upper and lower bounds of $\mathbf{I}(\theta)$.

1. Let $\alpha := 2\pi n_a/\lambda$. Substituting for Λ_θ^e and q in the integral expression of $[\mathbf{I}_\beta(\theta)]_{11}$ ($[\mathbf{I}_\beta(\theta)]_{22}$) that is given in Eq. 30, we have for $\beta \in \{u, l\}$,

$$[\mathbf{I}_\beta(\theta)]_{11} = [\mathbf{I}_\beta(\theta)]_{22} = \int_{t_0}^t \Lambda_0 \, d\tau \int_{\mathcal{B}_0\left(\frac{\beta}{M}\right)} \frac{1}{\frac{J_1^2(\alpha\sqrt{x^2 + y^2})}{\pi(x^2 + y^2)}} \left[\frac{\partial}{\partial x}\left(\frac{J_1^2(\alpha\sqrt{x^2 + y^2})}{\pi(x^2 + y^2)} \right) \right]^2 dxdy$$

$$= \frac{4\alpha^2}{\pi} \Lambda_0(t - t_0) \int_{\{(x,y)|\sqrt{x^2+y^2}<\beta/M\}} \frac{J_2^2(\alpha\sqrt{x^2 + y^2})}{(x^2 + y^2)^2} \, dxdy$$

$$= \frac{4\alpha^2}{\pi} \Lambda_0(t - t_0) \int_0^{2\pi} \cos^2 \phi \, d\phi \int_0^{\frac{\beta}{M}} \frac{J_2^2(\alpha\rho)}{\rho} \, d\rho$$

$$= 4\alpha^2 \Lambda_0(t - t_0) \int_0^{a\beta} \frac{J_2^2(w)}{w} \, dw = 4\alpha^2 \Lambda_0(t - t_0)$$

$$\times \left(\int_0^\infty \frac{J_2^2(w)}{w} \, dw - \int_{a\beta}^\infty \frac{J_2^2(w)}{w} \, dw \right)$$

$$= \frac{1 - \left(J_0^2(a\beta) + 2J_1^2(a\beta) + J_2^2(a\beta)\right)}{\lambda^2/((2\pi n_a)^2 \Lambda_0(t - t_0))},$$

where $x = \rho \cos \phi$, $y = \rho \sin \phi$, $a = \alpha/M$, the partial derivative of q with respect to x is given in Eq. 24, and the integral expressions in the final step are evaluated by using the integral identities $\int_0^\infty (J_n^2(t)/t) \, dt = (1/2n)$ (Watson, 1958, pp. 405) and $n \int_x^\infty (J_n^2(t)/t)$

$dt = (1/2)J_0^2(x) + J_1^2(x) + \cdots + J_{n-1}^2(x) + (1/2)J_n^2(x)$ (Bowman, 1968, pp. 95) with $n = 2$ and $x \in \mathbb{R}$. Similarly, for $\beta \in \{u, l\}$, we have

$$[\mathbf{I}_\beta(\theta)]_{33} = \frac{1}{\pi}\int_{t_0}^t \frac{1}{\Lambda_0}\,d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{J_1^2(\alpha\sqrt{x^2+y^2})}{x^2+y^2}\,dxdy = \frac{(t-t_0)}{\pi\Lambda_0}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M}}\frac{J_1^2(\alpha\rho)}{\rho}\,d\rho$$

$$= \frac{2(t-t_0)}{\Lambda_0}\int_0^{\alpha\beta}\frac{J_1^2(w)}{w}\,dw = \frac{1 - (J_0^2(a\beta)+J_1^2(a\beta))}{\Lambda_0/(t-t_0)}.$$

Using result 2 of Corollary 15 the result follows.

2. Substituting for Λ_θ^e and q in the integral expression of $[\mathbf{I}_\beta(\theta)]_{11}$ ($[\mathbf{I}_\beta(\theta)]_{22}$) that is given in Eq. 30, we have for $\beta \in \{u, l\}$,

$$[\mathbf{I}_\beta(\theta)]_{11} = [\mathbf{I}_\beta(\theta)]_{22} = \int_{t_0}^t \Lambda_0\,d\tau \int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{1}{\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}}\left(\frac{\partial}{\partial x}\left(\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\right)\right)^2\,dxdy$$

$$= \frac{\Lambda_0(t-t_0)}{2\pi\sigma^2}\int_{\mathcal{B}_0(\frac{\beta}{M})} \frac{x^2}{\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}\frac{dxdy}{\sigma^2} = \frac{\Lambda_0(t-t_0)}{2\pi\sigma^2}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M\sigma}}\rho^3 e^{-\frac{\rho^2}{2}}\,d\rho$$

$$= \frac{\Lambda_0(t-t_0)}{4\sigma^2}\int_0^{\frac{\beta}{M\sigma}}\rho^2 e^{-\frac{\rho^2}{2}}2\rho\,d\rho$$

$$= \frac{\Lambda_0(t-t_0)}{4\sigma^2}\int_0^{\left(\frac{\beta}{M\sigma}\right)^2}we^{-\frac{w}{2}}\,dw = \frac{1 - e^{-\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2}\left(1+\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2\right)}{\sigma^2/(\Lambda_0(t-t_0))},$$

$$[\mathbf{I}_\beta(\theta)]_{33} = \frac{1}{2\pi\sigma^2}\int_{t_0}^t\frac{1}{\Lambda_0}d\tau\int_{\mathcal{B}_0(\frac{\beta}{M})}e^{-\frac{x^2+y^2}{2\sigma^2}}\,dxdy$$

$$= \frac{(t-t_0)}{\Lambda_0 2\pi}\int_0^{2\pi}d\phi\int_0^{\frac{\beta}{M\sigma}}\rho e^{-\frac{\rho^2}{2}}\,d\rho = \frac{(t-t_0)}{2\Lambda_0}\int_0^{\frac{\beta}{M\sigma}}2\rho e^{-\frac{\rho^2}{2}}\,d\rho$$

$$= \frac{(t-t_0)}{2\Lambda_0}\int_0^{\left(\frac{\beta}{M\sigma}\right)^2}e^{-\frac{w}{2}}\,dw = \frac{1 - e^{-\frac{1}{2}\left(\frac{\beta}{M\sigma}\right)^2}}{\Lambda_0/(t-t_0)}.$$

Using result 2 of Corollary 15 we obtain the desired result. □

From the above Corollary we see that for both image functions, the localization accuracy bounds for x_0 (y_0) and the bounds for the limit of the accuracy of Λ_0 reduce to simple formulae. Note that the above results for the upper and lower bounds hold only if the point (Mx_0, My_0) is located on the detector $\mathcal{C}^{rd}$. In most experimental situations this condition is satisfied.

We now discuss the results derived in Corollary 16 by considering a finite-sized square detector. Figure 2a (Figure 2b) shows the behavior of the limit of the localization accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{11}}$ for x_0 for a square detector corresponding to the Airy profile (2D Gaussian profile) as a function of detector size. A reduced detector detects relatively lower number of photons than an infinite detector. Hence the limit of the localization accuracy for the reduced detector is greater (worse) than the fundamental limit of the localization accuracy that is calculated for the infinite detector (see Corollary 10). As the detector size increases, more photons are detected by the reduced detector and the limit of the localization accuracy approaches the fundamental limit. Figure 2a (Figure 2b) also shows the behavior of
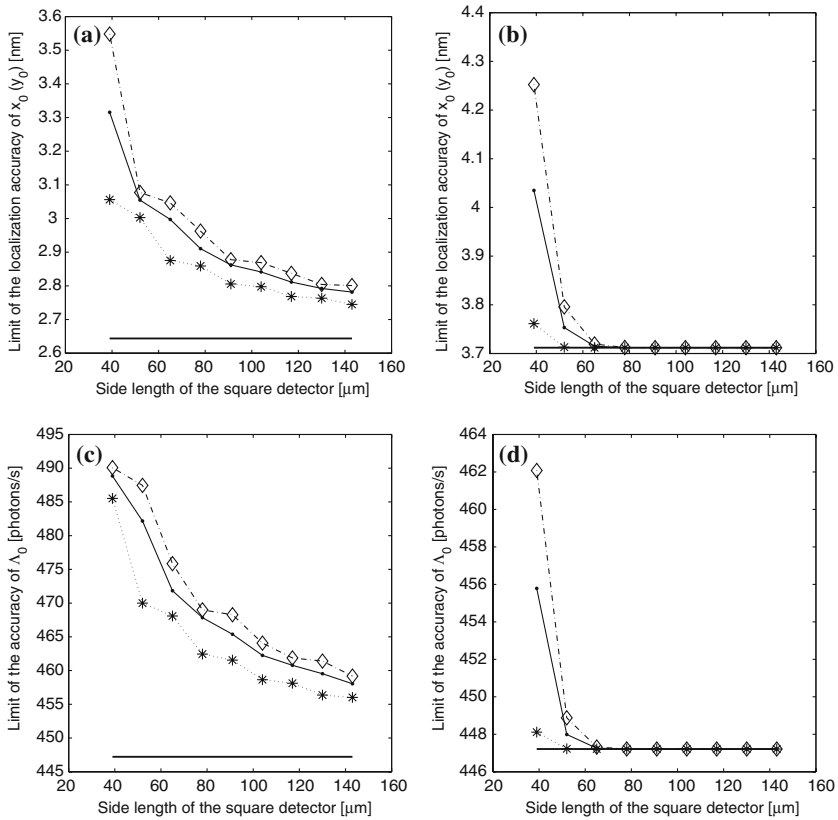
Fig. 2 **(a)** shows the behavior of the limit of the localization accuracy of x_0 (y_0) for a square detector corresponding to the Airy profile (•) as a function of detector size and **(b)** shows the same for the 2D Gaussian profile (•). **(c)** shows the behavior of the limit of the accuracy of Λ_0 for a square detector corresponding to the Airy profile (•) as a function of detector size and **(d)** shows the same for the 2D Gaussian profile (•). In all the panels the corresponding fundamental limit of the accuracy (—) and the upper ($\diamond$) and lower ($*$) bound to the limit of the accuracy for the square detector are shown. For a square detector with side length s, the position of the point source on the detector is set to be $r_c = (s/2, s/2)$, the radius of the lower circular bounding detector $\mathcal{B}_{r_c}(l)$ is $l = s/2$ and the radius of the upper circular bounding detector $\mathcal{B}_{r_c}(u)$ is $u = s/\sqrt{2}$. For the Airy profile, the numerical aperture is set to be $n_a = 1.4$ and the wavelength of the detected photons is set to be $\lambda = 0.52\,\mu m$. The parameter σ corresponding to the Gaussian profile is set to be $\sigma = 0.083\,\mu m$ and is determined by fitting a 2D Gaussian profile to an Airy profile ($n_a = 1.4$, $\lambda = 0.52\,\mu m$) through the least squares criterion. For all the plots, $\Lambda_0 = 10^4$ photons/sec, the acquisition time is set to be $t = 50\,msec$ (with $t_0 = 0$) and the magnification is set to be $M = 100$

the localization accuracy bounds given by Eq. 33 (Eq. 35) for a square detector corresponding to the Airy profile. Here, we see that the localization accuracy bounds provide a tight bound, as they are consistently close to the limit of the localization accuracy for the square detector.

Note that the behavior of the limit of the localization accuracy for a square detector also depends on the functional form of the image function. In the case of the 2D Gaussian profile (see Figure 2b), the limit of the localization accuracy for a square detector with side length $80\,\mu m$ is close to the fundamental limit of the localization accuracy. However, this is not

the case for the Airy profile (see Figure 2a), where even for a square detector with side length $140\,\mu m$, the limit of the localization accuracy does not come close to the fundamental limit. In a practical application such as single molecule data analysis, the above observation provides guidelines for choosing the optimal size of the region of interest on the acquired image. Moreover, it also shows the importance of using the correct image function, as this influences the behavior of the limit of the localization accuracy.

Figure 2c (Figure 2d) shows the variation of the limit of the accuracy $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$ of Λ_0 for a square detector corresponding to the Airy profile (2D Gaussian profile) as a function of detector size. Analogous to the behavior of the limit of the localization accuracy, the limit of the accuracy of Λ_0 approaches the fundamental limit of the accuracy of Λ_0 ($= \sqrt{\Lambda_0/(t - t_0)}$, see Eqs. 21 and 23) as the detector size increases. Moreover, the behavior of the limit of the accuracy of Λ_0 also depends on the functional form of the image function. Figure 2c (Figure 2d) also shows the behavior of the upper and lower bounds to the limit of the accuracy of Λ_0 that is given by Eq. 34 (Eq. 36) for a square detector corresponding to the Airy profile (2D Gaussian profile). Similar to the localization accuracy bounds (Figure 2a and 2b), we see that the upper and lower bounds for the limit of the accuracy of Λ_0 provide a tight bound.

Effects of pixelation

In all our results so far the detector $\mathcal{C}$ is such that the acquired data consist of the time points and the spatial coordinates of the detected photons. However, in the presence of pixelation the acquired data consist of the number of detected photons at each pixel. We next show how this data can be described in terms of the photon distribution profile and the photon detection rate of an image detection process. Here, we follow an approach that was introduced in Ober, Ram et al., (2004) to address the localization problem. This approach will be generalized and applied to the general parameter estimation problem that is discussed in this paper. Let $\mathcal{G}^1(\Lambda_\theta^1, \{f_{\theta,\tau}^1\}_{\tau \geq t_0}, \mathcal{C})$ denote an image detection process that models the detected photons from the object of interest. The pixelated version of the detector $\mathcal{C}$ is defined as a collection $\{C_1, \ldots, C_{N_p}\}$ of open, disjoint subsets of $\mathbb{R}^2$ such that $\bigcup_{k=1}^{N_p} C_k = \mathcal{C}$, where N_p denotes the total number of pixels. For $k = 1, \ldots, N_p$ and $t \geq t_0$, assume that n_k photons are detected in the pixel C_k during the time interval $[t_0, t]$. Let K denote the total number of detected photons from the object of interest, i.e., $\sum_{k=1}^{N_p} n_k = K$. Then it can be shown that for $k = 1, \ldots, N_p, n_k$ is independently Poisson distributed with mean $\mu_\theta(k, t) = \int_{t_0}^t \int_{C_k} \Lambda_\theta^1(\tau) f_{\theta,\tau}^1(r)\, dr d\tau, \theta \in \Theta$. Similarly, the number of detected photons at the kth pixel during the time interval $[t_0, t]$ from a background component $\mathcal{G}^2(\Lambda^2, \{f_\tau^2\}_{\tau \geq t_0}, \mathcal{C})$ is independently Poisson distributed with mean $\beta(k, t) = \int_{t_0}^t \int_{C_k} \Lambda^2(\tau) f_\tau^2(r)\, dr d\tau, \theta \in \Theta$. Hence the acquired data in the time interval $[t_0, t]$ from a pixelated detector can be described by a collection $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ of random variables given by

$$\mathcal{I}_{\theta,k} = S_{\theta,k} + B_k, \quad \theta \in \Theta, \quad k = 1, \ldots, N_p.$$

Using the standard expression for the Fisher information matrix of a Poisson distribution (snyder & Miller 1991), the Fisher information matrix for $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = \sum_{k=1}^{N_p} \frac{1}{\mu_\theta(k, t) + \beta(k, t)} \left(\frac{\partial \mu_\theta(k, t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k, t)}{\partial \theta}. \tag{37}$$

In a pixelated detector the acquired image contains measurement noise, which, for example, arises due to the readout process (Snyder, Helstrom, Lanterman, & White, 1995). At each pixel this can be modeled as a Gaussian random variable W_k with mean η_k and variance $\sigma_{w,k}^2$, $k = 1, \ldots, N_p$. The acquired image is then given by $\mathcal{I}_{\theta,k} = S_{\theta,k} + B_k + W_k$, $\theta \in \Theta$, $k = 1, \ldots, N_p$. To derive the Fisher information matrix for the present case, we first note that $\mathcal{I}_{\theta,k}$ is a sum of a Poisson and independent Gaussian random variable, and its probability density function is given by (see Snyder et al., 1995)

$$p_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\sigma_{w,k}} \sum_{l=0}^{\infty} \frac{[\nu_\theta(k,t)]^l e^{-\nu_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad z \in \mathbb{R}, \quad k = 1, \ldots, N_p,$$

where $\nu_\theta(k,t) := \mu_\theta(k,t) + \beta(k,t)$, $k = 1, \ldots, N_p$. If $\{n_1, \ldots, n_{N_p}\}$ denotes the acquired data, then the log likelihood function is given by $\mathcal{L}(\theta \mid n_1, \ldots, n_{N_p}) := \sum_{k=1}^{N_p} \ln[p_{\theta,k}(n_k)]$, $\theta \in \Theta$ and the partial derivative of the log-likelihood function with respect to θ is given by

$$\frac{\partial \mathcal{L}(\theta \mid n_1, \ldots, n_{N_p})}{\partial \theta} = \sum_{k=1}^{N_p} \left[\frac{\partial \mu_\theta(k,t)}{\partial \theta} (\zeta_{\theta,k}(n_k) - 1) \right], \quad \theta \in \Theta, \tag{38}$$

where

$$\zeta_{\theta,k}(z) := \frac{\sum_{l=1}^{\infty} \frac{l[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{k,w}}\right)}}{p_{\theta,k}(z)}, \quad \theta \in \Theta, \quad k = 1, \ldots, N_p, \quad z \in \mathbb{R}.$$

It can be shown that $E[\zeta_{\theta,k}(n_k)] = 1$ for $\theta \in \Theta$ and $k = 1, \ldots, N_p$. Further, it can be verified that the random variables $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ are mutually independent. Using these results and Eq. 38, the Fisher information matrix for $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$ corresponding to the time interval $[t_0, t]$ is given by

$$\mathbf{I}(\theta) = E\left[\left(\frac{\partial \mathcal{L}(\theta \mid n_1, \ldots, n_{N_p})}{\partial \theta} \right)^T \frac{\partial \mathcal{L}(\theta \mid n_1, \ldots, n_{N_p})}{\partial \theta} \right]$$

$$= E\left[\sum_{k=1}^{N_p} \sum_{m=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(m,t)}{\partial \theta} \left(\zeta_{\theta,k}(n_k)\zeta_{\theta,m}(n_m) - \zeta_{\theta,k}(n_k) - \zeta_{\theta,m}(n_m) + 1 \right) \right]$$

$$= \sum_{k=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \left(E[\zeta_{\theta,k}^2(n_k)] - 1 \right)$$

$$+ \sum_{k \neq m, k, m=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(m,t)}{\partial \theta} \left(E[\zeta_{\theta,k}(n_k)]E[\zeta_{\theta,m}(n_m)] - 1 \right)$$

$$= \sum_{k=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta} \right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \times$$

$$\left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2} \right)^2}{p_{\theta,k}(z)} \, dz - 1 \right), \quad \theta \in \Theta. \tag{39}$$

The above expressions are valid for general intensities $\mu_\theta(k, t)$ ($\nu_\theta(k, t)$), $k = 1, \ldots, N_p$, $t > t_0$. These intensities depend on the photon distribution profiles $f_{\theta,\tau}$, $\tau \geq t_0$, through the above identities. The approach is therefore generally applicable to a large class of photon distributions profiles/image functions. Note that in the above expressions no assumptions are made about the size or shape of the pixels. In Winick (1986) an expression was derived that is essentially a special case of Eq. 37, i.e. for a two-dimensional estimation problem of the location parameters for a Gaussian photon distribution profile in a stationary imaging scenario for square pixels in the absence of Gaussian noise.

Example-3

In this section we present a numerical comparison of the standard deviation of the maximum likelihood estimator and the performance limit for the problem of estimating the 2D location of a stationary point source. All calculations are carried out in the Matlab programming language (Coleman, Branch, & Grace, 2000). The image function of the point source is described by the Airy profile given in Eq. 20. The parameter vector θ is given by $\theta = (x_0, y_0)$ and it is assumed that the photon detection rate of the point source is a constant and is known, i.e., $\Lambda_\theta^1(\tau) := \Lambda_0^1$, $\theta \in \Theta$, $\tau \geq t_0$. The data consists of a sequence of 50 images of the point source that are simulated for a finite-sized detector in the presence of noise sources. Maximum likelihood estimation is carried out by using a gradient based search algorithm (Coleman, Branch, & Grace, 1999). Table 1 lists the standard deviations of the maximum likelihood estimates of the 2D location that are calculated for three different imaging conditions. The performance limit is calculated by using the expression for the Fisher information matrix given in Eq. 39. From the table we see that for all the data sets, the standard deviation of the maximum likelihood estimator comes close to the performance limit. For example, in data set 1 the performance limit predicts an accuracy not smaller than ± 5.01 nm to determine the x_0 and the y_0 coordinates of the point source. The corresponding standard deviations of the maximum likelihood estimator for the x_0 and the y_0 coordinates are ± 5.4 nm and ± 5.19 nm, respectively. Note that due to the finite number of location estimates, the standard deviations can be greater or smaller than the performance limit. With the increase in the number of location estimates, further agreement can be expected between the standard deviations and the performance limit. We note that the simulations presented here supplement the results given in Ober, Ram et al., (2004), where the performance of the maximum likelihood estimator has been compared with the performance limit for other imaging conditions.

Table 1 The table lists the performance limit to determining the 2D location of a point source along with the standard deviation of the maximum likelihood estimator for different imaging conditions

Data set No.	Pixel size ($\mu m \times \mu m$)	No. of pixels in the simulated image N_p	Photon detection rate Λ_0^1 (photon/s)	True location of point source in object space (x_0, y_0) (nm,nm)	Performance limit to determining x_0 and y_0 (nm)	Mean of the ML estimator for x_0 (nm)	Standard deviation of the ML estimator for x_0 (nm)	Mean of the ML estimator for y_0 (nm)	Standard deviation of the ML estimator for y_0 (nm)
1	9.6×9.6	15×15	1000	(720,720)	5.01	720.43	5.4	720.43	5.19
2	13×13	11×11	5000	(715,715)	1.47	715.05	1.57	715.08	1.57
3	16×16	9×9	5000	(720,720)	1.57	719.82	1.40	720.12	1.66

For all calculations the magnification is set to be $M = 100$, the numerical aperture is set to be $n_a = 1.4$, the emission wavelength is set to be $\lambda_{em} = 520$ nm, the mean of the background component is set to be $\beta(k, t) = 30$ photons/pixel/s, the acquisition time is set to be 1 s, the mean and the standard deviation of the measurement noise are set to be $\eta_{lk} = 0$, $k = 1, \ldots, N_p$ and $\sigma_{w,k} = 8\,e^-$/pixel (rms), $k = 1, \ldots, N_p$, respectively, where N_p denotes the total number of pixels in the simulated image. For all the data sets, the point source is positioned at the center of the pixel array, the performance limit is calculated for a pixelated detector in the presence of noise sources and the standard deviation of the maximum likelihood estimator for x_0 and y_0 are obtained from 50 estimates of the point source locations.

References

Bobroff, N. (1986). Position measurement with a resolution and noise limited instrument. *Review of Scientific Instruments, 57*(6), 1152–1157.

Born, M., & Wolf, E. (1999). *Principles of optics.* Cambridge, UK: Cambridge University Press.

Bowman, F. (1968). *Introduction to Bessel functions.* New York: Dover.

Coleman, T., Branch, M.A., & Grace, A. (1999). *MATLAB Optimization Toolbox user manual ver. 2.* Natick: The Mathworks, Inc.

Coleman, T., Branch, M.A., & Grace, A. (2000). *MATLAB programming reference manual, Version 6.* MA, The Mathworks Inc.

Goodman, J. W. (1996). *Introduction to Fourier optics* (2nd ed.). USA: Mc Graw Hill.

Kay, S. M. (1993). *Fundamentals of statistical signal processing.* New Jersey: Prentice Hall PTR.

Kubitscheck, U., Kückmann, O., Kues, T., & Peters, R. (2000). Imaging and tracking single GFP molecules in solution. *Biophysical Journal, 78*, 2170–2179.

Michalet, X., Kapanidis, A. N., Laurence, T., Pinaud, F., Doose, S., Pflughoefft, M., & Weiss, S. (2003). The power and prospects of fluorescence microscopies and spectroscopies. *Annual Review of Biophysics and Biomolecular Structure, 32*, 161–182.

Moerner, W. E., & Fromm, D. P. (2003). Methods of single-molecule fluorescence spectroscopy and microscopy. *Review of Scientific Instruments, 74*(8), 3597–3619.

Ober, R. J., Ram, S., & Ward, E.S. (2004). Localization accuracy in single molecule microscopy. *Biophysical Journal, 86*, 1185–1200.

Ober, R. J., Martinez, C., Lai, X., Zhou, J., and Ward, E. S. (2004). Exocytosis of IgG as mediated by the receptor, FcRn: An analysis at the single molecule level. *Proceeding of the National Academy of Sciences, 101*(30), 11076–11081.

Papoulis, A., & Pillai, S.U. (2002) *Probability, random variables and stochastic processes.* (4th ed.). New York: McGraw Hill.

Ram, S. (in preparation). Ph.D. Dissertation, University of Texas at Arlington/University of Texas Southwestern Medical Center at Dallas

Ram, S., Ward, E. S., & Ober, R. J. (2005). How accurately can a single molecule be localized in three dimensions using a fluorescence microscope? *Proceeding of SPIE, 5699*, 426–435.

Rudin, W. (1987). *Real and complex analysis.* New York: Mc Graw Hill.

Saleh, B. (1978). *Photoelectron statistics.* Berlin, Germany: Springer Verlag.

Santos, A., & Young, I.T. (2000). Model-based resolution: Applying the theory in quantitative microscopy. *Applied Optics, 39*, 2948–2958.

Saxton, M. J., & Jacobson, K. (1997). Single particle tracking: Applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure, 26*, 373–399.

Snyder, D. L., & Miller, M.I. (1999). *Random point processes in time and space* (2nd ed.). New York: Springer Verlag.

Snyder, D. L., Helstrom, C.W., Lanterman, A.D., & White, R. L. (1995). Compensation for read out noise in charge coupled device images. *Journal of the Optical Society of America A- Optics Image Science and Vision, 12*(2), 272–283.

Thompson, R.E., Larson, D.R., & Webb, W.W. (2002). Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal, 82*, 2775–2783.

Watson, G.N. (1958). *A treatise on the theory of Bessel functions.* Cambridge, UK: Cambridge University Press.

Weiss, S. (1999). Fluorescence spectroscopy of single biomolecules. *Science, 283*, 1676–1683.

Winick, K. A. (1986). Cramer-Rao lower bounds on the performance of charge coupled device optical position estimators. *Jornal of the Optical Society America A-Optics Image Science and Vision, 3*(11), 1809–1815.

Zacks, S. (1971). *The theory of statistical inference.* New York: John Wiley and Sons.

Zhang, F. (1999). *Matrix theory.* New York: Springer Verlag.

Raimund J. Ober received the Ph.D. degree in electrical engineering from the Cambridge University, Cambridge, U.K., in 1987. From 1987 to 1990, he was a Research Fellow at Girton College and the Engineering Department, Cambridge University. In 1990, he joined the University of Texas at Dallas, Richardson, where he is currently a Professor with the Department of Electrical Engineering. He is also Adjunct Professor at the University of Texas Southwestern Medical Center, Dallas. He is an associate editor of Systems and Control Letters and Mathematics of Control, Signals, and Systems. His research interests include system theory, signal and imaging processing, and their applications in bioengineering problems, such as fluorescence microscopy.

Sripad Ram received the B.Sc. degree in applied sciences from PSG College of Technology, Coimbatore, India, in 1999 and the M.Sc. degree in physics from Indian Institute of Technology, Chennai, in 2001. He is currently working toward the Ph.D. degree in the joint Biomedical Engineering graduate program at the University of Texas, Arlington, and the University of Texas Southwestern Medical Center, Dallas. His research interests include statistical image processing and fluorescence microscopy. He is a Student Member of the IEEE, the SPIE and the Biophysical Society.

E. Sally Ward completed her Ph.D. at the Department of Biochemistry at Cambridge University in 1985. From 1985 to 1987, she was a Research Fellow at Gonville and Caius College whilst working at the Department of Biochemistry, Cambridge University. From 1988 to 1990, she held the Stanley Elmore Senior Research Fellowship at Sidney Sussex College and carried out research at the MRC Laboratory of Molecular Biology, Cambridge. In 1990 she joined the Department of Microbiology and Cancer Immunobiology Center at the University of Texas Southwestern Medical Center at Dallas as an Assistant Professor. Since 2002 she has been a Professor at the same institution, and was recently appointed to the Paul and Betty Meek-FINA Professorship in Molecular Immunology. Her research interests include understanding the factors that lead to autoimmune disease and studying how antibody levels are maintained at different sites in the body.