Stochastic Framework

We consider a stochastic framework that is based on an optical microscope setup, which consists of the object of interest, the microscope optics, and a detector that captures the image of the object. Here, the object of interest is a pair of point sources (e.g. single molecules) and we consider imaging experiments that capture photons during a fixed acquisition time interval. Our approach to calculate the resolution measure is based on the theory concerning the Cramer-Rao lower bound (see e.g., refs. 1 and 2). The task of determining the distance of separation d between two point sources is a parameter estimation problem. The distance of separation is obtained by using an unbiased estimation procedure, and the performance of this estimator is given by the standard deviation of the distance estimates assuming repeated experiments. According to the Cramer-Rao lower bound (1, 2), the (co)variance of any unbiased estimator $\hat{\theta}$ of an unknown parameter θ is always greater than or equal to the inverse Fisher information matrix, i.e.,

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

An important property of the Fisher information matrix is that it is independent of how the parameter is estimated and only depends on the statistical description of the acquired data. Because the performance of an unbiased estimator is given by its standard deviation, the above inequality implies that the square root (of the corresponding leading diagonal entry) of the inverse Fisher information matrix provides a lower bound to the performance of any unbiased estimator of θ. Hence we define the resolution measure as the square root (of the corresponding leading diagonal entry) of the inverse Fisher information matrix that corresponds to the distance estimation problem.

Image Detection Process

The acquired data is assumed to consist of the time points and the spatial coordinates of the detected photons from the two point sources, which is modeled as a space-time random process (1) and is referred to as the image detection process. A detailed description of this process has been given elsewhere (3), and for completeness we give a brief description. The temporal part describes the time points of the detected photons and is modeled as an inhomogeneous Poisson process with intensity function Λ_θ. The spatial part describes the distribution of the detected photons over the detector and is modeled as a sequence of independent and identically distributed random variables with density function $f_{\theta,\tau}$. The temporal and spatial components are assumed to be mutually independent of each other. Here, τ denotes the time and θ denotes the unknown parameter that we want to estimate from the acquired data, which, in the present case, is the distance of separation d between the point sources.

Fisher Information Matrix for the Image Detection Process. The Fisher information matrix of the image detection process is given by (1, 3)

$$\mathbf{I}(\theta) = \int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\Lambda_\theta(\tau) f_{\theta,\tau}(r)} \left(\frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial\theta} \right)^T \frac{\partial[\Lambda_\theta(\tau) f_{\theta,\tau}(r)]}{\partial\theta} dr d\tau, \quad \theta \in \Theta, \qquad [\mathbf{1}]$$

where $[t_0, t]$ denotes the time interval during which the data is acquired. In Eq. **1**, we make no specific assumptions about the functional form of $f_{\theta,\tau}$ or Λ_θ. Therefore Eq. **1** provides a general expression for $\mathbf{I}(\theta)$ that is applicable to a wide variety of imaging conditions, such as polarized or unpolarized excitation and detection, total internal reflection mode of illumination, imaging under defocus, etc.

In the present case where our object of interest is a pair of point sources, Λ_θ and $f_{\theta,\tau}$ can be written as (see ref. 3)

$$\Lambda_\theta(\tau) := \Lambda_1(\tau) + \Lambda_2(\tau), \quad \tau \geq t_0, \quad \theta \in \Theta, \qquad [\mathbf{2}]$$

$$f_{\theta,\tau}(r) := \frac{\epsilon_\theta^1(\tau)}{M^2} q_1 \left(\frac{x}{M} + \frac{d}{2}, \frac{y}{M} \right) + \frac{\epsilon_\theta^2(\tau)}{M^2} q_2 \left(\frac{x}{M} - \frac{d}{2}, \frac{y}{M} \right), \quad r := (x, y) \in \mathbb{R}^2, \quad \theta \in \Theta, \quad \tau \geq t_0, [\mathbf{3}]$$

where Θ denotes the parameter space and $\epsilon_\theta^i(\tau) := \Lambda_i(\tau)/\Lambda_\theta(\tau)$, $\tau \geq t_0$, $\theta \in \Theta$, $i = 1, 2$. In Eqs. **2** and **3**, Λ_1 and Λ_2 denote the photon detection rate of the two point sources, M denotes the total lateral magnification of the microscope setup, d denotes the distance of separation between the point sources, and q_1 and q_2 denote the image functions of the two point sources. An image function q describes the image of an object at unit lateral magnification when the center of the object is located at the origin of the coordinate axes in the specimen plane. The function q is normalized to satisfy the integral identity $\int_{\mathbb{R}^2} q(x,y)dxdy = 1$ (see ref. 3 for details). In Eq. **3** we consider an arrangement (potentially after a suitable translation of the coordinate axes) in which the point sources lie along the x axis in the specimen plane and are equidistant from the origin of the coordinate axes. Substituting for $f_{\theta,\tau}$ and Λ_θ in Eq. **1**, the Fisher information matrix is given by

$$\begin{aligned}
I(\theta) = &\int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\frac{\Lambda_1(\tau)}{M^2} q_1\left(\frac{x}{M}+\frac{d}{2},\frac{y}{M}\right) + \frac{\Lambda_2(\tau)}{M^2} q_2\left(\frac{x}{M}-\frac{d}{2},\frac{y}{M}\right)} \times \\
&\qquad \left(\frac{\Lambda_1(\tau)}{M^2}\frac{\partial q_1\left(\frac{x}{M}+\frac{d}{2},\frac{y}{M}\right)}{\partial d} + \frac{\Lambda_2(\tau)}{M^2}\frac{\partial q_2\left(\frac{x}{M}-\frac{d}{2},\frac{y}{M}\right)}{\partial d}\right)^2 dxdyd\tau \\
= &\int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau) q_1\left(\frac{x}{M}+\frac{d}{2},\frac{y}{M}\right) + \Lambda_2(\tau) q_2\left(\frac{x}{M}-\frac{d}{2},\frac{y}{M}\right)} \times \\
&\qquad \left(\frac{\Lambda_1(\tau)}{2}\frac{\partial q_1\left(\frac{x}{M}+\frac{d}{2},\frac{y}{M}\right)}{\partial x} - \frac{\Lambda_2(\tau)}{2}\frac{\partial q_2\left(\frac{x}{M}-\frac{d}{2},\frac{y}{M}\right)}{\partial x}\right)^2 dxdyd\tau \\
= &\frac{1}{4}\int_{t_0}^{t} \int_{\mathbb{R}^2} \frac{1}{\Lambda_1(\tau) q_1(x+\frac{d}{2},y) + \Lambda_2(\tau) q_2(x-\frac{d}{2},y)} \times \\
&\qquad\qquad \left(\Lambda_1(\tau)\frac{\partial q_1(x+\frac{d}{2},y)}{\partial x} - \Lambda_2(\tau)\frac{\partial q_2(x-\frac{d}{2},y)}{\partial x}\right)^2 dxdyd\tau.
\end{aligned} \quad [\mathbf{4}]$$

Inverting Eq. **4** and taking the square root, we obtain the expression for the g-FREM.

Derivation of the FREM. Rayleigh's resolution criterion considers two identical, self-luminous, in-focus point sources that are imaged with a conventional wide-field optical microscope. Here we derive the expression of the fundamental resolution measure with similar assumptions. We assume the two point sources to have equal, constant intensities i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$, and identical image functions i.e., $q_1 = q_2$. According to optical diffraction theory (4), the image of an in-focus point source is described by the Airy profile, which is given by

$$q_i(x,y) := \frac{J_1^2(\alpha\sqrt{x^2+y^2})}{\pi(x^2+y^2)}, \quad (x,y) \in \mathbb{R}^2, \quad i = 1,2, \quad [\mathbf{5}]$$

where J_1 denotes the first order Bessel function of the first kind, $\alpha := 2\pi n_a/\lambda$, n_a denotes the numerical aperture of the objective lens, and λ denotes the wavelength of the detected photons. Using the well known recurrence relations for Bessel functions (see e.g., ref. 5, pp. 17 and 18), the partial derivative of q_i with respect to x is given by $\partial q_i(x,y)/\partial x = -2\alpha x J_1(\alpha\sqrt{x^2+y^2}) J_2(\alpha\sqrt{x^2+y^2})/(\pi(x^2+y^2)^{\frac{3}{2}})$, $(x,y) \in \mathbb{R}^2$, $i = 1,2$, where $\alpha = 2\pi n_a/\lambda$, and J_2 denotes the second-order Bessel function of the first kind. Substituting for $q_i(x,y)$ and $\partial q_i(x,y)/\partial x$ in Eq. **4** and setting $\Lambda_i(\tau) = \Lambda_0$, for $i = 1,2$, we get

$$\mathbf{I}(d) = \frac{\Lambda_0 \cdot (t-t_0)}{4} \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha\sqrt{(x+\frac{d}{2})^2+y^2})}{\pi((x+\frac{d}{2})^2+y^2)} + \frac{J_1^2(\alpha\sqrt{(x-\frac{d}{2})^2+y^2})}{\pi((x-\frac{d}{2})^2+y^2)}} \times$$

$$\Bigg[-2\alpha(x+\frac{d}{2}) \frac{J_1(\alpha\sqrt{(x+\frac{d}{2})^2+y^2}) J_2(\alpha\sqrt{(x+\frac{d}{2})^2+y^2})}{\pi\Big((x+\frac{d}{2})^2+y^2\Big)^{\frac{3}{2}}}$$

$$-(-2\alpha)(x-\frac{d}{2}) \frac{J_1(\alpha\sqrt{(x-\frac{d}{2})^2+y^2}) J_2(\alpha\sqrt{(x-\frac{d}{2})^2+y^2})}{\pi\Big((x-\frac{d}{2})^2+y^2\Big)^{\frac{3}{2}}} \Bigg]^2 dxdy$$

$$= \frac{\alpha^2 \Lambda_0 (t-t_0)}{\pi} \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha r_{01})}{r_{01}^2} + \frac{J_1^2(\alpha r_{02})}{r_{02}^2}} \left[(x+\frac{d}{2}) \frac{J_1(\alpha r_{01}) J_2(\alpha r_{01})}{r_{01}^3} - (x-\frac{d}{2}) \frac{J_1(\alpha r_{02}) J_2(\alpha r_{02})}{r_{02}^3} \right]^2 dxdy$$

$$= \frac{4n_a^2}{\lambda^2} \pi \cdot \Lambda_0 \cdot (t-t_0) \cdot \Gamma_0(d),$$

where $r_{01} := \sqrt{(x+d/2)^2+y^2}$, $(x,y) \in \mathbb{R}^2$, $r_{02} := \sqrt{(x-d/2)^2+y^2}$, $(x,y) \in \mathbb{R}^2$ and Γ_0 be given by

$$\Gamma_0(d) = \int_{\mathbb{R}^2} \frac{1}{\frac{J_1^2(\alpha r_{01})}{r_{01}^2} + \frac{J_1^2(\alpha r_{02})}{r_{02}^2}} \left((x+\frac{d}{2}) \frac{J_1(\alpha r_{01}) J_2(\alpha r_{01})}{r_{01}^3} - (x-\frac{d}{2}) \frac{J_1(\alpha r_{02}) J_2(\alpha r_{02})}{r_{02}^3} \right)^2 dxdy.$$

Note that the Fisher information matrix $\mathbf{I}(d)$ is a scalar quantity. The FREM is obtained by taking the square root of the inverse Fisher information matrix (i.e, $1/\mathbf{I}(d)$), and is given by

$$\delta_d := \frac{1}{\sqrt{\mathbf{I}(d)}} = \frac{1}{\sqrt{4\pi \cdot \Lambda_0 \cdot (t-t_0) \cdot \Gamma_0(d)}} \frac{\lambda}{n_a}.$$

Extension to Non-Poissonian Statistics. The derivation of the Fisher information matrix given in Eq. **4** assumes the time points of the detected photons in the acquired data to be Poisson distributed. We next consider the scenario in which the times points of the detected photons are described by a general counting process $\{N(\tau), \tau \geq t_0\}$ that has finite first and second moment, i.e.,

$0 \leq E[N(\tau)], E[N^2(\tau)] < \infty$. Analogous to eq. 4, the spatial and temporal components of the acquired data are assumed to be independent of each other. The general expression of the FREM for the case of non-Poissonian photon statistics is then given by

$$\left[\frac{E[N(t)]}{4}\int_{\mathbb{R}^2}\frac{1}{q_1(x+\frac{d}{2},y)+q_2(x-\frac{d}{2},y)}\left(\frac{\partial q_1(x+\frac{d}{2},y)}{\partial x}-\frac{\partial q_2(x-\frac{d}{2},y)}{\partial x}\right)^2 dxdy\right]^{-\frac{1}{2}}, \qquad [\mathbf{6}]$$

where q_1 and q_2 denote the image functions of the two point sources and d denotes the distance of separation.

Effects of Pixelation and Noise

In the Section *Image Detection Process*, it was assumed that the detector records the time points and the spatial coordinates of the detected photons, which was described by an image detection process. However, current imaging detectors have pixels, and the acquired data only consists of the number of detected photons at each pixel. For a pixelated detector $\{C_1, \ldots, C_{N_p}\}$ with N_p pixels, the photon count at the kth pixel is independently Poisson distributed. We consider two types of additive noise sources, namely Poisson and Gaussian noise source. Poisson noise is used to model the effect of spurious light sources such as autofluorescence, and Gaussian noise is used to model measurement noise such as readout noise in the detector (also see ref. 6).

Hence the data acquired by a pixelated detector during the time interval $[t_0, t]$ is described by a sequence of independent random variables $\{\mathcal{I}_{\theta,1}, \ldots, \mathcal{I}_{\theta,N_p}\}$, where $\mathcal{I}_{\theta,k} := S_{\theta,k} + B_k + W_k$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, and $S_{\theta,k}$, B_k and W_k are random variables such that $\{S_{\theta,1}, \ldots, S_{\theta,N_p}\}$, $\{B_1, \ldots, B_{N_p}\}$ and $\{W_1, \ldots, W_{N_p}\}$ are mutually independent and independent of each other. The random variable $S_{\theta,k}$ is Poisson distributed with mean $\mu_\theta(k,t)$ and describes the total number of detected photons at the kth pixel from the two point sources. The random variable B_k is Poisson distributed with mean $\beta(k,t)$ and describes the total number of detected photons at the kth pixel from spurious sources. The random variable W_k is Gaussian distributed with mean η_k and standard deviation $\sigma_{w,k}$ and describes the measurement noise at the kth pixel. We assume that $\beta(k,t)$, η_k and $\sigma_{w,k}$ are independent of θ, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$.

Fisher Information Matrix for a Pixelated Detector. In the absence of Gaussian noise (i.e., $W_k = 0$, $k = 1, \ldots, N_p$), the Fisher information matrix for a pixelated detector corresponding to the time interval $[t_0, t]$ is given by (see ref. 3 for details)

$$\mathbf{I}(\theta) := \sum_{k=1}^{N_p} \frac{1}{\mu_\theta(k,t) + \beta(k,t)} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta}, \quad \theta \in \Theta,$$

where $\mu_\theta(k,t)$ ($\beta(k,t)$) denotes mean number of detected photons at the kth pixel from the two point (spurious) sources. Setting $\beta(k,t) = 0$ in the above equation, we obtain an expression for the Fisher information matrix of a pixelated detector in the absence of additive noise sources. In the presence of Gaussian noise, the Fisher information matrix is given by (see ref. 3 for details)

$$\mathbf{I}(\theta) := \sum_{k=1}^{N_p} \left(\frac{\partial \mu_\theta(k,t)}{\partial \theta}\right)^T \frac{\partial \mu_\theta(k,t)}{\partial \theta} \left(\int_{\mathbb{R}} \frac{\left(\sum_{l=1}^{\infty} \frac{[\nu_\theta(k,t)]^{l-1} e^{-\nu_\theta(k,t)}}{(l-1)!} \cdot \frac{1}{\sqrt{2\pi}\sigma_{w,k}} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}\right)^2}{p_{\theta,k}(z)} dz - 1 \right),$$

where $\theta \in \Theta$, $\nu_\theta(k,t) := \mu_\theta(k,t) + \beta(k,t)$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, μ_θ and β are as given above, and

$$p_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\sigma_{w,k}} \sum_{l=0}^{\infty} \frac{[\nu_\theta(k,t)]^l e^{-\nu_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\eta_k}{\sigma_{w,k}}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \qquad [\mathbf{7}]$$

Analogous to Eq. **1**, the above equations of the Fisher information matrix for a pixelated detector are applicable to a wide variety of imaging conditions. To calculate the Fisher information matrix in the present context, we require the analytical expression for $\mu_\theta(k,t)$ (and $\partial \mu_\theta(k,t)/\partial \theta$), which is given in Eq. **8**. In addition, the numerical values of the noise parameters $\beta(k,t)$, η_k and $\sigma_{w,k}$ need to be known, which depend on the experimental setup.

The Generalized PREM. For the derivation of the g-PREM, we consider a geometry shown in Fig. 5, where the two point sources are located at arbitrary locations P_1 and P_2 on the specimen plane. Here (x_0, y_0) denotes the coordinates of the point P_1, d denotes the distance of separation between the point sources and ϕ denotes the angle of inclination of the line segment joining the two point sources with respect to the x axis. The coordinates of the point P_2 are given by $(x_0 + d\cos\phi, y_0 + d\sin\phi)$. In a practical situation, in addition to d, the other parameters, namely x_0, y_0 and ϕ are also unknown and therefore must be estimated along with d. Hence the unknown parameter vector is given by $\theta := (x_0, y_0, d, \phi)$. The general expression for μ_θ is given by

$$\mu_\theta(k,t) := \mu_\theta^1(k,t) + \mu_\theta^2(k,t), \quad \theta \in \Theta, \quad k = 1, \ldots, N_p, \qquad [\mathbf{8}]$$

where, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$,

$$\mu_\theta^1(k,t) := \frac{1}{M^2} \int_{t_0}^{t} \Lambda_1(\tau) d\tau \int_{C_k} q_1 \left(\frac{x}{M} - x_0, \frac{y}{M} - y_0 \right) dxdy, \qquad [\mathbf{9}]$$

$$\mu_\theta^2(k,t) := \frac{1}{M^2} \int_{t_0}^{t} \Lambda_2(\tau) d\tau \int_{C_k} q_2 \left(\frac{x}{M} - x_0 - d\cos\phi, \frac{y}{M} - y_0 - d\sin\phi \right) dxdy. \qquad [\mathbf{10}]$$

In Eqs. **8 - 10**, C_k denotes the kth pixel, $[t_0, t]$ denotes the acquisition time interval, M denotes the magnification of the microscope setup, $\Lambda_1(\tau)$ and $\Lambda_2(\tau)$, $\tau \geq t_0$, denote the photon detection rates of the point sources, and q_1 and q_2 denote the image functions of the point sources.

Because the parameter θ is a 1×4 vector, by definition the Fisher information matrix $\mathbf{I}(\theta)$ is a 4×4 matrix. The g-PREM is given by the square root of the leading diagonal entry in $\mathbf{I}^{-1}(\theta)$ that corresponds to the distance parameter d. In the present case, this is the third leading diagonal entry, i.e., $\sqrt{[\mathbf{I}^{-1}(\theta)]_{33}}$, as d is the third component of θ.

Derivation of the PREM. The PREM is a special case of the g-PREM in which the photon detection rates of the point sources are assumed to be a constant, i.e., $\Lambda_1(\tau) = \Lambda_2(\tau) = \Lambda_0$, $\tau \geq t_0$, and image functions of the point sources are assumed to be the Airy profile. In the Section *Analytical expression of μ_θ* (*Analytical expression of $\partial\mu_\theta/\partial\theta$*), we give the analytical expression for $\mu_\theta(k,t)$ ($\partial\mu_\theta(k,t)/\partial\theta$) in terms of the Airy profile, which is required for the calculation of the PREM.

Analytical expression of μ_θ. For the Airy profile, we have

$$\mu_\theta(k,t) := \mu_\theta^1(k,t) + \mu_\theta^2(k,t), \quad k = 1, \ldots, N_p, \quad \theta \in \Theta, \qquad [\mathbf{11}]$$

where, for $\theta \in \Theta$ and $k = 1, \ldots, N_p$,

$$\mu_\theta^1(k,t) := \Lambda_0(t - t_0) \int_{C_k} \frac{J_1^2(a\sqrt{(x - Mx_0)^2 + (y - My_0)^2})}{\pi((x - Mx_0)^2 + (y - My_0)^2)} dxdy, \qquad [\mathbf{12}]$$

$$\mu_\theta^2(k,t) := \Lambda_0(t-t_0)\int_{C_k}\frac{J_1^2(a\sqrt{(x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2})}{\pi((x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2)}dxdy, \qquad \textbf{[13]}$$

with $a = 2\pi n_a/(\lambda M)$, n_a denoting the numerical aperture of the objective lens and λ denoting the wavelength of the photons.

Analytical expression of $\partial\mu_\theta/\partial\theta$. For $\theta = (x_0, y_0, d, \phi) \in \Theta$, let μ_θ be given by Eq. **11**, $r_{01} := M(x_0, y_0)$, $r_{02} := M(x_0 + d\cos\phi, y_0 + d\sin\phi)$ and $a = 2\pi n_a/(\lambda M)$. For $M > 0$, define $||r - r_{01}|| := \sqrt{(x-Mx_0)^2+(y-My_0)^2}$
and $||r - r_{02}|| = \sqrt{(x-Mx_0-Md\cos\phi)^2+(y-My_0-Md\sin\phi)^2}$, where $r = (x, y) \in \mathbb{R}^2$. Then

$$\frac{\partial\mu_\theta(k,t)}{\partial\theta} := \left[\frac{\partial\mu_\theta(k,t)}{\partial x_0} \quad \frac{\partial\mu_\theta(k,t)}{\partial y_0} \quad \frac{\partial\mu_\theta(k,t)}{\partial d} \quad \frac{\partial\mu_\theta(k,t)}{\partial \phi}\right], \quad k = 1,\ldots,N_p, \quad \theta\in\Theta,$$

where the entries of the row vector in the above equation are given below:

$$\begin{aligned}
\frac{\partial\mu_\theta(k,t)}{\partial x_0} &= 2aM\Lambda_0(t-t_0)\left(\int_{C_k}(x-Mx_0)\frac{J_1(a||r-r_{01}||)J_2(a||r-r_{01}||)}{\pi||r-r_{01}||^3}dr\right.\\
&\qquad\left.+\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial y_0} &= 2aM\Lambda_0(t-t_0)\left(\int_{C_k}(y-My_0)\frac{J_1(a||r-r_{01}||)J_2(a||r-r_{01}||)}{\pi||r-r_{01}||^3}dr\right.\\
&\qquad\left.+\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial d} &= 2aM\Lambda_0(t-t_0)\left(\cos\phi\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right.\\
&\qquad\left.+\sin\phi\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),\\
\frac{\partial\mu_\theta(k,t)}{\partial \phi} &= 2aM\Lambda_0(t-t_0)\left(-d\sin\phi\int_{C_k}(x-Mx_0-Md\cos\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right.\\
&\qquad\left.+d\cos\phi\int_{C_k}(y-My_0-Md\sin\phi)\frac{J_1(a||r-r_{02}||)J_2(a||r-r_{02}||)}{\pi||r-r_{02}||^3}dr\right),
\end{aligned}$$

with $r = (x, y) \in \mathbb{R}^2$, $dr := dxdy$, $\theta \in \Theta$ and $k = 1,\ldots,N_p$.

Calculation of the Resolution Measure When Additional Spatial Information Is Present. This section discusses the calculation of the Fisher information matrix for the global estimation problem of determining the unknown parameter θ from data acquired before and after the first photobleaching event for a single-molecule pair that exhibits a double-step photobleaching behavior. Without loss of generality, the location coordinates of the single molecule that photobleaches first are set to be $(x_0 + d\cos\phi, y_0 + d\sin\phi)$ and the location coordinates of the single molecule that photobleaches last are set to be (x_0, y_0). The data acquired before and after the first photobleaching event are mutually independent and therefore the Fisher information matrix for the global estimation problem can be written as

$$\mathbf{I}_{tot}(\theta) := \mathbf{I}(\theta) + \mathbf{I}_a(\theta), \quad \theta \in \Theta.$$

In the above equation, $\mathbf{I}(\theta)$ and $\mathbf{I}_a(\theta)$ denote the Fisher information matrices that are calculated for the problem of estimating the unknown parameter θ from data acquired before and after the first photobleaching event, respectively. The expressions of $\mathbf{I}(\theta)$ for a pixelated detector in the presence and absence of noise sources are given in Section *Fisher information matrix for a pixelated detector*. The matrix $\mathbf{I}_a(\theta)$ is of the form $\mathbf{I}_a(\theta) := \begin{bmatrix} \mathbf{I}_l(\theta) & 0 \\ 0 & 0 \end{bmatrix}$, $\theta \in \Theta$, where $\mathbf{I}_l(\theta)$ denotes the Fisher information matrix for the problem of determining the location (x_0, y_0) of a single molecule. In the absence of Gaussian noise $\mathbf{I}_l(\theta)$ is given by

$$\mathbf{I}_l(\theta) = \sum_{k=1}^{N_p} \frac{1}{\mu_\theta^1(k,t) + \beta^1(k,t)} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \\ \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix} \left(\frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \quad \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \right), \quad \theta \in \Theta,$$

where $\mu_\theta^1(k,t)$ is given in Eq. **12**, $\beta^1(k,t)$ denotes the mean of the additive Poisson noise in the kth pixel, and the partial derivatives $\partial\mu_\theta^1(k,t)/\partial x_0$ and $\partial\mu_\theta^1(k,t)/\partial y_0$ are given by

$$\frac{\partial \mu_\theta^1(k,t)}{\partial \zeta_0} = 2aM\Lambda_0(t-t_0) \int_{C_k} (\zeta - M\zeta_0) \frac{J_1(a||r-r_{01}||)J_2(a||r-r_{01}||)}{\pi||r-r_{01}||^3} dr, \qquad [\mathbf{14}]$$

for $\zeta \in \{x, y\}$, $\theta \in \Theta$, $k = 1, \ldots, N_p$ and $||r - r_{01}|| := \sqrt{(x-Mx_0)^2 + (y-My_0)^2}$. In the presence of

Gaussian noise, $\mathbf{I}_l(\theta)$ is given by

$$\mathbf{I}_l(\theta) = \sum_{k=1}^{N_p} \begin{pmatrix} \frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \\ \\ \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \end{pmatrix} \left(\frac{\partial \mu_\theta^1(k,t)}{\partial x_0} \quad \frac{\partial \mu_\theta^1(k,t)}{\partial y_0} \right) \times$$

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

where $\nu_\theta^1(k,t) := \mu_\theta^1(k,t) + \beta^1(k,t)$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, μ_θ^1 is given in Eq. **12**, β^1 is given above, $\partial \mu_\theta^1(k,t)/\partial x_0$ and $\partial \mu_\theta^1(k,t)/\partial y_0$ are given in Eq. **14**, and the expression for $p_{\theta,k}^1$ is analogous to that given in Eq. **7**, but with the term $\nu_\theta(k,t)$ in Eq. **7** replaced by $\nu_\theta^1(k,t)$.

Maximum-Likelihood Estimation

The estimation of the unknown parameters is carried out on the data that is contained in a pixel array. We first consider the scenario when the pixel array is extracted from an individual image, which contains photons from both point sources. The log-likelihood function for the data in the pixel array is given by

$$\ln(\mathcal{L}(\theta \mid z_1, \ldots, z_{N_p})) := \ln \left(\prod_{k=1}^{N_p} p_{\theta,k}(z_k) \right) = \sum_{k=1}^{N_p} \ln(p_{\theta,k}(z_k)), \quad \theta \in \Theta, \qquad [\mathbf{15}]$$

where N_p denotes the total number of pixels in the pixel array, z_k denotes the detected photon count at the kth pixel in the pixel array, and $p_{\theta,k}$ denotes the probability density function of z_k that is given by Eq. **7**, $k = 1, \ldots, N_p$. For the distance estimation problem with the data acquired by a pixelated detector in the presence of noise sources, the vector of unknown parameters is set to be $\theta = (x_0, y_0, d, \phi)$ (see Fig. 5) and the image function of the point source is assumed to be the Airy profile. The maximum-likelihood estimate of θ is obtained by substituting the expression for μ_θ given by Eq. **11** in $p_{\theta,k}$ (Eq. **7**) and determining the value of θ that maximizes the log-likelihood function $\ln(\mathcal{L}(\theta))$.

We next consider the scenario when the pixel array is obtained by adding N_1 pixel arrays, which are extracted from N_1 individual images that contain photons from both point sources. The log-likelihood function for the data in the summed pixel array is given by

$$\ln(\tilde{\mathcal{L}}(\theta \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p})) := \sum_{k=1}^{N_p} \ln(\tilde{p}_{\theta,k}(\tilde{z}_k)), \quad \theta \in \Theta, \qquad [\mathbf{16}]$$

where $\tilde{z}_k$ denotes the detected photon count at the kth pixel in the summed pixel array and $\tilde{p}_{\theta,k}$ denotes the density function of $\tilde{z}_k$, $k = 1, \ldots, N_p$, which is given by

$$\tilde{p}_{\theta,k}(z) := \frac{1}{\sqrt{2\pi}\tilde{\sigma}_{w,k}} \sum_{l=0}^{\infty} \frac{[\tilde{\nu}_\theta(k,t)]^l e^{-\tilde{\nu}_\theta(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\tilde{\eta}_k}{\tilde{\sigma}_{w,k}}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \qquad [\mathbf{17}]$$

In Eq. **17**, $\tilde{\nu}_\theta(k,t) := N_1(\mu_\theta(k,t) + \beta(k,t))$, $\tilde{\eta}_k = N_1\eta_k$, and $\tilde{\sigma}_{w,k} = \sqrt{N_1}\sigma_{w,k}$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, where $\mu_\theta(k,t)$ ($\beta(k,t)$) denotes the mean photon count at the kth pixel from the point sources (scattering-noise sources) in the individual pixel array, η_k and $\sigma_{w,k}^2$ denote the mean and the variance of the readout noise at the kth pixel, respectively, in the individual pixel array. For the distance estimation problem, the maximum-likelihood estimate of $\theta = (x_0, y_0, d, \phi)$ is obtained by substituting the expression for μ_θ given by Eq. **11** in $\tilde{p}_{\theta,k}$ (Eq. **17**) and determining the value of θ that maximizes the log-likelihood function $\ln(\tilde{\mathcal{L}}(\theta))$.

If the point sources exhibit a double step photobleaching behavior, the images acquired after the first photobleaching event can also be used to estimate θ. Here, the experimental data that is used to estimate θ consists of two summed pixel arrays. One of the summed pixel arrays is obtained by adding N_1 pixel arrays that are extracted from N_1 individual images acquired before the first photobleaching event (i.e., images that contain photons from both point sources). In this summed pixel array the detected photon count at the kth pixel is denoted as $\tilde{z}_k$, $k = 1, \ldots, N_p$, where N_p denotes the total number of pixels. The other summed pixel array is obtained by adding N_2 pixel arrays that are extracted from N_2 individual images acquired after the first photobleaching event and $\tilde{z}_k^1$ denotes the detected photon count at the kth pixel in this summed pixel array. The log-likelihood function $\ln(\mathcal{L}_T(\theta))$ for the data contained in the two summed pixel arrays is given by

$$\ln(\mathcal{L}_T(\theta) \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p}; \tilde{z}_1^1, \ldots, \tilde{z}_{N_p}^1)) := \ln(\tilde{\mathcal{L}}(\theta \mid \tilde{z}_1, \ldots, \tilde{z}_{N_p})) + \ln(\tilde{\mathcal{L}}^1(\theta \mid \tilde{z}_1^1, \ldots, \tilde{z}_{N_p}^1)), \ \theta \in \Theta, [\mathbf{18}]$$

where $\ln(\tilde{\mathcal{L}})$ $(\ln(\tilde{\mathcal{L}}^1))$ denotes the log-likelihood function corresponding to $\{\tilde{z}_1, \ldots, \tilde{z}_{N_p}\}$ $(\{\tilde{z}_1^1, \ldots, \tilde{z}_{N_p}^1)\})$. The expression for $\ln(\tilde{\mathcal{L}})$ is given in Eq. **16**, and the expression for $\ln(\tilde{\mathcal{L}}^1)$ is given by $\ln(\tilde{\mathcal{L}}^1(\theta \mid \tilde{z}_1^1, \ldots, \tilde{z}_{N_p}^1)) := \sum_{k=1}^{N_p} \ln(\tilde{p}_{\theta,k}^1(\tilde{z}_k^1))$, $\theta \in \Theta$, where $\tilde{p}_{\theta,k}^1$ denotes the density function of $\tilde{z}_k^1$ that is given by

$$\tilde{p}_{\theta,k}^1(z) := \frac{1}{\sqrt{2\pi}\tilde{\sigma}_{w,k}^1} \sum_{l=0}^{\infty} \frac{[\tilde{\nu}_\theta^1(k,t)]^l e^{-\tilde{\nu}_\theta^1(k,t)}}{l!} e^{-\frac{1}{2}\left(\frac{z-l-\tilde{\eta}_k^1}{\tilde{\sigma}_{w,k}^1}\right)^2}, \quad \theta \in \Theta, \quad z \in \mathbb{R}. \qquad [\mathbf{19}]$$

In Eq. **19**, $\tilde{\nu}_\theta^1(k,t) := N_2(\mu_\theta^1(k,t) + \beta^1(k,t))$, $\tilde{\eta}_k^1 = N_2\eta_k^1$, and $\tilde{\sigma}_{w,k}^1 = \sqrt{N_2}\sigma_{w,k}^1$, $k = 1, \ldots, N_p$, $\theta \in \Theta$, where $\mu_\theta^1(k,t)$ $(\beta^1(k,t))$ denotes the mean photon count at the kth pixel from one of the point sources (scattering noise sources) in the pixel array that is extracted from the image acquired after the first photobleaching event, and η_k^1 and $(\sigma_{w,k}^1)^2$ denote the mean and the variance of the readout noise at the kth pixel, respectively in the pixel array that is extracted from the image acquired after the first photobleaching event. Thus, for the distance estimation problem, the maximum-likelihood estimate of $\theta = (x_0, y_0, d, \phi)$ from the two summed pixel arrays is obtained by substituting the expression for μ_θ given by Eq. **11** in $\tilde{p}_{\theta,k}$ (Eq. **17**) and substituting the expression for μ_θ^1 given by Eq. **12** in $\tilde{p}_{\theta,k}^1$ (Eq. **19**), and then determining the value of θ that maximizes the log-likelihood function $\ln(\mathcal{L}_T(\theta))$.

We note that Eqs. **15**, **16** and **18** can be used to obtain the maximum-likelihood estimate of θ in a wide variety of imaging conditions. For instance, consider the scenario when the image function of the point source is described by a profile that is different from the Airy profile. In this case, we use Eq. **8** (Eq. **9**) to obtain an expression for μ_θ (μ_θ^1) in terms of the desired image profile, and then maximize the corresponding log-likelihood function to obtain the maximum-likelihood estimate. In all of the above cases, the maximum-likelihood estimation is carried with the optimization toolbox of MATLAB in the MIATool software environment (software available on request).

References

1. Snyder, D. L. & Miller, M. I. (1991) *Random Point Processes in Time and Space* (Springer, Berlin).

2. Zacks, S. (1971) *The Theory of Statistical Inference* (Wiley, New York).

3. Ram, S., Ward, E. S. & Ober, R. J. (2006) *Multidim. Sys. Sig. Proc.* **17**, 27–58.

4. Born, M. & Wolf, E. (1999) *Principles of Optics* (Cambridge Univ. Press, Cambridge, U.K.).

5. Watson, G. N. (1958) *A Treatise on the Theory of Bessel Functions* (Cambridge Univ. Press, Cambridge, U.K.).

6. Ober, R. J., Ram, S. & Ward, E. S. (2004) *Biophys. J.* **86**, 1185–1200.