Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Localization precision of stochastic optical localization nanoscopy using single frames

Yi Sun

SPIE

Localization precision of stochastic optical localization nanoscopy using single frames

Yi Sun
City College of City University of New York, Electrical Engineering Department, Convent Avenue at 138th Street, New York 10031

Abstract. The Cramer-Rao lower bound (CRLB) on localization precision of unbiased estimators is analyzed for stochastic optical localization nanoscopy that localizes emitters frame by frame independently. It is found that the CRLB is a function of the mean number of detected photons per emitter, signal to Poisson noise ratio, signal to Gaussian noise ratio, point spread function (PSF), pixel size, and relationship of emitter locations. With a slight and practical approximation, effect of Gaussian noise is equivalent to increasing the mean photon count of Poisson noise by a number equal to the variance of Gaussian noise. Numerical examples demonstrate that the CRLB of emitters located on a curve increase fast as the distance of adjacent emitters increases. The mean CRLB of randomly uniformly distributed emitters in both two-dimensional and three-dimensional imaging increases exponentially fast as the emitter density increases. The effects of PSF, standard deviation of PSF, mean number of detected photons per emitter, signal to noise ratio, axial thickness, and pixel size on the CRLB are also numerically investigated. The analytical and numerical results provide a guideline for the design of location estimators and a benchmark for the achievable localization precision of stochastic optical localization nanoscopy. [DOI: 10.1117/1.JBO.18.11.111418]

Keywords: nanoscopy; superresolution; Cramer-Rao lower bound; full-width at half-maximum; single molecule imaging; microscopy; likelihood function.

Paper 130287SSRR received Apr. 27, 2013; revised manuscript received Sep. 7, 2013; accepted for publication Sep. 20, 2013; published online Oct. 25, 2013.

1 Introduction

Fluorescence microscopy has played a crucial role in the study of cellular structure. Due to the physics of light diffraction, a conventional fluorescence microscope can only achieve the highest spatial resolution of 250 to 500 nm, fundamentally limited by the wavelength of fluorescence emission. By using photoswitchable and photoactivatable probes, recent nanoscopy techniques[1–11] have broken the long-standing diffraction barrier to provide a spatial superresolution of ~20 nm and enabled the viewing of nanoscale cellular structures and dynamics that were unable to be seen before. The nanoscopy techniques have two categories. The first is called ensemble imaging approaches,[1] which employ patterned illumination to modulate spatially the fluorescence of probe molecules within the diffraction-limited region to break the diffraction barrier. This includes stimulated emission depletion (STED) microscopy,[2,3] reversible saturable optically linear fluorescence transitions (RESOLFT),[4] and saturated structured illumination microscopy (SSIM).[5] The second is usually called superresolution optical localization microscopy[12] or stochastic optical localization nanoscopy. It takes advantage of photoswitchable and photoactivatable probes that are stochastically activated so that single activated probes within their point spread functions (PSFs) are accurately localized by the single-emitter fitting. This includes stochastic optical reconstruction microscopy (STORM),[6] photoactivated localization microscopy (PALM),[7] fluorescence photoactivatable localization microscopy (FPALM),[8] superresolution optical fluctuation imaging (SOFI),[9] multicolor,[10] and three-dimensional (3-D)[11] STORM, double-helix PSF,[13] 3-D interferometric PALM,[14] and 3-D biplane FPALM.[15]

Localization nanoscopy has a number of advantages[1] that make it particularly useful and potentially powerful. In principle, localization nanoscopy employs photoswitchable fluorophores as probes to trade imaging time for a super localization resolution. In each image frame, only a sparse subset of densely distributed probes is randomly activated so that single emitters isolated within their PSFs are spatially resolvable.[6] Since the image data in the region where PSFs of multiemitters are overlapped have to be discarded, the probability of probe activation must be low to reduce the probability of overlapping. Therefore, a large number of image frames must be acquired to obtain one snapshot of biological structure, thus significantly increasing total imaging time or temporal resolution. One way to reduce the imaging time is to use the probes that can achieve high photoswitching rates under high laser intensities.[16] Another way is to exploit the location information of multiemitters embedded in the image data where their PSFs are overlapped. A number of estimators that can simultaneously localize multiemitters with overlapped PSFs have been developed in the literature. These estimators include the DAOPHOT/DAOSTORM,[17] high-density 3-D,[18] compressed sensing,[19] deconSTORM,[20] structured sparse model and Bayesian information criterion,[21] and parallel localization of multiple emitters via Bayesian information criterion recommendation.[22] It is expected that a number of multiemitter location estimators will be developed in the literature.

The spatial resolution is one of the most important figures of merit for a microscope. For the first category of nanoscopy

Address all correspondence to: Yi Sun, City College of City University of New York, Electrical Engineering Department, Convent Avenue at 138th Street, New York 10031. E-mail: ysun@ccny.cuny.edu

0091-3286/2013/$25.00 © 2013 SPIE

techniques, the spatial resolution is simply determined by the size of laser spot where fluorophores are activated. For the second category of nanoscopy techniques, in contrast, the localization precision of an emitter depends on a number of factors and is much more complicated to determine. In localization nanoscopy, as shown in Fig. 1, fluorescence of an emitter passes through an optical path with a PSF and forms a two-dimensional (2-D) image taken by a CCD or CMOS image sensor. Then an estimator of emitter locations—an algorithm—reconstructs emitter locations from the sensed image. Therefore, localization precision of an emitter depends not only on diffraction but also on the relationship of emitter locations, emitter power, noise power, PSF, and pixel size of image sensor as well as the location estimator. Calculated from the Fisher information of sensed image, the Cramer-Rao lower bound (CRLB)[23,24] provides a fundamental theoretical limit on localization precisions obtained by all unbiased estimators. While localization precision can be different and depends on a particular estimator, the CRLB is an intrinsic property of a localization nanoscopy system that is independent of a particular estimator. The CRLB of a single 2-D emitter with the Airy and Gaussian PSFs is analyzed in Ref. 25. The analysis is extended to a single 3-D emitter,[26] a single 3-D emitter with time-variant intensity,[27] a single 3-D emitter in multifocal plane microscopy,[28,29] two 2-D emitters,[30] two 3-D emitters,[31] and the distance between two emitters.[32] The CRLB with the PSF produced by a fluorescent dipole is analyzed in Ref. 33. By analyzing the likelihood function and CRLB, a fast single-molecule localization approach that achieves the theoretical minimum uncertainty is developed,[34] and the approach is extended to single-frame multiemitter localization.[35] By analyzing the CRLB, an information transfer function is derived based on three assumptions that enable the feasibility of analysis.[36]

In this paper, we analyze the CRLB[23,24] for the stochastic optical localization nanoscopy that localizes an arbitrary number of emitters with arbitrary locations independently from each single image frame. A rigorous analysis of the CRLB of multiemitters is important but is mostly missing in the literature. The CRLB is the intrinsic power of a localization nanoscopy system in localizing emitters. It provides a benchmark of the achievable localization precision for all estimators existing in the literature and to be developed. The CRLB of multiemitters analyzed in this paper is an extension of the CRLBs of a single emitter[25,26] and two emitters[30,31] and includes them as special instances. While an analysis of CRLB with multiemitters is considered in the development of the maximum likelihood estimator of multiemitter locations,[35] our analysis in this paper is far more in-depth. The aim of our analysis is to establish a theoretical foundation on which the performances of different systems, experiments, and estimators are comparable. To this end, the PSF, Poisson noise, Gaussian noise, signal to Poisson noise ratio (SPNR), and signal to Gaussian noise ratio (SGNR) are defined to be universal to all systems and experiments. It is found that the CRLB is a function of the mean number of detected photons per emitter, SPNR, SGNR, PSF, pixel size of image sensor, and relationship of emitter locations. With a slight and practical approximation, it is found that the effect of Gaussian noise is equivalent to increasing the mean photon count of Poisson noise by a number equal to the variance of Gaussian noise, and the mean of Gaussian noise does not come into play. An analysis of the Fisher information matrix points out that a cluster of emitters isolated from other clusters of emitters can be localized by using the data of this cluster alone without effect on the CRLB. By applying the analytical formula of CRLB, we investigate how the CRLB is affected by the placement of multiemitters in three cases: (1) a number of emitters are located on a curve with an increasing distance of adjacent emitters, (2) the number of emitters in a given square increases, (3) a number of emitters are randomly uniformly distributed in a 2-D square and in a 3-D cuboid, respectively. The first case considers 100 emitters on a curve and is different from the two-emitter case.[30–32] It is found that the CRLB of an emitter increases fast as the adjacent emitter distance decreases except for the emitters close to the denser end of the curve. In the second case, it is demonstrated that the mean CRLB increases exponentially fast as the number of emitters in a given square increases. In the third case, it is found that the mean CRLB increases exponentially fast as the density of emitters located in either a 2-D square or a 3-D cuboid increases. These studies indicate that a cluster of emitters mutually affects their CRLBs through their overlapped PSFs, and the CRLBs of a cluster of emitters increase as the distances between emitters decrease or the emitter density increases. A single emitter whose PSF is isolated from others reaches the lower bound of CRLB. We also numerically investigate the effects of PSF, standard deviation (SD) of PSF, mean number of detected photons per emitter, signal to noise ratio (SNR), axial thickness, and pixel size on the CRLB. The results are generally applicable to 2-D and 3-D stochastic optical localization nanoscopy with an arbitrary spatially variant or invariant PSF.

2 Method

2.1 *Cramer-Rao Lower Bound with Poisson Noise*

2.1.1 *Imaging model*

A flow chart of localization nanoscopy is illustrated in Fig. 1. Consider that in an image frame, M photoactivatable probes in a specimen at the cuboid of $[0, L_x] \times [0, L_y] \times [-L_z, L_z]$ are activated and, henceforth, called emitters throughout. The emitters are located at (x_m, y_m, z_m) for 3-D imaging or (x_m, y_m) for 2-D imaging, $m = 1, \ldots, M$. In the following, 3-D imaging will be considered, and all results are applicable to 2-D imaging. Each of the emitters emits I photons per second on average, and the point process of photon emission is a Poisson process. Due to diffraction and other optical effects, a photon emitted from the m'th emitter at (x_m, y_m, z_m) passes through the optical path and arrives at location (x, y) in the 2-D plane of image sensor with a probability density function $q_m(x, y)$, called PSF. It is worth pointing out that as a probability density function, the PSF $q_m(x, y)$ has a unit integral in the 2-D space $\mathbb{R}^2$ regardless of emitter location. A photon is detected by the image sensor with probability η. In other words, a photon is absorbed or vanished with probability $1 - \eta$. The mean number of detected photons of an emitter per second in the image plane is $I\eta$. It is notable that in the literature of optical nanoscopy, a PSF is defined in different ways[6,7,9,11,13,15,17–22,33,36] such that a PSF depends on the number of detected photons of an emitter, depends on the pixel size of image sensor, and/or has a nonunit integral. In contrast, the PSF defined here is independent of the pixel size and the number of detected photons of an emitter and therefore is universal to all systems and experiments.

Let Δ_t be the time of one frame in second during which an image sensor detects photons. The frame rate equals $1/\Delta_t$ Hz. The mean density function of detected photons in the image plane produced by the M emitters during Δ_t is then equal to

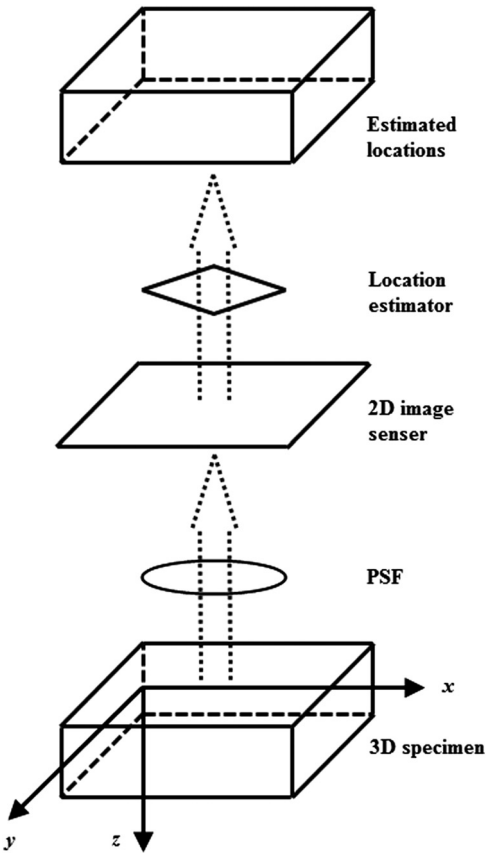

Fig. 1 A flow chart of localization nanoscopy. The 2-D image sensor can be CCD or CMOS. The estimator of emitter locations plays an important role in determination of localization precision.

$$s(x,y) = I\eta\Delta_t Q(x,y), \tag{1}$$

where

$$Q(x,y) = \sum_{m=1}^{M} q_m(x,y). \tag{2}$$

A 2-D image is taken by the image sensor with x and y pixel sizes of Δ_x and Δ_y, respectively. Assume that L_x and L_y are integer multiples $K_x = L_x/\Delta_x$ and $K_y = L_y/\Delta_y$ of Δ_x and Δ_y, respectively, and then K_x and K_y are image sizes in pixel. The mean number of photons in pixel $(k_x, k_y) \in \Omega = \{0, 1, \ldots, K_x - 1\} \times \{0, 1, \ldots, K_y - 1\}$ is equal to the spatial integral of $s(x,y)$ in the area of $[\Delta_x k_x, \Delta_x(k_x+1)] \times [\Delta_y k_y, \Delta_y(k_y+1)]$, given by $I\eta\Delta_t\Delta_x\Delta_y Q(k_x,k_y)$, where

$$Q(k_x,k_y) = \sum_{m=1}^{M} q_m(k_x,k_y), \tag{3}$$

with

$$q_m(k_x,k_y) = \frac{1}{\Delta_x\Delta_y}\int_{\Delta_y k_y}^{\Delta_y(k_y+1)}\int_{\Delta_x k_x}^{\Delta_x(k_x+1)} q_m(x,y)\mathrm{d}x\mathrm{d}y. \tag{4}$$

Here, for notation brevity, $Q(k_x,k_y)$ and $q_m(k_x,k_y)$ denote the averages of $Q(x,y)$ and $q_m(x,y)$ over pixel (k_x,k_y), respectively.

In each pixel (k_x,k_y), the number of detected photons $S(k_x,k_y)$ produced by the emitters is Poisson distributed with a mean of $I\eta\Delta_t\Delta_x\Delta_y Q(k_x,k_y)$. In addition, the background autofluorescence produces Poisson noise[37] $B(k_x,k_y)$ with a mean of b photons per second per nm^2 and therefore yields a mean of $\Delta_t\Delta_x\Delta_y b$ photons per pixel. Then the total mean photon count in pixel (k_x,k_y) is equal to

$$v(k_x,k_y) = I\eta\Delta_t\Delta_x\Delta_y[Q(k_x,k_y) + \gamma_p^{-1}], \tag{5}$$

where we define SPNR as

$$\gamma_p = \frac{I\eta}{b}\ (\mathrm{nm}^2/\mathrm{emitter}). \tag{6}$$

SPNR characterizes the emitter emission relative to the background emission and is independent of frame time and pixel size of a particular system. The mean of a Poisson process is equal to its variance; therefore, the mean number $I\eta$ of detected photons of an emitter per second in the image plane is the total power of the emitter with exclusion of its direct-current (DC) component. Moreover, b is the mean number of photons per second per nm^2, representing the power density of Poisson noise with exclusion of the DC component. Hence, SPNR can be considered as a signal to Poisson noise power ratio. As will be seen, γ_p is an independent parameter in determination of CRLB. Under γ_p, the CRLB is scalable to pixel size, PSF, and other parameters. In contrast, SNR for a nanoscopy system is usually defined as the ratio of the number of photons per emitter at the peak value of PSF-like image to the mean number of background photons per pixel.[8,9,13,21,22] Such a definition depends on and therefore is not scalable to the PSF and pixel size. Moreover, the peak value of PSF-like image is a realization of a random variable and therefore depends on a particular observation. SPNR defined in Eq. (6) is universal to all systems so that it can be used to compare the performances of different systems and experiments. Nevertheless, since $\gamma_p = \Delta_t\Delta_x\Delta_y I\eta/(\Delta_t\Delta_x\Delta_y b)$, given a particular nanoscopy system, SPNR can be estimated using the image data where a single emitter is presented and the data where no emitter is presented. To make SPNR a reasonable number in practice, we consider Poisson noise per μm^2 and therefore SPNR $= 10^{-6}\gamma_p$ (μm^2/emitter) is used.

Since the event that a photon emitted either from an emitter or from the background autofluorescence arrives at a pixel is independent of others, it follows from the property of classifying a Poisson number of events[38] that the number of detected photons $V(k_x,k_y) = S(k_x,k_y) + B(k_x,k_y)$ in pixel $(k_x,k_y) \in \Omega$ is an independent Poisson random variable with a mean of $v(k_x,k_y)$.

2.1.2 *Cramer-Rao lower bound*

An estimator estimates 3-D locations of emitters from an image $V(k_x,k_y)$, $(k_x,k_y) \in \Omega$. Since $V(k_x,k_y)$ is a realization of a stochastic process, estimates $\hat{x}_m$, $\hat{y}_m$, $\hat{z}_m$ of emitter location coordinates x_m, y_m, z_m are realizations of random variables. With different realizations of $V(k_x,k_y)$, estimates are different in general. Their SDs represent the estimation precision and performance of an estimator. The SD also provides a measurement on the power of distinguishing different emitters and, therefore, a measurement of localization precision. Depending on its capability of exploiting information of emitter locations in $V(k_x,k_y)$, an estimator may achieve a different localization precision. Nevertheless, the highest localization precision (i.e., the minimum SD) achieved by all unbiased estimators for a

localization nanoscopy system is determined by the Fisher information of emitter locations contained in $V(k_x, k_y)$ of the imaging model and is independent of a particular estimator. Specifically, the minimum SD is determined by the CRLB calculated from the Fisher information matrix.[24] In convention, the full width at half maximum (FWHM) of a probability distribution function is used as the spatial resolution of a microscope. It is known[39] that with some conditions, the maximum likelihood estimator is an asymptotically unbiased Gaussian random vector with a covariance matrix equal to the inverse of the Fisher information matrix. In this case, the FWHM is equal to $2\sqrt{2\ln 2} \cong 2.355$ times the SD. To be comparable to the FWHM of spatial resolution in the conventional microscopy, we consider $2\sqrt{2\ln 2}$ times the CRLB (called FWHM CRLB henceforth) on the SDs of unbiased estimators as the intrinsic power of stochastic optical localization nanoscopy in localizing emitters.

For 3-D imaging, let $\theta = (\theta_{1,1}, \theta_{1,2}, \theta_{1,3}, \ldots, \theta_{M,1}, \theta_{M,2}, \theta_{M,3})^T = (x_1, y_1, z_1, \ldots, x_M, y_M, z_M)^T$ be a vector of $3M$ emitter location coordinates. For 2-D imaging, let $\theta = (\theta_{1,1}, \theta_{1,2}, \ldots, \theta_{M,1}, \theta_{M,2})^T = (x_1, y_1, \ldots, x_M, y_M)^T$ be a vector of $2M$ emitter location coordinates. The FWHM CRLBs of x_m, y_m, z_m are denoted by $\Delta(\theta_{mj})$ for $j = 1, 2, 3$, respectively.

Based on the imaging model, the CRLBs of all emitters can be derived by analyzing a likelihood function of emitter locations. Since $V(k_x, k_y)$'s are spatially (pixel-wise) independent, a likelihood function of θ that is also the joint probability distribution function of all pixels in an image can be written as

$$f = \prod_{(k_x,k_y)\in\Omega} g[k_x, k_y, V(k_x, k_y)], \tag{7}$$

where

$$g[k_x, k_y, V(k_x, k_y)] = \frac{[v(k_x, k_y)]^{V(k_x,k_y)}}{V(k_x, k_y)!}\exp[-v(k_x, k_y)] \tag{8}$$

is the probability distribution function of $V(k_x, k_y)$.

The derivative of $\ln f$ with respect to a location coordinate θ_{ij} for $i = 1, \ldots, M$, $j = 1$, 2, or 3 is equal to

$$\frac{\partial \ln f}{\partial \theta_{ij}} = I\eta\Delta_t\Delta_x\Delta_y \sum_{(k_x,k_y)\in\Omega}\left[\frac{V(k_x, k_y)}{v(k_x, k_y)} - 1\right]\frac{\partial q_i(k_x, k_y)}{\partial \theta_{ij}}. \tag{9}$$

The Fisher information matrix is a symmetric matrix $\mathbf{F}$ whose $(\theta_{ij}, \theta_{ml})$'th element is

$$F(\theta_{ij}, \theta_{ml}) = E\left(\frac{\partial \ln f}{\partial \theta_{ij}}\frac{\partial \ln f}{\partial \theta_{ml}}\right),$$

where the expectation E is taken with respect to $V(k_x, k_y)$. Then,

$$F(\theta_{ij}, \theta_{ml}) = (I\eta\Delta_t\Delta_x\Delta_y)^2 \sum_{(k_x,k_y)\in\Omega} E\left\{\left[\frac{V(k_x, k_y)}{v(k_x, k_y)} - 1\right]^2\right\} \times \frac{\partial q_i(k_x, k_y)}{\partial \theta_{ij}}\frac{\partial q_m(k_x, k_y)}{\partial \theta_{ml}} = I\eta\Delta_t\Delta_x\Delta_y \sum_{(k_x,k_y)\in\Omega} \frac{1}{Q(k_x, k_y) + \gamma_p^{-1}}\frac{\partial q_i(k_x, k_y)}{\partial \theta_{ij}}\frac{\partial q_m(k_x, k_y)}{\partial \theta_{ml}}. \tag{10, 11}$$

To obtain Eq. (10), two properties are applied: (1) $V(k_x, k_y)$'s are mutually independent and (2) the score has a zero mean,[23] i.e., $E\{\partial \ln g[k_x, k_y, V(k_x, k_y)]/\partial\theta_{ij}\} = 0$. Equation (11) follows from the property that $V(k_x, k_y)$ has the same mean and variance of $v(k_x, k_y)$. Denote by $\tilde{F}(\theta_{ij}, \theta_{ml})$ the $(\theta_{ij}, \theta_{ml})$'th element of the inverse of $\mathbf{F}$. The variance of any unbiased estimate $\hat{\theta}_{ij}$ for θ_{ij} is lower bounded by the θ_{ij}'th diagonal element of the inverse of $\mathbf{F}$ as $\mathrm{Var}(\hat{\theta}_{ij}) \geq \tilde{F}(\theta_{ij}, \theta_{ij})$.[24] Then the FWHM CRLB of θ_{ij} is equal to

$$\Delta(\theta_{ij}) = 2\sqrt{2\ln 2\tilde{F}(\theta_{ij}, \theta_{ij})}. \tag{12}$$

Given an imaging model and emitter locations, the CRLBs of all emitters can be calculated.

The CRLB derived above can be extended to the limit of finest pixelation in which pixel sizes $\Delta_x \to 0$, $\Delta_y \to 0$, and an image sensor covers the entire 2-D plane $\mathbb{R}^2$. By the mean-value theorem, $Q(k_x, k_y) \to Q(x, y)$, $q_m(k_x, k_y) \to q_m(x, y)$, and then it follows from Eq. (11) that

$$F(\theta_{ij}, \theta_{ml}) = I\eta\Delta_t \int_{\mathbb{R}^2} \frac{1}{Q(x, y) + \gamma_p^{-1}}\frac{\partial q_i(x, y)}{\partial \theta_{ij}}\frac{\partial q_m(x, y)}{\partial \theta_{ml}}\mathrm{d}x\mathrm{d}y, \tag{13}$$

by which the FWHM CRLB is obtained through Eq. (12).

It is known[25] that the CRLB of a single 2-D emitter is independent of emitter location. However, a cluster of emitters with overlapped PSFs can mutually affect and make each other's CRLBs higher. A PSF $q_m(x, y)$ in optical nanoscopy usually has an effective region, say the set of all (x, y) with $\|(x, y) - (x_m, y_m)\| < \beta$ for some $\beta > 0$ for a 2-D Gaussian PSF, such that the PSF and its derivative with respect to the emitter location outside the effective region are equal to zero approximately. For a 2-D Gaussian PSF in Eq. (42) in Appendix A, the set of $\|(x, y) - (x_m, y_m)\| < \beta$ for $\beta = 2\sigma$ encompasses 95% detected photons emitted from an emitter. The emitters with overlapped PSFs (or overlapped effective regions) increase each other's CRLBs in two ways. First, when $M \geq 2$, we have $Q(x, y) \geq q_m(x, y)$ for all (x, y) belonging to the effect region of m'th PSF for any m where the equality is true if and only if the PSF of m'th emitter is not overlapped with others'. The j'th emitter whose PSF is overlapped with the PSF of m'th emitter generates a term of interemitter interference $q_j(x, y)$ in $1/[Q(x, y) + \gamma_p^{-1}]$ that affects similarly to the noise and therefore increases the CRLB of the m'th emitter. Conversely, the m'th emitter also generates a term of interemitter interference $q_m(x, y)$ that increases the CRLB of the j'th emitter. All the emitters whose PSFs are overlapped with the PSF of the m'th emitter generate the total interemitter interference $\sum_{j\neq m} q_j(x, y)$ to the m'th emitter. Clearly, the more the PSFs overlapped with the PSF of the m'th emitter, the severer the interemitter interference and the higher the CRLB of the m'th emitter. A single isolated emitter does not suffer from interemitter interference. Second, if the PSFs of the i'th and m'th emitters are not overlapped, then approximately $[\partial q_i(x, y)/\partial\theta_{ij}][\partial q_m(x, y)/\partial\theta_{ml}] = 0$ for all (x, y) and therefore

$F(\theta_{ij},\theta_{ml})=0$ by Eqs. (11) and (13). On the other hand, if the PSFs of the i'th and m'th emitters are overlapped, then $F(\theta_{ij},\theta_{ml})\neq 0$. By multiplying the Fisher information matrix with two proper permutation matrices to switch accordingly the rows and columns (which does not change the CRLB), a cluster of emitters whose PSFs are overlapped yields a submatrix $\mathbf{F}_1$ such that the new Fisher information matrix can be written as

$$\mathbf{F}=\begin{pmatrix}\mathbf{F}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{F}_2\end{pmatrix},$$

where $\mathbf{F}_2$ is yielded by other emitters. Consequently, the inverse of Fisher information matrix $\mathbf{F}$ is

$$\mathbf{F}^{-1}=\begin{pmatrix}\mathbf{F}_1^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{F}_2^{-1}\end{pmatrix}.$$

Clearly, the FWHM CRLB $\Delta(\theta_{ij})$ of one emitter depends on and only on the locations of all emitters in the same cluster. The closer two emitters in the cluster are, the larger the corresponding off-diagonal elements in $\mathbf{F}_1$. The more the emitters in the cluster are, the more the off-diagonal elements in $\mathbf{F}_1$. Both cases result in larger diagonal elements in $\mathbf{F}_1^{-1}$ and larger CRLBs. Hence, CRLB of an emitter is higher as its neighbor emitters are closer and/or as the number of emitters in the cluster is larger, both meaning a higher density of emitters. A single isolated emitter reaches the lower bound of CRLB. These properties will be illustrated by the numerical examples in Sec. 3. Two isolated clusters of emitters yield two submatrices without common nonzero elements in the Fisher information matrix $\mathbf{F}$. Therefore, emitters can be localized cluster by cluster independently without effect on localization precision. This has been a practice in the literature.[17,21,22] In a practical experiment, the probability of probe activation needs to be sufficiently low in order for an image to be split into a number of discrete clusters rather than one cluster spanning the entire image.

Equations (11) to (13) also indicate that the FWHM CRLB $\Delta(\theta_{ij})$ decreases as SPNR increases and is inversely proportional to the square root of the mean number of detected photons per emitter $(I\eta\Delta_t)^{0.5}$. $\Delta(\theta_{ij})$ is invariant to the lateral translation and rotation of all emitter locations as a whole. Therefore, $\Delta(\theta_{ij})$ depends on the lateral relationship of emitter locations regardless of their absolute locations.

2.2 *Cramer-Rao Lower Bound with Additional Gaussian Noise*

2.2.1 *Gaussian noise*

Electronic devices in an image sensor produce thermal readout noise $w(t,x,y)$ measured in voltage. A CMOS sensor[40] produces more significant readout noise than an electron-multiplying CCD (EMCCD) sensor that yields a readout noise much smaller than 1 photoelectron RMS.[41] It is common to consider the thermal noise as an additive, stationary, white, and Gaussian process.[42] Specifically, assume that the Gaussian noise process $w(t,x,y)$ is spatially and temporarily independent and is identically Gaussian distributed with mean μ and variance G. The autocorrelation function of $w(t,x,y)$ is equal to

$$E[w(t_1,x_1,y_1)w(t_2,x_2,y_2)]=(G+\mu^2)\delta(t_1-t_2)\delta(x_1-x_2)\delta(y_1-y_2). \quad (14)$$

By Eq. (14), the power spectral density function of $w(t,x,y)$ is a constant $G+\mu^2$, a white spectrum. G measures power in energy per second per nm^2 with exclusion of the DC component. The total Gaussian noise readout in pixel (k_x,k_y) is

$$W(k_x,k_y)=\int_0^{\Delta_t}\int_{\Delta_y k_y}^{\Delta_y(k_y+1)}\int_{\Delta_x k_x}^{\Delta_x(k_x+1)} w(t,x,y)\mathrm{d}x\mathrm{d}y\mathrm{d}t.$$

It is clear that $W(k_x,k_y)$ is spatially (pixel-wise) independent and identically Gaussian distributed with mean $\mu_w=\Delta_t\Delta_x\Delta_y\mu$ and variance

$$\sigma_w^2=\Delta_t\Delta_x\Delta_y G. \quad (15)$$

σ_w^2 is also the energy of Gaussian noise in a pixel with exclusion of DC component. With the additional Gaussian noise, the intensity of pixel (k_x,k_y) in an image sensor is $U(k_x,k_y)=V(k_x,k_y)+W(k_x,k_y)$, by which we assume that the photon count in an image sensor is measured in voltage.

We further define SGNR as

$$\gamma_g=\frac{I\eta}{G}\ (\mathrm{nm}^2/\mathrm{emitter}). \quad (16)$$

As a signal to noise power ratio, SGNR characterizes the emitter power relative to the Gaussian noise power density excluding the DC component. As will be seen, similar to SPNR, SGNR plays an independent role in CRLB and is scalable to PSF and pixel size. SGNR is universal to all systems and therefore can be used to compare the performances of different systems and experiments. However, since $\gamma_g=\Delta_t\Delta_x\Delta_y I\eta/(\Delta_t\Delta_x\Delta_y G)$, given a particular nanoscopy system, SGNR can be estimated using the image data where a single emitter is presented and the data where no emitter is presented. For a practical system, we use SGNR $=10^{-6}\gamma_g$ (μm^2/emitter).

The photon emission of emitters, Poisson noise, and Gaussian noise are mutually independent.

2.2.2 *Cramer-Rao lower bound*

Given a readout $U(k_x,k_y)$ of pixel (k_x,k_y), a likelihood function of emitter locations that is also the probability density function of $U(k_x,k_y)$ is

$$p[k_x,k_y,U(k_x,k_y)]=\sum_{l=0}^{\infty} g(k_x,k_y,l)h[U(k_x,k_y)-l], \quad (17)$$

where

$$h[W(k_x,k_y)]=\frac{1}{\sqrt{2\pi}\sigma_w}\exp\left\{-\frac{[W(k_x,k_y)-\mu_w]^2}{2\sigma_w^2}\right\}$$

is the probability density function of $W(k_x,k_y)$. Due to the independence of $V(k_x,k_y)$ and $W(k_x,k_y)$, the likelihood function in Eq. (17) is actually equal to $p[k_x,k_y,U(k_x,k_y)]=E\{h[U(k_x,k_y)-V(k_x,k_y)]\}$ with the expectation E taken over $V(k_x,k_y)$. Since $U(k_x,k_y)$'s are mutually independent, a likelihood function of θ obtained from the image is

$$f = \prod_{(k_x,k_y)\in\Omega} p[k_x, k_y, U(k_x, k_y)]. \tag{18}$$

Then

$$\frac{\partial \ln f}{\partial \theta_{ij}} = \sum_{(k_x,k_y)\in\Omega} \frac{1}{p[k_x, k_y, U(k_x, k_y)]} \frac{\partial p[k_x, k_y, U(k_x, k_y)]}{\partial \theta_{ij}},$$

where

$$\frac{\partial p[k_x, k_y, U(k_x, k_y)]}{\partial \theta_{ij}} = \frac{\partial v(k_x, k_y)}{\partial \theta_{ij}} \sum_{l=0}^{\infty} [g(k_x, k_y, l-1) - g(k_x, k_y, l)] h[U(k_x, k_y) - l]$$
$$= \frac{\partial v(k_x, k_y)}{\partial \theta_{ij}} \{p[k_x, k_y, U(k_x, k_y) - 1] - p[k_x, k_y, U(k_x, k_y)]\}. \tag{19}$$

Since $E\{\partial \ln p[k_x, k_y, U(k_x, k_y)]/\partial\theta_{ij}\} = 0$ (the mean of a score equals zero[23]), where the expectation E is taken with respect to $U(k_x, k_y)$, we obtain

$$F(\theta_{ij}, \theta_{ml}) = \sum_{(k_x,k_y)\in\Omega} E\left\{\frac{1}{p^2[k_x, k_y, U(k_x, k_y)]} \frac{\partial p[k_x, k_y, U(k_x, k_y)]}{\partial \theta_{ij}} \frac{\partial p[k_x, k_y, U(k_x, k_y)]}{\partial \theta_{ml}}\right\}$$
$$= \sum_{(k_x,k_y)\in\Omega} E\left(\frac{\{p[k_x, k_y, U(k_x, k_y) - 1] - p[k_x, k_y, U(k_x, k_y)]\}^2}{p^2[k_x, k_y, U(k_x, k_y)]}\right) \frac{\partial v(k_x, k_y)}{\partial \theta_{ij}} \frac{\partial v(k_x, k_y)}{\partial \theta_{ml}}$$
$$= (I\eta\Delta_t\Delta_x\Delta_y)^2 \sum_{(k_x,k_y)\in\Omega} \left[\int_{\mathbb{R}} \frac{p^2(k_x, k_y, u-1)}{p(k_x, k_y, u)} du - 1\right] \frac{\partial q_i(k_x, k_y)}{\partial \theta_{ij}} \frac{\partial q_m(k_x, k_y)}{\partial \theta_{ml}}, \tag{20}$$

which yields the FWHM CRLB through Eq. (12).

2.2.3 *Gaussian approximation*

The role that the mean and variance of Gaussian noise play in the Fisher information matrix of Eq. (20) is difficult to see. Moreover, the integral in Eq. (20) is computationally complex in a numerical evaluation. To overcome these difficulties, a slight approximation shall be applied to Eq. (20). $p(k_x, k_y, u)$ and $p(k_x, k_y, u-1)$ are probability density functions of $U(k_x, k_y)$ and $U(k_x, k_y) + 1$, respectively. When the mean number of detected photons $v(k_x, k_y)$ in pixel (k_x, k_y) is ≥ 10, a Poisson random variable $V(k_x, k_y)$ can be well approximated by a Gaussian random variable. This condition is generally true in a practical fluorescence microscope since Poisson noise alone generally satisfies $\Delta_t\Delta_x\Delta_y b \geq 10$.[17–19,21] Then $U(k_x, k_y)$ and $U(k_x, k_y) + 1$ can be well approximated as Gaussian random variables with means of $v(k_x, k_y) + \mu_w$ and $v(k_x, k_y) + \mu_w + 1$, respectively, and the same variance of $v(k_x, k_y) + \sigma_w^2$. Therefore, it can be proved that

$$\int_{\mathbb{R}} \frac{p^2(k_x, k_y, u-1)}{p(k_x, k_y, u)} du = \exp\left[\frac{1}{v(k_x, k_y) + \sigma_w^2}\right]. \tag{21}$$

Consequently, the integral in Eq. (20) can be well approximated by the right-hand side of Eq. (21). Furthermore, $\exp(x) \cong 1 + x$ for $|x| \ll 1$. Hence, from Eqs. (5), (15), (16), and (20), we obtain

$$F(\theta_{ij}, \theta_{ml}) = I\eta\Delta_t\Delta_x\Delta_y \sum_{(k_x,k_y)\in\Omega} \frac{1}{Q(k_x, k_y) + \gamma_p^{-1} + \gamma_g^{-1}} \times \frac{\partial q_i(k_x, k_y)}{\partial \theta_{ij}} \frac{\partial q_m(k_x, k_y)}{\partial \theta_{ml}}, \tag{22}$$

which yields the FWHM CRLB by Eq. (12).

Compared to Eq. (11) where there is no Gaussian noise, Eq. (22) indicates that the Gaussian noise effectively increases the Poisson noise; according to Eq. (15), the effectively increased mean photon count of Poisson noise is equal to the variance of Gaussian noise per pixel. Therefore, when both the Poisson and Gaussian noises exist, the total effective SNR is equal to

$$\gamma = \frac{\gamma_p\gamma_g}{\gamma_p + \gamma_g} = \frac{I\eta}{b + G} \text{ (nm}^2\text{/emitter)}. \tag{23}$$

In practice, we shall use SNR $= 10^{-6}\gamma$ (μm^2/emitter).

In the limit of finest pixelation,

$$F(\theta_{ij}, \theta_{ml}) = I\eta\Delta_t \int_{\mathbb{R}^2} \frac{1}{Q(x, y) + \gamma_p^{-1} + \gamma_g^{-1}} \times \frac{\partial q_i(x, y)}{\partial \theta_{ij}} \frac{\partial q_m(x, y)}{\partial \theta_{ml}} dxdy, \tag{24}$$

which is comparable with Eq. (13) where there is no Gaussian noise.

It follows from Eqs. (22) and (24) that the CRLB is determined by the mean number of detected photons of an emitter, SPNR, SGNR, PSF, pixel sizes, and relationship of emitter locations. The CRLB is inversely proportional to $(I\eta\Delta_t)^{0.5}$ and decreases as SPNR and/or SGNR increase. Equation (21) also indicates that the mean of Gaussian noise does not play a role in the CRLB. In practice, the mean of Gaussian noise can be always subtracted first from an image and the estimation accuracy of an estimator is not affected. Approximating the Gaussian noise as another Poisson noise process may also make a practical estimator computationally simpler. In this case, a new image can be produced by $\bar{U}(k_x, k_y) = U(k_x, k_y) + \sigma_w^2 - \mu_w$ so that the total noise is Poisson distributed with a mean of $b + G$ photons per second per nm^2. The CRLB obtained from the new image is unchanged and is given by Eqs. (22) and (24).

It is clear that even from the same image, a localization nanoscope yields a different CRLB for each emitter depending on the relationship of multiemitter locations in a cluster.

3 Results and Discussion

The analysis in the preceding section indicates that emitter locations can be estimated cluster by cluster without effect on the CRLB. It is desired to know when the image data for a cluster of emitters are worth using in estimation of emitter locations, and when the data shall be discarded. In this section, by numerical examples, we investigate how the CRLBs of emitters located on a curve are affected by the adjacent emitter distance and how the mean CRLB is affected by the number of emitters randomly uniformly distributed in a small square. We also investigate how the mean CRLB of emitters randomly uniformly distributed in a 2-D plane or in a 3-D cuboid is affected by the emitter density. Finally, we demonstrate how the mean CRLB is affected by the PSF, PSF SD, mean photon count per emitter, SNR, axial thickness, and pixel size. Three PSFs that are usually applied in localization nanoscopy are considered: 2-D Airy, 2-D Gaussian, and 3-D Gaussian with astigmatism. Their derivatives with respect to emitter locations, which are needed in the formulas of CRLB, are presented in Appendix A. Table 1 presents most parameters used in the simulations and figures. It is reported[16] that individual Alexa647 molecules are switched off in ~2 frames on average at a frame rate of 500 Hz with an average of 3500 detected photons per switching event. This means an Alexa647 molecule can yield averagely $I\eta = 875000$ detected photons per second. In Table 1, with the assumption of a frame rate of 100 Hz, it is practical to consider $I\eta\Delta_t = 3000$ and 5000 photons per frame per emitter. The specimen lateral size of $L_x = L_y = 4096$ nm is taken. In most cases, the pixel size is $\Delta_x = \Delta_y = 128$ nm, which is close to 166 nm used practically,[19] and then the image size is $K_x = K_y = 32$ pixels. Equations (22) and (12) are used in the calculation of FWHM CRLB. All simulations are carried out in MATLAB (The MathWorks Inc.) using the customer codes. The Poisson and Gaussian noise samples are generated using the functions embedded in MATLAB.

3.1 *Effect of Adjacent Distance of Emitters on a Curve*

To study the effect of adjacent emitter distance on CRLB, consider that a hundred emitters are located on a 2-D curve as shown in Fig. 2(b). The most inside emitter is indexed with 1, and the rest of the emitters are indexed in order on the curve and so the most outside emitter is indexed with 100. As the index increases, the distance between adjacent emitters linearly increases with the increment of ~2.65 nm per emitter. The minimum distance of adjacent emitters is 24.80 nm and the maximum is 287.63 nm. The 2-D Gaussian PSF (see Appendix A) with SD of 78.26 nm is used. The SD is equivalent to that of Airy PSF with numerical aperture of 1.4 and wavelength of 520 nm as calculated by Eq. (49) in Appendix B. The other parameters are presented in Table 1.

Shown in Figs. 2(a) to 2(c) are an image with the noise added according to the SPNR and SGNR in Table 1, the original and estimated emitter locations, and the FWHM CRLB versus the adjacent emitter distance, respectively. Figure 2(b) demonstrates how the reconstructed image of emitter locations looks, compared with the original emitter locations, if the same pattern of emitter locations appears five times in an acquired movie. As shown in Figs. 2(b) and 2(c), the FWHM CRLB of an emitter increases as the adjacent emitter distance decreases. Though having smaller distances of adjacent emitters, the emitters close to the inner end of the curve have less interemitter interference; consequently, their CRLBs with respect to the distance of adjacent emitters are convex down in Fig. 2(c). For a desired

Table 1 Parameters used in the numerical examples.

Figure	Dim	PSF	σ (nm)	$L_x = L_y$ (nm)	$2L_z$ (nm)	$\Delta_x = \Delta_y$ (nm)	$K_x = K_y$	$I\eta\Delta_t$ (ph/emt)	Signal to Poisson noise ratio (μm^2/emt)	Signal to Gaussian noise ratio (μm^2/emt)	Density (emt/μm^2)
Figs. 2(a) to 2(c)	2-D	Gauss	78.26	4096	0	128	32	3000	0.4	0.6	NA
Figs. 3(a) to 3(c)	2-D	Gauss	78.26	4096	0	128	32	3000	0.4	0.6	Vary
Fig. 4(a)	2-D	Airy	78.26	4096	0	128	32	5000	0.4	0.6	20[b]
Fig. 4(b)	2-D	Airy/ Gauss	78.26	4096	0	128	32	5000	0.4	0.6	Vary
Fig. 4(c)	2-D	Gauss	Vary	4096	0	128	32	5000	0.4	0.6	20, 16, 12, 8[b]
Fig. 5(a)	3-D	Gauss	[a]	4096	512	128	32	5000	0.4	0.6	4[c]
Fig. 5(b)	3-D	Gauss	[a]	4096	512	128	32	5000	2	2	Vary
Fig. 6(a)	3-D	Gauss	[a]	4096	512	128	32	Vary	2	2	3[c]
Fig. 6(b)	3-D	Gauss	[a]	4096	512	128	32	5000	Vary		3[c]
Fig. 6(c)	3-D	Gauss	[a]	4096	Vary	128	32	5000	2	2	3[c]
Fig. 6(d)	3-D	Gauss	[a]	4096	512	Vary		5000	2	2	3[c]

[a]The x and y SDs of three-dimensional (3-D) Gaussian point spread function (PSF) in Appendix A are determined by $c = 205$ nm, $d = 290$ nm, $\sigma_{x0} = 140$ nm, $A_x = 0.05$, $B_x = 0.03$, $\sigma_{y0} = 135$ nm, $A_y = -0.01$, $B_y = 0.02$.
[b]In Figs. 4(a) and 4(c), the emitter density of 20, 16, 12, and 8 emitters/μm^2 corresponds to 230, 184, 138, and 92 emitters, respectively.
[c]In Figs. 5(a) and 6(a) to 6(d), the lateral emitter density of 3 and 4 emitters/μm^2 corresponds to 34 and 46 emitters, respectively.

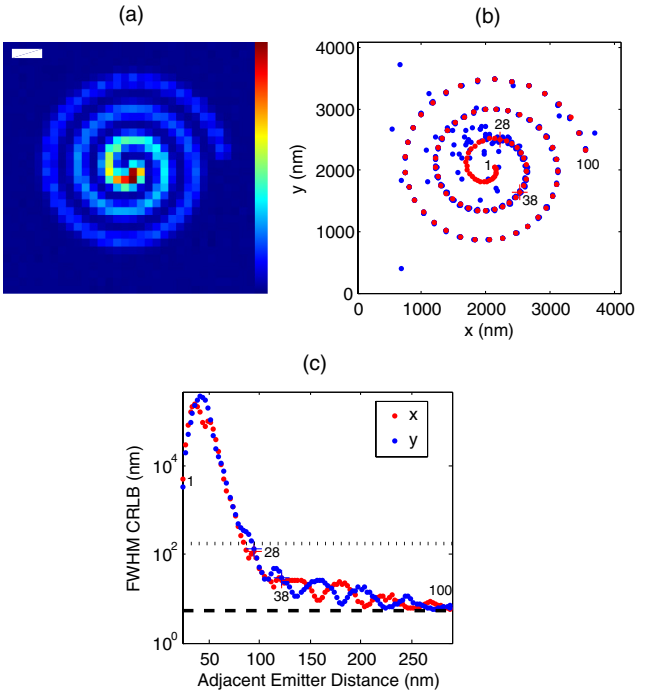

Fig. 2 How the Cramer-Rao lower bounds (CRLBs) of emitters located on a curve are affected by the distance of adjacent emitters. (a) An image of 100 emitters located on a 2-D curve with the noise added according to the signal to Poisson noise ratio (SPNR) and signal to Gaussian noise ratio (SGNR) in Table 1. The bar size is 500 nm. (b) The original (red) and estimated (blue) emitter locations. The locations of emitters 1, 28, 38, and 100 are indicated, and the locations of emitters 28, 38 are also denoted by +. The five estimated locations for each emitter are generated based on the full width at half maximum (FWHM) CRLB [calculated by Eqs. (22) and (12)] with an assumption of Gaussian estimates. (c) The FWHM CRLB for x (red) and y (blue) versus the adjacent emitter distance. The FWHM CRLBs of emitters 1, 28, 38, and 100 are indicated, and the FWHM CRLBs of emitters 28, 38 are also denoted by +. The single emitter bound (--) and point spread function (PSF) FWHM (···) are shown. For a desired FWHM CRLB, the allowable minimum distance of adjacent emitters can be obtained from (c).

FWHM CRLB, the allowable minimum distance of adjacent emitters can be determined from Fig. 2(c). For example, to obtain a localization precision >30 nm in both x and y directions, the adjacent emitter distance must be >121.88 nm, which is satisfied by the emitters indexed with 38 through 100. This means that no unbiased estimator can achieve a localization precision >30 nm if the adjacent emitter distance is <121.88 nm in this case. In practice, to obtain a 30 nm or greater localization precision, the image data for a cluster of emitters located on a curve should not be used in emitter localization if the adjacent emitter distance is <122 nm. Similarly, to achieve a localization precision better than the PSF FWHM of $2\sqrt{2\ln 2} \times 78.26 = 184.29$ nm, the adjacent emitter distance must be >94.87 nm, which is satisfied by the emitters indexed with 28 through 100. In practice, the image data for a cluster of emitters located on a curve should be discarded if the adjacent emitter distance is <94.87 nm in order to obtain a localization precision greater than the PSF FWHM. While increasing with the adjacent emitter distance decreasing, the FWHM CRLB of an emitter increases slowly for the last 70 emitters of which the adjacent emitter distance is >102.96 nm, and then suddenly deteriorates and increases exponentially fast for the remaining 30 emitters with the adjacent emitter distance <102.96 nm.

3.2 *Effect of the Number of Emitters in a Small Square*

The system setup is the same as in Sec. 3.1 as presented in Table 1. The 2-D region of $L_x \times L_y = 4096 \times 4096$ nm^2 is partitioned into 16 small squares each of size 1024×1024 nm^2. In each square, M emitters are randomly generated with a uniform distribution in a smaller centered square of size 512×512 nm^2, where $M = 1, 2, \ldots, 16$, respectively. The corresponding emitter density is equal to 3.81 (single emitter), 7.63, 11.44, 15.26, 19.07, 22.89, 26.70, 30.52, 34.33, 38.15, 41.96, 45.78, 49.59, 53.41, 57.22, 61.04 emitters/μm^2, respectively.

A realization of emitter locations in a square is randomly generated according to the uniform distribution. The FWHM CRLBs of all emitters in the square are calculated by Eqs. (22) and (12). Then the mean FWHM CRLB is obtained by averaging the FWHM CRLBs over all emitters in the square and over 1000 realizations. The mean FWHM CRLB is an appropriate measure on the quality of the reconstructed image of emitter locations. Figure 3(a) shows the image of one realization of emitter locations with the noise added according to the SPNR and SGNR in Table 1. Figure 3(b) illustrates the corresponding original and five estimated locations for each emitter. Figure 3(c) demonstrates the mean FWHM CRLB in each square versus the number of emitters (or emitter density accordingly).

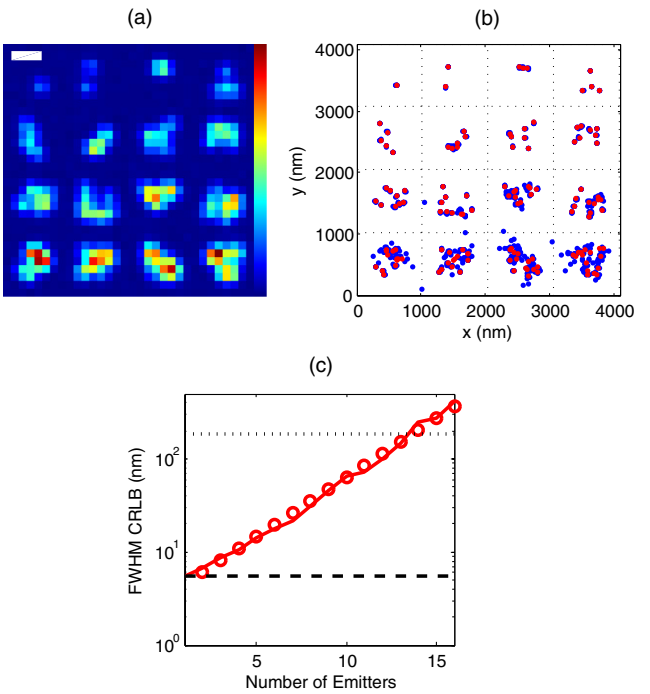

Fig. 3 How the mean CRLB of emitters located in a square is affected by the number of emitters (or emitter density). (a) An image for the realization of emitter locations in (b) with the noise added according to the SPNR and SGNR in Table 1. The bar size is 500 nm. (b) The original (red) and estimated (blue) emitter locations for one realization of emitter locations. The five estimated locations for each emitter are generated based on the FWHM CRLB [calculated by Eqs. (22) and (12)] with an assumption of Gaussian estimates. (c) The mean FWHM CRLB (-) in each square, averaged over 1000 realizations of emitter locations, versus the number of emitters. The mean FWHM CRLB well fits to an exponential function (o). The single emitter bound (--) and PSF FWHM (···) are shown.

It is interesting that the mean FWHM CRLB in the scale of 10-based logarithm is presented as a straight line and therefore is phenomenally presented as an exponential function of emitter number M as

$$\bar{\Delta} = \rho_0 a^M. \tag{25}$$

It is found that the mean FWHM CRLB fits well to $\bar{\Delta} = 3.38 \times 1.34^M$ in the least square sense, implying that the mean FWHM CRLB exponentially increases as the number of emitters (or emitter density) in a square increases. Figure 3(c) indicates that to obtain an FWHM CRLB <30 nm, the number of emitters in a square must satisfy $M \leq 7$ (or emitter density ≤26.70) on average. Hence, in practice, the image data for a cluster of $M \geq 8$ emitters in a square of size 512×512 nm^2 shall be discarded if an unbiased estimator is applied. It also indicates that when $M \geq 13$ (or emitter density ≥49.59), no unbiased estimator can produce on average a localization precision greater than the PSF FWHM of 184.29 nm.

3.3 *Effect of Emitter Density in a Large Square or Cuboid*

To further investigate the effect of emitter density, we consider the mean FWHM CRLB of emitters that are randomly uniformly distributed in a relatively large square or cuboid. A realization of emitter locations in a 2-D square or 3-D cuboid is randomly generated according to the uniform distribution. The FWHM CRLBs of all emitters in the realization are then calculated by Eqs. (22) and (12). The mean FWHM CRLB is obtained by averaging the FWHM CRLBs over all emitters in the square or the cuboid and then over a large number of realizations.

3.3.1 *2-D imaging*

For 2-D imaging, the mean FWHM CRLB is computed with both the 2-D Airy and 2-D Gaussian PSFs presented in Appendix A. Given numerical aperture n_a and emission wavelength λ, the Airy PSF is determined by Eq. (41) with $\alpha = 2\pi n_a/\lambda$. To fairly compare the performances of the two PSFs, the SD of an Airy PSF is computed and then used in the Gaussian PSF. In the simulation, we consider $n_a = 1.4$ and $\lambda = 520$ nm, which are normal in a practical system.[25] The Airy PSF SD computed from Eq. (49) is equal to $\sigma = 78.26$ nm, which is used in the Gaussian PSF. The emitters are randomly uniformly distributed in a square of $[350,3746] \times [350,3746]$ nm^2, where the 350 nm margins are used to eliminate the boundary effect. The other parameters are given in Table 1.

Figure 4(a) shows an image of one realization of emitter locations with the noise added according to the SPNR and SGNR in Table 1. Figure 4(b) illustrates the mean FWHM CRLB for both the Airy and Gaussian PSFs versus the emitter density. With the Gaussian PSF, at each point, the mean FWHM CRLB is obtained by averaging over 500 realizations of randomly generated emitter locations. The double integral in Eq. (4) with the Gaussian PSF can be decomposed into two single integrals in Eqs. (32), (45), and (46) with which the computational time is acceptable in a PC implemented with a CPU of Core i7. However, with the Airy PSF, the double integral in Eq. (4) is time-consuming. To make the computational time acceptable, the mean FWHM CRLB with the Airy PSF is

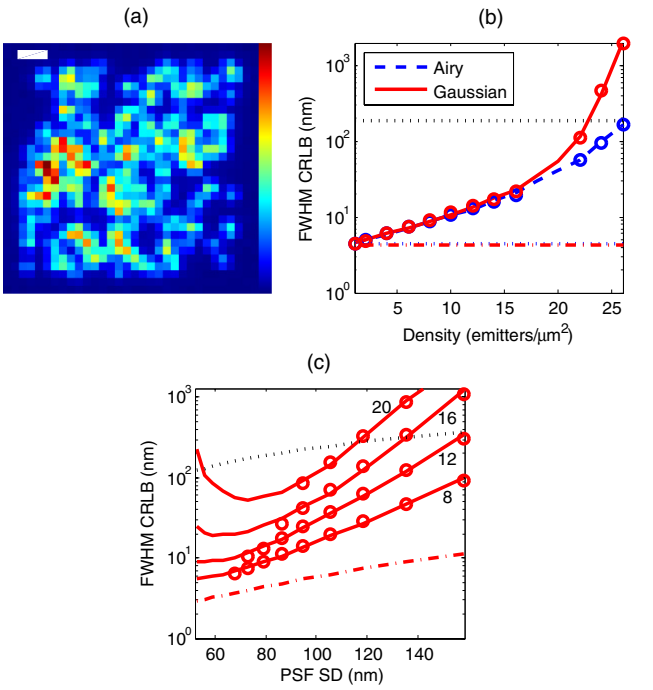

Fig. 4 How the mean CRLB is affected by the Airy and Gaussian PSFs, emitter density, and PSF standard deviation (SD) in 2-D imaging. (a) An image of a realization of emitter locations for 230 emitters (or emitter density of 20 emitters/μm^2) with the noise added according to the SPNR and SGNR in Table 1. The bar size is 500 nm. (b) The mean FWHM CRLB of multiemitters versus the emitter density with the Airy (- -) and Gaussian (–) PSFs. Both well fit to an exponential function (o) for the emitter density ≤16 emitters/μm^2 and for the emitter density ≥22 emitters/μm^2, respectively. The single emitter bounds with the Airy PSF (-···-) and with the Gaussian PSF (lower ···) and the PSF FWHM (upper ···) are shown. (c) The mean FWHM CRLB with the Gaussian PSF versus the PSF SD for the emitter densities (–) of 20, 16, 12, and 8 emitters/μm^2, respectively, all fitting well to an exponential function (o) when the PSF SD is approximately larger than a half of the pixel size (>128/2 = 64 nm). The single emitter bound (-···-) and the PSF FWHM (···) are shown.

obtained by averaging over 100 realizations of emitter locations for the emitter density ≤12, 50 realizations for the emitter density equal to 14 and 16, and 25 realizations for the emitter density ≥ 18 emitters/μm^2. Since the Airy and Gaussian PSFs are symmetric to the origin, the mean FWHM CRLBs in the x and y directions are analytically identical. To save the computational time, the mean FWHM CRLB in the calculation is also averaged over both the x and y directions.

As shown in Fig. 4(b), in the low-density regime where the emitter density is ≤16 emitters/μm^2, the mean FWHM CRLBs using the Airy and Gaussian PSFs are almost equal and both increase exponentially fast. They well fit to an exponential function of emitter density D as

$$\bar{\Delta} = \rho_0 a^D, \tag{26}$$

where $\bar{\Delta} = 4.04 \times 1.10^D$ with the Airy PSF, and $\bar{\Delta} = 3.96 \times 1.11^D$ with the Gaussian PSF. Both are <23 nm, representing a superresolution. In the high-density regime where the emitter density is ≥22 emitters/μm^2, the mean FWHM CRLB increases much faster and exponentially and well fits to $\bar{\Delta} = 0.17 \times 1.30^D$ with the Airy PSF and

$\bar{\Delta} = (1.51 \times 10^{-5}) \times 2.05^D$ with the Gaussian PSF. Therefore, in the regime of low emitter density, practically using a Gaussian PSF with the SD calculated by Eq. (49) is expected to obtain almost the same localization precision as using the Airy PSF. According to Fig. 4(b), no unbiased estimator can produce a mean localization precision < 30 nm if the emitter density is greater than ~18.2 emitters/μm^2. In practice, the image data with an emitter density >18.2 emitters/μm^2 should be discarded in order for an unbiased estimator to obtain a localization precision < 30 nm on average. Moreover, no unbiased estimator can yield a mean localization precision greater than the PSF FWHM of 184.29 nm when the emitter density is ≥26.43 with the Airy PSF or ≥ 22.70 emitters/μm^2 with the Gaussian PSF.

Figure 4(c) further illustrates how the Gaussian PSF SD σ affects the mean FWHM CRLB under different emitter densities. By Eq. (49), the range of $\sigma \in [52.67,158.02]$ nm in the figure is equivalent to the range of $n_a \in [1.0, 1.4]$ and $\lambda \in [350,750]$ nm with the Airy PSF, which covers most practical values.[43] A mean FWHM CRLB is obtained by averaging over 1000 realizations of emitter locations. When the PSF SD σ is sufficiently large, say $2\sigma > \Delta_x = \Delta_y = 128$ (or $\sigma > \Delta_x/2 = \Delta_y/2 = 64$), the mean FWHM CRLB increases exponentially fast as σ increases. The increase is more significant when the emitter density is larger, incurred by the severer overlapping of multiemitter PSFs. The mean FWHM CRLB well fits to an exponential function $\bar{\Delta} = \rho_0 a^\sigma$ in the regime of large PSF SD, that is, $\bar{\Delta} = 0.382 \times 1.059^\sigma$, $\bar{\Delta} = 0.320 \times 1.053^\sigma$, $\bar{\Delta} = 0.585 \times 1.040^\sigma$, and $\bar{\Delta} = 0.864 \times 1.030^\sigma$ for the emitter density of 20, 16, 12, and 8 emitters/μm^2, respectively. It is notable that for a large emitter density and a small PSF SD, the mean FWHM CRLB decreases as the PSF SD increases. This is due to the fact that with the given pixel size, when the PSF SD is small, say $\sigma \cong \Delta_x/2$, two neighbor emitters located within one pixel are difficult to localize since their emitted photons might be mostly located in the same pixel. It is known[25] that without pixelation, the FWHM CRLB of a single emitter is linearly proportional to the Gaussian PSF SD σ. The PSF FWHM $2\sqrt{2 \ln 2}\sigma$ is also linear in σ. Therefore, both of them are presented as a logarithm function in the scale of 10-based logarithm in Fig. 4(c).

3.3.2 *3-D imaging*

In 3-D imaging, the 3-D Gaussian PSF in Eqs. (27) to (31) in Appendix A is employed with the parameters presented in Table 1. The x and y parameters are slightly asymmetric, which is possible in practice.[34] The parameters are close to those[11] estimated from a practical system. A specimen is assumed to be located in a cuboid of size $L_x = 4096$, $L_y = 4096$, and L_z. Emitters are randomly uniformly distributed in a smaller cuboid of $[350,3746] \times [350,3746] \times [-L_z, L_z]$ nm^3. Other parameters are presented in Table 1. The mean FWHM CRLB is computed in the same way as in the 2-D imaging.

Figure 5(a) shows an image of one realization of emitter locations with the noise added according to the SPNR and SGNR in Table 1. There are 46 emitters corresponding to the lateral emitter density of 4 emitters/μm^2. Figure 5(b) illustrates the mean FWHM CRLB versus the lateral emitter density. The mean FWHM CRLB is obtained by averaging the FWHM CRLBs over all emitters in 1000 realizations of emitter locations. Since the x and y FWHM CRLBs are slightly different with the slight asymmetry of PSFs in the x and y directions, only the x and z mean FWHM CRLBs are shown. In the single-emitter case, the lateral emitter density is ~0.0876 emitters/μm^2.

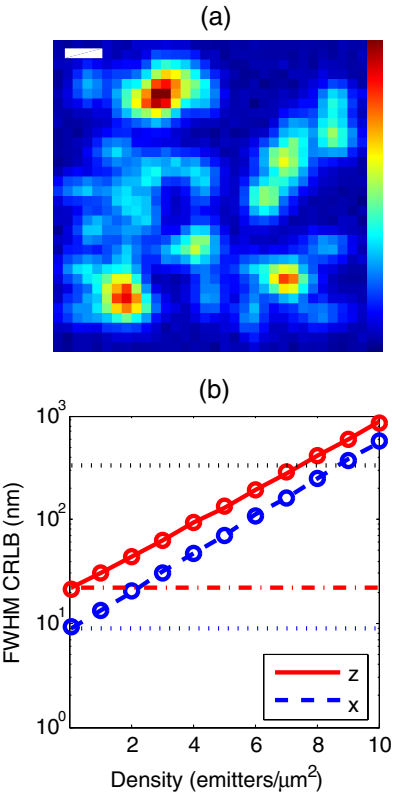

Fig. 5 How the mean CRLB of three-dimensional (3-D) emitters that are randomly uniformly distributed in a cuboid is affected by the lateral emitter density. (a) An image of one realization of emitter locations for 46 emitters (or lateral emitter density of 4 emitters/μm^2) with the noise added according to the SPNR and SGNR in Table 1. The bar size is 500 nm. (b) The mean FWHM CRLB versus the lateral emitter density in the z direction with multiemitters (–), in the x direction with multiemitters (--), in the z direction with a single emitter (-·), and in the x direction with a single emitter (···). The upper line (···) is the PSF FWHM in the x direction. The mean FWHM CRLB of multiemitters perfectly fits to an exponential function (o) in both the x and z directions.

In both the lateral plane and axial direction, the mean FWHM CRLB in the scale of 10-based logarithm is presented as a straight line, implying an exponential function of lateral emitter density. As shown in Fig. 5(b), the mean FWHM CRLB well fits to $\bar{\Delta} = 8.86 \times 1.52^D$ in the x direction and $\bar{\Delta} = 20.91 \times 1.45^D$ in the z direction. In terms of the exponential functions, in practice, the image data with the lateral emitter density >2.91 emitters/μm^2 on average should be discarded in order for an unbiased estimator to obtain a mean localization precision < 30 nm in x direction or < 61.65 nm in z direction. At the focal plane of x direction, the PSF SD of $\sigma_{x0} = 140$ nm is equivalent to an FWHM of 329.67 nm. By the exponential function, for the lateral emitter density >8.64 emitters/μm^2, no unbiased estimator can obtain on average a lateral localization precision greater than the PSF FWHM.

Compared with Fig. 4(b) for 2-D imaging with the same emitter density, the mean FWHM CRLBs in Fig. 5(b) for 3-D imaging are much higher. The reason is that the 3-D PSF SD of 140 nm is much larger than the 2-D PSF SD of 78.26 nm, although the effect of pixelation in the 2-D case is

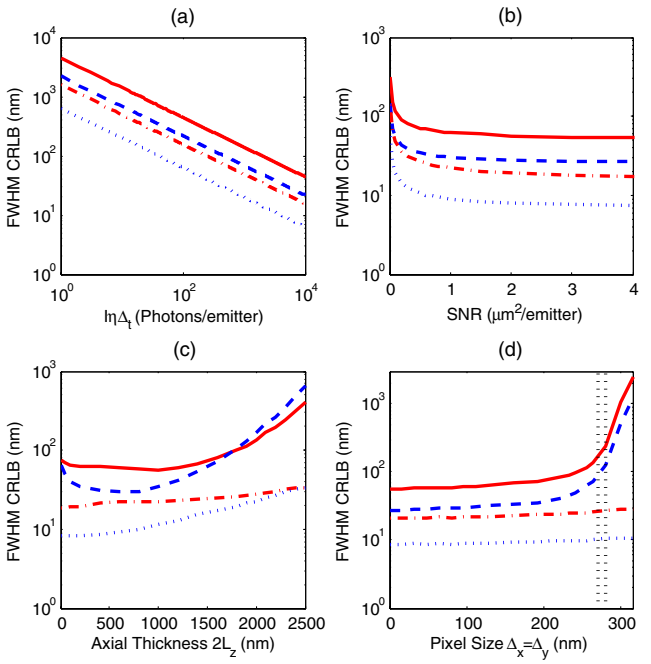

Fig. 6 The mean FWHM CRLB of 3-D emitters in the same setup of Fig. 5 versus (a) the mean number $I\eta\Delta_t$ of detected photons per emitter, (b) SNR, (c) axial thickness $2L_z$, and (d) pixel size $\Delta_x = \Delta_y$. The curves are for multiemitters in the z direction (–) and x direction (- -), and for a single emitter in the z direction (-·) and x direction (···). The vertical lines (···) in (d) indicate $2\sigma_{y0} = 270$ nm and $2\sigma_{x0} = 280$ nm, respectively.

relatively severer. Hence, the mean FWHM CRLB is more sensitive to the emitter density than the pixelation.

3.4 *Effect of Other Parameters in 3-D Imaging*

With the same setup of the above 3-D imaging, the effect of other parameters on the mean FWHM CRLB is investigated as follows. The parameters are presented in Table 1.

Mean photon count per emitter $I\eta\Delta_t$: Equations (11), (13), (22), and (24) indicate that given SPNR and SGNR, the FWHM CRLB monotonically decreases with the mean number of detected photons per emitter at the rate of $(I\eta\Delta_t)^{-0.5}$. This is verified in Fig. 6(a).

SNR: Figure 6(b) illustrates that the mean FWHM CRLB decreases monotonically as SNR increases. Moreover, SNR is divided into two regimes. In the regime of low SNR ≤ 0.5 μm^2/emitter, the mean FWHM CRLB is high and decreases fast as SNR increases. In the regime of high SNR > 0.5 μm^2/emitter, the mean FWHM CRLB is low and holds approximately a constant. This property is true in both the lateral plane and axial direction and in both the multiemitter and single-emitter cases.

Axial thickness $2L_z$: Figure 6(c) demonstrates that the mean FWHM CRLB is convex up with respect to axial thickness $2L_z$. In the regime of $100 \leq 2L_z \leq 1200$ nm, the mean FWHM CRLB is low. For $2L_z < 100$ nm, the x and y PSF SDs are approximately equal, and therefore a z location is slightly more difficult to estimate and so are lateral locations. In the regime of $2L_z > 1200$ nm, both the x and y PSF SDs are large; therefore, stronger interemitter interference and more noise enter the region of a PSF. Consequently, the mean FWHM CRLB of multiemitters increases exponentially as the axial thickness increases. In contrast, in the single-emitter case, increasing axial thickness only increases the noise in the region of a PSF and then slightly increases the mean FWHM CRLB.

Pixel size $\Delta_x = \Delta_y$: Figure 6(d) demonstrates that as the pixel size increases, the mean FWHM CRLB increases slowly in the regime of pixel size smaller than two times the PSF SDs at the focal planes (i.e., $\Delta_x = \Delta_y < 2\sigma_{x0} = 280$ nm or $2\sigma_{y0} = 270$ nm), and then increases exponentially fast in the regime of large pixel size where the interemitter interference is severe. In sharp contrast, the mean FWHM CRLB of a single emitter barely increases in the entire range of pixel size considered.

4 Conclusions

We analyzed the CRLB of unbiased estimators for stochastic optical localization nanoscopy that estimates locations of an arbitrary number of emitters frame by frame independently. It is found that the CRLB is a function of the mean number of detected photons per emitter, SPNR, SGNR, PSF, and relationship of emitter locations. With a slight and practical approximation, the effect of Gaussian noise is equivalent to increasing the mean photon count of Poisson noise by a number equal to the variance of Gaussian noise; the mean of Gaussian noise does not come into play. The PSF, Poisson noise, Gaussian noise, SPNR, and SGNR are defined to be universal to all systems and therefore can be used to compare performances of different systems, experiments, and estimators. A cluster of emitters through their overlapped PSFs causes interemitter interference and increases each other's CRLBs. An isolated single emitter reaches the lower bound of CRLB. An isolated cluster of emitters can be localized using only the image data of the cluster without effect on CRLB.

The CRLBs of emitters located on a curve increase fast as the adjacent emitter distance decreases. The mean CRLB of emitters randomly located in a small square with a uniform distribution increases exponentially fast as the number of emitters or the emitter density in the square increases. In order to achieve a desired mean CRLB, the image data for a cluster of emitters in a small region shall be discarded if the number of emitters or the emitter density is greater than a threshold. As the emitter density increases, the mean CRLB of emitters randomly uniformly distributed in a relatively large region increases exponentially fast in all cases of 2-D, 3-D, Airy PSF, and Gaussian PSF. In the regime where the PSF SD is greater than a half of pixel size, the mean CRLB increases exponentially fast as the PSF SD increases. The optimum pixel size is about two times the PSF SD. The mean CRLB decreases linearly as the square root of the mean photon count per emitter increases. In the regime of low SNR, the mean CRLB decreases fast as the SNR increases; in contrast, in the regime of high SNR, the mean CRLB is low and decreases slightly as the SNR increases. The mean CRLB is convex up with respect to the axial thickness and is low and flat in a range of axial thickness of practical interest. In the regime where the pixel size is smaller than two times the PSF SD, the mean CRLB increases slowly as the pixel size increases; but in the regime of large pixel size, it increases exponentially fast. In contrast, the CRLB of a single emitter in both 2-D and 3-D barely changes in the entire considered ranges of pixel size and PSF SD.

The analytical formulas of CRLB provide a theoretical foundation on which the performances of different systems, experiments, and estimators are comparable. The analytical and numerical results provide a guideline for the design of emitter location estimators and a benchmark for the achievable

localization precision. All the results are generally extensively applicable to 2-D and 3-D stochastic optical localization nanoscopy with an arbitrary spatially variant or invariant PSF.

Acknowledgments

The author thanks the anonymous reviewers for their constructive comments that helped improve the quality of this paper. The author declares no competing financial interest.

Appendix A: Derivatives of Three PSFs

The formulas of CRLB derived in Sec. 2 are applicable to a regular PSF $q_m(x,y)$ in 2-D and 3-D nanoscopy systems. The PSF can be spatially invariant, i.e., identical to all emitters, or spatially variant so that each emitter has a different PSF. A spatially variant PSF can be produced by aberration of an optical lens. To numerically evaluate the CRLB, the derivatives of PSF with respect to emitter locations need to be given in Eqs. (11), (13), (22), and (24). We consider three spatially invariant PSFs that are usually employed in localization nanoscopy: 3-D Gaussian, 2-D Airy, and 2-D Gaussian. Their derivatives are utilized in the numerical examples in Sec. 3. The derivatives of the 2-D Airy and 2-D Gaussian PSFs are derived in Ref. 25 and the derivative of 3-D Gaussian PSF is derived in Ref. 34. For completeness, the derivatives of these three PSFs are presented in this appendix in the context of this paper.

A.1 3-D PSF with Astigmatism

In 3-D nanoscopy with astigmatism, a cylindrical lens is added to a light path to produce an astigmatic PSF[11,44,45] such that the x and y SDs depend on z location of an emitter. In general, an astigmatic PSF can be approximated by a Gaussian density function as[34]

$$q_m(x,y) = q_x^{(m)}(x) q_y^{(m)}(y), \tag{27}$$

where

$$q_x^{(m)}(x) = \frac{1}{\sqrt{2\pi}\sigma_x(z_m)} \exp\left[-\frac{(x-x_m)^2}{2\sigma_x^2(z_m)}\right], \tag{28}$$

$$q_y^{(m)}(y) = \frac{1}{\sqrt{2\pi}\sigma_y(z_m)} \exp\left[-\frac{(y-y_m)^2}{2\sigma_y^2(z_m)}\right], \tag{29}$$

$$\sigma_x(z) = \sigma_{x0}\sqrt{1+(z+c)^2/d^2 + A_x(z+c)^3/d^3 + B_x(z+c)^4/d^4}, \tag{30}$$

$$\sigma_y(z) = \sigma_{y0}\sqrt{1+(z-c)^2/d^2 + A_y(z-c)^3/d^3 + B_y(z-c)^4/d^4} \tag{31}$$

are the x and y SDs, respectively. σ_{x0} and σ_{y0} are the x and y SDs at their focal planes, i.e., $z=-c$ and $z=c$, respectively. $\pm c$'s are offsets of the y and x focal planes from the average focal plane $z=0$. A_x, B_x, A_y, and B_y are coefficients for the higher-order distortion of SDs. By adding an extra cylindrical objective lens, the x and y SDs in Eqs. (28) to (31) change in accordance with z_m. Moreover, the peak value of PSF is also determined by z_m. Hence, both the SDs and peak value contain the information of location z_m. This is due to the fact that the integral of $q_m(x,y)$ as a probability density function is equal to one regardless of emitter location (x_m, y_m, z_m).

It follows from Eq. (27) that the mean photon count of pixel (k_x, k_y) produced by the m'th emitter in Eq. (4) is equal to

$$q_m(k_x,k_y) = q_x^{(m)}(k_x) q_y^{(m)}(k_y), \tag{32}$$

where

$$q_x^{(m)}(k_x) = \frac{1}{\Delta_x}\int_{\Delta_x k_x}^{\Delta_x(k_x+1)} q_x^{(m)}(x)\mathrm{d}x = \frac{1}{\Delta_x}\left\{\Phi\left[\frac{\Delta_x(k_x+1)-x_m}{\sigma_x(z_m)}\right] - \Phi\left[\frac{\Delta_x k_x - x_m}{\sigma_x(z_m)}\right]\right\}, \tag{33}$$

$$q_y^{(m)}(k_y) = \frac{1}{\Delta_y}\int_{\Delta_y k_y}^{\Delta_y(k_y+1)} q_y^{(m)}(y)\mathrm{d}y = \frac{1}{\Delta_y}\left\{\Phi\left[\frac{\Delta_y(k_y+1)-y_m}{\sigma_y(z_m)}\right] - \Phi\left[\frac{\Delta_y k_y - y_m}{\sigma_y(z_m)}\right]\right\}, \tag{34}$$

and $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x}\exp(-t^2/2)\mathrm{d}t$ is the accumulative probability distribution of a standard Gaussian random variable with zero mean and unit variance. The derivatives of PSF with respect to emitter locations can be written as

$$\frac{\partial q_i(k_x,k_y)}{\partial x_i} = \frac{\partial q_x^{(i)}(k_x)}{\partial x_i} q_y^{(i)}(k_y), \tag{35}$$

$$\frac{\partial q_i(k_x,k_y)}{\partial y_i} = q_x^{(i)}(k_x)\frac{\partial q_y^{(i)}(k_y)}{\partial y_i}, \tag{36}$$

$$\frac{\partial q_i(k_x,k_y)}{\partial z_i} = \frac{\partial q_x^{(i)}(k_x)}{\partial z_i} q_y^{(i)}(k_y) + q_x^{(i)}(k_x)\frac{\partial q_y^{(i)}(k_y)}{\partial z_i},$$

where

$$\frac{\partial q_x^{(i)}(k_x)}{\partial x_i} = \frac{1}{\Delta_x}\{-q_x^{(i)}[\Delta_x(k_x+1)] + q_x^{(i)}(\Delta_x k_x)\}, \tag{37}$$

$$\frac{\partial q_y^{(i)}(k_y)}{\partial y_i} = \frac{1}{\Delta_y}\{-q_y^{(i)}[\Delta_y(k_y+1)] + q_y^{(i)}(\Delta_y k_y)\}, \tag{38}$$

$$\frac{\partial}{\partial z_i} q_x^{(i)}(k_x) = \frac{\partial \sigma_x^2(z_i)}{\partial z_i}\frac{1}{2\Delta_x\sigma_x^2(z_i)\sqrt{2\pi}}\int_{(\Delta_x k_x - x_i)/\sigma_x(z_i)}^{[\Delta_x(k_x+1)-x_i]/\sigma_x(z_i)} \times (x^2-1)\exp\left(-\frac{x^2}{2}\right)\mathrm{d}x,$$

$$\frac{\partial}{\partial z_i} q_y^{(i)}(k_y) = \frac{\partial \sigma_y^2(z_i)}{\partial z_i} \frac{1}{2\Delta_y \sigma_y^2(z_i)\sqrt{2\pi}} \int_{(\Delta_y k_y - y_i)/\sigma_y(z_i)}^{[\Delta_y(k_y+1)-y_i]/\sigma_y(z_i)} \times (y^2-1)\exp\left(-\frac{y^2}{2}\right) dy$$

with

$$d\sigma_x^2(z_i)/dz_i = \sigma_{x0}^2[2(z_i+c)/d^2 + 3A_x(z_i+c)^2/d^3 + 4B_x(z_i+c)^3/d^4], \quad (39)$$

$$d\sigma_y^2(z_i)/dz_i = \sigma_{y0}^2[2(z_i-c)/d^2 + 3A_y(z_i-c)^2/d^3 + 4B_y(z_i-c)^3/d^4]. \quad (40)$$

In the limit of finest pixelation, the derivatives of PSF with respect to emitter locations in Eqs. (13) and (24) are equal to

$$\frac{\partial q_i(x,y)}{\partial x_i} = \frac{x-x_i}{\sigma_x^2(z_i)} q_i(x,y),$$

$$\frac{\partial q_i(x,y)}{\partial y_i} = \frac{y-y_i}{\sigma_y^2(z_i)} q_i(x,y),$$

$$\frac{\partial q_i(x,y)}{\partial z_i} = \left\{ \frac{1}{2\sigma_x^2(z_i)} \left[\frac{(x-x_i)^2}{\sigma_x^2(z_i)} - 1 \right] \frac{d\sigma_x^2(z_i)}{dz_i} + \frac{1}{2\sigma_y^2(z_i)} \left[\frac{(y-y_i)^2}{\sigma_y^2(z_i)} - 1 \right] \frac{d\sigma_y^2(z_i)}{dz_i} \right\} q_i(x,y)$$

with the notions of Eqs. (39) and (40).

A.2 2-D Airy PSF

Denote numerical aperture by n_a and fluorescence emission wavelength by λ. The 2-D Airy PSF at the focal plane of circular aperture can be expressed as[46]

$$q_m(x,y) = \frac{J_1^2\left[\alpha\sqrt{(x-x_m)^2+(y-y_m)^2}\right]}{\pi[(x-x_m)^2+(y-y_m)^2]}, \quad (41)$$

where $\alpha = 2\pi n_a/\lambda$ and J_1 is the first-order Bessel function of the first kind. The Airy PSF cannot be decomposed into a product of x and y PSFs. In this case, it follows from Eq. (4) that

$$\frac{\partial q_i(k_x,k_y)}{\partial x_i} = -\frac{1}{\Delta_x\Delta_y} \int_{\Delta_y k_y}^{\Delta_y(k_y+1)} \{q_i[\Delta_x(k_x+1), y] - q_i(\Delta_x k_x, y)\} dy,$$

$$\frac{\partial q_i(k_x,k_y)}{\partial y_i} = -\frac{1}{\Delta_x\Delta_y} \int_{\Delta_x k_x}^{\Delta_x(k_x+1)} \{q_i[x, \Delta_y(k_y+1)] - q_i(x, \Delta_y k_y)\} dx.$$

In the limit of finest pixelation,[25]

$$\frac{\partial q_i(x,y)}{\partial x_i} = \frac{2\alpha(x-x_i)J_1\left[\alpha\sqrt{(x-x_i)^2+(y-y_i)^2}\right] J_2\left[\alpha\sqrt{(x-x_i)^2+(y-y_i)^2}\right]}{\pi[(x-x_i)^2+(y-y_i)^2]^{3/2}},$$

$$\frac{\partial q_i(x,y)}{\partial y_i} = \frac{2\alpha(y-y_i)J_1\left[\alpha\sqrt{(x-x_i)^2+(y-y_i)^2}\right] J_2\left[\alpha\sqrt{(x-x_i)^2+(y-y_i)^2}\right]}{\pi[(x-x_i)^2+(y-y_i)^2]^{3/2}}.$$

A.3 2-D Gaussian PSF

The 2-D PSF of a wide-field fluorescence microscope can be well approximated by a 2-D Gaussian probability density function.[47] It can be decomposed into the product of two one-dimensional Gaussian probability density functions

$$q_m(x,y) = q_x^{(m)}(x) q_y^{(m)}(y), \quad (42)$$

where

$$q_x^{(m)}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-x_m)^2}{2\sigma^2}\right], \quad (43)$$

$$q_y^{(m)}(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y-y_m)^2}{2\sigma^2}\right] \quad (44)$$

with the assumption that the x and y SDs are identical to σ. Then Eq. (32) can be applied with

$$q_x^{(m)}(k_x) = \frac{1}{\Delta_x}\left\{\Phi\left[\frac{\Delta_x(k_x+1)-x_m}{\sigma}\right] - \Phi\left(\frac{\Delta_x k_x - x_m}{\sigma}\right)\right\}, \quad (45)$$

$$q_y^{(m)}(k_y) = \frac{1}{\Delta_y}\left\{\Phi\left[\frac{\Delta_y(k_y+1)-y_m}{\sigma}\right] - \Phi\left(\frac{\Delta_y k_y - y_m}{\sigma}\right)\right\}. \quad (46)$$

Note that Eqs. (35) to (38) are applicable with the expressions of Eqs. (43) to (46).

In the limit of finest pixelation,[25]

$$\frac{\partial q_i(x,y)}{\partial x_i} = \frac{x - x_i}{\sigma^2} q_i(x,y), \tag{47}$$

$$\frac{\partial q_i(x,y)}{\partial y_i} = \frac{y - y_i}{\sigma^2} q_i(x,y). \tag{48}$$

Appendix B: Standard Deviation of an Airy PSF

To fairly compare the performances of the Airy and Gaussian PSFs, the SD of an Airy PSF is computed and then used in the Gaussian PSF. Suppose an emitter is located at the origin. Accordingly, the Airy PSF is

$$q_a(x,y) = \frac{J_1^2\left(\alpha\sqrt{x^2+y^2}\right)}{\pi(x^2+y^2)},$$

and the Gaussian PSF is

$$q_g(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right).$$

Since $\sigma = 2\sqrt{2\pi}\int_0^\infty x^2 q_g(x,0)\mathrm{d}x$, the equivalent SD of the Airy PSF would be obtained by replacing $q_g(x,0)$ with $q_a(x,0)$ in the integral. However, the SD of an Airy PSF does not exist since $q_a(x,0)$ diminishes in the way[46] similar to $[\sin(\alpha x)/(\alpha x)]^2$ as x increases. To obviate the divergence, the SD is estimated by integral up to the second zero point of $q_a(x,0)$, i.e., $\alpha x = 7.016$.[46] Hence, the equivalent SD is calculated as

$$\sigma = \frac{2\sqrt{2}}{\sqrt{\pi}\alpha}\int_0^{7.016} J_1^2(x)\mathrm{d}x. \tag{49}$$

References

1. B. Huang, H. Babcock, and X. Zhuang, "Breaking the diffraction barrier: super-resolution imaging of cells," *Cell* **143**(7), 1047–1058 (2010).
2. S. W. Hell, "Far-field optical nanoscopy," *Springer Ser. Chem. Phys.* **96**(Pt 7), 365–398 (2010).
3. J. N. Farahani, M. J. Schibler, and L. A. Bentolila, "Stimulated emission depletion (STED) microscopy: from theory to practice," in *Microscopy: Science, Technology, Applications and Education*, pp. 1539–1547, Formatex, Badajoz, Spain (2010).
4. M. Hofmann et al., "Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins," *Proc. Natl. Acad. Sci. U. S. A.* **102**(49), 17565–17569 (2005).
5. M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution," *Proc. Natl. Acad. Sci. U. S. A.* **102**(37), 13081–13086 (2005).
6. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–796 (2006).
7. E. Betzig et al., "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642–1645 (2006).
8. S. R. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophys. J.* **91**(11), 4258–4272 (2006).
9. T. Dertinger et al., "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Natl. Acad. Sci. U. S. A.* **106**(52), 22287–22292 (2009).
10. M. Bates et al., "Multicolor super-resolution imaging with photo-switchable fluorescent probes," *Science* **317**(5845), 1749–1753 (2007).
11. B. Huang et al., "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science* **319**(5864), 810–813 (2008).
12. C. Cremer and B. R. Masters, "Resolution enhancement techniques in microscopy," *Eur. Phys. J. H.*, 1–64 (2013).
13. S. R. P. Pavani and R. Piestun, "Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system," *Opt. Express* **16**(26), 22048–22057 (2008).
14. G. Shtengel et al., "Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure," *Proc. Natl. Acad. Sci. U. S. A.* **106**(9), 3125–3130 (2009).
15. M. F. Juette et al., "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples," *Nat. Methods* **5**(6), 527–529 (2008).
16. S. A. Jones et al., "Fast, three-dimensional super-resolution imaging of live cells," *Nat. Methods* **8**(6), 499–505 (2011).
17. S. J. Holden, S. Uphoff, and A. N. Kapanidis, "DAOSTORM: an algorithm for high-density super-resolution microscopy," *Nat. Methods* **8**(4), 279–280 (2011).
18. H. Babcock, Y. M. Sigal, and X. Zhuang, "A high-density 3D localization algorithm for stochastic optical reconstruction microscopy," *Opt. Nanoscopy* **1**(1), 1–10 (2012).
19. L. Zhu et al., "Faster STORM using compressed sensing," *Nat. Methods* **9**(7), 721–723 (2012).
20. E. A. Mukamel, H. Babcock, and X. Zhuang, "Statistical deconvolution for superresolution fluorescence microscopy," *Biophys. J.* **102**(10), 2391–2400 (2012).
21. T. Quan et al., "High-density localization of active molecules using structured sparse model and Bayesian information criterion," *Opt. Express* **19**(18), 16963–16974 (2011).
22. Y. Wang et al., "PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy," *Opt. Express* **20**(14), 16039–16049 (2012).
23. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., John Wiley and Sons, New York (2006).
24. S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*, Prentice Hall PTR, Upper Saddle River, NJ (1993).
25. R. Ober, S. Ram, and E. S. Ward, "Localization accuracy in single-molecule microscopy," *Biophys. J.* **86**(2), 1185–1200 (2004).
26. S. Ram, E. S. Ward, and R. J. Ober, "How accurately can a single molecule be localized in three dimensions using a fluorescence microscope?," *Proc. SPIE* **5699**, 426–435 (2005).
27. S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of performance limits for optical microscopes," *Multidimens. Syst. Signal Process.* **17**(1), 27–57 (2006).
28. S. Ram et al., "A novel approach to determining the three-dimensional location of microscopic objects with applications to 3D particle tracking," *Proc. SPIE* **6443**, 64430D (2007).
29. S. Ram et al., "High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells," *Biophys. J.* **95**(12), 6025–6043 (2008).
30. S. Ram, E. S. Ward, and R. J. Ober, "Beyond Rayleigh's criterion: a resolution measure with application to single-molecule microscopy," *Proc. Natl. Acad. Sci. U. S. A.* **103**(12), 4457–4462 (2006).
31. J. Chao et al., "A resolution measure for three-dimensional microscopy," *Opt. Commun.* **282**(9), 1751–1761 (2009).
32. S. Ram, E. S. Ward, and R. J. Ober, "A stochastic analysis of distance estimation approaches in single molecule microscopy: quantifying the resolution limits of photon-limited imaging systems," *Multidimens. Syst. Signal Process.* **24**(3), 503–542 (2013).
33. K. I. Mortensen et al., "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nat. Methods* **7**(5), 377–381 (2010).
34. C. S. Smith et al., "Fast, single-molecule localization that achieves theoretically minimum uncertainty," *Nat. Methods* **7**(5), 373–375 (2010).

35. F. Huang et al., "Simultaneous multiple-emitter fitting for single molecule super-resolution imaging," *Biomed. Opt. Express* **2**(5), 1377–1393 (2011).
36. E. A. Mukamel and M. J. Schnitzer, "Unified resolution bounds for conventional and stochastic localization fluorescence microscopy," *Phys. Rev. Lett.* **109**(16), 168102 (2012).
37. M. Elbaum and P. Diament, "SNR in photocounting images of rough objects in partially coherent light," *Appl. Opt.* **15**(9), 2268–2275 (1976).
38. S. M. Ross, *Stochastic Processes*, 2nd ed., John Wiley and Sons, New York (1996).
39. S. Zacks, *The Theory of Statistical Inference*, John Wiley and Sons, New York (1971).
40. F. Huang et al., "Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms," *Nat. Methods* **10**, 653–658 (2013).
41. M. S. Robbins and B. J. Hadwen, "The noise performance of electron multiplying charge-coupled devices," *IEEE Trans. Electron Devices* **50**(5), 1227–1232 (2003).
42. J. G. Proakis, *Digital Communications*, 4th ed., McGraw Hill, New York (2000).
43. Invitrogen, *Alexa Fluor Dyes—Simply the Best and Brightest Fluorescent Dyes and Conjugates*, Invitrogen Detection Technology (2011).
44. L. Holtzer, T. Meckel, and T. Schmidt, "Nanometric three-dimensional tracking of individual quantum dots in cells," *Appl. Phys. Lett.* **90**(5), 053902 (2007).
45. H. P. Kao and A. S. Verkman, "Tracking of single fluorescent particles in three dimensions: use of cylindrical optics to encode particle position," *Biophys. J.* **67**(3), 1291–1300 (1994).
46. M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, Cambridge, UK (1999).
47. B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Appl. Opt.* **46**(10), 1819–1829 (2007).