Supplementary online material: Unified performance limits for conventional and stochastic localization microscopy

Eran A. Mukamel[1,2,3] and Mark J. Schnitzer[4,5]

[1] *Department of Physics, Stanford University, Stanford, California 94305, USA,*

[2] *Center for Brain Science, Harvard University, Cambridge, MA 02138, USA;*

[3] *Center for Theoretical Biological Physics,*

University of California, San Diego, 92093, USA,

[4] *Departments of Applied Physics and Biology,*

Stanford University, Stanford, California 94305, USA,

[5] *Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA.*

Correspondence: eran@post.harvard.edu, mschnitz@stanford.edu

Contents

I. FISHER INFORMATION MATRIX FOR TWO EMITTERS

We calculate the Cramér-Rao lower bound on localization error for the four degrees of freedom specifying the location of two emitters in 2D [1] [Fig. 1(a) in main text]. The separation between emitters is $2d_x$. If the instrument's PSF is approximated as a Gaussian, $h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$, and assuming both emitters are equally bright, the Fisher information matrix is diagonal with entries:

$$\begin{aligned}
\mathbf{J}_{y_0,y_0} &= \frac{N}{\sigma^2}, \\
\mathbf{J}_{x_0,x_0} &= \frac{N}{\sigma^2}\left(1 - \frac{d_x^2}{\sigma^2} e^{-\frac{d_x^2}{2\sigma^2}} \int_{-\infty}^{\infty} f_{d_x}(x)dx\right), \\
\mathbf{J}_{d_y,d_y} &= \frac{N}{\sigma^2}\left(1 - e^{-\frac{d_x^2}{2\sigma^2}} \int_{-\infty}^{\infty} f_{d_x}(x)dx\right), \\
\mathbf{J}_{d_x,d_x} &= \frac{N}{\sigma^2}\left(1 - e^{-\frac{d_x^2}{2\sigma^2}} \int_{-\infty}^{\infty} \frac{x^2}{\sigma^2} f_{d_x}(x)dx\right),
\end{aligned} \tag{1}$$

where $f_{d_x}(x) \equiv \frac{1}{\sqrt{2\pi}\sigma} \frac{e^{-x^2/2\sigma^2}}{\cosh(d_x x/\sigma^2)}$. Numerical results for an Airy disk PSF, as well as for a scene containing two emitters with unequal intensities, are shown in Fig. S1.

A surprising result of our analysis is the dip in $\mathbf{J}_{x_0,x_0}$ as a function of d_x [Fig. 1(a), red]. $\mathbf{J}_{x_0,x_0}$ approaches N/σ^2 for $d_x \gg \sigma$, as well as for $d_x \ll \sigma$. In the first case, ample separation between emitters makes it easy to assign each photon to its source. Averaging the photon locations from each emitter provides just as much information about the four degrees of freedom as if the emitters were imaged separately. In the opposite limit of overlapping PSFs, assigning photons correctly to one emitter is impossible. However, when emitters are very close together their mean position may be estimated by averaging the arrival locations of all observed photons, without regard to each photon's source. The minimum value of $\mathbf{J}_{x_0,x_0}$ is $\simeq 0.52N/\sigma^2$, which occurs at an intermediate separation, $d_x \simeq 1.23\sigma$. Unlike this non-monotonic behavior, $\mathbf{J}_{y_0,y_0}$ is independent of d_x [Fig. 1(a), cyan].

II. MAXIMUM LIKELIHOOD FITTING PROCEDURES

We tested the ability of a maximum likelihood fitting algorithm to achieve the information-theoretic limit on emitter localization for scenes containing two emitters with a Gaussian PSF [Fig. 1(a)], as well as for single emitters and an Airy disk PSF [Fig. 2(b)]. The Airy disk PSF is defined by $h(x,y) = \frac{J_1\left(\sqrt{x^2+y^2}/\sigma\right)^2}{\pi(x^2+y^2)}$, where J_1 is a Bessel function of the first kind. For both types of PSF, the radial symmetry implies that the Fisher information per photon about the emitter's

x, y coordinates satisfies $\mathbf{J}^{(1)}_{y,y} = \mathbf{J}^{(1)}_{x,x}$ and $\mathbf{J}^{(1)}_{x,y} = 0$. We find for the Airy PSF,

$$\begin{aligned}
\mathbf{J}^{(1)}_{x,x} &\equiv -\iint dxdy \frac{\partial^2 \log h(x,y)}{\partial x^2} h(x,y) \\
&= \iint dxdy \frac{1}{h(x,y)} \left(\frac{\partial h(x,y)}{\partial x} \right)^2 \\
&= \int_0^\infty dr \frac{4 J_2(r/\sigma)^2}{r\sigma^2} \\
&= \frac{1}{\sigma^2}.
\end{aligned} \tag{2}$$

The Airy disk PSF thus has the same Fisher information matrix as the Gaussian PSF, despite having infinite variance.

We generated simulated photon observation data by sampling from the two-dimensional distribution defined by the PSF. Using the simulated photon locations, $\{z_j\}$, we estimated the average positions of the emitters by maximizing the appropriate likelihood function using a simplex search method (MATLAB function `fminsearch`). This procedure gave an unbiased estimate of each emitter's location, and the error variance in each case saturated the CRLB for $n \gtrsim 100$ [Figs. 1(a) and 2(b)].

III. EXACT RESULTS AND GENERAL BOUNDS FOR THE INFORMATION TRANSFER FUNCTION (ITF) FOR GENERIC SCENES

In this section we quantify the estimation accuracy for a scene examined by stochastic localization microscopy, assuming no overlap between the images of neighboring emitters. This corresponds to working with a low density of simultaneously active emitters, $\rho_a \ll \sigma^{-2}$. In this regime, we calculate the information transfer function, $F(k)$, and its dependence on two key parameters: the overall density of emitters, ρ_e, and the number of photons collected per emitter, n. Here, the emitter density ρ_e is the density of all the emitters that are activated during many rounds of imaging. To avoid a high-variance, grainy scene estimate, it is generally desirable to achieve $\rho_e \gg \sigma^{-2}$ even though the instantaneous density of activated emitters is much lower. The number of image frames acquired, and thus the time required to attain a desired accuracy in reconstructing the scene, scales as ρ_e/ρ_a.

The information transfer function (ITF) is defined by the diagonal of the inverse Fisher information matrix:

$$F(k) \equiv \frac{1}{[\mathbf{J}^{-1}]_{I(k),I^*(k)}} \tag{3}$$

where

$$\mathbf{J}_{I(k),I^*(k')} \equiv \left\langle -\frac{\delta^2 \log P(\{N_l\}|I)}{\delta I(k)\delta I^*(k')} \right\rangle, \tag{4}$$

where $\{N_l\}$ is the observed fluorescence data. The angled brackets here indicate averaging over the distribution of the data, $P(\{N_l\}|I)$. To calculate the ITF we must specify this distribution.

As in Eq. 2 in the main text, we will parameterize the fluorescence data in terms of the locations of observed photons. Let $\mathbf{z}_{ij}$ be the observed location in the image plane of the i'th photon emitted by fluorophore j. The fluorophore's true position in the sample plane, $\mathbf{x}_j^{sample}$, corresponds via geometric optics to a location in the image plane, $\mathbf{x}_j$. The scene is represented by a non-negative intensity function, $B(\mathbf{x})$, and the total intensity, $B_0 = \int d\mathbf{x} B(\mathbf{x})$, is the expectation value of the number of activated emitters, m. The normalized image is $I(\mathbf{x}) \equiv B(\mathbf{x})/m$. The observed fluorescence data are completely specified by $\{m, \mathbf{z}_{ij}\}$. If h is the microscope's point spread function (PSF), the statistical model is $P(\{\mathbf{z}_{ij}\}, m|B) = P(m|B)P(\{\mathbf{z}_{ij}\}|m, B)$, where:

$$\begin{aligned} P(m|B) &= e^{-B_0} B_0^m / m! \\ P(\{\mathbf{z}_{ij}\}|m, B) &= \prod_{j=1}^{m} \int d\mathbf{x}_j P(\{\mathbf{z}_{ij}\}|\{\mathbf{x}_j\}) P(\mathbf{x}_j|B) \\ &= \prod_{j=1}^{m} \int d\mathbf{x}_j \prod_{i=1}^{n_j} h(\mathbf{z}_{ij} - \mathbf{x}_j) B(\mathbf{x}_j)/B_0. \end{aligned} \tag{5}$$

This equation assumes that each photon has been assigned to an emitter, whereas Eq. 1 and Eq. 2 in the main text describe a situation in which the observer does not know which emitter gave rise to each photon.

The log likelihood simplifies:

$$\begin{aligned} \log P(\{\mathbf{z}_{ij}\}, m|B) &= \log P(m|B) + \log P(\{\mathbf{z}_{ij}\}|m, B) \\ &= -B_0 + m \log B_0 - \log m! + \sum_{j=1}^{m} \log \left[\int d\mathbf{x}_j \prod_{i=1}^{n_j} h(\mathbf{z}_{ij} - \mathbf{x}_j) B(\mathbf{x}_j)/B_0 \right] \\ &= -B_0 + \sum_{j=1}^{m} \log \left[\int d\mathbf{x}_j \prod_{i=1}^{n_j} h(\mathbf{z}_{ij} - \mathbf{x}_j) B(\mathbf{x}_j) \right] - \log m!. \end{aligned} \tag{6}$$

In what follows we first derive the exact expression for $\mathbf{J}$ in case the PSF is Gaussian. We also derive an exact result for any PSF in the limit of conventional microscopy ($n = 1$). Finally we derive a bound on $F(k)$ that is valid for general PSFs and shows how the ITF scales with spatial frequency.

A. Exact calculation of the ITF for a Gaussian point spread function

For notational convenience, in this section we work in one spatial dimension, but the generalization to 2D is straightforward. To do an exact calculation, we assume the microscope can be effectively described by a Gaussian PSF with width σ, $h(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}}$. This form allows us to directly compute the integral over all possible locations, x_j, for the j'th emitter:

$$\int dx_j B(x_j) \prod_{i=1}^{n_j} h(z_{ij} - x_j) = \int dx_j B(x_j) \frac{1}{(\sqrt{2\pi}\sigma)^{n_j}} e^{-\sum_{i=1}^{n_j} \frac{(z_{ij}-x_j)^2}{2\sigma^2}} \tag{7}$$

It is convenient to write this in terms of the maximum likelihood estimate of the location of the j'th emitter, $\hat{x}_j \equiv \frac{1}{n_j} \sum_{i=1}^{n_j} z_{ij}$:

$$\begin{aligned} &= \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_j}(z_{ij}-\hat{x}_j)^2}}{(\sigma\sqrt{2\pi})^{n_j-1}} \int \frac{dx_j}{\sqrt{n_j}} B(x_j) \sqrt{\frac{n_j}{2\pi\sigma^2}} e^{-\frac{n_j}{2\sigma^2}(x_j-\hat{x}_j)^2} \\ &= Z_j \frac{\tilde{B}_{n_j}(\hat{x}_j)}{\sqrt{n_j}} \end{aligned} \tag{8}$$

where the factor $Z_j \equiv \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_j}(z_{ij}^2-\hat{x}_j^2)}}{(\sigma\sqrt{2\pi})^{n_j-1}}$ is independent of B, and $\tilde{B}_{n_j}(x) \equiv \int dx' B(x') h_{n_j}(x'-x)$ is a blurred version of the scene with Gaussian blurring kernel $h_{n_j}(x) \equiv e^{-\frac{n_j}{2\sigma^2}x^2} \sqrt{\frac{n_j}{2\pi\sigma^2}}$. Note that the blurring kernel width scales as $\sigma/\sqrt{n_j}$, becoming sharper as more photon locations are averaged to estimate the position of the j'th emitter. We can now write

$$\log P(\{z_{ij}\}, m|B) = -B_0 + \sum_{j=1}^{m} \log \tilde{B}_{n_j}(\hat{x}_j) + \sum_{j=1}^{m} \log Z_j/\sqrt{n_j} - \log m!. \tag{9}$$

Only the first two terms depend on the image and thus contribute to the derivatives with respect to B:

$$\begin{aligned} \frac{\delta \log P(\{z_{ij}\}, m|B)}{\delta B(y)} &= -1 + \sum_{j=1}^{m} \frac{h_{n_j}(\hat{x}_j - y)}{\tilde{B}_{n_j}(\hat{x}_j)}, \\ -\frac{\delta^2 \log P(\{z_{ij}\}, m|B)}{\delta B(y) \delta B(z)} &= \sum_{j=1}^{m} \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{\tilde{B}_{n_j}(\hat{x}_j)^2} \end{aligned} \tag{10}$$

We thus have

$$\mathbf{J}_{B(y),B(z)} = \sum_{m=0}^{\infty} P(m|B) \sum_{j=1}^{m} \int dz_{1j} \dots dz_{n_j j} \frac{Z_j}{\sqrt{n_j} B_0} \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{\tilde{B}_{n_j}(\hat{x}_j)} \tag{11}$$

The integral with respect to $\{z_{1j}, \dots, z_{n_j j}\}$ can be reduced to a one-dimensional integral by a change of variables. Define $w_{lj} \equiv \frac{1}{\sqrt{n_j}} \sum_{p=1}^{n_j} e^{2\pi i (l-1)(p-1)/n_j} z_{pj}$, $l = 1 \dots n_j$. This change of variables is an orthogonal transformation, so the volume element for integration is $\prod_{ij} dz_{ij} =$

$\prod_{lj} dw_{lj}$. We also have $\sum_{p=1}^{n_j} z_{pj}^2 = \sum_{l=1}^{n_j} |w_{lj}|^2$. In terms of these coordinates, $\hat{x}_j = \frac{1}{\sqrt{n_j}} w_{1j}$ and $\sum_{i=1}^{n_j} (z_{ij}^2 - \hat{x}_j^2) = \sum_{l=1}^{n_j} |w_{lj}|^2 - w_{1j}^2 = \sum_{l=2}^{n_j} |w_{lj}|^2$. We thus find

$$\begin{aligned} \int dz_{1j} \dots dz_{n_j j} \frac{Z_j}{\sqrt{n_j} B_0} \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{\tilde{B}_{n_j}(\hat{x}_j)} = \\ \int dw_{2j} \dots dw_{n_j j} \frac{e^{-\frac{1}{2\sigma^2} \sum_{l=2}^{n_j} |w_{lj}|^2}}{(\sigma\sqrt{2\pi})^{n_j - 1}} \int \frac{dw_{1j}}{\sqrt{n_j}} \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{B_0 \tilde{B}_{n_j}(\hat{x}_j)} \end{aligned} \tag{12}$$

The first term on the right hand side of this equation is just 1, so we have

$$= \int d\hat{x}_j \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{B_0 \tilde{B}_{n_j}(\hat{x}_j)} \tag{13}$$

Thus

$$\mathbf{J}_{B(y),B(z)} = \sum_{m=0}^{\infty} P(m|B) \sum_{j=1}^{m} \int d\hat{x}_j \frac{h_{n_j}(\hat{x}_j - y) h_{n_j}(\hat{x}_j - z)}{B_0 \tilde{B}_{n_j}(\hat{x}_j)} \tag{14}$$

We now simplify the Fisher information by assuming each fluorophore emits an equal number of photons, $n_j = n$. To verify that this approximation does not affect our results, we carried out numerical calculations of the ITF using a more realistic model with Poisson distributed n_j; the results (Fig. S2) show that the approximation is valid as long as $n \gtrsim 500$. Under the assumption of equal photon counts, we have

$$\begin{aligned} \mathbf{J}_{B(y),B(z)} &= \sum_{m=0}^{\infty} P(m|B) \frac{m}{B_0} \int d\hat{x} \frac{h_n(\hat{x} - y) h_n(\hat{x} - z)}{\tilde{B}_n(\hat{x})} \\ &= \int d\hat{x} \frac{h_n(\hat{x} - y) h_n(\hat{x} - z)}{\tilde{B}_n(\hat{x})}. \end{aligned} \tag{15}$$

We now transform this to an expression for the Fisher information about image spatial frequency components, $B(k) \equiv \int dx B(x) e^{i2\pi kx}$. We have:

$$\begin{aligned} \mathbf{J}_{B(k_1),B^*(k_2)} &= \int dy dz \frac{\delta B(y)}{\delta B(k_1)} \mathbf{J}_{B(y),B(z)} \frac{\delta B(z)}{\delta B^*(k_2)} \\ &= \int dy dz e^{-i2\pi(k_1 y - k_2 z)} \int d\hat{x} \frac{h_n(\hat{x} - y) h_n(\hat{x} - z)}{\tilde{B}_n(\hat{x})} \\ &= h_n(k_1) h_n(k_2) \int d\hat{x} \frac{e^{-i2\pi(k_1 - k_2)\hat{x}}}{\tilde{B}_n(\hat{x})} \end{aligned} \tag{16}$$

The inverse of $\mathbf{J}$ is simply

$$\mathbf{J}^{-1}_{B(k_1),B^*(k_2)} = \frac{1}{h_n(k_1) h_n(k_2)} \int d\hat{x} e^{-2\pi i(k_1 - k_2)\hat{x}} \tilde{B}_n(\hat{x}) \tag{17}$$

We can verify directly that this is the inverse:

$$\begin{aligned}\int dk_2 \mathbf{J}_{B(k_1),B^*(k_2)}\mathbf{J}^{-1}_{B(k_2),B^*(k_3)} =& \int dk_2 h_n(k_1)h_n(k_2)\int dx \frac{e^{-2\pi i(k_1-k_2)x}}{\tilde{B}_n(x)}\times \\ & \frac{1}{h_n(k_2)h_n(k_3)}\int dx' e^{-2\pi i(k_2-k_3)x'}\tilde{B}_n(x') \\ =& \frac{h_n(k_1)}{h_n(k_3)}\int dx dx' e^{-2\pi i(k_1 x - k_3 x')}\frac{\tilde{B}_n(x')}{\tilde{B}_n(x)}\int dk_2 e^{-2\pi i k_2(x'-x)}. \\ =& \frac{h_n(k_1)}{h_n(k_3)}\int dx dx' e^{-2\pi i(k_1 x - k_3 x')}\frac{\tilde{B}_n(x')}{\tilde{B}_n(x)}\delta(x-x') \\ =& \frac{h_n(k_1)}{h_n(k_3)}\int dx e^{-2\pi i(k_1-k_3)x}. \\ =& \delta(k_1-k_3).\end{aligned} \tag{18}$$

Finally, we express this as a bound on the estimate of a normalized image, $I(x) = \frac{1}{m}B(x)$: $\mathbf{J}^{-1}_{I(k_1),I^*(k_2)} = \frac{1}{m^2}\mathbf{J}^{-1}_{B(k_1),B^*(k_2)}$. The diagonal elements of $\mathbf{J}^{-1}$ give the CRLB on error variance, $\mathbf{E}_{I(k),I^*(k)} = \langle|\hat{I}(k) - I(k)|^2\rangle \geq [\mathbf{J}^{-1}]_{I(k),I^*(k)}$. We have

$$\mathbf{J}^{-1}_{I(k),I^*(k)} = \frac{1}{m^2 h_n(k)^2}\int dx \tilde{B}_n(x) = \frac{h_n(0)B_0}{m^2 h_n(k)^2} = \frac{B_0}{m^2}e^{\frac{(2\pi\sigma)^2}{n}k^2}, \tag{19}$$

where we used the normalization of the PSF ($\int dx h_n(x) = h_n(k=0) = 1$). This expression for the inverse Fisher information depends on the total image brightness, B_0, which is unknown to the observer. To obtain an expression in terms of the known parameters (m, n, σ, k) we replace B_0 with its maximum likelihood estimate, $\text{argmax}_{B_0} P(m|B_0) = m$. Finally, then, the ITF for Gaussian PSF, $F_0(k)$, is:

$$\begin{aligned}F_0(k) &= 1/[\mathbf{J}^{-1}]_{I(k),I^*(k)} \\ &= A\rho_e e^{-\frac{(2\pi k\sigma)^2}{n}},\end{aligned} \tag{20}$$

where $A\rho_e = m$. Surprisingly, the ITF is independent of the particular scene, $I(x)$, and thus provides a universal measure of performance that is valid for any scene. Eq. 20 shows how the resolution at each spatial frequency, k, depends on both the number of photons per emitter, n, and the emitter density, ρ_e.

B. Exact calculation of the ITF for conventional microscopy

Conventional microscopy can be treated as a special case of this framework when only a single photon is observed for each emitter, $n = 1$. In this case, Eq. 5 becomes $P(\{\mathbf{z}_j\}|m, B) =$

$\prod_{j=1}^{m} \int d\mathbf{x}_j B(\mathbf{x}_j) h(\mathbf{z}_j - \mathbf{x}_j)/B_0 = \prod_{j=1}^{m} \tilde{B}(\mathbf{z}_j)/B_0$, where $\tilde{B}$ is a blurred version of the scene given by convolution with the PSF. We have:

$$\mathbf{J}_{B(\mathbf{k}_1),B^*(\mathbf{k}_2)} = h(\mathbf{k}_1)h^*(\mathbf{k}_2) \int d\mathbf{z} \frac{e^{-i2\pi(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{z}}}{\tilde{B}(\mathbf{z})}. \tag{21}$$

By the same argument used in the previous section, we can find the diagonal element of the inverse Fisher information matrix:

$$\mathbf{J}^{-1}_{I(\mathbf{k}),I^*(\mathbf{k})} = \frac{1}{m^2|h(\mathbf{k})|^2} \int d\mathbf{z} \tilde{B}(\mathbf{z}) = \frac{1}{m|h(\mathbf{k})|^2}, \tag{22}$$

where we invoked the normalization of the scene and the PSF. As in the case of a Gaussian PSF discussed in the previous section, the ITF is independent of the particular scene. Because $n = 1$ here, the total number of photons is the same as the number of emitters, $N = m$. We thus find that the ITF for conventional microscopy is simply the square of the MTF, scaled by the number of observed photons:

$$F(\mathbf{k}) = N|h(\mathbf{k})|^2. \tag{23}$$

C. Bound on the ITF for stochastic localization microscopy with non-Gaussian PSFs

We now use the exact result, $F_0(k)$, to obtain a general bound on the ITF that applies for a non-Gaussian PSF. Using the convolution theorem, the statistical model can be written in terms of the Fourier transform of the image, $B(k) \equiv \int dx B(x) e^{i2\pi kx}$:

$$\begin{aligned} P(\{z_{ij}\}|m, B) &= \prod_{j=1}^{m} \int dk_j B(k_j) G(k_j; \{z_{ij}\})/B_0, \\ G(k_j; \{z_{ij}\}) &\equiv \int dk_{ij} \prod_{i=1}^{n_j} h(k_{ij}) e^{-i2\pi \sum_{i=1}^{n_j} k_{ij} z_{ij}} \delta\left(k_j - \sum_i k_{ij}\right) \end{aligned} \tag{24}$$

We thus have

$$\begin{aligned} \frac{\delta \log P(\{z_{ij}\}, m|B)}{\delta B(k)} &= -\delta(k) + \sum_{j=1}^{m} \frac{G(k; \{z_{ij}\})}{\int dk_j B(k_j) G(k_j; \{z_{ij}\})}, \\ -\frac{\delta^2 \log P(\{z_{ij}\}, m|B)}{\delta B(k_1) \delta B^*(k_2)} &= \sum_{j=1}^{m} \frac{G(k_1; \{z_{ij}\}) G^*(k_2; \{z_{ij}\})}{[\int dk_j B(k_j) G(k_j; \{z_{ij}\})]^2} \end{aligned} \tag{25}$$

The Fisher information matrix is

$$\mathbf{J}_{B(k_1),B^*(k_2)} = \sum_{m=0}^{\infty} P(m|B) \sum_{j=1}^{m} \int dz_{1j} \dots dz_{n_j j} \frac{1}{B_0} \frac{G(k_1; \{z_{ij}\}) G^*(k_2; \{z_{ij}\})}{\int dk_j B(k_j) G(k_j; \{z_{ij}\})} \tag{26}$$

Diffraction limits the spatial frequencies transmitted by any microscope, so that $h(k) \sim 0$ for k larger than some cutoff [2]. For example, a microscope with a circular limiting aperture is characterized by an Airy pattern PSF, $h(x,y) = \frac{J_1\left(\sqrt{x^2+y^2}/\sigma\right)^2}{\pi(x^2+y^2)}$, for which $h(k) = (1 - 2k\sqrt{1-k^2\pi^2} - \frac{2}{\pi}\sin^{-1}(k\pi))$ if $k < \frac{1}{\pi\sigma}$ and $h(k) = 0$ otherwise.

Here we make the general assumption that the PSF envelope is dominated by a Gaussian, $|h(k)| \leq e^{-(2\pi k\sigma)^2/2}$; this is true, in particular, for the Airy disk PSF, defined above. We will use the triangle inequality, which implies $|\int f(x)dx| \leq \int |f(x)|dx$. We have

$$\begin{aligned} |G(k_j;\{z_{ij}\})| &= \left| \int \prod_{i=1}^{n_j} dk_{ij} h(k_{ij}) e^{-i2\pi \sum_i k_{ij} z_{ij}} \delta\left(k_j - \sum_i k_{ij}\right) \right| \\ &\leq \int \prod_{i=1}^{n_j} dk_{ij} |h(k_{ij})| \delta\left(k_j - \sum_i k_{ij}\right) \\ &\leq \int \prod_{i=1}^{n_j} dk_{ij} e^{-\frac{(2\pi\sigma)^2}{2}\sum_i k_{ij}^2} \delta\left(k_j - \sum_i k_{ij}\right). \end{aligned} \tag{27}$$

To evaluate the Gaussian integral, we use a similar change of variables as before, $w_{lj} \equiv \frac{1}{\sqrt{n_j}} \sum_{p=1}^{n_j} e^{2\pi i(l-1)(p-1)/n_j} k_{pj}$, $l = 1 \ldots n_j$. Note that $w_{1j} = \frac{1}{\sqrt{n_j}} \sum_i k_{ij}$, and $\sum_i k_{ij}^2 = \sum_l |w_{lj}|^2$. We thus have

$$\begin{aligned} |G(k_j;\{z_{ij}\})| &\leq \int \prod_{l=1}^{n_j} dw_{lj} e^{-\frac{(2\pi\sigma)^2}{2}\sum_{l=1}^{n} |w_{lj}|^2} \delta\left(k_j - \sqrt{n_j} w_{1j}\right). \\ &= \frac{1}{\sqrt{n_j}} e^{-\frac{(2\pi\sigma)^2}{2n_j} k_j^2} (\sqrt{2\pi}\sigma)^{-n_j+1}. \end{aligned} \tag{28}$$

As we did above, we assume that each emitter produces the same number of photons, $n_j = n$. The Fisher information matrix is then bounded by

$$\begin{aligned} |\mathbf{J}_{I(k_1),I^*(k_2)}| &= m^2 \left| \sum_{m=0}^{\infty} P(m|B) m \int dz_1 \ldots dz_n \frac{G(k_1;\{z_i\}) G^*(k_2;\{z_i\})}{B_0 \int dk' B(k') G(k';\{z_i\})} \right| \\ &\leq m^2 \int dz_1 \ldots dz_n \frac{|G(k_1;\{z_i\})||G(k_2;\{z_i\})|}{|\int dk' B(k') G(k';\{z_i\})|} \\ &\leq m e^{-\frac{(2\pi\sigma)^2}{2n}(k_1^2+k_2^2)} \int dz_1 \ldots dz_n \frac{1}{|\int dk' I(k') G(k';\{z_i\})|} \\ &\leq mC e^{-\frac{(2\pi\sigma)^2}{2n}(k_1^2+k_2^2)}, \end{aligned} \tag{29}$$

where C is a constant, scene-dependent coefficient that does not depend on k_1, k_2 or m.

To relate this result to the CRLB, we need to compute the inverse of the Fisher information matrix. Here we can make use of a general relationship between the diagonal elements of a matrix, and the corresponding diagonal elements of the matrix inverse [3]: $\frac{1}{[\mathbf{J}^{-1}]_{kk}} \leq \mathbf{J}_{kk}$. Applying this

bound in our case, we find

$$F(k) = \frac{1}{[\mathbf{J}^{-1}]_{I(k),I(k)^*}} \leq mCe^{-(2\pi k\sigma)^2/n} = CF_0(k). \tag{30}$$

This proves the bound stated as Eq. 7 in the main text.

We can derive another bound on the ITF in the case of strictly band-limited PSFs, such as the Airy disk, for which $h(k) = 0$ for all $|k| \geq \frac{1}{\pi\sigma}$. In this case we can prove that $|G(k_j; \{z_{ij}\})| = 0$ for $|k_j| > \frac{n_j}{\pi\sigma}$. To show this, note that

$$\begin{aligned} G(k_j; \{z_{ij}\}) &= \int_{-\infty}^{\infty} dk_{ij} \prod_{i=1}^{n_j} h(k_{ij}) e^{-i2\pi \sum_i k_{ij} z_{ij}} \delta\left(k_j - \sum_i k_{ij}\right) \\ &= \int_{-1/\pi\sigma}^{1/\pi\sigma} dk_{ij} \prod_{i=1}^{n_j} h(k_{ij}) e^{-i2\pi \sum_i k_{ij} z_{ij}} \delta\left(k_j - \sum_i k_{ij}\right), \end{aligned} \tag{31}$$

where the second line follows from the fact that the PSF is band-limited. In the integrand we therefore have $|k_{ij}| \leq \frac{1}{\pi\sigma}$, so that $|\sum_{i=1}^{n_j} k_{ij}| \leq \sum_{i=1}^{n_j} |k_{ij}| \leq \frac{n_j}{\pi\sigma}$. Thus for $|k_j| > \frac{n_j}{\pi\sigma}$ the integrand vanishes due to the delta function. Frequencies greater than this hard cutoff are not transmitted at all, and the ITF, $F(k)$, vanishes if $|k| \geq \frac{n_j}{\pi\sigma}$. Note, however, that this frequency cutoff grows linearly with n, whereas the previously derived bound shows that the ITF is exponentially suppressed above a cutoff that scales with $\sqrt{n}$. The inequality, Eq. 30, is thus a stronger result when $n \gg 1$.

IV. ITF FOR SCENES CONSISTING OF A SET OF DISCRETE EMITTERS

Here we calculate the ITF, $F(\mathbf{k})$, for a scene composed of discrete emitters, $I(\mathbf{x}) = \frac{1}{m}\sum_{j=1}^{m} \delta(\mathbf{x} - \mathbf{x}_j)$ in real space or $I(\mathbf{k}) = \frac{1}{m}\sum_{j=1}^{m} e^{-i2\pi\mathbf{k}\cdot\mathbf{x}_j}$ in the frequency domain. More generally, the model applies to scenes of M emitters obtained in M/m distinct rounds with exactly m active emitters per round. Each round contributes independent information, so the total ITF is the sum of the ITF for each. We first treat the super-resolution limit, $\rho_a \ll \sigma^{-2}$, and then turn to the limit of conventional microscopy, $\rho_a \gg \sigma^{-2}$. Finally we use numerical simulations of 2D scenes to calculate the ITF for intermediate values of the active emitter density.

A. Super-resolution limit

In the super-resolution limit, emitters are sufficiently separated that their images can be treated as non-overlapping. In this case, and assuming an isotropic PSF, the Fisher information matrix for the location of a single emitter, $\mathbf{J}_{ij}$, is diagonal: $\mathbf{J}_{ij} = nJ^{(1)}\delta_{ij}$. Here we have defined $J^{(1)}$ to be the

Fisher information per photon for each of an emitter's two location coordinates. n is the number of detected photons assigned to each emitter. For the Gaussian and Airy disk PSFs defined above (see "Maximum likelihood fitting procedures"), $J^{(1)} = \frac{1}{\sigma^2}$.

To compute the ITF, we use the change of variable formula [4]:

$$1/F(\mathbf{k}) = \sum_{i,j=1}^{2M} \frac{\partial I(\mathbf{k})}{\partial q_i} \mathbf{J}_{ij}^{-1} \frac{\partial I^*(\mathbf{k})}{\partial q_j}. \tag{32}$$

Noting that $\frac{\partial I(\mathbf{k})}{\partial \mathbf{x}_j} = \frac{-i2\pi\mathbf{k}}{M} e^{-i2\pi\mathbf{k}\cdot\mathbf{x}_j}$, we have

$$\begin{aligned} 1/F(\mathbf{k}) &= \frac{1}{nJ^{(1)}} \sum_{i,j=1}^{2M} \frac{\partial I(\mathbf{k})}{\partial q_i} \delta_{ij} \frac{\partial I^*(\mathbf{k})}{\partial q_j} \\ &= \frac{1}{nJ^{(1)}} \sum_{j=1}^{M} \left| \frac{\partial I(\mathbf{k})}{\partial x_j} \right|^2 + \left| \frac{\partial I(\mathbf{k})}{\partial y_j} \right|^2 \\ &= \frac{1}{nJ^{(1)}} M \frac{(2\pi)^2 (k_x^2 + k_y^2)}{M^2} \\ &= \frac{(2\pi k)^2}{nA\rho_e J^{(1)}}. \end{aligned} \tag{33}$$

Here, $\rho_e = \frac{M}{A}$ is the total emitter density, summed over each separate round of imaging. This result shows that the ITF declines as $1/k^2$ in the limit of super-resolution microscopy. This exact result pertains to any isotropic PSF.

B. Conventional microscopy

In conventional microscopy, the PSFs of neighboring emitters overlap extensively. We approximate the Poisson photon count distribution in each camera pixel using a Gaussian noise model. (Note this assumption has nothing to do with the shape of the PSF, which may be Gaussian or Airy or, indeed, any other functional form). This noise model is valid if the mean number of photons per pixel is large enough, $N \gtrsim 10$. The approximation thus approaches the original Poisson model as the emitter density, and hence the total photon flux, increases. Denoting the PSF of the optical instrument used to collect the photons by $h(\mathbf{x})$, the observed fluorescence intensity (number of photons) at location $\mathbf{x}$ is

$$\nu(\mathbf{x}) = N \int d\mathbf{x}' h(\mathbf{x} - \mathbf{x}') I(\mathbf{x}') + \eta(\mathbf{x}), \tag{34}$$

where $\nu(\mathbf{x})$ is the number of detected photons at $\mathbf{x}$, N is the mean total number of detected photons, and the noise term, η, is Gaussian distributed with $\langle \eta(\mathbf{x}) \rangle = 0$ and $\langle \eta(\mathbf{x}) \eta(\mathbf{x}') \rangle = \langle \nu(\mathbf{x}) \rangle \delta(\mathbf{x} - \mathbf{x}')$.

We set the noise variance equal to the mean signal strength in each pixel to mimic Poisson shot noise, including independent noise statistics in different pixels. The Fourier transform of the observed data is

$$\tilde{\nu}(\mathbf{k}) = Nh(\mathbf{k})I(\mathbf{k}) + \tilde{\eta}(\mathbf{k}), \tag{35}$$

where $\tilde{\eta}$ is a Gaussian random process, with $\langle\tilde{\eta}(\mathbf{k})\rangle = 0$ and $\langle\tilde{\eta}(\mathbf{k})\tilde{\eta}^*(\mathbf{k}')\rangle = \langle\tilde{\nu}(\mathbf{k}-\mathbf{k}')\rangle$. Note that, because the noise term $\eta(\mathbf{x})$ lacks translational invariance, the noise components at different spatial frequencies are correlated. The probability distribution for the observations, $\tilde{\nu}$, is thus Gaussian with mean $\mu(\mathbf{k}) \equiv \langle\tilde{\nu}(\mathbf{k})\rangle = Nh(\mathbf{k})I(\mathbf{k})$ and covariance $\mathbf{C}_{\mathbf{k},\mathbf{k}'} = \mu(\mathbf{k}-\mathbf{k}')$.

The Fisher information for a Gaussian distributed random process is $\mathbf{J} = \mathbf{J}^{(mean)} + \mathbf{J}^{(var)}$, where [4]

$$\begin{aligned}\mathbf{J}^{(mean)}_{\mathbf{k},\mathbf{k}'} &= \frac{\partial\mu^T}{\partial I(\mathbf{k})}\mathbf{C}^{-1}\frac{\partial\mu}{\partial I^*(\mathbf{k}')},\\ \mathbf{J}^{(var)}_{\mathbf{k},\mathbf{k}'} &= \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial I(\mathbf{k})}\mathbf{C}^{-1}\frac{\partial\mathbf{C}}{\partial I^*(\mathbf{k}')}\right].\end{aligned} \tag{36}$$

The first term, $\mathbf{J}^{(mean)}$, represents the information carried by the mean photon count whereas the second term accounts for information conveyed by the variance in photon count. Note that μ and $\mathbf{C}$ are both proportional to the total photon count, N, so that $\mathbf{J}^{(mean)} \propto N$ but $\mathbf{J}^{(var)}$ is independent of N. Thus in the limit of a large number of photons, we expect $\mathbf{J}^{(mean)}$ will make the dominant contribution to $\mathbf{J}$. We therefore neglect the information in the variance and approximate $\mathbf{J} \approx \mathbf{J}^{(mean)}$.

We next invert the Fisher information,

$$\left[\mathbf{J}^{(mean)}\right]^{-1}_{\mathbf{k},\mathbf{k}'} = \frac{1}{N^2}\frac{1}{h(\mathbf{k})}\mathbf{C}_{\mathbf{k},\mathbf{k}'}\frac{1}{h^*(\mathbf{k}')} = \frac{1}{N}\frac{1}{h(\mathbf{k})}h(\mathbf{k}-\mathbf{k}')I(\mathbf{k}-\mathbf{k}')\frac{1}{h^*(\mathbf{k}')}. \tag{37}$$

The fact that the PSF and the scene are both normalized, $h(\mathbf{k}=0) = I(\mathbf{k}=0) = 1$, implies that

$$F(\mathbf{k}) = \frac{1}{\left[\mathbf{J}^{(mean)}\right]^{-1}_{\mathbf{k},\mathbf{k}}} = N|h(\mathbf{k})|^2. \tag{38}$$

This result agrees with the previous exact calculation for a generic scene, Eq. 23, which was derived using a different model for the class of images and assuming $n = 1$ detected photons per emitter.

C. Numerical calculation of ITF for scenes with a discrete set of simultaneously active emitters

To study the dependence of the ITF on the density of simultaneously activated emitters, ρ_a, we numerically calculated the Fisher information matrix, $\mathbf{J}_{ij}$, for simulated 2D scenes comprising

a set of discrete emitters (Fig. 3 in the main text). Each simulated scene had $M = 150$ emitters of equal brightness within a square field of view of area $A = M/\rho_a$. Emitter locations were chosen randomly from a uniform distribution over the field of view. We binned the field of view using a fine grid of pixels at locations $\mathbf{z}_k$, whose spacing was much smaller than σ to avoid aliasing. We used Eq. 2 in the main text, with an Airy disk PSF, to calculate the probability distribution, $P(\mathbf{z}_k|\{\mathbf{x}_j\})$, for observing a photon at location $\mathbf{z}_k$ [analogous to Fig. 1(b)].

We next computed the $2M \times 2M$ Fisher information matrix for the parameters $\{q_i, i = 1, \ldots, 2M\} = \{x_1, \ldots, x_M, y_1, \ldots, y_M\}$. We used an analytic formula for the derivative of h to compute the gradient of the probability distribution with respect to each emitter location parameter, $\frac{\partial P(\mathbf{z}_k)}{\partial q_i}$, at each of the pixels, $\mathbf{z}_k$. The Fisher information matrix for the emitter locations [Fig. 1(c)] is then given by

$$\mathbf{J}_{ij} = N \sum_k \Delta\mathbf{z}_k \frac{\partial P(\mathbf{z}_k|\{q_i\})}{\partial q_i} \frac{\partial P(\mathbf{z}_k|\{q_i\})}{\partial q_j} \frac{1}{P(\mathbf{z}_k|\{q_i\})}, \tag{39}$$

where the sum is over all pixels, $\Delta\mathbf{z}_k$ is the pixel area, and N is the total number of detected photons.

To compute the ITF for the simulated scene, we used the change of variable formula, Eq. 32. We averaged over angles in $\mathbf{k}$-space to define the ITF, $1/F(k)$, as a function of the magnitude, k, for each scene. Finally we repeated this procedure for 40 randomly generated scenes at each value of ρ_a to determine the median and standard deviation of $1/F(k)$.

V. CRAMÉR-RAO BOUND FOR BIASED ESTIMATORS AND THE PERFORMANCE OF PRACTICAL SCENE ESTIMATORS

The Cramér-Rao theorem provides a fundamental lower bound on the variance of an unbiased estimator. In this section we consider situations in which there is no unbiased estimator, or in which there is no estimator that achieves the lower bound on error variance.

We first state a generalization of the Cramér-Rao theorem which applies to biased and unbiased estimators [5]:

$$\mathbf{E} - \mathbf{A}^\dagger \mathbf{J}^{-1} \mathbf{A} \geq 0. \tag{40}$$

Here, $\mathbf{A}_{ij} = \frac{\partial \langle \hat{q}_i \rangle}{\partial q_j}$ is a measure of the bias which, for an unbiased estimator, is simply $\mathbf{A}_{ij} = \delta_{ij}$. $\mathbf{A}$ depends on the estimator, and so Eq. 40 does not provide a universal bound on scene estimation for all data processing procedures.

In practice there may be no data processing procedure with variance equal to the information-theoretic limit for all the parameters. If such a procedure does exist, it is called an efficient estimator. In fact, an efficient estimator exists if and only if the derivative of the log likelihood function can be expressed in the following form:

$$\frac{\delta}{\delta I(k)} \log P(\{z_{ij}\}|I) = \int dk' \mathbf{J}_{I(k),I^*(k')} \left[g(\{z_{ij}\}) - I^*(k') \right], \tag{41}$$

where $\mathbf{J}_{I(k),I^*(k')}$ is the Fisher information matrix and g is a function of the observed data [4]. However, for a Gaussian PSF we have (see Eq. 25):

$$\frac{\delta \log P(\{z_{ij}\}, m|B)}{\delta I(k)} = -m\delta(k) + \sum_{j=1}^{m} \frac{G(k;\{z_{ij}\})}{\int dk_j I(k_j) G(k_j;\{z_{ij}\})}. \tag{42}$$

Because the dependence on $I(k)$ in this expression does not have the form of Eq. 41, no efficient scene estimator exists. However, the situation may not be as dire as this result suggests.

In fact, simple estimators, including those that have been used in stochastic localization microscopy experiments [6, 7], are unbiased and efficient over a broad range of spatial frequencies in which the ITF is non-zero. The inefficiency of these estimators only applies to high spatial frequencies, above the cutoff at which the ITF vanishes.

To illustrate with a simple example, consider a 1D scene containing a single emitter ($m = 1$) at location x_0, so that $I(x) = \delta(x - x_0)$ and $I(k) = e^{i2\pi k x_0}$. Assuming the PSF is Gaussian, the maximum-likelihood estimate of x_0 given N observed photons is simply $\hat{x}_0 = \frac{1}{N}\sum_{i=1}^{N} z_i$. In fact, this location estimator is both unbiased, $\langle \hat{x}_0 \rangle = x_0$, and efficient, $\langle (\hat{x}_0 - x_0)^2 \rangle = J^{-1}_{x_0,x_0} = \sigma^2/N$. A simple strategy for estimating the scene using $\hat{x}_0$ is the *point estimator*, $\hat{I}_1(x) \equiv \delta(x - \hat{x}_0)$, or $\hat{I}_1(k) \equiv e^{i2\pi k \hat{x}_0}$. However, this estimate is biased since $\langle \hat{I}_1(k) \rangle = e^{-(2\pi k\sigma)^2/2N} I(k)$. As shown in Fig. S3 (solid curves), the bias grows with k. However, for $k < \sqrt{N}/2\pi\sigma$ the bias is small and the variance is very close to the Cramér-Rao bound, $[J^{-1}]_{k,k} = (2\pi k\sigma)^2/N$ (Fig. S3, black lines).

To avoid the increasing error variance at high spatial frequencies, a *blob estimator* can be used, such as $\hat{I}_2(x) = \sqrt{\frac{N}{2\pi\sigma^2}} e^{-(x-\hat{x}_0)^2 N/2\sigma^2}$, or $\hat{I}_2(k) = e^{-(2\pi k\sigma)^2/2N} e^{i2\pi k\hat{x}_0}$. This is a commonly used estimator for displaying the results of stochastic localization microscopy experiments [6, 7]. This estimator reduces the variance of the high-frequency components of the scene, at the cost of slightly increased bias (Fig. S3, dashed curves).

To produce an unbiased estimator would require an opposite approach, in which the point estimator is sharpened rather than smoothed. In the spatial frequency domain we can define such a *sharp estimator*, $\hat{I}_3(k) = e^{(2\pi k\sigma)^2/2N} e^{i2\pi k\hat{x}_0}$. Although this frequency-domain estimator is unbiased, it blows up at high frequencies and has no inverse Fourier transform; it is therefore not

practical for estimating the scene in real space. Nevertheless, we can compute the variance of this estimator, $\mathrm{var}(\hat{I}_3) = \langle|\hat{I}_3(k) - I(k)|^2\rangle = e^{(2\pi k)^2\sigma^2/N} - 1$. The sharp estimator indeed obeys the CRLB corresponding to Eq. 33, $\mathrm{var}(\hat{I}_3) \geq \frac{(2\pi k\sigma)^2}{N}$, and its variance is very close to the lower bound for frequencies lower than the cutoff, $k < \sqrt{N}/2\pi\sigma$.

All three of the estimators considered above are approximately unbiased and efficient for spatial frequencies in the range $k < \sqrt{N}/(2\pi\sigma)$. Bias and inefficiency become problematic for these estimators only for very high spatial frequencies, $k \gtrsim \sqrt{N}/(2\pi\sigma)$. For these high spatial frequencies, the ITF, $F(k) = N/(2\pi k\sigma)^2$, becomes smaller than unity, meaning that the information about these scene components is negligible.

---

[1] R. J. Ober, S. Ram, and E. S. Ward, Biophysical Journal **86**, 1185 (2004).

[2] J. W. Goodman, *Introduction to Fourier optics* (Roberts and Co., Englewood, Colo., 2005), 3rd ed.

[3] H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, J Opt Soc Am A Opt Image Sci Vis **12**, 834 (1995).

[4] S. M. Kay, *Fundamentals of statistical signal processing*, Prentice Hall signal processing series (Prentice-Hall PTR, Englewood Cliffs, N.J., 1993).

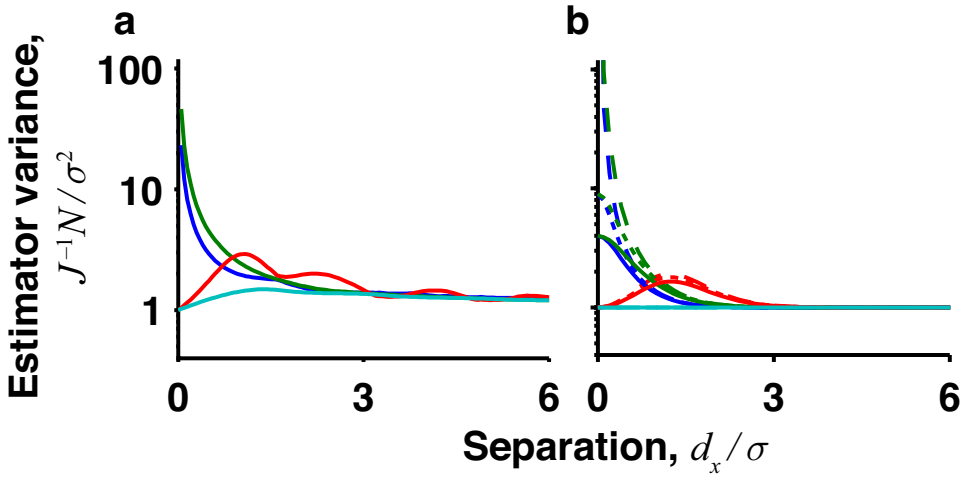
[5] T. M. Cover and J. A. Thomas, *Elements of information theory* (Wiley-Interscience, Hoboken, N.J., 2006), 2nd ed.

[6] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, Science **313**, 1642 (2006).

[7] M. J. Rust, M. Bates, and X. Zhuang, Nature Methods **3**, 793 (2006).

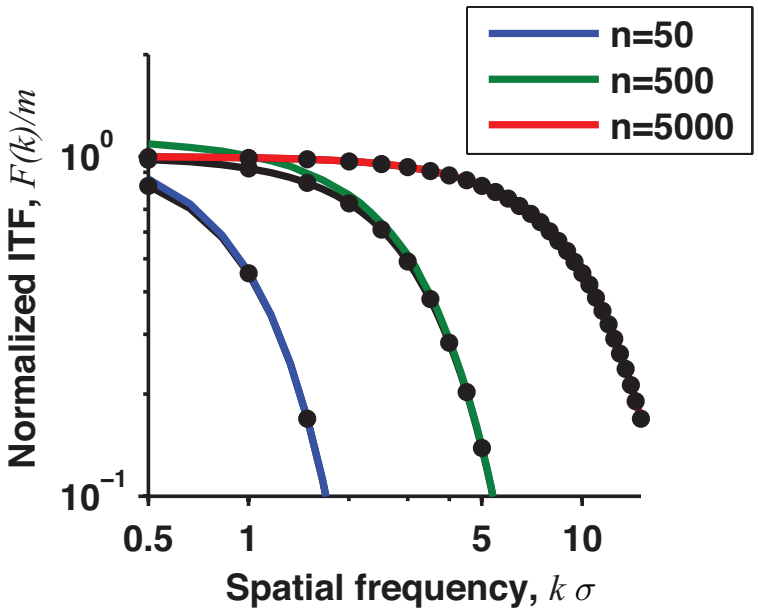
VI. SUPPLEMENTAL FIGURES

Figure S1



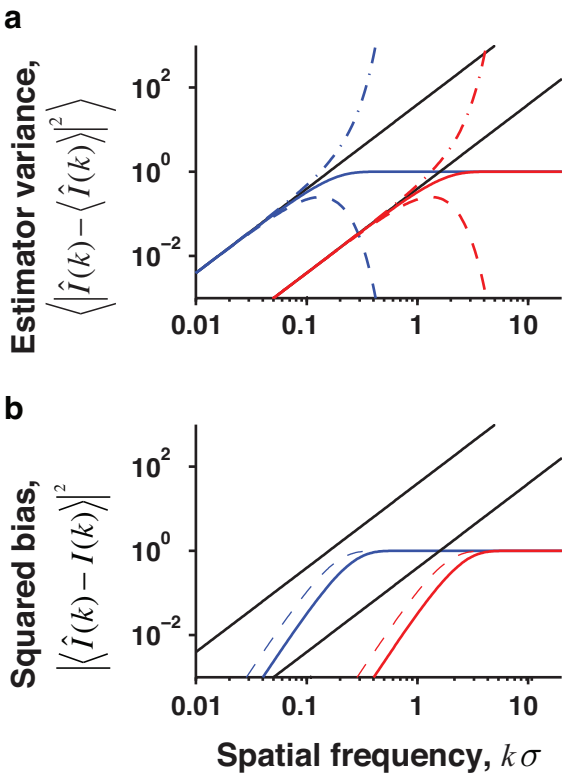
Estimator variance for the localization of two emitters as a function of spatial separation assuming an Airy disk PSF (a) or a Gaussian PSF with unequal emitter brightness (b). In (a) and (b), colors denote the four degrees of freedom illustrated in the inset of Fig. 1(a) in the main text (red, x_0; cyan, y_0; green, d_x; blue, d_y). In (b), curves correspond to brightness ratio 1:1 (dashed), 1:2 (dotted) and 1:3 (solid).

Figure S2



Comparison of ITF for a one-dimensional scene calculated using the assumption that all emitters produce the same number of photons ($n_j = n$, colored curves) or by choosing n_j from a Poisson distribution with mean n (black dots). Simulations used $m = 200$ total emitters. The results are in very close quantitative agreement, particularly for $n \geq 500$.

Figure S3



Variance (a) and squared magnitude of bias (b) of practical scene estimators for a single emitter in 1D. The Cramér-Rao bound is given by the black curve. Colored curves show performance for N =1 (blue) or 100 (red) photons using three estimators: 1. the point estimator, $\hat{I}_1(k) = e^{i2\pi k\hat{x}}$ (solid); 2. the blob estimator, $\hat{I}_2(k) = e^{-(2\pi k\sigma)^2/2N} e^{i2\pi k\hat{x}}$ (dashed); and 3. the sharp estimator, $\hat{I}_3(k) = e^{(2\pi k\sigma)^2/2N} e^{i2\pi k\hat{x}}$ (dot-dashed). Note that only $\hat{I}_3(k)$ is unbiased, and therefore the dot-dashed curve does not appear on the log scale in (b). Within the frequency range $k \lesssim \sqrt{n}/(2\pi\sigma)$, all three estimators have low bias and achieve the minimum variance limit set by the CRLB.