Optimized localization analysis for single-molecule tracking and super-resolution microscopy

Kim I Mortensen[1,4], L Stirling Churchman[2–4], James A Spudich[2] & Henrik Flyvbjerg[1]

We optimally localized isolated fluorescent beads and molecules imaged as diffraction-limited spots, determined the orientation of molecules and present reliable formulas for the precision of various localization methods. Both theory and experimental data showed that unweighted least-squares fitting of a Gaussian squanders one-third of the available information, a popular formula for its precision exaggerates beyond Fisher's information limit, and weighted least-squares may do worse, whereas maximum-likelihood fitting is practically optimal.

Microscopy is limited in resolution by fundamental diffraction effects. To resolve two objects, they must be separated by

$$\Delta x \geq \lambda / 2\mathrm{NA} \qquad (1)$$

(Abbe's law[1]), in which λ is the wavelength of the light used, and NA is the numerical aperture of the microscope objective. However, if an isolated nanoscale emitter (fluorophore, quantum dot or fluorescent bead) is imaged, the photons forming the image are distributed as described by the point spread function (PSF) of the microscope for the chosen source. The center of this image can be determined with much higher precision than its width. This is exploited in single-molecule tracking and localization microscopy, and spatial resolution of a few nanometers, a hundred times less than Abbe's limit (equation 1), are achieved[2–4]. Also, far-field super-resolution techniques now exist such as photoactivated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM), which sequentially isolate each probe in a densely labeled sample to resolve intracellular protein localization patterns to within a few nanometers[5,6]. Additional development and application of super-resolution microscopy is likely to provide insight into cellular processes at both systems and mechanistic levels.

The localization analysis used in these super-resolution schemes is immature, however. Commonly, a two-dimensional (2D) Gaussian (plus a constant background) is least-squares fitted to the distribution of intensity in each spot. This function is not the true PSF, however (and neither is the Airy distribution[7]). So results obtained by fitting a Gaussian have never been compared with the ultimate precision that can be achieved with the true PSF for a given number of photons, although this immutable benchmark is provided by information theory. The information inequality, the Cramér-Rao lower bound[8,9], states that once an isolated probe has been imaged digitally, the precision with which a parameter θ of the PSF can be estimated, is given by

$$\Delta\theta \geq \frac{\sqrt{2}}{\sqrt{N}\sqrt{i(\theta)}} \qquad (2)$$

in which N is the number of photons in the image, and θ can, for example, be a position coordinate. The function $i(\theta)$ is the information content of a single photon and is calculated from the PSF. The factor $2^{1/2}$ is included to account for the excess noise[10] of the commonly used electron-multiplying charge-coupled device (EMCCD) camera. As N is limited by photobleaching of fluorophores, it is of great practical interest to achieve the equality in equation 2. Information theory is clear on this point: maximum likelihood estimation (MLE) with the true PSF does this, and other unbiased estimators can only perform with lower precision.

With this in mind, we analyzed the cases of (i) fluorophores with fixed spatial orientation and (ii) isotropic distributions of fluorophores. We assumed perfect imaging conditions and that the imaged probes were in focus. We showed that theoretical PSFs, derived under these ideal conditions, provided accurate descriptions of measured PSFs and used them in MLE to show that in practice one achieves the lower bound of equation 2 for all parameters of interest. In case i, we optimally estimated location and orientation of probes in focused images. In case ii, we optimally estimated probe locations and compared our precision of localization to those of often used estimators.

First, we considered a fluorescent molecule with fixed spatial orientation of its dipole, which we modeled as a point-source dipole emitter. Diffraction of the electromagnetic dipole field results in a PSF that depends on the dipole's orientation in space[11] and is in general asymmetric (**Fig. 1a–d**). The PSF for a dipole oriented with azimuthal angle α and polar angle β is

$$\begin{aligned} p(x',y') \propto \sin^2\beta I_{\parallel}(p_{\parallel} + \cos(2\phi' - 2\alpha)\Delta p_{\parallel}) \\ + \sin\beta\cos\beta\cos(\phi' - \alpha)I_{\times}p_{\times} + \cos^2\beta I_{\perp}p_{\perp} \end{aligned} \qquad (3)$$

in which $(x', y') = \rho'(\cos\phi', \sin\phi')$ are Cartesian and polar coordinates of a point in the image plane relative to the dipole's location in that plane and all dependencies on angles are explicit (**Supplementary Note 1**). The component $p_{\perp}$ is the PSF of a

[1]Department of Micro- and Nanotechnology, Technical University of Denmark, Kongens Lyngby, Denmark. [2]Department of Biochemistry and [3]Department of Physics, Stanford University School of Medicine, Stanford, California, USA. [4]These authors contributed equally to this work. Correspondence should be addressed to H.F. (henrik.flyvbjerg@nanotech.dtu.dk).

RECEIVED 14 OCTOBER 2009; ACCEPTED 22 FEBRUARY 2010; PUBLISHED ONLINE 4 APRIL 2010; DOI:10.1038/NMETH.1447

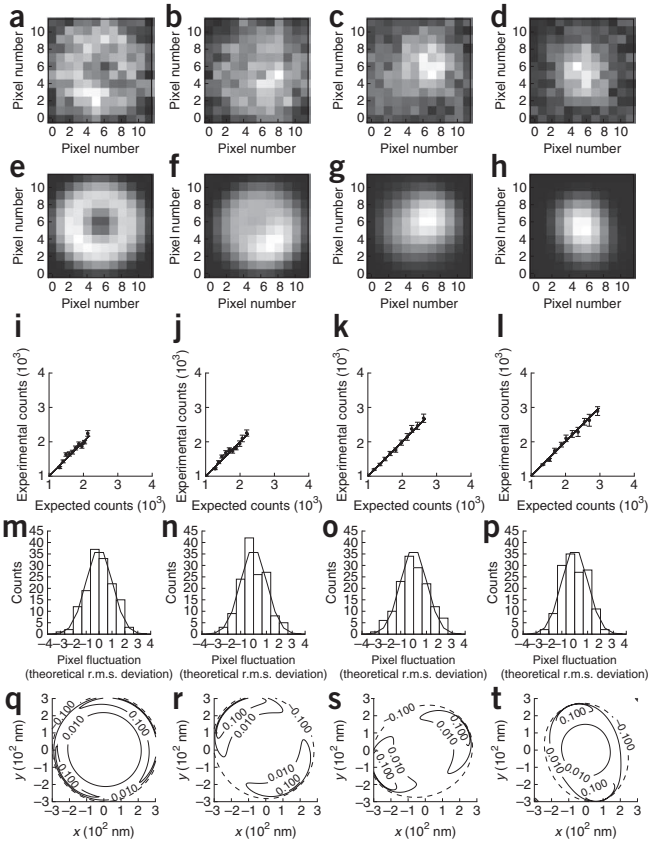

Figure 1 | Point spread functions for four fixed fluorophores with different spatial orientations. (**a**–**d**) TIRF microscopy images of four different fixed rhodamine molecules (Online Methods). (**e**–**h**) Theoretical images with parameter values obtained by applying MLEwT to fixed probe images in **a**–**d**, respectively, and using rhodamine's peak emission wavelength of 578 nm. We found orientations in polar angles of 0.28, 0.60, 0.89 and 1.18 radians, respectively, for **e**–**h** and in azimuthal angles of 0.68, 2.23, 3.61 and 5.03. (**i**–**l**) Measured signal values compared to expected values. We binned the expected signals and associated pixels with a bin if the expected signal fell in it. For each bin, the mean experimental signal is plotted against the expected signal. Error bars indicate the theoretical s.e.m. with which data should scatter about the straight line through the origin with unit slope. (**m**–**p**) Histograms of pixel fluctuations around their expected values. For each pixel in the experimental images in **a**–**d**, the fluctuation is the measured signal value minus its expected value, scaled by its theoretical r.m.s. deviation. For sufficiently large expected pixel signals, theory predicts the standard normal distribution (solid line) for the fluctuations. (**q**–**t**) Comparisons of the theoretical PSF to its analytical approximation. At (0,0) the two functions coincide. For each fit in **e**–**h**, we show the deviation of the analytical approximation from the theoretical PSF as a contour plot. Contour lines are plotted for deviations of 0.01 (1%), 0.1 (10%) (both solid lines) and –0.1 (–10%) (dashed lines).

fixed dipole orthogonal to the image plane, and $p_{||}$ is the PSF of an isotropic distribution of dipoles parallel to the image plane, with all of these dipoles located at the same position. These two PSFs and the functions $\Delta p_{||}$ and $p_{\times}$ depend only on the distance ρ' to the location of the molecule in the image plane. All components except $\Delta p_{||}$ are normalized by virtue of the definitions of $I_{||}$ and $I_{\times}$ (**Supplementary Note 1**). We evaluated these functions as numerical integrals (**Supplementary Software 1**) because no closed exact expressions exist for them. To evaluate these functions faster, we also derived accurate analytical approximations to them (**Supplementary Note 1**). We used these in equation 3 to fit simultaneously the center coordinates of the dipole probe and its angles, the total photon number in the spot and a background level to experimental images (**Fig. 1a**–**h**). We used MLE to fit and refer to this estimator as MLE with the theoretical PSF (MLEwT), with the accurate analytical approximation understood. Thus estimates of the fluorophore's location and spatial orientation were obtained simultaneously and directly from focused images, unlike existing methods[12–14]. The PSF in equation 3 accurately describes the data to which we fitted it with MLE (**Fig. 1e**–**p**), and the analytical approximation was accurate compared to the full evaluation of equation 3 (**Fig. 1q**–**t**).

To test the performance of MLEwT, we photographed single rhodamine fluorophores at 1 Hz, generating a time-lapse movie. Rhodamine is bifunctional, and we used it to cross-link two cysteines engineered into chicken calmodulin[12]. This kept the flurophore's orientation fairly fixed. The cross-linked calmodulins adsorbed nonspecifically to the coverslip. For analysis, we chose seven fluorophores that were sufficiently fixed in location and orientation during the imaging period. In each image of the time-lapse movie, we estimated the probe's position and orientation using MLEwT. The fluctuations of the polar and azimutal angles around their respective mean values agreed with the size of their theoretical error bars (**Fig. 2a**,**b**). The compounded statistics of these fluctuations demonstrated full agreement between experiment and theory (**Fig. 2c**). We compared fluctuations in distances between all pairs of probes to the r.m.s. deviations they should have according to theory, if caused only by shot noise (**Fig. 2d**). Because each probe contributed to several distance estimates, the latter were not fully independent statistically, but this 'oversampling' of data merely reduced statistical noise in the histogram (**Fig. 2d**). We conclude that MLEwT estimates positions and angles with the ultimate precision possible according to Fisher's information limit, which we calculated using the full evaluation of equation 3 to ensure a rigid test. Note in particular that precision was not compromised discernibly by our analytical approximation to the theoretical PSF.

Next, we considered the case of a point-size isotropic distribution of dipoles. If the exciting light is isotropic, the PSF of such a distribution is an isotropic superposition of PSFs for dipoles with fixed orientations, that is,

$$p(x', y') = N_{||} p_{||} + N_{\perp} p_{\perp} \qquad (4)$$

in which $N_{||}$ and $N_{\perp}$ are normalization constants (**Supplementary Note 1**). The same PSF results if the exciting light is polarized, provided thermal motion can rotate the dipoles freely and far between excitation and emission. A single freely and quickly rotating dipole has the same PSF. However, if the light source is polarized, and the dipoles are not free to rotate, then even an isotropic distribution of dipoles, as in a fluorescent bead, will have an asymmetric PSF (**Supplementary Note 1**). Only if the exciting light is incident at the critical angle or along the coverslip, does the maximum of the PSF in this case coincide with the location of its dipole source (**Supplementary Fig. 1**).

We fitted such an asymmetric theoretical PSF to the experimental distribution from a total internal reflection fluorescence (TIRF)-illuminated 40-nm fluorescent bead (**Fig. 3a**,**b**). We observed similar asymmetric distributions owing to the polarized

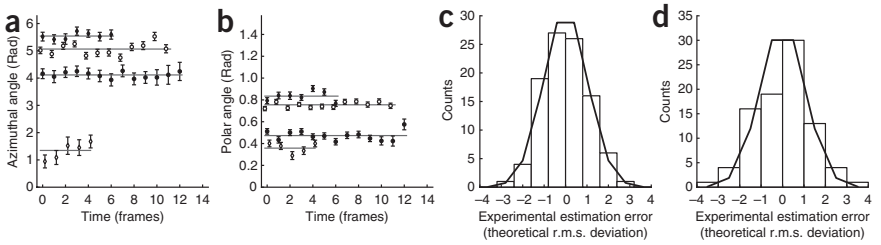

Figure 2 | Demonstration of MLEwT on fixed fluorophores. (**a**,**b**) Time series of repeatedly estimated azimuthal (**a**) and polar (**b**) angles of four fixed fluorophores (data for each marked by a different symbol) using MLEwT. Error bars indicate the theoretical r.m.s. deviation of each estimate, as caused by shot noise only. Lines represent the weighted mean value of the angle. (**c**) Histogram of fluctuations about the mean values of azimuthal and polar angles in seven time series consisting of 68 estimates analyzed using MLEwT. Each fluctuation was rescaled using the theoretical covariance matrix for it. Theory dictates a normal distribution with variance 1 (solid line). (**d**) Histogram of fluctuations about the mean values in time series of distance estimates with MLEwT. For each of the seven probes used in **c**, time series of Euclidean distances between the probe and all other probes fluorescing simultaneously were calculated. Fluctuations are given in units of their theoretical r.m.s. deviation, found by assuming that shot-noise is the only source of statistical fluctuations. If this assumption is correct, theory dictates the normal distribution with unit variance (solid line).

illumination of TIRF using single, freely rotating Cy3 fluorophores (data not shown). We used left-handed circular polarization of the initial excitation light and thus were left with five fitting parameters: the x and y coordinates of the bead's center, the expected number of photons emitted by the bead, a constant background and the super-critical angle of the incident light. The width of the PSF was not fitted directly but followed from the emission wavelength, the known properties of the optics and the characteristics of the incident light. We approximated the PSF as we did for a fixed dipole (**Supplementary Software 2**). The near-perfect agreement between theory and data (**Fig. 3c**,**d**) ensures that in practice the information limit is reached by MLE with this PSF (MLEwT).

We compared MLEwT to two commonly used[4–6,15] estimators, the so-called Gaussian mask estimator (GME) and weighted (or full) least-squares Gaussian fit (WLS)[16,17]. GME uses least-squares fitting with constant weights of a 2D Gaussian plus a constant background. Constant weights are incorrect weights because photon counts in any pixel follow a Poisson distribution. Consequently, some researchers use WLS, which approximates the weight of each pixel with 1/(the experimental count in the pixel). If pixels with low expected counts occur, however, some of them will have actual counts near zero, yielding artificially large weights that throw the estimate off target, unless one patches WLS by adding counts *ad hoc* in such pixels. Fortunately, there is an easy alternative to GME and WLS: a maximum likelihood fit of a 2D Gaussian plus a constant to the data (MLEwG) (**Supplementary Software 3**). MLE weighs data correctly and is the optimal fitting procedure when the fitted function describes the data and one has a large total number of counts (here, photons), as we did.

A 2D Gaussian plus constant does approximate the experimentally measured PSF very well in a properly chosen region (**Supplementary Fig. 1**). But for light incident at angles above the critical angle, the center of the theoretical PSF differs from its maximum. The 2D Gaussian does not capture this feature, which results in a small bias in its estimates of center coordinates. One is rarely interested in absolute coordinates, however, and owing to identical illumination of all probes, the asymmetry of the PSF of a bead is the same anywhere in the field of view, so distance estimates using 2D Gaussians are unbiased, that is, accurate. As for precision of estimated distances, GME has finite variance because its unweighted least-squares fit effectively ignores the photon counts in the $(1/\rho')^3$ power-law tail of the experimental data (**Supplementary Note 1**). MLEwG is mathematically identical to the centroid and hence would locate with infinite variance (**Supplementary Note 1**) if no background was assumed. But if we model the shoulders of the theoretical PSF as a constant background and truncate the fitted part of the image at those shoulders, then MLEwG achieves the precision of Fisher's information limit because a Gaussian plus a constant approximates the theoretical PSF almost perfectly there (**Supplementary Fig. 1**).

To test the performance of estimators, we melted 40-nm fluorescent beads onto a coverslip and photographed them 500 times at 10 Hz. In each image of this time-lapse movie, we localized the same beads with MLEwT, MLEwG, GME and WLS (**Supplementary Note 1**). This revealed nonconstant drift between camera and coverslip. Distances between beads were unaffected by drift but showed thermal motion (**Fig. 4a–d**). The radius of each bead was 20 nm, so the bead center wiggled measurably over the spot it

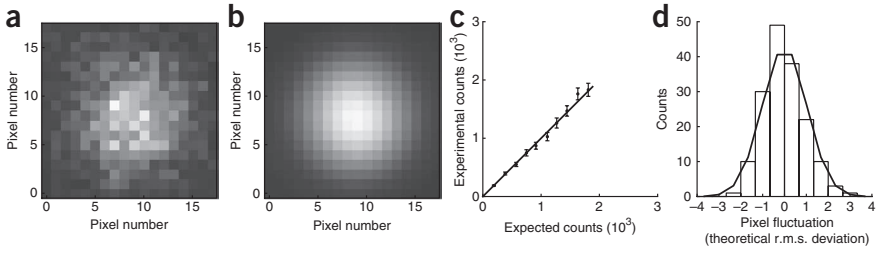

Figure 3 | The point spread function of a 40-nm fluorescent bead. (**a**) TIRF microscopy image of a 40 nm fluorescent bead (Online Methods). (**b**) Theoretical image with parameter values obtained by applying MLEwT to the experimental image in **a**. We assumed circular polarization of the incident light. (**c**) Measured signal values compared to expected values. We binned the expected signals and associated pixels with a bin having expected pixel values within it. For each bin, the mean experimental signal was plotted against the expected signal. The error bars indicate the theoretically predicted s.e.m. with which an experimental mean value is allowed to scatter about the straight line drawn through the origin with unit slope. (**d**) Histogram of pixel fluctuations around their expected values. For each pixel in the experimental image in **a**, the fluctuation is the measured signal value minus its expected value, scaled by its theoretical r.m.s. deviation. For a sufficiently large expected pixel signal, the actual signal value is approximately normally distributed. For such fluctuations, theory dictates the normal distribution with variance 1 (solid line).

was melted onto[18], but slowly, damped by the high internal friction of its polystyrene material. This thermal motion gave rise to excess power at the lowest frequencies of the power spectrum of measured distances (**Fig. 4e**). The simplest possible model for this thermal motion around a fixed position assumes a linear restoring force of the center's coordinate to its average value. As inertia is negligible, the motion then has a characteristic Lorentzian power spectrum[19] to which the localization error resulting from shot noise adds its white-noise spectrum. The constant power of the latter stands out at large frequencies, where the Lorentzian has vanished (**Fig. 4e**) and gives directly twice the time-averaged localization variance, which we observed to be the same for MLEwT and MLEwG, a factor 1.5 larger for GME and 1.7-fold larger for WLS. These results agree perfectly with two theoretical predictions given in equations 5 and 6 below. Alternatively, thermal noise can be removed by high-pass filtering the time series of bead-bead separations. When we did this, we were left with (correlated) fluctuations in bead-bead separations that originated solely in shot noise. We compared these errors to the expected errors and found excellent agreement between experiment and theory for the three reliable estimators (**Fig. 4f–i**). This demonstrated that the theoretically predicted advantage of MLEwT and MLEwG was realized experimentally. Furthermore, for the data presented here, we found that the light is incident close to the critical angle, so coordinates determined by fitting 2D Gaussians should be nearly unbiased. We confirmed this experimentally by comparing this with coordinates estimated with MLEwT (data not shown).

The theoretical r.m.s. deviations of estimators with which experimental data were compared (**Fig. 4**) are the square roots of twice the following expressions for the variance of our estimate μ_x for the x coordinate of the center of the fluorescent bead (**Supplementary Note 1**). The doubling accounts for the 'excess noise' of the electron multiplication process of the EMCCD (**Supplementary Note 1**).

$$\mathrm{Variance}(\mu_x) = \frac{\sigma_a^2}{N}\left(1 + \int_0^1 \frac{\ln t}{1 + Na^2 t/2\pi\sigma_a^2 b^2}\,dt\right)^{-1} \quad (5)$$

$$\mathrm{Variance}(\mu_x) = \frac{\sigma_a^2}{N}\left(\frac{16}{9} + \frac{8\pi\sigma_a^2 b^2}{Na^2}\right) \quad (6)$$

In both expressions $\sigma_a{}^2 = \sigma^2 + a^2/12$, making them correct to order $a^2/(12\sigma^2)$ in the pixel area a^2, which is an excellent approximation for $a \le \sigma$. Both expressions treat the assumed background of b^2 expected photons per pixel exactly and replace equation 17 in reference 16, which underestimates error bars systematically and to such an extent that it violates the information limit when applied to the experimental data discussed here (**Fig. 4e** and **Supplementary Fig. 2**). When pixel counts are sufficiently high to stabilize WLS, its variance is also given by equation 5 (**Supplementary Note 1**). The factor 16/9 in equation 6 distinguishes it from the widely used error formula given in equation 17 in reference 16 and explains why "30% excess error" was observed in reference 16, when true errors of GME applied to computer simulated Gaussian-distributed 'photons' were compared to the errors predicted by that equation 17 (ref. 16): $(16/9)^{1/2} = 133\%$. Note furthermore that GME, MLEwG and WLS when WLS works well, all underestimate the number N of photons in a spot to ~60% of its true value (**Supplementary Fig. 2**) because they all treat shoulders and tails in the measured PSF as background. The reason MLEwG can ignore 40% of the photons recorded by MLEwT yet have the same localization precision is because it ignores photons in the slowly varying tail, which contribute negligibly to localization precision. Note finally that computer algorithms for

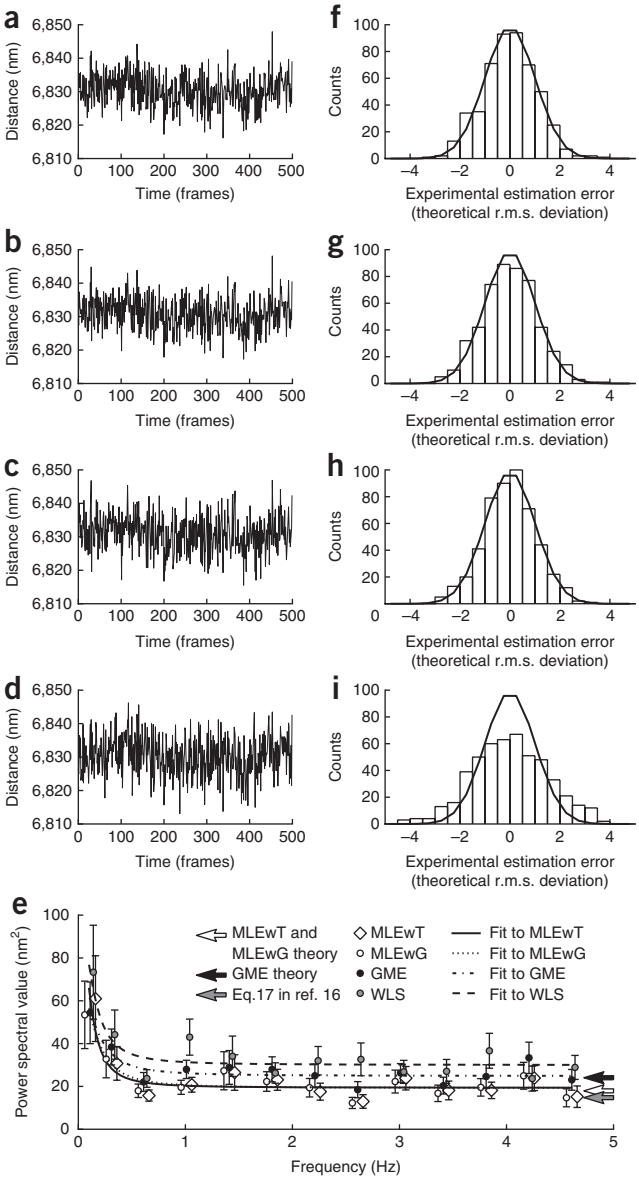

Figure 4 | Comparing four estimators applied to the same data set. (**a–d**) Time series of repeatedly measured distances, between two fluorescent beads melted onto a coverslip, obtained with MLEwT (**a**), MLEwG (**b**), GME (**c**) and WLS (**d**). (**e**) Power spectra of the four time series. The plateau value at larger frequencies equals twice the variance of the point-source localization scheme. Curves are fits of a Lorentzian plus a constant to experimental power spectra. Arrows indicate plateau values according to theory. The white arrow (MLEwT and MLEwG, undistinguishable) marks the information limit: all unbiased estimators have variances larger than or equal to this limit. (**f–i**) Histograms of fluctuations about the mean value in time series of distance estimates with MLEwT (**f**), MLEwG (**g**), GME (**h**) and WLS, rescaled using equation 5 (**i**). Each fluctuation was divided by the theoretical r.m.s. deviation for it. Consequently, theory dictates a normal distribution with variance 1 (solid line). The discrepancy between the histogram and theory in **i** is due to low experimental values in some pixels.

least-squares fitting typically converge faster than algorithms for MLE. Consequently, when computational speed is an issue, one can use least-squares fitting for a speedy near-optimal fit and then use that fit as starting point for a truly optimal fit with MLE.

The simple and optimal analysis provided by MLEwG combined with equation 5 should satisfy a practical need for precision in localization of isotropic probes. Where it does not satisfy this need, for counting photons or for analyzing fluorophore molecules with fixed or time-resolved dipole orientation or for absolute localization of TIRF-illuminated fluorescent beads, one can use MLEwT, optimally with our analytical approximate PSF, and with the assurance that doing better is impossible because this unbiased estimator achieves the information limit.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank S.M. Block, W.E. Moerner and R.S. Rock for discussions; Z.D. Bryant for allowing us to use his microscope for some of the data collection and M.W. Elting and J.M. Sung for assisting us. This work was supported by the European Union (FP7-HEALTH-F4-2008-201418, Revolutionary Approaches and Devices for Nucleic Acid Analysis to H.F.), by the US National Institutes of Health (GM33289 to L.S.C. and J.A.S.), by the Human Frontier Science Program (GP0054/2009-C to J.A.S. and H.F.) and the Damon Runyon Cancer Research Foundation (DRG-1997-08 to L.S.C.).

AUTHOR CONTRIBUTIONS

H.F., K.I.M. and L.S.C. designed research; K.I.M. and H.F. performed the theoretical calculations and analyzed data; J.A.S. supervised the experiments; L.S.C. conducted experiments; K.I.M. did numerical simulations; H.F., K.I.M., L.S.C. and J.A.S. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Born, M. & Wolf, E. *Principles of Optics* (Cambridge University Press, New York, 1999).
2. Barak, L.S. & Webb, W.W. *J. Cell Biol.* **95**, 846–852 (1982).
3. Yildiz, A. *et al. Science* **300**, 2061–2065 (2003).
4. Okten, Z., Churchman, L.S., Rock, R.S. & Spudich, J.A. *Nat. Struct. Mol. Biol.* **11**, 884–887 (2004).
5. Betzig, E. *et al. Science* **313**, 1642–1645 (2006).
6. Moerner, W.E. *Proc. Natl. Acad. Sci. USA* **104**, 12596–12602 (2007).
7. Abraham, A.V., Ram, S., Chao, J., Ward, E.S. & Ober, R.J. *Opt. Express* **17**, 23352–23373 (2010).
8. Rao, C.R. *Linear Statistical Inference and Its Applications* (Wiley, New York, New York, 1973).
9. Ober, R.J., Ram, S. & Ward, E.S. *Biophys. J.* **86**, 1185–1200 (2004).
10. Robbins, M.S. & Hadwen, B.J. *IEEE Trans. Electron. Dev.* **50**, 1227–1232 (2003).
11. Enderlein, J., Toprak, E. & Selvin, P. *Opt. Express* **14**, 8111–8120 (2006).
12. Forkey, J.N., Quinlan, M.E., Shaw, M.A., Corrie, J.E.T. & Goldman, Y.E. *Nature* **422**, 399–404 (2003).
13. Toprak, E. *et al. Proc. Natl. Acad. Sci. USA* **103**, 6495–6499 (2006).
14. Aguet, F., Geissbühler, S., Märki, I., Lasser, T. & Unser, M. *Opt. Express* **17**, 6829–6848 (2009).
15. Toprak, E. & Selvin, P.R. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 349–369 (2007).
16. Thompson, R.E., Larson, D.R. & Webb, W.W. *Biophys. J.* **82**, 2775–2783 (2002).
17. Bobroff, N. *Rev. Sci. Instrum.* **57**, 1152–1157 (1986).
18. Carter, A.R. *et al. Appl. Opt.* **46**, 421–427 (2007).
19. Berg-Sørensen, K. & Flyvbjerg, H. *Rev. Sci. Instrum.* **75**, 594–612 (2004).

ONLINE METHODS

Imaging of fluorescent beads. We melted 40-nm fluorescent beads (FluoSpheres, 580/605; Invitrogen) onto a coverslip by loading a 1:10,000 bead dilution into a flow cell made from a coverslip, a microscope slide and double-sided tape. The flow cell was placed on a heat block at 98 °C (coverslip side down) for ~2 min. Water or buffer was added to the flow cell before imaging. The sample was imaged using a home-built objective-type TIRF microscope using a 1.65 NA 100× objective (Olympus) to create evanescent excitation and to collect the emission of the fluorescent probe, and an EMCCD camera (Andor Technology; iXon DV 887 EMCCD) was used for detection[4]. Images were taken with a pixel size of ~28 nm.

Imaging of fixed rhodamine molecules. We used bis-((*N*-iodo acetyl)piperazinyl)sulfonerhodamine (Invitrogen) to cross-link two engineered cysteines (P66C and A73C) in chicken calmodulin, as previously described[11,14]. Briefly, bifunctional rhodamine was introduced in a 1:4 ratio to calmodulin, which had been exchanged into labeling buffer (25 mM phosphate buffer (pH 7.4) 100 mM $NaCl_2$, 1 mM $CaCl_2$). These conditions yielded ~90% monomers whereas other conditions produced more dimers. The reaction was allowed to proceed for 45 min at room temperature (23–27 °C) in the dark. To quench the reaction, DTT was added to a final concentration of 2 mM. Buffer exchange was then performed using Micro Bio-Spin 6 chromatography columns (Bio-Rad) to remove excess dye. Separation of the labeled calmodulin by denaturing PAGE and imaging on a Typhoon scanner (GE Healthcare Biosciences) allowed measurement of the dimer:monomer ratio.

The labeled calmodulin adsorbed nonspecifically to the coverslip and was imaged using a Nikon TIRF microscope (1.49 NA, 100× objective) to create evanescent excitation and to collect the emission of the fluorophore. The emission was imaged with an EMCCD camera (Andor Technology; iXon DV 887 EMCCD) with pixel size of approximately 44 nm. In all data acquisitions, no special effort was made to ensure precise focus. The focus was determined by eye.

Summary of theoretical methods. Detailed derivations and analyses are available in **Supplementary Note 1**. Briefly, we modeled the diffraction-limited image of an isolated in-focus fluorescent probe systematically for two cases: (i) a single dipole-emitter with fixed orientation in space and (ii) an isotropic superposition of such dipoles with homogeneous distribution in space within a sphere with 40 nm diameter, TIRF-illuminated[20]. Building on the known pattern of diffraction in a circular aperture of a monochromatic, in-focus, on-axis fixed dipole emitter[11] (not to be confused with the textbook Airy pattern formed by a monochromatic plane wave with on-axis wave vector), we accounted for finite pixel size, Poisson statistics of photons from source and background, excess noise from the on-chip electron multiplication process of the EMCCD[21], read-out noise of the EMCCD[21] and the emission spectrum of the fluorophore. Read-out noise was negligible for our purposes. The 20-nm radius of the fluorescent sphere and the width of the emission spectrum both resulted in a just discernible broadening of the diffraction-limited image compared to the PSF from a monochromatic point emitter. When we ignored this broadening, we found a negligible effect on the precision with which the fluorescent source (either fixed dipole or fluorescent sphere) can be localized as compared with the combined errors owing to photon shot noise and excess noise. We consequently treated these sources as monochromatic and point-like. The theoretical PSF in this approximation was further approximated precisely, using the cumulant expansion, to an analytical expression that reduces the need for numerical integration to once for each experimental setting. This analytical approximation was used as the theoretical PSF for fixed dipoles, and a weighted superposition was used as the theoretical PSF for TIRF-illuminated[20] fluorescent beads. These theoretical PSFs describe the expectation value for experimental data in each pixel as functions of the location of the point source and other parameters describing it. We also calculated the expected variance of experimental data with respect to this expectation value in each pixel to test our theory against experimental data. As all effects accounted for by our theory are well-proven physics, the experimental variance can only exceed or equal our theoretical variance and will exceed it if our modeling ignores discernible effects, by definition of the latter. Consequently, agreement between theoretical variance and experimental variance is strong confirmation that our choices of what to leave in and what to leave out were correct.

Information theory[8] states a limit to the precision with which one can localize an isolated point source in a diffraction-limited image of that source. Information theory also states that MLE with the true PSF achieves this lower bound[8] in the limit of high photon numbers. We calculated the information-theoretical limit to the precision with which one can determine the location and other parameters of a fixed dipole emitter, using its theoretical PSF. We did the same for a TIRF-illuminated sphere using its theoretical PSF. Finally, we assumed a 2D Gaussian plus a constant as PSF and calculated the information-theoretical limit with which its center can be located by fitting a function of the same form to data with MLE (MLEwG). This we did analytically, as a function of photon count, the Gaussian's width, background photon level and pixel size. For comparison, we also calculated analytically how well that center can be determined by fitting the 2D Gaussian plus a constant to data with least-squares method. This amounts to the so-called Gaussian mask estimator scheme of reference 16, and our formula for the expected error replaces the approximate interpolation formula given there. Finally, we demonstrated that the expected error on locations estimated by weighted least-squares fitting of a 2D Gaussian plus a constant to similarly distributed data (WLS)[16] equals the result for MLEwG. All these analytical calculations are made possible by an approximation that converts apparently intractable sums over pixels to doable integrals over the image plane, doable because of the rotational symmetry possessed by the assumed problem, the Gaussion-plus-constant–distributed photons, once pixels are removed. The diffraction-limited image of a TIRF-illuminated fluorescent sphere is not quite rotationally symmetric in the plane, so fitting a symmetric PSF to it, a Gaussian plus a constant, may cause a biased estimate, as vividly illustrated by a simple one-dimensional problem in reference 22. However, because the asymmetry is caused by asymmetric illumination of symmetric objects, we found the bias is the same for all fluorescent spheres in a given image, and hence does not affect their relative positions.

Our theoretical results for expected errors were obtained by the usual analytical propagation of errors by linearization of functions

doi:10.1038/nmeth.1447

around expectation values. The quality of this approximation was tested by finding the real error of each estimator, by applying it to synthetic data produced by Monte Carlo simulation of photons distributed according to the theoretical PSF and processed by the EMCCD. The analytical results did not differ from the Monte Carlo simulation results, thus confirming that linearization for error propagation is an excellent approximation for our realistic photon counts, even the lowest among them.

The distance measurements that we used as experimental tests of the performance of various localization estimators, were all done for distances much larger than the widths of the diffraction-limited images of the probes whose separations we measured. Consequently, we could safely assume that our distance estimates had Gaussian-distributed errors because our localization estimates did. The bias on distance estimates introduced with this assumption, which was derived, described and demonstrated in reference 22, is negligible in our case.

20. Axelrod, D., Burghardt, T.P. & Thompson, N.L. *Annu. Rev. Biophys. Bioeng.* **13**, 247–268 (1984).
21. Ulbrich, M.H. & Isacoff, E.Y. *Nat. Methods* **4**, 319–321 (2007).
22. Churchman, L.S., Flyvbjerg, H. & Spudich, J.A. *Biophys. J.* **90**, 668–671 (2006).