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

Kim I. Mortensen[1,4], **L. Stirling Churchman**[2,3,4], **James A. Spudich**[2], and **Henrik Flyvbjerg**[1]

[1]Department of Micro- and Nanotechnology, Technical University of Denmark - DTU, Building 345E, DK-2800 Kgs. Lyngby, Denmark

[2]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305

[3]Department of Physics, Stanford University School of Medicine, Stanford, CA 94305

Abstract

We optimally localize isolated fluorescent beads and molecules imaged as diffraction-limited spots, determine the orientation of molecules, and present reliable formulae for the precisions of various localization methods. For beads, theory and experimental data both show 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, while maximum likelihood fitting is practically optimal.

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

$$\Delta x \geq \frac{\lambda}{2NA} \quad (\text{Abbe's law}^1). \quad (1)$$

Here 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, 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 resolutions of a few nanometers, a hundred times less than Abbe's limit, are achieved[2,3,4]. Also, far-field super-resolution techniques now exist (PALM, STORM, FPALM, PALMIRA) which sequentially isolate each probe in a densely labeled sample to resolve intracellular protein localization patterns to within a few nanometers[5,6]. Further development and application of this super-resolution microscopy is likely to provide insight into nearly every cellular process at both the systems level and the mechanistic level.

The localization analysis used in these super-resolution schemes is immature, however: Commonly, a 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

$$\Delta\theta \geq \frac{\sqrt{2}}{\sqrt{N}\,\sqrt{i(\theta)}}. \quad (2)$$

Here N is the number of photons in the image, and θ can, e.g., 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-employed EMCCD camera. Since N is limited by photobleaching of fluorophores, it is of great practical interest to achieve the equality in Eq.(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, provide accurate descriptions of measured PSFs and used them in MLE to show that one in practice achieves the lower bound of Eq.(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 estimators popular in the literature.

First we considered a fluorescent molecule with fixed spatial orientation of its dipole, which was 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 (Supplementary Note)

$$p(x', y') \propto \sin^2\beta I_\parallel (p_\parallel + \cos(2\phi' - 2\alpha)\Delta p_\parallel) + \sin\beta\cos\beta\cos(\phi' - \alpha)I_\times p_\times + \cos^2\beta I_\perp p_\perp, \quad (3)$$

where $(x', y') = \rho'(\cos\varphi', \sin\varphi')$ are Cartesian, respectively polar coordinates of a point in the image plane relatively to the dipole's location in that plane, and all dependencies on angles are explicit. The component $p_\perp$ is the PSF of a fixed dipole orthogonal to the image plane, while $p_\parallel$ is the PSF of an isotropic distribution of dipoles parallel to the image plane and all located at the same position. These two PSFs and the functions $p_\parallel$ and $p_\times$ depend only on the distance ρ' to the location of the molecule in the image plane. All components but $p_\parallel$ are normalized by virtue of the definitions of $I_\parallel$ and $I_\times$ (Supplementary Note) We evaluated these functions as numerical integrals (Supplementary Software 1) since no closed exact expressions exist for them. To evaluate these functions faster, we also derived accurate analytical approximations to them (Supplementary Note). These were used in Eq.(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 methods12–14. The PSF in Eq.(3) accurately describes the data to which it was fitted with MLE (Fig. 1e–p) and the analytical approximation was shown accurate compared to the full evaluation of Eq.(3) (Fig. 1q–t).

To demonstrate the performance of MLEwT, we photographed single Rhodamine fluorophores at 1 Hz. This dye is bifunctional and was crosslinked to calmodulin proteins, which had been chemically engineered to have two binding cysteine sites12. This kept the flurophore's orientation fairly fixed. The cross linked calmodulins were adsorbed non-specifically to the coverslip. We chose seven fluorophores, which were sufficiently fixed in location and orientation during the imaging period, for analysis. In each image of this time-lapse movie, we estimated the probe's position and orientation using MLEwT. 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. 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–d). We conclude that MLEwT estimates positions and angles with the ultimate precision possible according to Fisher's information limit, which was calculated using the full evaluation of Eq.(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, i.e.,

$$p(x', y') = N_\parallel p_\parallel + N_\perp p_\perp, \quad (4)$$

where $N_\parallel$ and $N_\perp$ are normalization constants (Supplementary Note). 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 fast rotating dipole has the same PSF. If the light source is polarized, however, 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). 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 TIRF-illuminated 40-nm fluorescent bead (Fig. 3a–b). Similar asymmetric distributions due to the polarized illumination of TIRF were observed using single, freely rotating Cy3 fluorophores (data not shown). We used left-handed circular polarization of the initial excitation light, and so 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[4,5,6,15] estimators, the so-called Gaussian Mask Estimator (GME) and Weighted (or Full) Least-Squares Gaussian Fit (WLS) [16,17]. GME employs least-squares fitting with constant weights of a 2D Gaussian plus a constant background. Constant weights are incorrect weights, since photon counts in any pixel are Poisson distributed. Consequently, some practitioners 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, hence 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). However, 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 due 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, i.e., accurate. As for precision of estimated distances, GME has finite variance because its unweighted least-squares fit effectively ignores the photon counts in the $1/r^3$ power-law tail of the experimental data (Supplementary Note). MLEwG is mathematically identical to the centroid, hence would locate with infinite variance (Supplementary Note), were no

background assumed. If the shoulders of the theoretical PSF are modeled as a constant background, however, and the fitted part of the image is truncated 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 demonstrate the performance of estimators, we melted 40-nm fluorescent beads onto a cover slip 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). This revealed non-constant drift between camera and cover slip. Distances between beads were unaffected by drift, but showed thermal motion (Fig. 4a–d). The radius of each bead was 20 nm, so its center wiggled measurably over the spot it was attached to18. 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. Since inertia is negligible, the motion then has a characteristic Lorentzian power spectrum19 to which the localization error due to 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 is seen to be the same for MLEwT and MLEwG, a factor 1.5 larger for GME, and 1.7 for WLS. These results agree perfectly with two theoretical predictions given in Eqs. (5) and (6) below. Alternatively, thermal noise can be removed by high-pass filtering the time series of bead-bead separations. We did that and were left with (correlated) fluctuations in bead-bead separations that originate solely in shot noise. We compared these errors to the expected errors and demonstrated 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 is 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. This was confirmed experimentally by comparing 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 (Supplementary Note). (The doubling accounts for the "excess noise" of the electron multiplication process of the EMCCD (Supplementary Note).)

$$\text{Var}(\mu_x) = \frac{\sigma_a^2}{N} \left(1 + \int_0^1 \frac{\ln t}{1 + \frac{Na^2}{2\pi\sigma_a^2 b^2} t} dt \right)^{-1} \qquad \text{(MLEwG)} \quad (5)$$

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

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 $a \ll \sigma$. Both expressions treat the assumed background of b^2 expected photons per pixel exactly, and replace Ref.16's Eq.(17), which underestimates error bars systematically (Fig. 4e and Supplementary Fig. 2). When pixel counts are sufficiently high to stabilize WLS, its variance is also given by Eq.(5)

(Supplementary Note). The factor 16/9 in Eq.(6) distinguishes it from the widely used error formula given in Eq.(17) in Ref.16 and explains why "30% excess error" was observed in Ref.16, when true errors of GME applied to computer simulated Gaussian distributed "photons" were compared to the errors predicted by that Eq.(17): $(16/9)^{1/2} = 133\%$. Note furthermore that GME, MLEwG, and WLS when it 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 least-squares fitting typically converge faster than algorithms for MLE. Consequently, when computational speed an issue, one can use least-squares fitting for a speedy near-optimal fit, then use that fit as starting point for a truly optimal fit with MLE.

In summary, given the exponential rise in reports on analysis of isolated fluorescent probes in vitro and in vivo, the simple and optimal analysis provided by MLEwG combined with Eq.(5) should satisfy a practical need for precision in localization of isotropic probes. Where it doesn't, or for counting photons, or for fluorophore molecules with fixed or time-resolved dipole orientation, 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

Imaging of fluorescent beads

We melted 40 nm fluorescent beads (FluoSpheres, 580/605, Invitrogen, F8770) onto a coverslip by loading a 1:10000 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 98C (coverslip-side down) for ~2 minutes. 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, Melville, NY) to create evanescent excitation and to collect the emission of the fluorescent probe and an EMCCD camera was used for detection (Andor Technology, Belfast, Ireland, iXon DV 887 EMCCD)4. Images were taken with a pixel size of ~28 nm.

Imaging of fixed rhodamine molecules

We used bis-((N-iodoacetyl)piperazinyl)sulfonerhodamine (Invitrogen, B-10621) to crosslink two engineered cysteines in chicken calmodulin (P66C, A73C) as previously described11,14. 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 NaCl2, 1 mM CaCl2). These conditions were found to produce ~90% monomers while other conditions produced higher levels of dimers. The reaction was allowed to occur for 45 minutes at room temperature in the dark. To quench the reaction, 2 mM of DTT was added. Buffer exchange was then performed using Micro Bio-Spin 6 chromatography columns (Bio-Rad, Hercules, CA) to remove excess dye. Electrophoresis of the labeled calmodulin

through a denaturing PAGE gel and imaging on a Typhoon scanner (GE Healthcare Biosciences, Pittsburgh, PA) allowed the measurement of the dimer:monomer ratio.

The labeled calmodulin adsorbed non-specifically 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, Belfast, Ireland, iXon DV 887 EMCCD) with pixel size 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 may be found in the Supplementary Note. Briefly, **w**e 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. (ii) An isotropic superposition of such dipoles with homogeneous distribution in space within a sphere with 40 nm diameter and 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 text-book 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 result 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 due 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, in order to test our theory against experimental data. Since 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 Ref. 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 gotten rid of. 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 Ref. 22. However, since 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, hence not affecting their relative positions.

Our theoretical results for expected errors were obtained by the usual analytical propagation of errors by linearization of functions 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 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 use 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 have Gaussian distributed errors because our localization estimates do. The bias on distance-estimates introduced with this assumption, which was derived, described, and demonstrated in Ref. 22, is negligible in our case.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Born, M.; Wolf, E. Principles of Optics. New York: Cambridge University Press; 1999.

2. Barak LS, Webb WW. J. Cell Biol. 1982; 95:846–852. [PubMed: 6296157]

3. Yildiz A, et al. Science. 2003; 300:2061–2065. [PubMed: 12791999]

4. Okten Z, Churchman LS, Rock RS, Spudich JA. Nat. Struct. Mol. Biol. 2004; 11:884–887. [PubMed: 15286724]

5. Betzig E, et al. Science. 2006; 313:1642–1645. [PubMed: 16902090]

6. Moerner WE. Proc. Natl. Acad. Sci. USA. 2007; 104:12596–12602. [PubMed: 17664434]

7. Abraham AV, Ram S, Chao J, Ward ES, Ober RJ. Optics Express. 2010; 17:23352–23373. [PubMed: 20052043]

8. Rao, CR. Linear statistical inference and its applications. New York, NY: Wiley; 1973.

9. Ober RJ, Ram S, Ward ES. Biophys. J. 2004; 86:1185–1200. [PubMed: 14747353]

10. Robbins MS, Hadwen BJ. IEEE Transactions on Electron Devices. 2003; 50:1227–1232.

11. Enderlein J, Toprak E, Selvin P. Optics Express. 2006; 14:8111–8120. [PubMed: 19529183]

12. Forkey JN, Quinlan ME, Shaw MA, Corrie JET, Goldman YE. Nature. 2003; 422:399–404. [PubMed: 12660775]

13. Toprak E, et al. Proc. Natl. Acad. Sci. USA. 2006; 103:6495–6499. [PubMed: 16614073]

14. Aguet F, Geissbühler S, Märki I, Lasser T, Unser M. Optics Express. 2009; 17:6829–6848. [PubMed: 19365511]

15. Toprak E, Selvin PR. Ann. Rev. Biophys. Biomol. Struct. 2007; 36:349–369. [PubMed: 17298239]

16. Thompson RE, Larson DR, Webb WW. Biophys. J. 2002; 82:2775–2783. [PubMed: 11964263]

17. Bobroff N. Rev. Sci. Instrum. 1986; 57:1152–1157.

18. Carter AR, et al. Applied Optics. 2007; 46:421–427. [PubMed: 17228390]

19. Berg-Sørensen K, Flyvbjerg H. Rev. Sci. Instrum. 2004; 75:594–612.

20. Axelrod D, Burghardt TP, Thompson NL. Ann. Rev. Biophys. Bioeng. 1984; 13:247–268. [PubMed: 6378070]

21. Ulbrich MH, Isacoff EY. Nat. Methods. 2007; 4:319–321. [PubMed: 17369835]

22. Churchman LS, Flyvbjerg H, Spudich JA. Biophys. J. 2006; 90:668–671. [PubMed: 16258038]
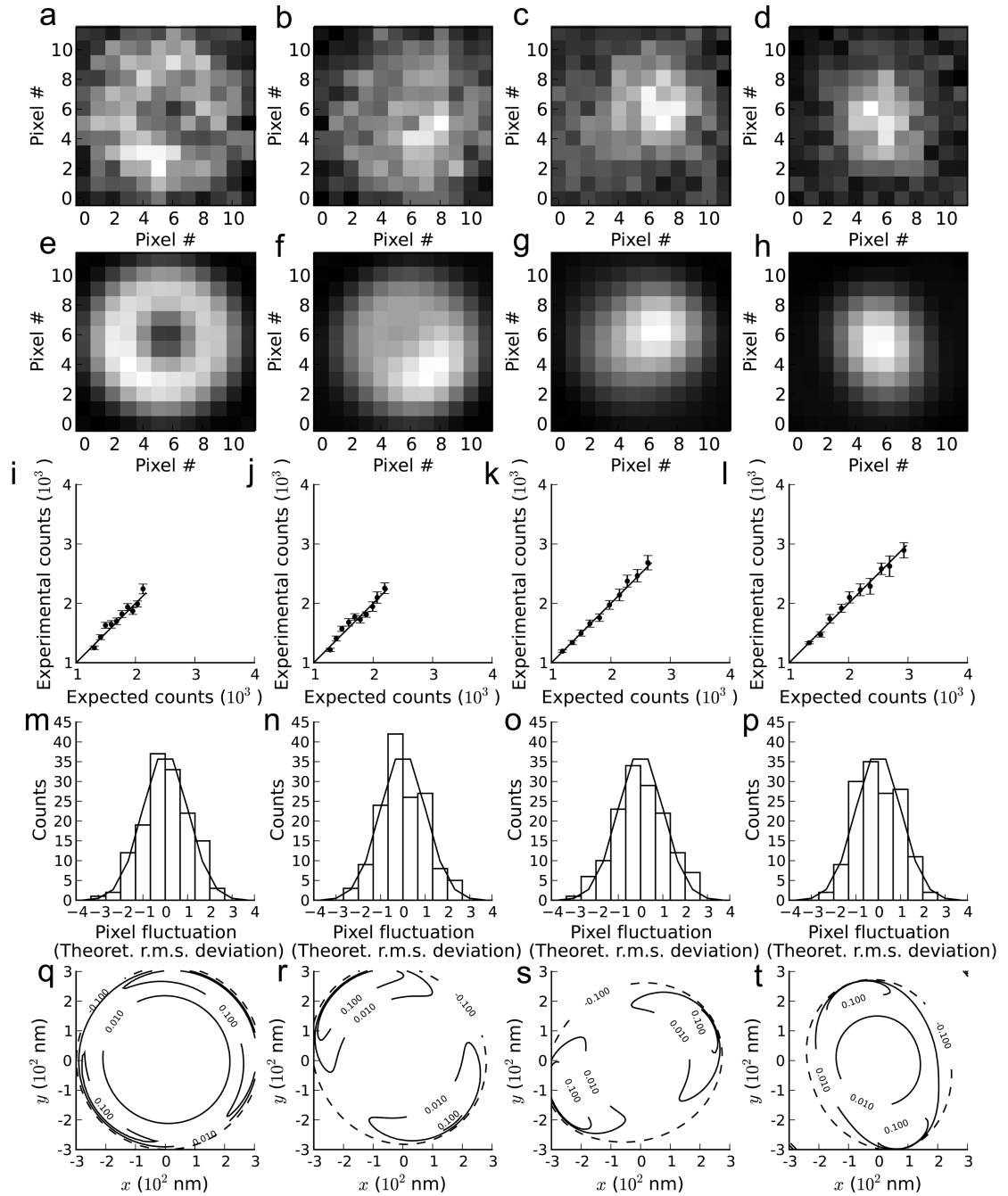
Figure 1.
Point spread functions for four fixed fluorophores with different spatial orientations. (**a–d**): Fixed rhodamine molecules were imaged as described in **Methods**. (**e–h**): MLEwT applied to fixed probes (**a–d**) with 578 nm peak emission wavelength. We found orientations in polar angles, 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 falls within 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 (**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**): Comparing the theoretical PSF to its analytical approximation (Supplementary Data). At (0,0) the two functions coincide. For each fit in (**e–h**), we show the percentile deviation of the analytical approximation from the theoretical PSF as a contour plot. Contour lines are plotted for 0.01, 0.1 (both solid lines), and −0.1 (dashed lines).
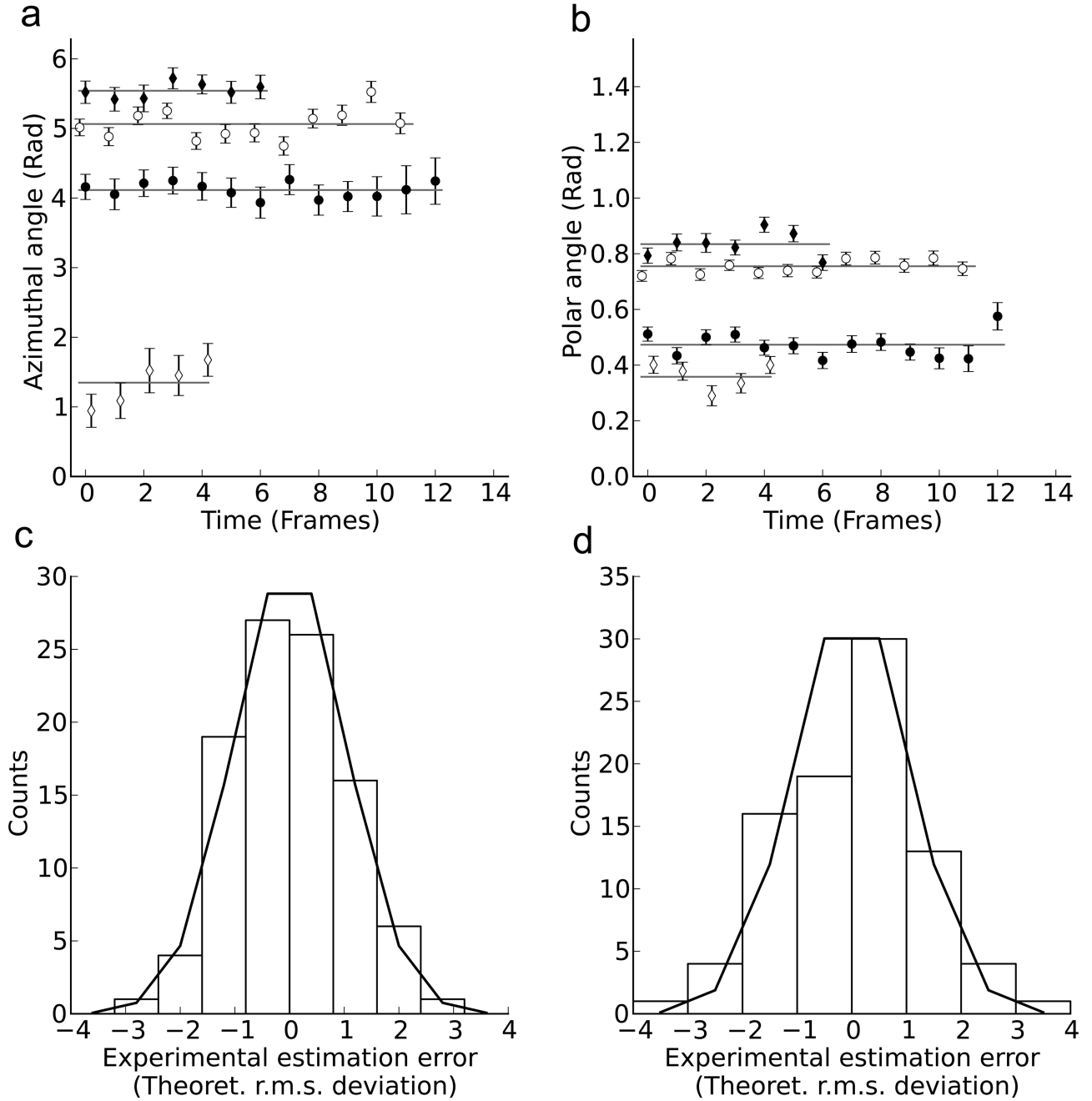
Figure 2.
Demonstration of MLEwT on fixed fluorophores. (**a**): Time-series of repeatedly estimated azimuthal angles of four fixed fluorophores using MLEwT. Error bars indicate the theoretical r.m.s. deviation of each estimate, as caused by shot-noise only. For each time-series, the full lines represent the weighted mean value of the angle. (**b**): Polar angles for the same probes, obtained as in (**a**). (**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 re-scaled using the theoretical covariance matrix for it.

Consequently, theory dictates a normal distribution with variance one, shown with the 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 shot-noise is the only source of statistical fluctuations. If this assumption is correct, theory dictates the normal distribution with unit variance, shown here as a solid line. (Because each probe contributes to several distance estimates, the latter are not fully independent statistically, but this "oversampling" of data merely reduces statistical noise in the shown histogram.)
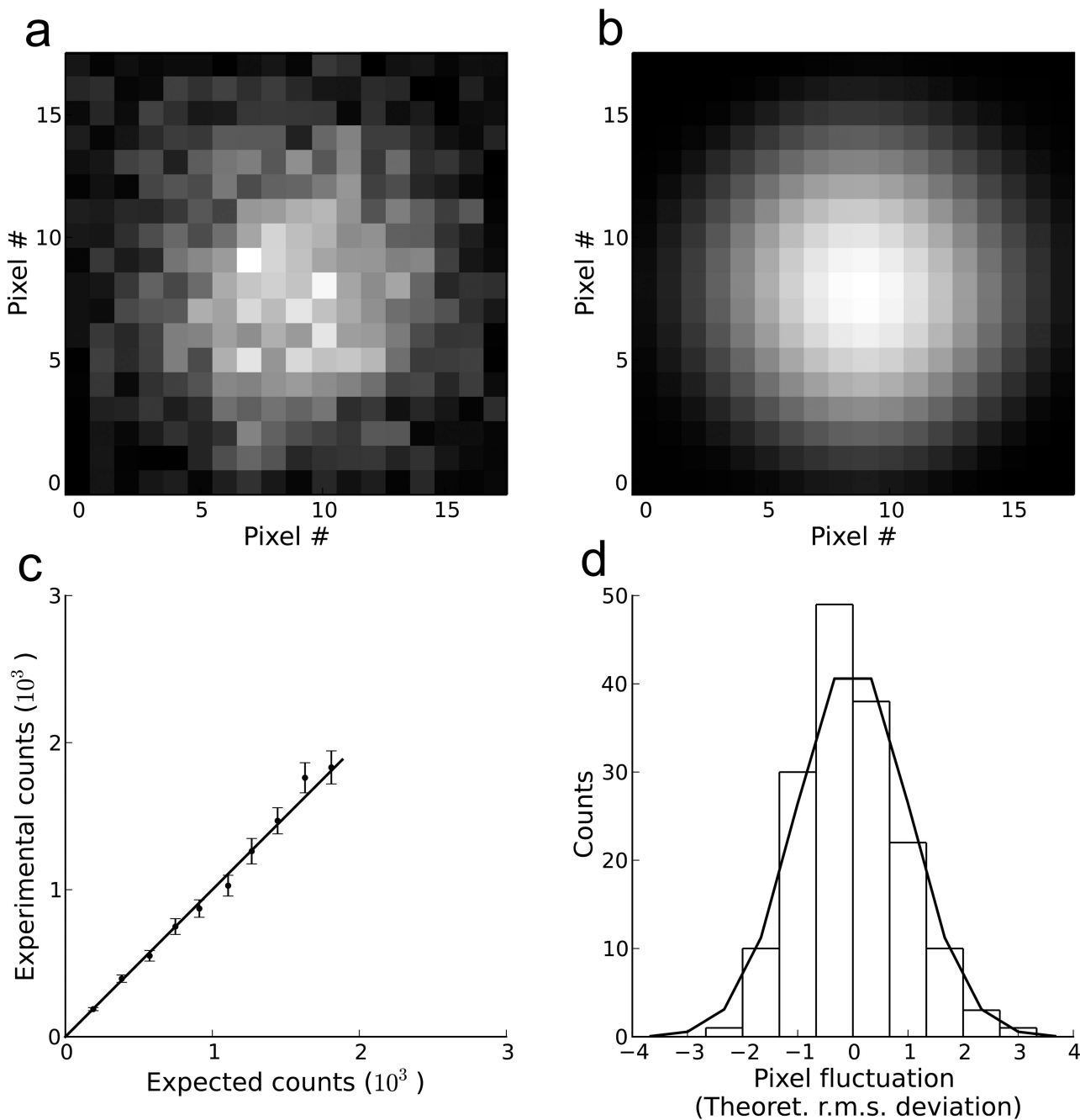
Figure 3.

The point spread function of a 40-nm fluorescent bead. (**a**): 40 nm fluorescent beads were imaged as described in **Methods**. (**b**): Theoretical image: Obtained by applying MLEwT to the experimental image (**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 is 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 (**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, the theory dictates the normal distribution with variance one, shown with the solid line.
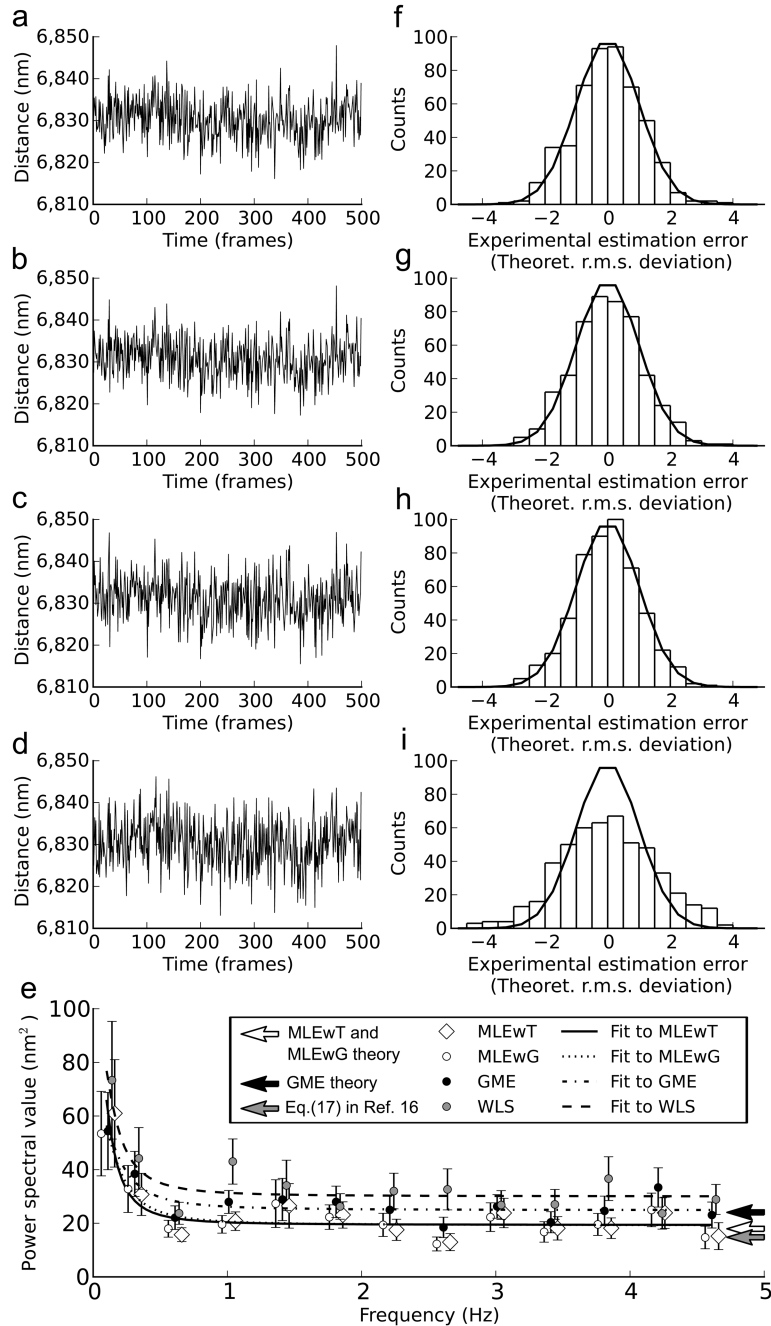
Figure 4.
Comparing four estimators. (**a–d**): Time series of repeatedly measured distance between two fluorescent beads melted onto a cover slip. Notice the larger scatter in (**d**) compared to (**c**) and (**c**) compared to (**a**) and (**b**). (**a**): Distances obtained with MLEwT. (**b**): Distances obtained with MLEwG applied to the very same data. (**c**): Distances obtained with GME. (**d**): Distances obtained with WLS. (**e**): Power spectra of the four time series. The excess power at lowest frequencies may be due to Brownian motion of the beads18. The plateau value at larger frequencies equals twice the variance of the point-source localization scheme.

Curves: Fits of a Lorentzian plus a constant to experimental power spectra. Arrows: Plateau values according to theory (Supplementary Note). Note that the white arrow (MLEwT and MLEwG, undistinguishable) marks the information limit: All unbiased estimators have variances larger than or equal to this limit. Note also that Eq.(17) in Ref. 16 violates the information limit when applied to the experimental data discussed here, hence must be wrong. (**f**): Histogram of fluctuations about the mean value in time-series of distance estimates with MLEwT. Each fluctuation was divided by the theoretical r.m.s. deviation for it. Consequently, theory dictates a normal distribution with variance one, shown with the solid line. (**g**): Same for MLEwG. (**h**): Same for GME. (**i**): Same for WLS, rescaled using Eq.(5). The width of the histogram compared to theory is due to low experimental values in some pixels.