Top-Hat and Asymmetric Gaussian-Based Fitting Functions for Quantifying Directional Single-Molecule Motion

David J. Rowland and Julie S. Biteen*[a]

Single-molecule fluorescence permits super-resolution imaging, but traditional algorithms for localizing these isolated fluorescent emitters assume stationary point light sources. Proposed here are two fitting functions that achieve similar nanometer-scale localization precision as the traditional symmetric Gaussian function, while allowing, and explicitly accounting for, directed motion. The precision of these methods is investigated through Fisher information analysis, simulation and experiments, and the new fitting functions are then used to measure, for the first time, the instantaneous velocity and direction of motion of live bacteria cells. These new methods increase the information content of single-molecule images of fast-moving molecules without sacrificing localization precision, thus permitting slower imaging speeds, and our new fitting functions promise to improve tracking algorithms by calculating velocity and direction during each image acquisition.

1. Introduction

Fluorescence microscopy is traditionally limited to several hundred nanometers by the diffraction of light. However, centroid fitting of isolated molecule fluorescent emission enables the locations of single molecules to be determined with very high precision.[1,2] The ability to accurately locate individual molecules afforded by single-molecule fluorescence (SMF) microscopy can address fundamental questions in biology and chemistry that cannot be approached with bulk measurements or kinetics. Since SMF microscopy can be applied even to mobile molecules, it has been used to directly observe motions as diverse as molecular motors, proteins in live bacterial cells, DNA dynamics, and nanocargo translocation.[3–6] Here, SMF experiments record images of isolated fluorescent emitters, and the position of each emitter is considered to be the center of that emitter's image (point spread function; PSF). This position determination can be achieved by calculating the center-of-mass,[7] by least-squares fitting to a Gaussian function,[8] or by maximum likelihood estimation (MLE).[9] If the emitter is small enough to be considered a point source, if the fluorophore emits isotropically,[10] and if the object's motion is negligible during the imaging frame, then these PSF-fitting methods provide reasonable estimates for the emitter position.

For point light sources that are stationary and in-focus during fluorescent imaging, the resulting PSF is an Airy disk that can be reasonably approximated by a symmetric Gaussian function with standard deviation, σ, where σ depends on the numerical aperture of the microscope objective and the wavelength of light. Thompson et al. reported closed-form expressions for the localization precision of immobile fluorescent molecules fit with this symmetric Gaussian fitting function.[11] Later, the theoretical limits of PSF fitting were determined from a more rigorous and general treatment of these fitting methods and their statistics based on Fisher information theory,[12] and using the true image of a diffracted immobile point light source corrupted by noise.[13] Furthermore, this MLE-based method was implemented as a fast, iterative algorithm.[14] The simple symmetric Gaussian fitting function has been adapted for many applications, including three-dimensional imaging,[15] and Wang and co-workers added precision estimates for localization along the vertical axis by making use of the relation of standard deviation of the molecule image to the molecule's distance from the objective focal plane.[16]

Furthermore, in many situations, even mobile molecules can be described by fits to the symmetric Gaussian function, as a sufficiently fast imaging speed can compensate for molecular motion. However, fast-moving molecules can no longer be considered essentially immobile since fast motion will blur the PSF extensively. Since the accuracy of the position determination depends on the quality of the fitting function, a fitting function that accounts explicitly for molecular motion is desirable. For the case of diffusive motion, no net direction of motion is expected: the molecule PSF will be blurred in all directions, and the convolution of the normal distribution of diffusive motion with the Gaussian-approximated photon distribution function of an immobile molecule results simply in another symmetric Gaussian with a larger standard deviation. This broadened σ has been related to the diffusion coefficient by empirical calibration[17] and by assuming the blurring to be represented by a path distribution function characterized by a normal distribution with a standard deviation that can be related directly to the molecular diffusion coefficient.[18] Error propagation and the results from these earlier works allow the error in the diffusion coefficient estimate to be predicted.

[a] D. J. Rowland, Prof. J. S. Biteen
    Department of Chemistry
    University of Michigan
    Ann Arbor, MI 48109 (USA)
    E-mail: jsbiteen@umich.edu

Though diffusion is ubiquitous, many biological processes are characterized by directed motion. For instance, bacterial chemotaxis and particle flow in a current. Rather than producing isotropic PSF widening, such directional motion causes the image PSF to be blurred in the direction of motion, x. In this case, a fitting function that is derived considering molecular mobility will have the added benefit of providing information about the speed and directionality of the motion. To quantify directional motion, we have developed two methods that explicitly account for motion at a fixed velocity. The first method more exactly models a moving molecule PSF as the convolution of a stationary molecule PSF (approximated by a symmetric Gaussian function) and motion at a fixed speed and direction (represented by a top-hat function). The second method applies the asymmetric Gaussian, which has been previously used for three-dimensional imaging through an astigmatic lens[19]##, to a moving molecule; here the long axis of the asymmetric Gaussian corresponds to the direction of motion. Though Yüce et al. recently treated moving single-molecule fluorophores with a maximum likelihood approach,[20] the two fitting functions presented here have the benefit of being easily incorporated into established data processing routines.

1.1. SErf Fitting Function

The distribution of photons emitted from a fluorophore moving at a constant in-plane velocity during an imaging frame is approximated by the analytical convolution [Eq. (1)] of a top-hat function, h, with a symmetric Gaussian function, g:

$$h(x,y) = \delta(y) \times \begin{cases} 0 : & x < -\frac{D}{2} \\ 1 : & -\frac{D}{2} \quad x > \frac{D}{2} \\ 0 : & x > \frac{D}{2} \end{cases}$$

$$g(x,y) = \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\left(\frac{\sigma'}{a}\right)^2}\right) \tag{1}$$

$$q_{SErf} = h \otimes g$$

Here, D is the displacement (along the x axis), σ' is the PSF standard deviation in nm, and the magnification factor, a, which is dependent on the dimensions of the image sensor and magnification of the microscope, is in the range of 50–200 nm pixel^{-1} for single-molecule fluorescence experiments. The center position of the molecule during the imaging frame is (x_0, y_0). We consider 2D motion restricted to the xy plane, where the axes are free to rotate through an angle, φ, such that the Cartesian coordinates x' and y' become [Eq. (2)]:

$$x = x'\cos(\varphi)\hat{i} - y'\sin(\varphi)\hat{j}; \quad y = x'\sin(\varphi)\hat{i} + y'\cos(\varphi)\hat{j} \tag{2}$$

where $\hat{i}$ and $\hat{j}$ are the Cartesian unit vectors. All quantities x, y, x_0, y_0, D, and (σ'/a) are in units of pixels. Throughout the remainder of this discussion, the variable (σ'/a) is simplified to σ, the standard deviation in units of pixels, i.e. the number of pixels in one standard deviation of the point spread function

(PSF) of the microscope. A higher σ means the PSF is spread across more pixels.

q_{SErf} [Eq. (3)] is the Sum of Error Functions (SErf) function that arises from the convolution in Equation (1):

$$q_{SErf}(x,y) = \frac{N}{D\sigma\sqrt{8\pi}}\exp\left(-\frac{(y-y_0)}{2\sigma^2}\right)\left(\text{erf}\frac{(x-x_0)+D/2}{\sigma\sqrt{2}}\right)$$
$$-\text{erf}\frac{(x-x_0)-D/2}{\sigma\sqrt{2}} + b \tag{3}$$

Equation (3) has been normalized to include the photon count, N, and the function is offset by a background level, b. Figure 1 (a) shows g and h for typical single-molecule imaging conditions. Figure 1 (b), (c) show the results of this convolution
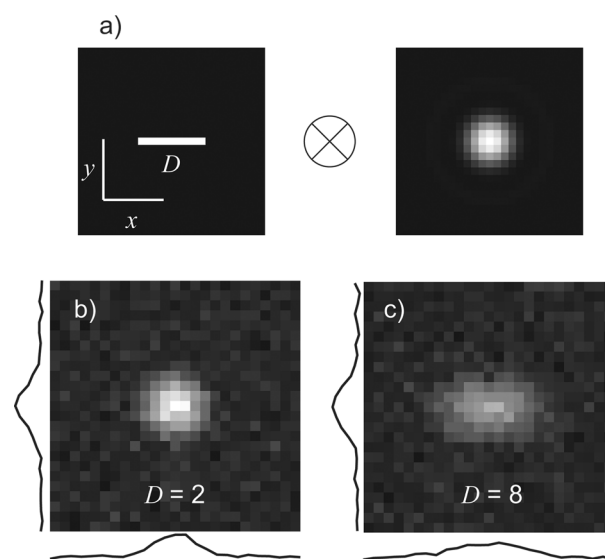

Figure 1. a) Graphical representation of the mathematical convolution of a top-hat function of width D (in-frame displacement length) in the x direction (direction of motion) and a symmetric Gaussian function that gives rise to the SErf function. b,c) Simulated SErf functions for $D=2$ pixels (b) and $D=8$ pixels (c). The integral of the images over dy and dx are shown below and to the left, respectively, of each image. Here, $N=1\times10^5$ photons, $b=200$ photons pixel^{-1}, and $\sigma=2$ pixels.

for $D=2$ pixels and $D=8$ pixels, respectively, with $N=1\times10^5$ photons, $b=200$ photons pixel^{-1}, and $\sigma=2$ pixels. In this paper, noise is modeled as a Poisson distribution with mean and variance equal to b^2.[11] The traces below and to the left of Figure 1 (b), (c) show the integrals of the respective PSF images over dy and dx, respectively. The PSF widening is nearly imperceptible in Figure 1 (b), and becomes more obvious in Figure 1 (c).

1.2. Asymmetric Gaussian Fitting Function

As an alternative to the SErf function introduced in Section 1.1, we consider the asymmetric Gaussian function, q_{AG} [Eq, (4)]:

$$q_{AG}(x,y) = \frac{N}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{(y-y_0)^2}{2\sigma_y^2} - \frac{(x-x_0)^2}{2\sigma_x^2}\right) + b \qquad (4)$$

Here, the standard deviations (in units of pixels) both along the direction of motion and perpendicular to it, σ_x and σ_y, respectively, are considered, and x and y are again free to rotate in the xy plane about an angle φ according to Equation (2). The asymmetric Gaussian function, q_{AG}, becomes the conventional symmetric Gaussian function, q_{SG}, when $\sigma_x = \sigma_y = \sigma$.

Since the aspect ratio (σ_x/σ_y) of a fit to Equation (4) will increase monotonically with increasing in-frame displacement, D, simulated data can be used to calibrate Equation (4). In Section 2.2, the simulated data is fit to the parabola [Eq. (5)]:

$$\frac{\sigma_x}{\sigma_y} = cD^2 + 1 \qquad (5)$$

This parabolic fit of the calibration curve provides a single empirical fit constant, c, which allows for error propagation.

In this paper, we use Fisher information theory[12] to analytically evaluate the precision (Section 2.1), we consider simulated data to numerically determine the precision (Section 2.2), and we fit experimental measurements of scattering from gold nanoparticles moving at a constant velocity to experimentally measure the precision (Section 2.3) for each of the fitting functions, q_{SErf} and q_{AG}, introduced above. We then apply the mobile single-molecule asymmetric Gaussian and SErf functions to fluorescently labeled *Vibrio cholerae* cells swimming in the imaging plane of a microscope in Section 2.4.

2. Results and Discussion

2.1. Analytical Precision

The smallest possible variance, or Cramér-Rao lower bound, of a fitting parameter, θ_i, is equal to the inverse of the Fisher information matrix of the fitting function, $I(\theta)$.[13] Here, for both q_{SErf} and q_{AG}, the 7×7 square matrix [Eq. (6)]:

$$[I(\vec{\theta})]_{i,j} = \int \frac{1}{b^2+q}\frac{\partial q}{\partial\theta_i}\frac{\partial q}{\partial\theta_j}dxdy \qquad (6)$$

In Equation (6), q is the image function q_{SErf} or q_{AG} [Eq. (3) or (4)], here used as a likelihood function. b^2 is the mean of the Poissonian noise corrupting the image, in photons per pixel. Cross-terms of $[I]_{i,j}$ are neglected because symmetry causes them to be near-zero or identically zero. The derivatives of q_{SErf} and q_{AG} are taken with respect to each fit parameter: $\{x,y,D,\sigma,\varphi,N,b\}$ for q_{SErf} and $\{x,y,\sigma_x,\sigma_y,\varphi,N,b\}$ for q_{AG}. For the diagonal elements of $[I]_{i,j}$, each derivative is squared, divided by the original function plus a noise term, (b^2+q), and then integrated over all space per Equation (6). The inverse square root of

this quantity is the standard deviation (RMS error), $\Delta\theta_i$, of the corresponding parameter, θ_i.

Equation (6) is applied to determine the variances $(\Delta x)^2$, $(\Delta y)^2$, and $(\Delta D)^2$ of x, y, and D, respectively, for the asymmetric Gaussian and SErf functions, and the results are given in Table 1. As described in the Experimental Section below, solu-

Table 1. Closed-form analytical solutions to the Fisher information for three parameters of the asymmetric Gaussian and SErf fitting functions. Each solution is given as the sum of the solution in the high-noise regime and the solution in the low-noise regime, and Taylor series expansions are used for some of the SErf results, as described in the text.

	Gaussian	SErf
$(\Delta x)^2$	$\frac{8b^2\pi\sigma_x^3\sigma_y}{N^2} + \frac{\sigma_x^2}{N}$	$\frac{2b^2D^2\pi\sigma^2}{N^2\left(1-\exp\left(-\frac{D^2}{4\sigma^2}\right)\right)} + \frac{24\sigma^2}{N(24\sigma^2-D^2)}$
$(\Delta y)^2$	$\frac{8b^2\pi\sigma_y^3\sigma_x}{N^2} + \frac{\sigma_y^2}{N}$	$\frac{192b^2\pi\sigma^6}{N^2(24\sigma^2-D^2)} + \frac{\sigma^2}{N}$
$(\Delta D)^2$	$\frac{4b^2\pi\sigma_x\left(1+\sigma_x^2\right)}{3cN^2\left(\sigma_x-\sigma_y\right)} + \frac{1+\sigma_x^2}{8cN\sigma_y\left(\sigma_x-\sigma_y\right)}$	$\frac{768b^2\pi\sigma^6}{N^2D^2} + \frac{72\sigma^4}{ND^2}$

tions for the asymmetric Gaussian function are exact except for an initial separation of the integrals into high- and low-noise regimes (a simplification used for all solutions). To find ΔD for the asymmetric Gaussian function, the calibration fit [Eq. (5)] is used in the error propagation according to Equation (8) in the Experimental Section below. Furthermore, for the SErf function, Taylor expansions around $D=0$ were required to determine $(\Delta D)^2$ (high- and low-noise limits), $(\Delta y)^2$ (high-noise limit only), and $(\Delta x)^2$ (low-noise only); those solutions in Table 1 are therefore best when the displacements are small. The variance in the estimation of the displacement does not noticeably differ between numerical and closed-form evaluations of the integrals.

Plotted in Figure 2(a),(b) are the RMS errors (Δx, Δy and ΔD) found by numerical integration of the Fisher information of x, y, and D (solid lines), and the RMS errors from by the closed-form solutions of Table 1 (dotted lines), for the asymmetric Gaussian function [Figure 2(a)] and the SErf function [Figure 2(b)], as a function of number of detected photons, N. These Cramér-Rao lower bounds depend on D, b and σ in addition to N; here, $D = 5.1$ pixels, $b = 200$ photons pixel^{-1}, $\varphi = 0$ and $\sigma = 2$ pixels.

As expected, all the errors decrease monotonically with increasing N, and we find consistently $\Delta D \gg \Delta x \geq \Delta y$. Importantly, Δy is identical (in the high-photon-count regime) to errors in the position estimators, Δx and Δy, for fitting immobile molecules with a symmetric Gaussian.[11] Figure 2(a),(b) show only slight discrepancies in the errors of the two new fitting functions, so we expect the two functions to perform similarly. Good agreement is observed between closed-form and numerical solutions, so one may use the analytical expressions for the error in Table 1 to design and troubleshoot experiments. In Figure 2(b), there is a significant discrepancy between the closed-form and numerical solution for ΔD in the S_{Erf} function;
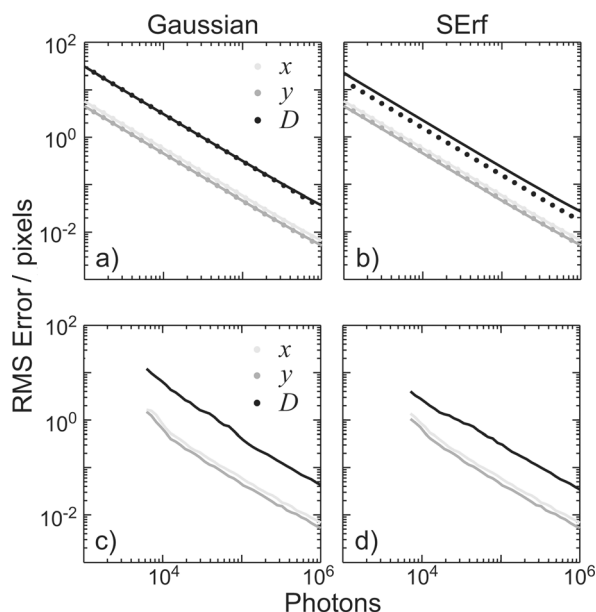
Figure 2. Precision for *x*, *y*, and *D*. a,b) Precision computed using Fisher information theory. Solid lines (——): numerical solutions to the integrals in Equation (6); dotted lines (•••••): approximate closed-form analytical solutions from Table 1. c,d) Precision calculated from variance of fits to 1000 simulated images. Here, $D = 5.1$ pixels, $b = 200$ photons pixel^{-1}, $\varphi = 0$ and $\sigma = 2$ pixels.
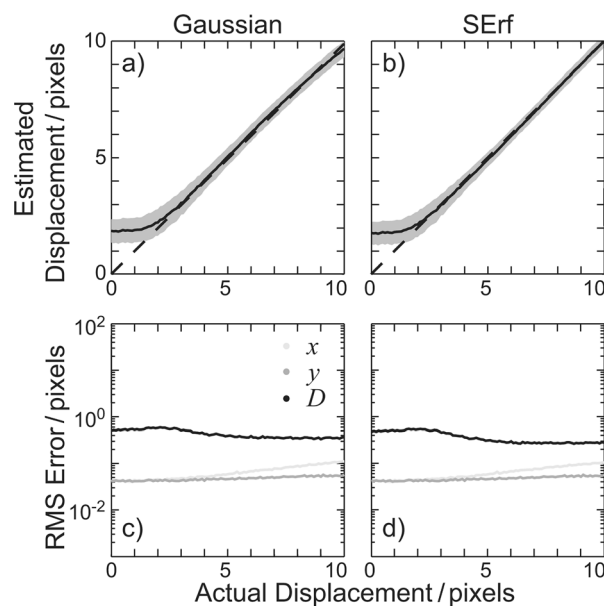
Figure 3. Effect of changes in the actual simulated displacement, *D*, on fit parameters computed based on fits to simulated data. a,b) Fit displacement, D_{fit}, versus actual displacement, *D*. c,d) Precision for *x*, *y*, and *D* computed from fits to simulated data. Here, $N = 1 \times 10^5$ photons, $b = 200$ photons pixel^{-1}, $\varphi = 0$ and $\sigma = 2$ pixels.

this discrepancy arises from the limit of small *D* imposed by the Taylor expansions, and the deviation grows with *D* (Supporting Figure 1). For both fitting functions, as the displacement is increased, Δx increases while Δy stays fairly constant, and ΔD decreases. At the $D = 5.1$ pixels case considered in Figure 2(a), (b), Δx and Δy have already diverged.

Though *D* cannot be determined from a fit to a conventional symmetric Gaussian function, the theoretical localization error for the symmetric Gaussian function can be found by considering the closed-form solutions for Δx and Δy for the asymmetric Gaussian function in Table 1 under the condition $\sigma_x = \sigma_y = \sigma$. The error for the symmetric Gaussian function is thus $(\Delta x)^2 = (\Delta y)^2 = 8b^2\pi\sigma^4/N^2 + \sigma^2/N$, as was found by Thompson et al. in their consideration of the symmetric Gaussian fitting function.[11] The theoretical localization precisions of the symmetric and asymmetric Gaussian functions are therefore identical when measuring stationary sources (Supporting Figure 3).

2.2. Numerical Precision

The analytical treatment in Section 2.1 gives the lower bounds for RMS error in *x*, *y* and *D* based only on the choice of fitting function. To provide an estimate of the RMS errors in the localization and displacement estimation for realistic experimental conditions, simulated data was created to span a reasonable range of photon count, *N*, and displacement, *D*. For each condition, 1000 simulated images were fit by least-squares minimization to each of the fit functions, q_{SErf} and q_{AG}, and the variance in the fit parameters was measured. Figure 2(c), (d) show Δx, Δy, and ΔD for a typical displacement, $D = 5.1$ pixels. The precision of all parameters improves with higher photon

count, *N*, and when $D > 0$, Δx and Δy diverge and we consistently find $\Delta D \gg \Delta x > \Delta y$.

At the limit of $D = 0$, the localization precision of the new fitting functions, q_{SErf} and q_{AG}, are the same as that of the traditional symmetric Gaussian function, q_{SG}. Furthermore, even as *D* becomes finite, the localization errors are still nearly identical (Supporting Figure 2). Finally, as in-frame displacement, *D*, increases, q_{SErf} and q_{AG} begin to outperform q_{SG} in terms of localization accuracy (Supporting Figure 3). Since the traditional symmetric Gaussian function, q_{SG} cannot measure *D*, the new functions have the added benefit of providing an instantaneous measure of velocity.

The errors Δx and Δy agree between the two new fitting functions, but we find in the low *N* regime that ΔD is smaller for the SErf function, i.e. the SErf function performs better than the asymmetric Gaussian function for estimating *D*. For the $b = 200$ photons pixel^{-1} level considered here, at low photon count, neither fitting function is able to distinguish the particle from background noise, giving rise to random guesses within the imaged area. The data in Figure 2(c), (d) is therefore limited to the $N > 6000$ regime, in which the parameter variances depend on *N*.

Because the fitting functions proposed in this paper are not perfect estimators of the PSF, there exists a bias in the estimation of *D*, in part because q_{SErf} and q_{AG} cannot distinguish between forward and reverse motion and is therefore always taken to be positive. We compute this bias from fits to simulated data by averaging over 1000 simulated data sets for each condition. Figure 3(a), (b) compare the estimated displacement, D_{fit}, from fits to the asymmetric Gaussian and SErf functions, respectively, to the actual displacement, *D*. Here, the solid lines

are the average D_{fit} as a function of D, and the shaded region shows $\pm \Delta D_{fit}$. For both fitting functions, the correspondence is worst (large bias in D) at low D, and this bias decreases as D increases, finally disappearing altogether for $D > 3$. The bias in the asymmetric Gaussian fit, Figure 3(a), is attributed to errors in the $c = 0.0104$ fit of the calibration curve, Equation (5). This parabolic function is mathematically simple, but gives rise to a large bias at small D. This bias could be removed with a more elaborate calibration curve. On the other hand, Figure 3(b) shows that for the SErf function, at small D, the bias depends on D with a slope of -1. This bias in the estimation of D arises because least-squares fitting to the SErf function never returns a D_{fit} value smaller than some minimum value (~2 pixels), while the true displacement may be as small as 0. Displacement estimations below that value are no longer dependent on the real displacement, and instead always return the same value on average. Overall, Figure 3(a),(b) show that the two fitting functions have very similar biases in D.

Figure 3(c),(d) show the errors, Δx, Δy, and ΔD for the asymmetric Gaussian and SErf functions, respectively, as a function of D. For both functions, Δy is mostly independent of D, and indeed this relationship, $\Delta y \neq f(D)$, is used in Section 2.1 to derive the closed-form precision for y in the high photon count regime. Δx and Δy are nearly identical for the two fitting functions, so the localization precision is unaffected by choice of function. Furthermore, though D was found to be biased at small displacements, no bias in x, y, or φ was identified (data not shown). The asymmetric Gaussian and SErf functions can therefore both make accurate estimations of position, (x, y), and angle, φ, at all values of D.
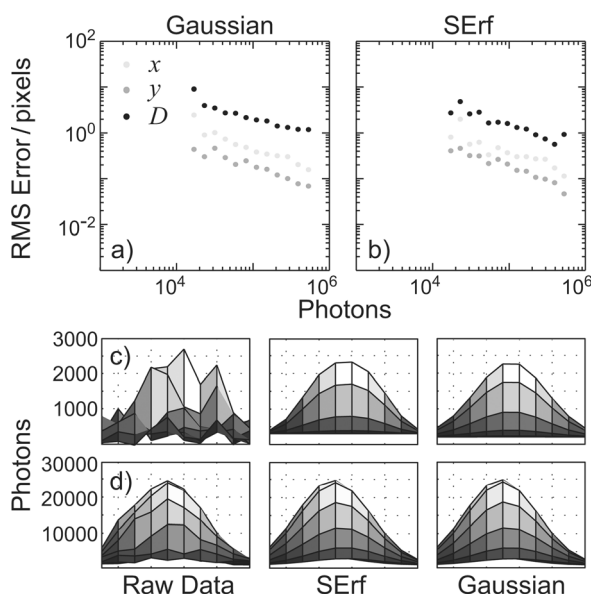

Figure 4. Experimental precision for x, y, and D measured by translating immobilized gold nanoparticles on a motorized stage. a,b) Precision using asymmetric Gaussian and SErf fits, respectively. $b > 100$ photons pixel^{-1}, $c = 0.0104$. (c). d) Example data at low photon counts ($N = 4.56 \times 10^4$) and at high photon counts ($N = 5.13 \times 10^5$), respectively. Left panels: experimental data, center panels: SErf fits, right panels: asymmetric Gaussian fits.

2.3. Controlled Motion

Scattering from 26 nm gold nanoparticles immobilized on a coverslip was imaged under 532 nm laser excitation and rastered at a constant velocity across the microscope viewing area with a piezo scanner. Figure 4(a),(b) show the precision in x, y, and z from fitting the recorded movies to the asymmetric Gaussian and SErf functions, respectively. Though the actual experimental velocity was in principle known, the piezo stage motion was very non-uniform over the course of the experiment. These significant variations in speed introduced some additional systematic error to the experiment. To minimize this error in D, the estimated D from the fits were compared to the average velocity of the stage over segments of only five consecutive frames (see Experimental Section). Those average values were then subtracted from the position measurements to measure the variance in the localization estimates.

As was found analytically and numerically, at all N, the errors followed the trend $\Delta D \gg \Delta x > \Delta y$, and the differences in the localization precision for both Δx and Δy between the two methods were subtle. The asymmetric Gaussian function, however, gave greater errors in D. This additional ΔD may be due to the fact that D is not a fit parameter for q_{AG} but rather related to the ratio of two fit parameters and dependent on another parameter in the calibration curve, Equation (5). Furthermore, for both fitting functions, ΔD depends only weakly on N. This may be related to the fact that D can only be well estimated if φ is accurately determined.

The results from this experiment are similar to the results from simulation (Figure 2(c),(d)), and only slightly worse than the lower bounds suggested by Fisher information analysis (Figure 2(a),(b)). The shallower slopes of error versus N in Figure 4 can be attributed to the background, which is no longer constant as it was in Figure 2; rather, in these experiments, there is a non-negligible increase in b as N increases.

Figure 4(c),(d) show typical experimental data from the low-photon-count ($N = 4.56 \times 10^4$) and high-photon-count ($N = 5.13 \times 10^5$) regimes of this experiment, respectively, as well as the fits of each data set to the asymmetric Gaussian and SErf functions. The asymmetric Gaussian fits are too sharply peaked at its center, whereas the SErf function fits with a smaller residual.

2.4. Cellular Motion

The fit functions q_{SErf} and q_{AG} proposed in this paper can be used to characterize directional motion of small objects such as bacteria cells. Here, we excite fluorescent markers on the surface of *V. cholerae* cells with a 488 nm laser and record images as the cells move in two dimensions within a thin layer of media between a coverslip and an agarose pad. Figure 5(a) presents the results of one typical trajectory (see Supporting Movie 1). Here, a cell marked with a single Alexa-488/antibody label is tracked for 65 imaging frames, during which time, it diffuses randomly (frames 1–7), then moves consistently at an angle of $\varphi \sim 135°$ from the horizontal (frames 8–41), and then finally resumes diffusive motion (frames 42–65). Though the
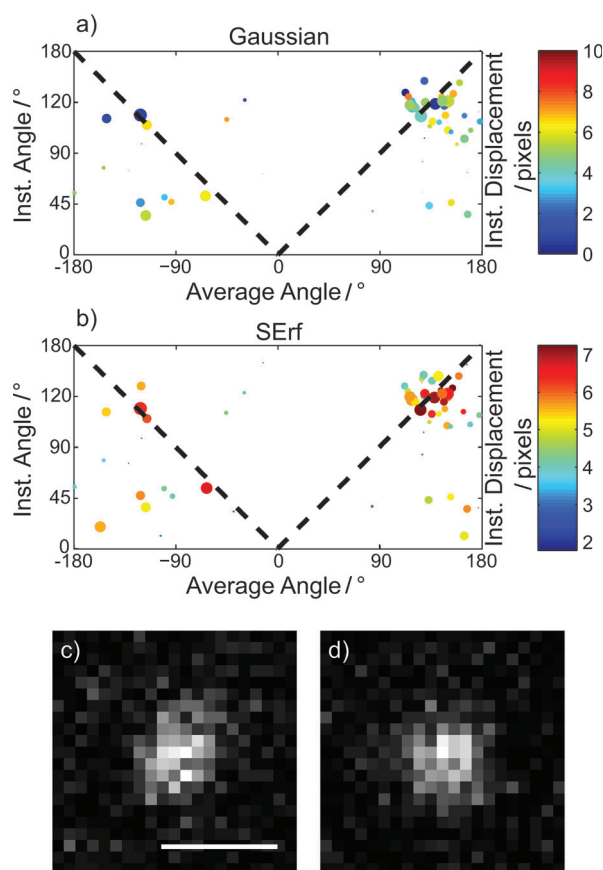
Figure 5. Analysis of experimental images of a *V. cholerae* cell in motion in the 2D plane. a,b) Comparison of instantaneous direction of motion, φ, from a fit to the asymmetric Gaussian function (a) and the SErf function (b), to average direction of motion from the center-to-center average angle over three-frame segments. The dashed lines (-----) show perfect correspondence, the color scale indicates the instantaneous displacement for each measurement from the fit, and the size of the points is inversely related to the square of the 95% confidence interval for the fit to the angle, φ. c) Representative image of a fluorophore undergoing directed motion (frame 40 of Supporting Movie 1). Scale bar: 500 nm. d) Representative image of a fluorophore undergoing diffusive motion (frame 56 of Supporting Movie 1).

average displacement, D, and angle, φ, can be estimated from the trajectories, fits to q_{SErf} and q_{AG} allowed for instantaneous measurements of the velocity and directionality.

During the intermediate frames, the cell moved fast enough that fits to the SErf function, q_{SErf}, and the asymmetric Gaussian function, q_{AG}, yielded meaningful data. Here, we expect that the instantaneous fit D is reasonable; unfortunately the average displacement calculated from the average of changes in position was a bad approximation and so the comparison between instantaneous fit D and average D was poor (Supporting Figure S4). Fortunately, the average angle of motion, φ, was more accurately estimated from the center-to-center average angle over three-frame segments, and a comparison of the instantaneous angle from fit to q_{AG} and q_{SErf} and the average angle from frame-to-frame changes is presented in Figure 5(a),(b), respectively. Here, the dashed lines show perfect correspondence, the color scale indicates the instantaneous displacement from the fit for each measurement, and the size of the points is inversely related to the square of the 95% con-

fidence interval for the fit to the angle, φ. For both functions, there is good agreement between instantaneous and frame-to-frame angle measurements for the directed motion at $\varphi \sim \pm 135°$, where D is large. As expected, in the D small case (random cellular diffusion), isotropic blurring dominates and the methods described here produce incorrect instantaneous angle measurement. Accordingly, Figure 5(a),(b) show no correspondence for the fits where D is small.

Figure 5(c) shows the fluorescent image of the cell during the directional motion portion of its trajectory, frame 40 of Supporting Movie 1. The fluorescent label on the cell appears to be asymmetric and elongated in the direction of motion (~135° from the horizontal), and the displacement and angle can both be estimated from fits to q_{AG} and q_{SErf}. Figure 5(d) shows the fluorescent image of the cell during the random portion of its trajectory, corresponding to frame 56 of Supporting Movie 1. Here, the asymmetry vanishes and the resulting PSF is a symmetric Gaussian with increased standard deviation, as per Schuster et al.[17] The directional motion fitting functions, q_{AG} and q_{SErf}, cannot accurately determine D or φ here.

3. Conclusions

In this paper, we have introduced two new fit functions, the asymmetric Gaussian function and the Sum of Error Functions (SErf) function, that can each measure the instantaneous position, directionality and velocity (in-frame displacement) from a fit to the fluorescent image of a single molecule undergoing directional motion. Both functions have the benefit of being easily integrated into standard single-molecule fluorescence microscopy fitting routines. The results from Fisher information analysis provide a lower bound to the precision of each parameter in the fitting function. Though deriving the closed-form solutions in Table 1 required several simplifications, the true precision of each parameter was easily attained in Section 2.1 by numerically solving the Fisher information integrals. Overall, we find that the two fitting functions perform very similarly, and that the errors consistently follow the trend $\Delta y \leq \Delta x \ll \Delta D$, where x is the direction of motion. Here, the errors in the position estimator for the y direction (perpendicular to the motion) are identical to those derived for fitting stationary molecules to a symmetric Gaussian function,[11] indicating that no localization precision is lost with the new functions relative to standard single-molecule fitting algorithms.

The simulations in Section 2.2 show that the two fitting functions perform similarly in nonlinear least squares minimization, though in the low photon count regime, the SErf function provides more precise estimates of position and displacement. Unfortunately, though estimations of position and angle are unbiased, we have found a systematic bias when estimating in-frame displacement by fitting to either function for small D. This regime could be avoided in applications by choosing experimental conditions that result in $D > 3$ pixels/imaging frame and by carefully calibrating the asymmetric Gaussian function. The results from simulations are similar to the Cramér-Rao lower bounds, and the slight increase in error in the simulations can be attributed to our use of nonlinear least squares

minimization, which is superior to maximum likelihood estimation (MLE) in speed and simplicity, but inferior to MLE in precision. Though these simulations were done with very bright objects in a very noisy background, applications to lower photon count and noise cases does not alter the results qualitatively. Results for this regime, which was recently considered by Yüce et al.,[20] are presented in Supporting Figure 5 and 6.

The application of the fitting functions to the motion of gold nanoparticles in Section 2.3 verified our method in the context of a controlled experiment, and indeed, we find that the new fitting method provides the velocity and directionality of each nanoparticle in each image without degrading the localization precision relative to the symmetric Gaussian fitting function. Though these nanoparticle tracking experiments are not perfect controls due to variations in the stage velocity, the experiments validate our analytical framework, yielding experimental errors similar to the errors in simulations, and only slightly worse than the Cramér-Rao lower bounds.

In Section 2.4, we applied our new fitting functions to the motion of a live *V. cholerae* cell moving in the imaging plane of a microscope. To the best of our knowledge, this is the first example of instantaneous direction and velocity measurement of a freely moving bacteria cell. We observed both random, diffusive motion and steadily directional motion, and found that the instantaneous displacement and angle were only well measured by the fitting functions in the case of directional motion, yet the localization precision was not compromised in any case. The mobile single-molecule fitting functions proposed in this paper are therefore appropriate for any situation where there is directional motion for some part of a trajectory.

In conclusion, presented here are two methods for measuring positions and in-frame velocities of single point light sources. By explicitly accounting for the motion of the fluorophore, these fitting functions permit even mobile single molecules to be fit in standard post-processing routines. These methods promise to improve single-molecule tracking as fewer data sets must be rejected, and tracking algorithms can also be refined to use the instantaneous displacement and directionality as a predictor for the next data point. In a crowded environment, these two factors will make tracking algorithms less ambiguous. These methods do not require unique experimental design, and can even be applied to pre-existing data to recover information previously unnoticed. The precision of the two new fitting functions is nearly identical under all of the conditions we have considered, though least-squares fitting with the SErf function is 4.5% slower than least-squares fitting with the asymmetric Gaussian function.

The directional-motion fitting functions presented in this paper have the potential to inform on the dynamics of living, functioning cells. In addition, we envision that these methods can enable single-molecule imaging of bio-molecules within live, moving cells, since cellular imaging with the new fitting functions, like in Section 2.4, can provide a moving frame of reference for concurrent intracellular single-molecule imaging. This is particularly promising for studies of cellular chemotaxis and signaling, though it can be applied in any situation where directional motion is expected.

Experimental Section

Fitting Function: SErf

The distributive property of the derivative of convolutions is used to convolve a top-hat function, h, with a symmetric Gaussian function, g, as in Equation (1). This distribution gives rise to Equation (7), which is convolved, integrated over all space and normalized to produce Equation (3), the SErf function, $q_{SErf}(x, y)$:

$$\frac{\partial}{\partial_x} q_{SErf}(x,y) =$$

$$\left[\delta(y) \left(\delta\left(x + \frac{D}{2}\right) - \delta\left(x - \frac{D}{2}\right) \right) \right] \otimes \exp\left(- \frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2} \right)$$

$$(7)$$

Fitting Function: Asymmetric Gaussian

To adapt the asymmetric Gaussian function to measuring instantaneous directed motion, a calibration curve is introduced that relates the ratio of the standard deviations parallel and perpendicular to the direction of motion, that is, σ_x/σ_y, to D. The curve is populated by simulated data (see Simulated Data Section below) with varying in-frame displacement, D, and the ratio σ_x/σ_y is obtained from an asymmetric Gaussian fit. The simulated data had $N = 1 \times 10^5$ photons, $b = 200$ photons pixel^{-1}, $\sigma_y = 2$ pixels, and 1000 simulated data sets were fit for each D. This calibration curve is fit by a parabola, Equation (5). The parabolic fit serves both as a look-up table for estimating displacement, D, from fitting parameters $\sigma_x(D)$ and $\sigma_y(D)$, as well as the method of estimating, by error propagation, the precision of D.

Fisher Information

Exact solutions for the asymmetric Gaussian function Fisher information integrals, Equation (6), were found by splitting these integrals into high- and low-noise regimes ($b \gg q$ and $b \ll q$, respectively). For the less analytically tractable SErf function, further simplification was required and Taylor expansions at $D = 0$ were used. Because numerical integration showed Δy to be independent of D at high photon count, the integral for y was further simplified by taking $\lim_{D \to 0} \left\{ \frac{1}{b^2 + q} \left(\frac{\partial q}{\partial y} \right)^2 \right\}$. Because the variance of each parameter is reciprocal to its Fisher information, the square roots of the results in Table 1 give the RMS error. The Fisher information for the two functions was numerically integrated in Mathematica for comparison to the closed-form solutions. Error propagation, Equation (8), using the calibration curve in Equation (5), was used to find the precision of D as estimated by the asymmetric Gaussian function. This treatment was not necessary when evaluating ΔD for the SErf function because D is explicitly included in the function. Values of parameters for the results in this section were as follows: $D = 5.1$, $\sigma = 2$, $b = 200$ for the SErf integrals, and: $\sigma_x = 2.54$, $\sigma_y = 2$, $b = 200$ for the asymmetric Gaussian integrals. This value of σ_x was calculated from Equation (5) using the calibration constant from simulation, $c = 0.0104$, $D = 5.1$, and $\sigma_y = 2$ [Eq. (8)]:

$$(\Delta D)^2 = \left(\frac{\partial D(\sigma_x, \sigma_y)}{\sigma_x} \right)^2 (\Delta \sigma_x)^2 + \left(\frac{\partial D(\sigma_x, \sigma_y)}{\sigma_y} \right)^2 (\Delta \sigma_y)^2 \qquad (8)$$

Pixelation, which adds noise due to the finite size of a, is ignored in our solutions to the Fisher information (i.e. Figure 2(a),(b) and Table 1) because our results for RMS error were found to depend

only weakly on pixel size within parameter regimes relevant to single-molecule experiments. For the SErf function and the asymmetric Gaussian function, pixelation can be accounted for by adding the variance of a top-hat distribution with width a, $(\Delta a)^2 = a^2/12$, to the variance of the point spread function, that is, σ^2 or σ_x^2 and σ_y^2.

Simulated Data

Simulated images of point-source emitters were generated by weighting random numbers (generated by the Matlab routine *rand*) by a distribution function given by the numerical convolution of a top-hat function and an Airy disk. The convolution was done at high resolution, 1000×1000 pixels, and the distributions were then down-sampled by linear interpolation to 25×25 pixels. The number of photons for each simulated image was chosen from a Poisson distribution with mean equal to the desired photon count. Poissonian noise with standard deviation of 200 counts per pixel was added to the simulated image. The computed precision in Figure 2(c),(d) was determined by fitting simulated data with the asymmetric Gaussian and SErf functions with the built-in bounded non-linear least squares minimization Matlab routine *lsqcurvefit*. The variance of the parameter estimations was calculated based on fitting 1000 simulated data sets for each value of N and D.

Controlled Nanoparticle Motion Imaging

Glass microscope coverslips were cleaned using an oxygen plasma etch (10 min at 200 mTorr; Plasma Etch, Inc. PE-50). Gold nanoparticle substrates were prepared on the cleaned coverslips using a spin-assisted layer-by-layer technique[21] based on polyelectrolyte films of positively charged poly(diallyldimethyl ammonium chloride) solution (PDADMAC, Sigma Aldrich) and negatively charged poly(sodium 4-styrene) solution (PSS, Sigma Aldrich). 26 nm diameter spherical gold nanoparticles (Nanopartz, Inc.) were used as received. Polyelectrolyte solutions (20% by weight in water) were diluted in distilled deionized (DDI) water. 20 mM PDADMAC (calculated using monomeric weights) was spun onto coverslips (300 μL, 15 s, 4000 rpm), then washed three times with DDI water (300 μL, 15 s, 4000 rpm). A mixture of nanoparticles and 20 mM PSS was then spun onto the PDADMAC-coated coverslips (100 μL NPs, 200 μL PSS, 15 s, 4000 rpm), followed by three washes with DDI water.

The resulting sparse gold nanoparticle samples were imaged using with a 60× 1.49-NA oil-immersion objective (APON60XOTIRFM) in an Olympus IX81 inverted microscope. Wide-field epifluorescence single-molecule microscopy was performed using 532 nm light (Crystalaser CL532-150mW-L) as an excitation source via fiber input. Scattered light was imaged on a 512×512 pixel EMCCD (Andor iXon 897) at a frame rate of 2 Hz for 3 min. This setup had a magnification factor $a = 50$ nm pixel^{-1}. The samples were translated during imaging at a constant velocity of 0.5 μm s^{-1} with a capacitive piezoelectric *xyz* stage (Physik Instrumente).

Controlled Motion Analysis

Movies of translating gold nanoparticles were fit by the Matlab routine *lsqcurvefit* to the asymmetric Gaussian and SErf functions, yielding instantaneous values for all fit parameters for each localized nanoparticle in each imaging frame. To determine the Δx and Δy as a function of photons detected in Figure 4, particle tracks were separated into five-frame segments. The average center-to-center displacement/frame for each segment was calculated then

subtracted from the position measurements, producing five stationary images. Then the measured x, y, and D values were binned by N and the variance of the parameter estimations were calculated for each bin.

Cellular Motion Imaging

Cells of the *Vibrio cholerae* classical strain O395 were grown in LB rich medium at 37 °C, then grown to turbidity (OD 0.3) at 30 °C in M9 minimal medium. *V. cholerae* cells were incubated first with intact whole-cell anti-*V. cholerae* polyclonal rabbit primary antibodies (1/2000, Abcam) for 30 min at RT, rinsed 3 times in M9, then incubated with Alexa-488 goat anti-rabbit secondary antibodies (1/1000, Life Technologies) for 30 min at RT, and again rinsed three times in M9. This produced cells with 0–3 fluorescent labels on their surface. 2.0 μL of cells in M9 media were sandwiched between a 1.5% agarose in M9 pad and a glass coverslip. The agarose pad was wet enough to allow for two-dimensional motion within the focal plane.

The labeled *V. cholerae* cells were imaged using with a 100× 1.40-NA oil-immersion objective in an Olympus IX71 inverted microscope. Wide-field epifluorescence single-molecule microscopy was performed, using 488 nm light (Sapphire 488-50) as an excitation source with excitation powers of 100–500 μW. Scattered light was eliminated with a dichroic and a long-pass filter (Semrock Di01-R488 and Semrock BLP01-488), emitted signal was magnified by a 3.3× beam expander to a final magnification of $a = 49$ nm/pixel, and the emission was recorded on a 512×512 pixel EMCCD (Photometrics Evolve EMCCD) at a frame rate of 10 Hz for 5 min.

Cellular Motion Analysis

Swimming cells with a single visible fluorophore were identified for analysis, and the fluorescent label was fit in each imaging frame with the SErf and asymmetric Gaussian functions. Each fit yielded a measured instantaneous D and φ for each imaging frame, i, and these values were compared to an estimate for D and φ determined from the trajectory between the center positions of that fluorophore in frames $(i-1)$ and $(i+1)$.

[1] J. Gelles, B. Schnapp, M. Sheetz, *Nature* **1988**, *331*, 450–453.
[2] M. K. Cheezum, W. F. Walker, W. H. Guilford, *Biophys. J.* **2001**, *81*, 2378–2388.
[3] A. Yildiz, J. N. Forkey, S. A. Mckinney, T. Ha, Y. E. Goldman, P. R. Selvin, P. R. Selvinl, *Science* **2003**, *300*, 2061–2065.
[4] S. Y. Kim, Z. Gitai, A. Kinkhabwala, L. Shapiro, W. E. Moerner, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10929–10934.
[5] S. Semrau, A. Pezzarossa, T. Schmidt, *Biophys. J.* **2011**, *100*, L19–L21.

[6] L. Xiao, L. Wei, C. Liu, Y. He, E. S. Yeung, *Angew. Chem.* **2012**, *124*, 4257 – 4260; *Angew. Chem. Int. Ed.* **2012**, *51*, 4181 – 4184.

[7] R. N. Ghosh, W. W. Webb, *Biophys. J.* **1994**, *66*, 1301 – 1318.

[8] C. M. Anderson, G. N. Georgiou, I. E. Morrison, G. V. Stevenson, R. J. Cherry, *J. Cell Sci.* **1992**, *101*, 415 – 425.

[9] K. I. Mortensen, L. S. Churchman, J. A. Spudich, H. Flyvbjerg, *Nat. Methods* **2010**, *7*, 377 – 381.

[10] J. Engelhardt, J. Keller, P. Hoyer, M. Reuss, T. Staudt, S. W. Hell, *Nano Lett.* **2011**, *11*, 209 – 213.

[11] R. E. Thompson, D. R. Larson, W. W. Webb, *Biophys. J.* **2002**, *82*, 2775 – 2783.

[12] S. Kay, *Fundamentals of Statistical Signal Processing*, Vol. 1, Prentice Hall PTR, Upper Saddle River, **1993**, p.30.

[13] R. J. Ober, S. Ram, E. S. Ward, *Biophys. J.* **2004**, *86*, 1185 – 1200.

[14] C. S. Smith, N. Joseph, B. Rieger, K. A. Lidke, *Nat. Methods* **2010**, *7*, 373 – 375.

[15] F. Aguet, D. Van De Ville, M. Unser, *Opt. Express* **2005**, *13*, 10503 – 10522.

[16] M. C. DeSantis, S. K. Zareh, X. Li, R. E. Blankenship, Y. M. Wang, *Opt. Express* **2012**, *20*, 3057 – 3065.

[17] J. Schuster, F. Cichos, C. von Borczyskowski, *J. Phys. Chem. A* **2002**, *106*, 5403 – 5406.

[18] S. K. Zareh, M. C. DeSantis, J. M. Kessler, J.-L. Li, Y. M. Wang, *Biophys. J.* **2012**, *102*, 1685 – 1691.

[19] B. Huang, W. Wang, M. Bates, X. Zhuang, *Science* **2008**, *319*, 810 – 813.

[20] M. Y. Yüce, A. Jonás, A. Kiraz, A. T. Erdog, *Microsc. Microanal.* **2012**, *18*, 781 – 792.

[21] P. Chiarelli, M. Johal, D. Holmes, J. Casson, J. Robinson, H. Wang, *Langmuir* **2002**, *18*, 168 – 173.