Supporting Material for

Classification of dynamical diffusion states in single molecule tracking microscopy

P.J. Bosch[†], J.S. Kanger[‡] and V. Subramaniam^{†‡#}

[†] Nanobiophysics, MESA+ Institute for Nanotechnology, University of Twente, The Netherlands

[‡] MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands

[#] Present address: FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands

Contents

Supporting Figures	2
Supporting Sections	12
Quantification measures	12
Mean squared displacement (MSD)	12
Windowed MSD	13
Maximum Likelihood Estimation	13
Relative confinement	14
Radius of gyration evolution	14
Live cell experiments methodology	16
Cell culture	16
Microscopy	16
Tracking	16
Accuracy of MSD methods and CDF fitting to obtain a one population diffusion coefficient	17
Settings file for SPT tracking software	19
Supporting References	20

Supporting Figures



FIGURE S1: A conventional full-trajectory MSD analysis of a two-population diffusion systems leads to an apparent distribution of diffusion values, easily resulting in erroneous conclusions for this motion system. The true motion system is a two-population diffusion system with diffusion coefficients of 0.15 and 0.01 μ m²/s, as indicated by the arrows in the graph. For this graph, 1000 trajectories consisting of 500 frames (hence relatively long trajectories) were simulated. The trajectories are analyzed using the conventional full-trajectory MSD method, where the first 4 points of the MSD curve were used following the "rule of thumb" rule to fit the diffusion constant. The found diffusion constants and the apparent spread both incorrectly describe this two-population situation. The MSD analysis should therefore only be used for homogeneous (one-population) motion. The discussion further in the supplementary material discusses the accuracy of determining the diffusion constant for a one-population system using the full-trajectory MSD method.



FIGURE S2: Correctness dependence of window length and of the number of points in the MSD curve used to fit the diffusion value. (*A*) Simulation case A has a localization inaccuracy σ_{xy} of 40nm and short state lifetimes ($\tau_1 = \tau_2 = 300 \text{ ms}$). (*B*) This simulation case differs from case A only by a lowered localization inaccuracy $\sigma_{xy} = 20 \text{ nm}$. (*C*) This simulation case differs from case A only by having longer fast state lifetimes ($\tau_1 = 900 \text{ ms}$; $\tau_2 = 300 \text{ ms}$; $\sigma_{xy} = 40 \text{ nm}$). (*D*) This simulation case also has longer slow state lifetimes ($\tau_1 = 900 \text{ ms}$; $\sigma_{xy} = 40 \text{ nm}$). The color bar aids in reading the classification correctness percentage.





B







80 Relative confinement MLE Gyration 2 2 2 61.1% 58.8% 75 69.2% 71.1% 72.0% 72.3% 72.1° <mark>69.9%</mark> 72.3% 73.3% 73.3% 73.0% 64.7% 65.0% 63.8% 66.0% 66.1% 65.6% 4 70.8% 72.3% 73.2% 73.0% 72.6% 4 72.1% 72.9% 73.5% 73.2% 72.6% 67.9% 68.3% 68.1% 6 72.9% 74.9% 74.9% 74.4% 73.5% 75.3% 75.2% 74.8% 73.8% 69.1% 69.1% 68.5% 70 73.3% 75.1% 74.7% 74.2% 73.4% 6 6 75.3% 74.7% 74.1% 73.5% 75.2% -70.2% 70.5% 70.3% 8 window size vindow size window size 71.0% 71.0% 70.4% 74.3% 76.6% 76.2% 75.3% 74.3% 76.6% 76.1% 75.3% 74.5% 10 .71.7% 71.9% 71.7% 8 74.2% 76.2% 75.7% 74.6% 73.7% 8 76.0% 75.2% 74.3% 73.5% 65 72.2% 72.2% 71.8% 76.7% 75.3% 74.3% 75.3% 74.3% 12 72 6% 72.9% 72.6% 73.0% 73.0% 72.6% 10 74.3% 76.4% 76.0% 74.8% 73.6% 10 75.9% 75.2% 74.2% 73.3% 60 73.3% 73.4% 73.3% 14 74.2% 76.7% 76.5% 75.5% 74.1% 76.1% 75.7% 74.9% 74.0% 73.4% 73.5% 73.2% 12 73.8% 75.9% 75.6% 74.8% 73.6% 12 <mark>75.8% 75.4%</mark> 74.8% 73.9% 73.0% 16 73.6% 73.7% 73.7% 73.7% 73.8% 73.6% 73.6% 76.0% 76.0% 75.2% 74.0% 75.6% 75.4% 75.1% 74.4% 73.6% 55 73.8% 74.0% 74.0% 18 14 -73.2% 75.2% 75.3% 74.5% 73.6% 14 <mark>75.0% 74.7% 74.2% 73.5% 72.7%</mark> 74 0% 73.9% 73.9% 72 8% 75 1% 75 5% 74 8% 73 8% 74 8% 74 7% 74 3% 73 8% 73 2% 74 0% 74 1% 74 1% 20 50 2 3 4 5 2 3 4 5 1 2 3 1 1 minimum state duration (tc) minimum state duration (tc) minimum state duration (tc) D 80 Relative confinement Gyration MLE 2 2 2 64.99 66.0% 66 7% 64.8% 65.4% 67.2% 67.0% 65.1% 66.3% 68.2% 69.3% 69.0% 66.6% 75 67 7% 68.6% 69.0% 4 69.5% 69.7% 69.4% 4 67.6% 69.0% 70.2% 69.6% 67.7% 4 70.1% 70.8% 71.8% 71.0% 69.6% .71.1% 71.5% 71.4% 6 70.5% 72.9% 73.1% 72.7% 70.7% 73.4% 74.1% 74.5% 74.3% 72.5% 72.2% 72.2% 71.8% 70 6 71.7% 73.9% 73.9% 72.9% 71.2% 6 74 9% 74.8% 74.6% 74.1% 73.2% 73.6% 73.3% 8 73.3% size vindow size vindow size 74.0% 74.0% 73.5% 73.3% 76.0% 75.9% 74.6% 72.9% 76.4% 76.6% 76.3% 75.7% 74.8% 10 74.7% 75.0% 74.7% window 8 73 8% 76 1% 75 8% 74 2% 72 5% 8 75.8% 74 9% 73 9% 65 75.2% 75.2% 74.7% 74.6% 77.4% 77.2% 75.3% 73.5% 76.0% 75.0% 12 75.6% 75.8% 75.5% 75.9% 75.9% 75.5% 10 74.6% 76.9% 76.6% 74.9% 72.9% 10 76.1% 75.1% 73.9% 60 76.3% 76.2% 76.1% 14 74 9% 77 5% 77 4% 75 8% 73 6% 77 3% 76 8% 76 0% 74 8% 76.0% 12 74.7% 76.8% 76.6% 75.2% 73.2% 12 76.0% 75.0% 73.8% 16 74.7% 77.0% 77.1% 75.9% 73.9% **76.8% 76.4% 75.6% 74.6%** 55 76.7% 76.7% 18 14 74.5% 76.5% 76.4% 75.0% 73.3% 14 76.2% 75.6% 74.7% 73.6% 76.8% 76.5% 74.3% 76.4% 76.7% 75.6% 73.9% 76.2% 75.8% 75.1% 74.2% 20 50 2 3 4 5 2 3 4 5 1 2 3 1 1 minimum state duration (tc) minimum state duration (tc) minimum state duration (tc)

FIGURE S3: Effect of a minimum state duration filter on classification correctness. Since all the above methods make use of segments, the state classification does not allocate exact time points of a trajectory to a state, but to a window of several time points. The state of this window is classified at the center time point of the window, and thereby each time point has its own state allocation based on different but correlated data sets. Now the state classification is defined at all timepoints except for the beginning and end of a (sub)trajectory. Since the window is several frames long, state durations shorter than the segment length (window size) may appear illogical, and therefore we also tested whether filtering out short state durations leads to better correctness for different classification methods. In most simulation cases (as described in the main text) this was not the case, and we have therefore looked only at unfiltered state classifications in the main text. (*A-D*) Classification correctness for simulation case A-D. The color bar aids in reading the classification correctness percentage.

С



FIGURE S4: Correctness of the two-population classification for different ratios for D_1 and D_2 . D_2 was kept fixed at 0.015 μ m²/s. Shown are the results for a gyration and MLE based classification for simulations of cases A and D (with changing D_1). We find that ratio of 4 in our simulation cases was indeed a challenging situation, and that with ratios larger than 10 the best achievable classification are obtained. The classification also depends on other factors as shown by the different results for the two cases shown.



FIGURE S5: The distribution of lifetimes found by the gyration based classification method simulation case A. The simulation consisted of 20 trajectories and was repeated 1,000 times to yield the distribution of lifetimes of the slow state (*A*) and the fast state (*B*). The gyration method used a segment length of 7 frames. The arrows indicate the true lifetime of both states. (*C*) Example of a lifetime fit of the fast state from one simulation. This fit is only performed over state events larger than 5 frames. Although the lifetimes found are not too far off in this case, in cases with longer state durations (such as case D), the lifetimes found are significantly underestimated.



FIGURE S6: Image reconstruction of the state classification distribution of simulation case A. In this case there were as many slow as fast states. The dimension of the image is $55x55\mu$ m, and the image is reconstructed at 60nm/pixel.



FIGURE S7: Image reconstructions of the state classification distribution of a simulation with spatially defined diffusion states. (*A*) Simulation (actual) image of fast state. (*B*) Simulation (actual) image of slow state. (*C*) Fast state image found by MLE method. (*D*) Slow state image found by MLE method. (*E*) Colour image of states map as found by the MLE method. Green represents fast state, and red represents slow state. (*F*) Colour image of states map as found by the MLE method. Green represents fast state, and red represents slow state. A segment length of 4 frames was used in both classification methods (this value yielded the highest correctness). The dimension of the image is $15x15\mu$ m, and the image is reconstructed at 30nm/pixel.



FIGURE S8: The proposed approach for diffusion state classification applied to experimental data. (*A*) The CDF fit, with corresponding PDF (*B*) and residuals (*C*), for the motion of EGF receptor in an MCF7 cell shows that the model of two-population Brownian diffusion is a suitable motion model. (D) An example of a single molecule fluorescence frame recording of EGF receptor in an MCF7 cell. The image has not been modified or filtered.



FIGURE S9: Histograms of unliganded EGF receptor state lifetimes for the fast state (*upper*) and the slow state (*lower*) in frames (video was recorded at 25 fps), determined by gyration analysis. The characteristic lifetime is determined from an exponential fit of state durations longer than 5 frames, since the use of a segment length of 7 frames does not correctly resolve shorter lifetimes. The state duration will be underestimated because of random incorrect state classification.



FIGURE S10: Image reconstructions of the distribution of states exhibited by liganded EGF receptor proteins in an MCF7 cell. The resolution of the reconstructed images is 30nm/pixel. (*A*) Image showing the areas travelled by receptors in the fast diffusion state (green), and areas where receptors in the slow diffusion state were detected (red). (*B*) Zoomed image of the indicated area (white box) in *A*.

Supporting Sections

Quantification measures



FIGURE S11: Illustrations of various segment analysis methods as quantification measures for the proposed diffusion state classification approach. (*A*) Simulated two-population diffusion trajectory and examples of selected information for the relative confinement and the gyration method on a trajectory segment. Relative confinement detection takes the variance of distances from the center of the segment (indicated by the grey arrows). The gyration radius depends on the variance and covariance of the coordinates (segment indicated in grey with a dashed line). (*B*) Cumulative distribution function (CDF) values and fit of squared displacements from a set of trajectories with two diffusion coefficients (*upper*), and the corresponding probability density function which is the derivative of the CDF (*lower*). The simulated values are in grey, the fit is drawn in black. (*C*) A conventional full trajectory MSD curve from pure one-population diffusion (first 500 frames shown). (*D*) A windowed MSD provides an instantaneous diffusion coefficient for all timepoints of a trajectory by performing an MSD analysis only on a segment (the window, as indicated with dashed grey lines in panel *A*) of the trajectory, and sliding this window through the whole trajectory.

Mean squared displacement (MSD)

The most straightforward way to determine the diffusion coefficient is by using the mean squared displacement (MSD) versus time lag curve (1). This provides an estimate of the diffusion coefficient, and also confinement (2), but the method requires that within the complete trajectory there is only one type of homogeneous motion. In short, the MSD is usually defined as the average of all squared distances between points within a certain lag time $\tau = n \cdot \Delta t$, with Δt the time-delay between consecutive frames, and *n* the interval of frames over which the distance is measured and averaged. For intervals larger than 1 frame, usually all available distances of a given duration $n \cdot \Delta t$ are included, such that the distances are not statistically independent. Yet this way of averaging gives less variance to the average squared displacement value compared to taking only the independent distances (3). For pure Brownian motion, the relation between squared displacements (ΔR)² and the diffusion coefficient is a linear relation:

$$MSD(\tau) = \langle (\Delta R_{\tau})^2 \rangle = 4 D \tau + 2(\sigma_x^2 + \sigma_y^2) = 4 D \tau + 4 \sigma_{xy}^2$$
(1)

where σ_{xy} is the standard deviation of the localization inaccuracy in one dimension, which is independent of the time lag. The estimated diffusion coefficient *D* is found from fitting a line through the points at the different lag times in the MSD curve. We emphasize that it is not straight forward how to perform the fitting of the MSD curve to obtain diffusion values. A more detailed insight on this is given in the section "Accuracy of MSD methods and CDF fitting to obtain a one population diffusion coefficient" in the Supporting Materials. Moreover, conclusions from MSD curves must always be tested against unconstrained diffusion, as the randomness of normal diffusion may result in apparent anomalous diffusion (4).

A recurring question is which points in the MSD curve can still be considered reliable. Certainly the variance of larger time lags gets increasingly larger, such that the points of larger time lags do not provide any reliable information. In the literature the first 10% points of the curve are often assumed to have not too much variance in their values (5). However, the analytical expression for one population Brownian motion for the variances has been derived (3, 6). Following this expression, Michalet discussed what the optimal number of points is to be taken into the fit for determining the diffusion coefficient (7). The optimal number of points depends on the ratio $\beta = \sigma^2/(D \Delta t)$, with σ the standard deviation of the localization inaccuracy. In the limit of no (or relatively small) localization inaccuracy, i.e. for small β , it was shown that the most accurate value for *D* is obtained by fitting with only the first two population diffusion systems which have both high and low diffusion constants and correspondingly both low and high β values, we do not readily know the optimal number of points of the MSD curve used using simulations (Fig. S2).

Windowed MSD

Typically, the MSD curve is made up from all positions in a trajectory, which cannot resolve local changes in the diffusion coefficient. Windowed MSD tries to give the local or instantaneous diffusion coefficient at each timepoint of a trajectory by performing the MSD analysis on small segments of the trajectory. First an MSD curve is composed for w subsequent positions in a trajectory, and the estimated D value is obtained from the first three points in the curve for this segment. This value is taken as the measure W. Then the MSD curve is made for the next subsequent positions, until the full trajectory has been slid through, and D values have been obtained for each segment, see also Fig. S11D. The use of a moving window makes it possible to detect temporal changes in the mode of motion on the order of the segment length (window size). The resolution is limited by the averaging nature of the method, since reducing the segment length means that the MSD curve is made up from fewer points, therefore increasing the statistical uncertainty of the fitted diffusion coefficient.

Maximum Likelihood Estimation

We have used a likelihood estimation approach here by comparing a window of measured squared displacements, a set of a few single steps $\{(\Delta R)_i\}$, to the expectation value thereof given the distribution function of squared displacements originating from motion with a diffusion constant *D*. For one step of length ΔR , we use $P((\Delta R)^2 | D)$ to express the chance to find a certain squared displacement given Brownian motion with diffusion coefficient *D*. Since the expectation value of one squared displacement is independent of its predecessors, the chances for a tested *D* can be multiplied for each squared displacement (ΔR)_{*i*}², hence the likelihood is given by:

$$L(\{(\Delta R)_{i}\} \mid D) = \prod_{i=1}^{N} \frac{1}{\sqrt{4\pi (D\tau + \sigma_{xy}^{2})}} \cdot \exp\left(-\frac{(\Delta R)_{i}^{2}}{4(D\tau + \sigma_{xy}^{2})}\right)$$
(2)

where τ is the time lag, which is 1 frame, and N is the total number of steps in the window. The values for D are taken from the earlier CDF fit. In practice, the localization inaccuracy σ_{xy} must be

determined by other means first. Here we assumed that this value can be precisely obtained, and we used the true value as used in the generation of the trajectories. Here we determine the likelihood of both states, $L_1(\{(\Delta R)_i\} | D_1)$ and $L_2(\{(\Delta R)_i\} | D_2)$, and if $L_1 > L_2$, the segment is classified as state 1. We could write this as a measure W by:

$$W(track, frame) = \frac{L(\{(\Delta R)_i\} \mid D_1)}{L(\{(\Delta R)_i\} \mid D_2)}$$
(3)

We have not taken exposure blur into account (8). Note that the MLE can also be used to estimate the value of the diffusion constant itself, by maximizing the expectation value by varying the tested D value; the maximum gives the most likely D value (9).

Relative confinement

Inspired by the confinement detection method of Simson (10), Meilhac used a slightly altered way to detect confinement (11), which we also use here. The relative confinement is defined by the parameter L as:

$$L\left(t_0 + \frac{1}{2}\delta t\right) = \frac{\delta t}{\text{variance}(s)}$$
(4)

$$s = \left| r(t) - r\left(t_0 + \frac{1}{2}\delta t\right) \right| \text{ on interval } t = [t_0 \dots t_0 + \delta t]$$
(5)

An illustration of the distances s is given by arrows in Fig. S11A. Here we use the inverse of the value L for the motion quantification measure.

Radius of gyration evolution

The use of the radius of gyration has been first proposed by Saxton to measure asymmetry in single molecule trajectories (4), and it was demonstrated by Elliott et al. that it could also be used to detect confinement (12). The gyration radius is a measure of the space that is explored (defined by radius R_g) by the molecule within the segment, hence the radius will have a lower value for slow diffusion than for fast diffusion. Therefore the gyration radius is a local measure of the diffusion of a molecule, and can be used as a differentiation criterion in classification. We note that the expression in reference 12 contains a typographical error, as the radius of gyration is defined as the square root of the non-squared sum of the eigenvalues of the covariance matrix. However we followed Elliot et al. in an alternative measure, also called R_g . This alternative gyration radius R_g is defined as:

$$R_g^2 = \sqrt{R_1^2 + R_2^2} \tag{6}$$

where R_1 and R_2 are the eigenvalues of the gyration tensor T:

$$\boldsymbol{T} = \begin{pmatrix} \frac{1}{N} \sum_{i=1}^{N} (x_i - \langle \boldsymbol{x} \rangle)^2 & \frac{1}{N} \sum_{i=1}^{N} (x_i - \langle \boldsymbol{x} \rangle) (y_i - \langle \boldsymbol{y} \rangle) \\ \frac{1}{N} \sum_{i=1}^{N} (x_i - \langle \boldsymbol{x} \rangle) (y_i - \langle \boldsymbol{y} \rangle) & \frac{1}{N} \sum_{i=1}^{N} (y_i - \langle \boldsymbol{y} \rangle)^2 \end{pmatrix}$$
(7)

with *i* enumerating all subsequent positions (x_i, y_i) in a segment of length *N*. We will use the value R_g as a motion quantification measure.

Supporting Material

Table S1: Motion quantification measures

Method	Quantification measure W
Windowed MSD	fit of a MSD curve of a segment using first two points in curve
Confinement	$\frac{1}{L(t_0 + \frac{1}{2}\delta t)}$
Gyration	R_g
MLE	$L(\{(\Delta R)_i\} \mid D_1)$
	$\overline{L(\{(\Delta R)_i\} \mid D_2)}$



FIGURE S12 Distribution of found quantification measure values for pure one-population diffusion. The histograms of three different diffusion constants are shown, where in all cases we added a localization inaccuracy σ_{xy} of 40 nm to the positions in the simulations. (*A*) Histograms of values found using a windowed MSD. The broadening in the slower diffusion distributions are due to the convolution with the localization inaccuracy. (*B*) Histogram of values found using relative confinement. (*C*) Histogram of values found using gyration.

Live cell experiments methodology

Cell culture

All cell culture materials were obtained from PAA Laboratories (Pasching, Austria) unless stated otherwise. MCF7 cells, a human breast cancer cell line, and plasmid coding for SNAP-EGFR were a gift from Jenny Ibach (Max Planck Institute in Dortmund, Germany). Cells were cultured in Dulbecco's Modified Eagle's medium supplemented with 10% FBS and penicillin/streptomycin at 37°C with 5% CO₂. Before measurements, the cells were transferred to CellView dishes product #627870 (Greiner Bio-one, Alphen aan den Rijn, The Netherlands), grown overnight, transfected with SNAP-EGFR using Effectene (Qiagen, Venlo, The Netherlands), and then starved overnight the day after transfection in medium without FBS. Labeling of the SNAP-EGFR proteins was done by incubating the cells for 1 minute with 400nM of SNAP-Surface 549 (New England BioLabs, Ipswich, MA, USA) in 0.5% BSA. Measurements were performed in PBS buffer with added magnesium and calcium (PAA Laboratories, product H15-001).

Microscopy

Measurements were performed on a microscope with an Olympus PlanApo 100x/1,45 Oil TIRF objective using TIRF illumination. For excitation a 532nm laser (400mW) from Pegasus Shanghai Optical Systems (Pegasus Lasersysteme, Wallenhorst, Germany) was used. All the light filters were obtained from SemRock (Rochester, NY). The infrared light produced by the laser was not sufficiently suppressed, therefore the green laser light passed an FF01-543/22 filter. The excitation and emission is split by an FF494/540/650-Di01 dichroic mirror. The emission light is filtered with an NF03-532/1064E notch filter and an FF01-580/60 bandpass filter. Fluorescence images were acquired using an Andor iXon EM+ DU-897 back illuminated EMCCD with an acquisition time of 9ms and a kinetic cycle time of 38ms (25.8 fps). The microscope stage was heated with a sample heating plate and the objective was heated with a ring heater to 35-37°C.

Tracking

To obtain the trajectories from the raw videos, we used tracking software developed by others (13, 14). The settings used for the cost matrices in this software can be found at the end of the Supporting Materials.

Accuracy of MSD methods and CDF fitting to obtain a one population diffusion coefficient

It might seem, and it is often stated, that the CDF method is more accurate in determining the diffusion coefficient for a one population diffusion system compared to simply averaging the stepsizes as in MSD methods (15), as it considers the whole distribution of stepsizes. In practice however, this is not always correct. Also the number of points from an MSD curve taken into the fit to determine the diffusion coefficient are often based on a "rule of thumb" concept, such as taking the first three or four or the first 10% of the curve. However the accuracy to find the diffusion coefficient can simply be found by simulation and also by calculation (3). We show a simulation approach here to determine the spread of found diffusion coefficients from CDF and MSD methods.

We simulated one-population unconstrained diffusion for 100 trajectories of various lengths, with a relatively small localization error compared to the diffusion coefficient, so for $\beta = \sigma^2/(D \Delta t)$ ratio smaller than 1, see (7). We found that, for all lengths of trajectories, a CDF fit with only 1 stepsize is indeed, but only slightly, more accurate compared to the best MSD based fit; the value is of course wrong when not corrected for the added localization inaccuracy to the real diffusion coefficient. In practice this means we have to use the CDF of 2 steps too, and use the difference for CDF 2 steps and CDF 1 step to determine the diffusion coefficient. This 2 steps CDF methods has been described in detail in the methods section. Using this last method however, we found to be less accurate compared to the best MSD based fit where we take only the first two points in the curve (also the 1-steps and 2-steps). Using only the first two points in the MSD curve was the best MSD based fit for this ratio of β . Therefore the CDF was not taken as a method for classification, as the MSD is preferred for one population diffusion constants unlike the windowed MSD distribution, so this is still a straight forward method to find the global diffusion constant values and fractions when there are enough datapoints to build a reliable CDF.



FIGURE S13 Error and standard deviation of MSD methods to obtain a one population diffusion coefficient. For 100 simulated trajectories exhibiting one-population Brownian motion (D=0.1 um²/s, 25fps) of various lengths (3,5,8,..,100 steps) as plotted on the x-axis, the diffusion value was determined from fitting the MSD curves of all trajectories. We added localization inaccuracy of σ_{xy} = 40nm to the simulated trajectories (*C*,*D*). This fit was done using: the full curve (a), full curve weighted using the variance of each point (b), the first 10% (c), only the first two points (d), and using cumulative distribution function (CDF) fitting of steps (e), and using CDF of one-step and two-step distances (f). We repeated this 1,000 times, and looked at the standard deviation σ (*A*,*C*), and the average mismatch $\langle \epsilon \rangle \langle B, D \rangle$ in the fitted diffusion value when there is a localization inaccuracy as in practice. The most accurate way of using the points in the MSD curve, is to only use the first two points of the MSD curve.

Settings file for SPT tracking software

```
Dat.PixelSize = .119;
Dat.TimeStep = 0.03868;
Dat.ch bin = [1];
Params.verbose = 1;
Params.frames = [];
Params.psf = [0.84034 \ 0.84034];
Params.imMask = [];
Params.wvMask = [];
Params.CCDGain = 63.8298;
Params.CCDOffset = 0;
Params.Intensity = 1.90;
Params.FitBoxSize = [7];
Params.Iterations = 10;
Params.MaxCudaFits = 30000;
Params.MinCRLBSigma = 0.5;
Params.MinPValue = 0.01;
Params.MinPhotons = [10];
Params.ConnectParams.costMatF2Fparams = costMatFrame2FrameSetOptions;
Params.ConnectParams.costMatGCparams = costMatCloseGapsSetOptions;
%%% set parameters for frame 2 frame connections %%%
Params.ConnectParams.costMatF2Fparams.funcName =
'costMatFrame2FrameDensity';
Params.ConnectParams.costMatF2Fparams.density = [];
Params.ConnectParams.costMatF2Fparams.D =
[0.06*Dat.TimeStep/Dat.PixelSize^2 0.06*Dat.TimeStep/Dat.PixelSize^2 ];
Params.ConnectParams.costMatF2Fparams.maxSearchDist = [4 4];
Params.ConnectParams.costMatF2Fparams.kon = 0.1;
Params.ConnectParams.costMatF2Fparams.koff = 0.0001;
Params.ConnectParams.costMatF2Fparams.maxWvSearchDist = [];
Params.ConnectParams.costMatF2Fparams.wvJump = [];
%%% set parameters for gap closing %%%
Params.ConnectParams.costMatGCparams.timeWindow = 10;
Params.ConnectParams.costMatGCparams.funcName = 'costMatCloseGapsDensityM';
Params.ConnectParams.costMatGCparams.density = [];
Params.ConnectParams.costMatGCparams.D = [0.01 0.01];
Params.ConnectParams.costMatGCparams.maxSearchDistPerFrame = [3 3];
Params.ConnectParams.costMatGCparams.maxSearchDist = [10 10];
Params.ConnectParams.costMatGCparams.minTrackLen = 2;
Params.ConnectParams.costMatGCparams.kon = 0.1;
Params.ConnectParams.costMatGCparams.koff = 0.0001;
Params.ConnectParams.costMatGCparams.maxWvSearchDist = [];
Params.ConnectParams.costMatGCparams.wvJump = [];
Params.TrackFunction = 'obj.makeTrack'; % standard two stage tracking call.
```

Supporting References

- Holtzer, L., and T. Schmidt. 2010. The Tracking of Individual Molecules in Cells and Tissues. In: Bräuchle C, DC Lamb, J Michaelis, editors. Single Particle Tracking and Single Molecule Energy Transfer. WILEY-VCH Verlag GmbH.
- 2. Wieser, S., and G.J. Schütz. 2008. Tracking single molecules in the live cell plasma membrane-Do's and Don't's. Methods. 46: 131–40.
- 3. Qian, H., M.P. Sheetz, and E.L. Elson. 1991. Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. Biophys. J. 60: 910–21.
- 4. Saxton, M.J. 1993. Lateral diffusion in an archipelago. Single-particle diffusion. Biophys. J. 64: 1766–80.
- De Keijzer, S., A. Sergé, F. van Hemert, P.H.M. Lommerse, G.E.M. Lamers, H.P. Spaink, T. Schmidt, and B.E. Snaar-Jagalska. 2008. A spatially restricted increase in receptor mobility is involved in directional sensing during Dictyostelium discoideum chemotaxis. J. Cell Sci. 121: 1750–7.
- Saxton, M.J. 1997. Single-particle tracking: the distribution of diffusion coefficients. Biophys. J. 72: 1744–53.
- 7. Michalet, X. 2010. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. Phys. Rev. E. 82: 041914.
- 8. Berglund, A. 2010. Statistics of camera-based single-particle tracking. Phys. Rev. E. 82: 011917.
- 9. Montiel, D., H. Cang, and H. Yang. 2006. Quantitative characterization of changes in dynamical behavior for single-particle tracking studies. J. Phys. Chem. B. 110: 19763–70.
- Simson, R., E.D. Sheets, and K. Jacobson. 1995. Detection of temporary lateral confinement of membrane proteins using single-particle tracking analysis. Biophys. J. 69: 989–93.
- 11. Meilhac, N., L. Le Guyader, L. Salomé, and N. Destainville. 2006. Detection of confinement and jumps in single-molecule membrane trajectories. Phys. Rev. E. 73: 011915.
- Elliott, L.C.C., M. Barhoum, J.M. Harris, and P.W. Bohn. 2011. Trajectory analysis of single molecules exhibiting non-brownian motion. Phys. Chem. Chem. Phys. 13: 4326–34.
- Low-Nam, S.T., K. a Lidke, P.J. Cutler, R.C. Roovers, P.M.P. van Bergen En Henegouwen, B.S. Wilson, and D.S. Lidke. 2011. ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. Nat. Struct. Mol. Biol. 18: 1244–1249.
- 14. Smith, C.S., N. Joseph, B. Rieger, and K.A. Lidke. 2010. Fast, single-molecule localization that achieves theoretically minimum uncertainty. Nat. Methods. 7: 373–5.
- 15. Brauchle, C., D.C. Lamb, and J. Michaelis. Single Particle Tracking and Single Molecule Energy Transfer. Weinheim: WILEY-VCH Verlag GmbH.