

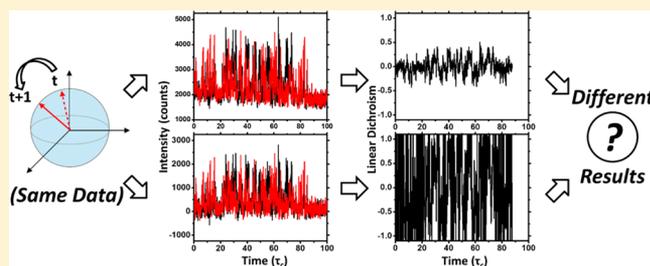
Extraction of Rotational Correlation Times from Noisy Single Molecule Fluorescence Trajectories

Dat Tien Hoang, Keewook Paeng, Heungman Park, Lindsay M. Leone, and Laura J. Kaufman*

Department of Chemistry, Columbia University, New York, New York 10027, United States

Supporting Information

ABSTRACT: Monitoring single molecule probe rotations is an increasingly common approach to studying dynamics of complex systems, including supercooled liquids. Even with advances in fluorophore design and detector sensitivity, such measurements typically exhibit low signal to noise and signal to background ratios. Here, we simulated and analyzed orthogonally decomposed fluorescence signals of single molecules undergoing rotational diffusion in a manner that mimics experimentally collected data of probes in small molecule supercooled liquids. The effects of noise, background, and trajectory length were explicitly considered, as were the effects of data processing approaches that may limit the impact of noise and background on assessment of environmental dynamics. In many cases, data treatment that attempts to remove noise and background were found to be deleterious. However, for short trajectories below a critical signal to background threshold, a thresholding approach that successfully removed data points associated with noise and spared those associated with signal allowed for assessment of environmental dynamics that was as accurate and precise as would be achieved in the absence of noise.



Systems in which contiguous nanoscale regions display heterogeneity in structure and/or dynamics are common in biology and in materials science. Single molecule approaches have proven powerful in directly investigating and detailing the distribution of structure and/or dynamics in such environments.^{1–6}

The most commonly employed single molecule techniques rely on fluorescence from high quantum yield endogenous or exogenous probes. Despite significant advances in available fluorophores and detector sensitivity, single molecule fluorescence measurements typically display low signal intensities relative to noise and background. This problem is particularly acute in frontier areas for single molecule measurements, such as in live cells and in systems in which total internal reflection approaches cannot be implemented. Poor signal to noise may affect the ability to localize a single molecule's fluorescence as is required in localization and super-resolution microscopy and may also complicate analysis of time-dependent data, including in determining evolving position and orientation of mobile molecules.

Single molecule fluorescence microscopy has previously been implemented to image probe molecules in molecular supercooled liquids.⁷ Supercooled liquids are thought to be dynamically heterogeneous, with molecules just nanometers apart displaying dynamics that differ substantially despite the absence of overt structural heterogeneity.^{8–12} To detail the degree to which supercooled liquids are dynamically heterogeneous and the length and time-scales over which these heterogeneities exist and persist, techniques that probe molecular length scales are critical. As such, single molecule

studies hold promise for elucidating key questions about supercooled liquids.

Most small molecule supercooled liquids have refractive indices and exist in temperature ranges such that implementing total internal reflection microscopy (TIRF) is impractical. As such, standard epi-fluorescence microscopy has typically been employed. For investigated systems thicker than ≈ 100 nm, experimental signals collected via epi-fluorescence will display poorer signal to background relative to those collected via TIRF. Additionally, in such experiments, a wide-field rather than a confocal approach may be employed to allow simultaneous data collection from many single molecules. Although this allows for good statistics and access to rare events, it may further degrade signal to background in thick samples through loss of confocality.

In molecular and polymeric supercooled liquids as well as in other complex materials such as mesoporous silica materials, single molecule experiments to explore host dynamics and/or structure have primarily monitored probe rotation.^{4,13–32} Such measurements record probe fluorescence in two orthogonal polarizations that fluctuate in an anticorrelated manner in time as the probe rotates in the sample plane. Autocorrelation analysis of those varying intensities is typically used to assess whether individual probe molecules experience multiple dynamic environments, as well as to extract individual probe

Received: July 13, 2014

Accepted: August 23, 2014

Published: August 23, 2014

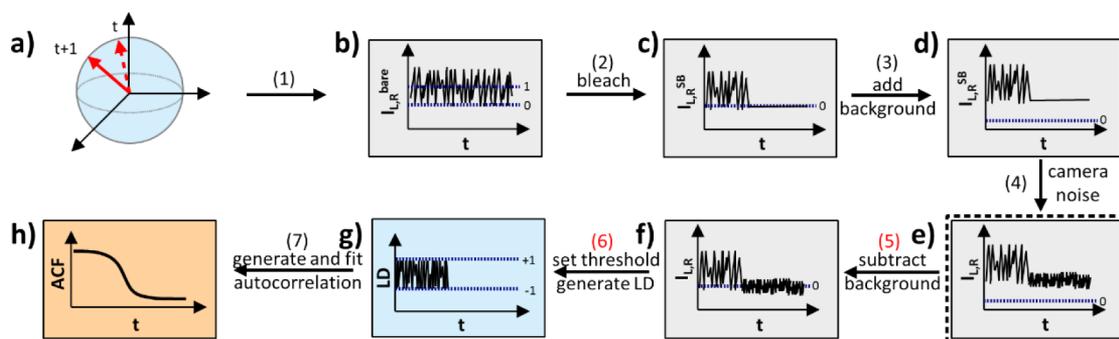


Figure 1. Approach to simulation, data construction, and data analysis. Molecules demonstrating homogeneous rotational diffusion were simulated as an ensemble of vectors, representing molecular transition dipole moments, performing random walks on a sphere. (1) From a vector's orientation, relative fluorescence intensity in two orthogonal polarizations was calculated. (2) A simulated bleach was added at 25 or 75 τ_{set} to the simulated trajectory by setting intensities in both channels to zero after this point. (3) A background signal B was added by vertically shifting the trajectories in the y -direction. (4) Camera noise was added to the signal and background. At this point, as represented by the trajectory in the dashed rectangle, data construction was complete and steps forward represent choices in data processing depicted by red numbers and discussed in this paper. These steps include (5) background subtraction and (6) setting a threshold before calculating the linear dichroism and (7) its autocorrelation.

molecule's average rotational correlation times that reflect the local dynamics of the surrounding host.³³

Previously, simulations have been employed to assess how various experimental parameters, such as frame rate and trajectory length, affect the accuracy of rotational correlation time determination.^{17,34–36} Such studies revealed that for typical experimental parameters, length of the recorded trajectory was the dominant factor in the precision and accuracy of rotational correlation time determination. Indeed, for typical experimental trajectories, it was found that individual molecules experiencing homogeneous rotational diffusion could be mistaken for molecules experiencing heterogeneous dynamics, and an ensemble of molecules experiencing identical homogeneous rotational dynamics could be mistaken for a heterogeneous population.^{34–36}

In some previous studies of the effect of experimental parameters on extracted correlation times, noise-free trajectories were employed whereas in others, experimentally reasonable noise was added to the simulated trajectories. However, no study to date has considered how particular types and degree of noise may affect extracted information. Here, we explicitly considered a number of sources of noise and data processing effects on the extraction of dynamics through autocorrelation function analysis. The optimal manner to treat experimentally realistic single molecule fluorescence data is presented. In particular, we find that for short trajectories with low signal to background, applying a specific thresholding technique yields optimal results. Moreover, this approach does not have deleterious effects on any trajectories explored. These techniques are relevant for any SM measurement in which signal to noise is low and autocorrelation analysis is employed.

EXPERIMENT DETAILS

Although this work largely explored simulated data, to validate the simulated data, comparisons with experiment were made. The experiment was performed on *N,N'*-bis(2,5-*tert*-butylphenyl)-3,4,9,10-perylene dicarboximide (tbPDI, MW = 799.96 g/mol) in an *ortho*-terphenyl (OTP) host as previously described.¹⁸ In brief, the sample was spin-coated onto a native oxide covered silicon wafer from a solution of 5.0 mg/mL OTP in toluene impregnated with tbPDI such that the resultant concentration of tbPDI in OTP was $\approx 10^{-10}$ M. The sample was then placed into a cryostat and cooled to 254.5 K, 11.5 K above

the glass transition temperature of OTP. Data was collected using a standard microscope in epi-fluorescence configuration and an air-objective (NA = 0.75). On the detection side, a Wollaston prism (Karl Lambrecht MWQ12-2) was used to resolve orthogonal polarizations of fluorescence. An electron-multiplying charge coupled device (EMCCD) camera (Andor iXon DV887) recorded data at a frame rate of 5 Hz, resulting in ≈ 12 frames per median probe rotational correlation time.

Simulation Details. An ensemble of 500 molecules displaying homogeneous rotational dynamics was the subject of these simulations. Figure 1 displays the way in which simulated data was constructed to best match experimental data and then analyzed as if it were experimental data.

Rotational diffusion for each molecule was modeled as a random walk of a vector on a sphere whose direction of rotation was chosen arbitrarily and magnitude of rotation was chosen from a Rayleigh distribution (Figure 1a). The vector represents a fluorescent probe molecule's transition dipole. The characteristics of this random walk were given by a diffusion coefficient related to the rotational correlation time, τ_{set} set equal to 100 steps. Trajectories of 10 000 time steps (or 100 τ_{set}) at either 100 points and 20 points per τ_{set} were simulated, with the latter more similar to most reported experimental conditions.

The vector orientation was then used to calculate the expected fluorescence intensity in two orthogonal polarizations as collected by an objective lens and measured by separate detectors following relations proposed by Fourkas.³⁷ For simulations shown here, the numerical aperture was set to 0.75 to reflect experimental conditions. The two generated trajectories were then scaled by dividing each by its mean value, giving $I_L^{\text{bare}}(t)$ and $I_R^{\text{bare}}(t)$, as shown in Figure 1b.

To investigate the importance of trajectory length in these simulations, as well as to test methods to identify photobleaching in experimental data, an explicit bleach was added to the simulations (Figure 1c). Fluorescence intensities of these simulated molecules were left as nonzero until they had progressed to either 25% or 75% from the beginning of the 100 τ_{set} length trajectories. After this time, the fluorescence intensity was set to zero. This results in trajectory lengths of 25 τ_{set} or 75 τ_{set} respectively, which are consistent with previously reported experimental results.³³

Setting Signal to Background and Adding Noise. In ideal conditions, when a probe rotates perpendicular to the sample plane, blinks, or bleaches, a signal of zero is expected; however, this is not seen experimentally. Instead it is nonzero, due to photons from sources other than the probes of interest, camera readout noise, and dark current. We define this intensity as background (B). Because intensity of fluorescence of a particular polarization for a molecule with a 1D transition dipole undergoing rotation will vary between the background level for a molecule fully out-of-plane and a maximal value associated with an in-plane molecule of particular orientation, we defined signal to background (SB) as the average intensities before and after photobleaching via

$$SB = \frac{\overline{I}(t < t_{\text{bleach}})}{\overline{I}(t > t_{\text{bleach}})} \quad (1)$$

To generate a simulated signal with similar noise and background characteristics to the experimental signal, several steps were taken. First, the bare intensity trace was divided into two orthogonal components $I_{L,R}^{\text{bare}}(t)$ for left and right channels via the Fourkas calculation³⁷ (Figure 1b), and then both intensity trajectories were scaled point-by-point by a constant A (Figure 1c). The background constant B was then added to the entirety of the trajectory (Figure 1d).

$$I_{L,R}^{\text{SB}}(t) = AI_{L,R}^{\text{bare}}(t) + B \quad (2)$$

The scaling constant A is determined by eq 1 to be $B(SB - 1)$. Therefore, the noise-free intensity trajectories, $I_{L,R}^{\text{SB}}(t)$, were expressed as

$$I_{L,R}^{\text{SB}}(t) = B(SB - 1)I_{L,R}^{\text{bare}}(t) + B \quad (3)$$

At this point, the proper signal to background ratio was in place but no noise had been added. As an approximation, the camera noise (encompassing uncertainty due to photon counting, EM gain and readout noise) was generated by adding to each point in the trajectory a value picked from a Gaussian distribution, $P(\sigma)$ with a mean of 0 and standard deviation (σ) of the square-root of twice that point's intensity, as shown in eq 4³⁸

$$I_{L,R}(t) = I_{L,R}^{\text{SB}}(t) + P(\sigma) \quad (4)$$

where $\sigma = (2I_{L,R}^{\text{SB}}(t))^{1/2}$. The generated intensity trajectories were then scaled to be in units of electrons, as typically used in data acquisition programs for EMCCD cameras.

To achieve intensity trajectories that were similar to those obtained experimentally, initial guess B and SB values were chosen based on examination of experimental trajectories, and B and SB values were iterated until the distribution of linear dichroisms,

$$LD = \frac{I_L - I_R}{I_L + I_R}, \quad (5)$$

the observable typically analyzed in measurements of this type, from simulation matched those from experiment. A B value of 180 and SB values of 1.2–1.3 matched experiment very well in intensity and linear dichroism trajectories as well as in linear dichroism distribution (Figure 2). We note that although SB is a measure sometimes used to quantify the quality of single molecule data,³¹ the values of B and SB used here can also be translated to the more traditional measure of signal to noise via eq 6:

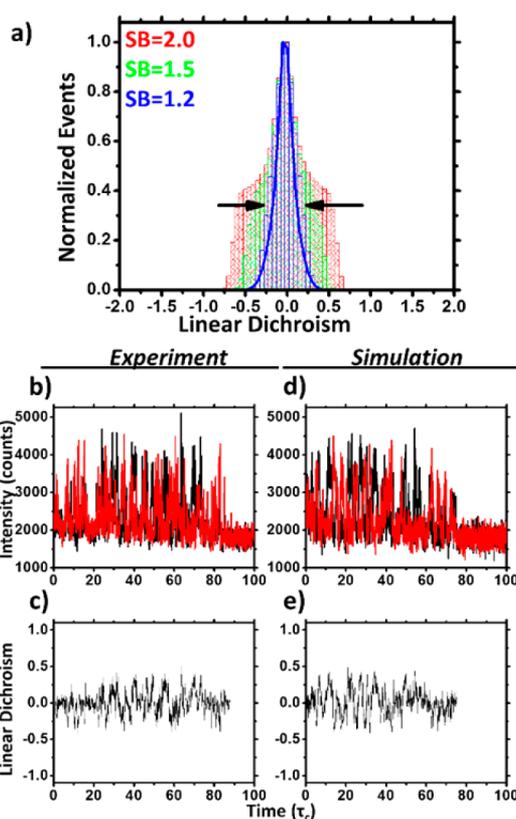


Figure 2. (a) Linear dichroism distributions for nonbackground subtracted trajectories as a function of SB for $B = 180$. The distribution becomes narrower as SB is decreased. The solid blue line corresponds to an experimental linear dichroism distribution. (b) Intensity trajectories in L and R polarizations and (c) linear dichroism trajectory for a representative tbPDI molecule in OTP at 254.5 K. (d,e) Analogous data generated by simulation with $B = 180$ and $SB = 1.3$.

$$SN = SB \sqrt{\frac{B}{2}} \quad (6)$$

For the values of $B = 180$ and $SB = 1.2$, $SN = 11$.

As can be seen in Figure 2a, decreasing SB values at fixed B led to a clear decrease in LD distribution width. Conversely, for a fixed SB, changes in B did not strongly affect the LD distribution (Figure S1, Supporting Information). This occurs because the SB value plays the dominant role in setting the signal relative to the background that constricts the LD compared to its full theoretical range spanning -1 to $+1$.

RESULTS AND DISCUSSION

Data Treatment: Background Subtraction. Following the addition of camera noise, the simulated trajectories bore striking resemblance to measured trajectories, as can be seen in Figure 2. Steps forward from this point (as shown in Figure 1f–h and Figure 3) represent possible approaches to data treatment and analysis to maximize accuracy and precision of extracted rotational dynamics.

In treating wide-field epi-fluorescence data in the past, background subtraction has been employed.³³ Specifically, average intensity in a circle surrounding the imaged feature to be analyzed was used to approximate the background. Removing this background allowed for the calculated linear dichroism to span the range of -1 to $+1$ expected and removed possible effects due to slowly fluctuating background over time

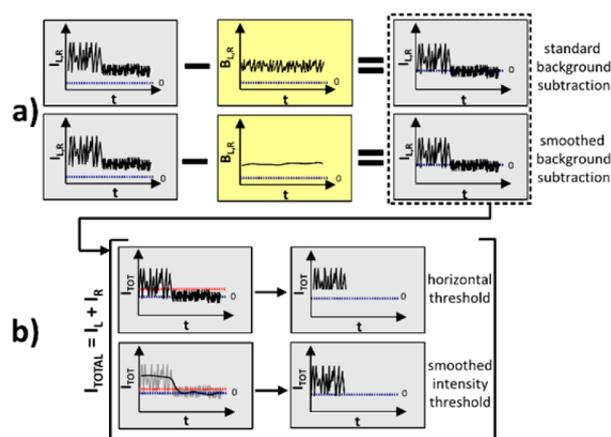


Figure 3. Data treatment approaches. (a) Some generated simulated intensity trajectories were subjected to background subtraction, as also depicted in Figure 1f. Background signal was separately generated and then subtracted from signal or smoothed and subtracted from signal. (b) After possible background subtraction, some trajectories had a horizontal or smoothed thresholding approach applied, as described in the text.

and/or inhomogeneous illumination of the sample. However, background subtraction following this scheme was imperfect, as was evidenced by the presence of background-subtracted intensities that were negative and LD values outside the range of $-1 \leq LD \leq 1$. To approximate the inaccuracies inherent in background subtraction as performed on experimental data, in previous simulations noise including that “due to background subtraction” was added to simulated trajectories, with 30% noise on the mean intensity considered appropriate.¹⁷

To explicitly reflect the experimental procedure for background subtraction, in this work, background signal was simulated separately, having the same camera noise and EM gain as the signal. This background trajectory was then subtracted from the signal, which includes its own background (Figures 1f and 3a). The added and subtracted background trajectories differed because camera noise was chosen from a Gaussian distribution at each point in both the signal and background trajectories. In this paper, data was processed either with no background subtraction, standard background subtraction, or smoothed background subtraction, in which a running average of 50 frames was generated before subtracting the background signal.

Figure 4 shows standard background subtracted LD distributions for the same trajectories presented without background subtraction in Figure 2. As with the nonbackground subtracted versions in Figure 2a, simulations with $B = 180$ and $SB = 1.2$ reproduced the LD distribution seen in a typical experiment as shown in Figure 4a. As expected, imperfect background subtraction led to LD values outside the expected range of $-1 \leq LD \leq 1$. For $B = 180$ and SB varied as in Figure 2, the general trend was opposite that seen in the absence of background subtraction: here the smaller SB signals at fixed B gave rise to the broader distributions with an increased number of unphysical LD values. This occurred because when a small signal sits on a large noise floor, errors in removing that floor, which can only be approximated, are more likely. At the B and SB values that most closely resembled our experimental data (blue histogram in Figure 4a), we note that in addition to a broad distribution of LD values, there was also an accentuation of the peak at $LD = 0$ relative to that of the

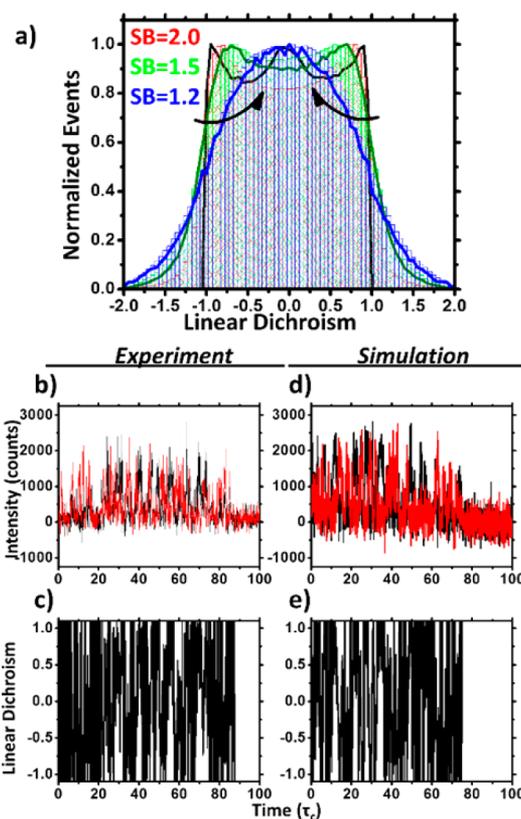


Figure 4. (a) Linear dichroism distributions for background subtracted trajectories as a function of SB and B . Colors of histograms correspond to those in the main panel of Figure 2a. The blue line is the background subtracted LD distribution of experimental data and, as in Figure 2a, matches with the $B = 180$, $SB = 1.2$ simulation very well. The black line is the analytical result for a noise-free simulation, and the solid green line corresponds to a situation in which noise due to background subtraction was estimated as 30% of the mean intensity of the signal. Results represented by black and green lines were also reported in reference 17. (b) Background-subtracted intensity trajectories in L and R polarizations and (c) linear dichroism trajectories for a representative tbPDI molecule in OTP at 254.5 K. (d,e) Analogous data generated by simulation with $B = 180$ and $SB = 1.3$.

noise-free analytical signal (black line in Figure 4a). This is reflective of the intensities being distributed over a more confined range when SB is small.

The noise-free analytical result in Figure 4a (as was also shown in ref 17) looks similar to the highest SB situation shown ($SB = 2.0$, red histogram), but does not closely resemble either background subtracted (blue line in Figure 4a) or nonbackground subtracted (blue line in Figure 2a) experimental signal. The LD distribution for a set of simulations previously explored, in which background noise was not considered explicitly but was instead approximated, is also shown in Figure 4a (green line).¹⁷ This LD distribution looks similar to the $B = 180$ and $SB = 1.5$ situation (green histogram in Figure 4a) but distinct from the lower SB situation that closely resembles data presented here.

Data Treatment: Thresholding. Following photobleaching, experimentally measured trajectories contain only background signal and noise. Automated identification of the time of bleaching is important because it allows facile calculation of the prebleaching trajectory length, information that is important in the assessment of accuracy of extracted dynamics

from LD ACFs.^{34–36} Identifying bleaching time also allows data points after the bleach, composed only of background and noise that may degrade ACF quality, to be excluded from the ACF. To identify bleaching time in an automated fashion, intensity thresholding, which eliminates points from trajectories below a certain intensity level, has been performed. While this is meant to exclude points after bleaching, it may exclude other points in the trajectory as well. In this work, we investigated the effects of thresholding as depicted in Figures 1g and 3b.

When no thresholding was employed, the entire intensity trajectories were used to form the linear dichroism trajectory. In horizontal thresholding (upper panels in Figure 3b), the total intensity from both fluorescence polarizations acted as the reference trajectory. A horizontal value greater than the highest value of this reference trajectory after photobleaching as determined by visual inspection was chosen as the threshold. The end of the trajectory was judged to occur at the last 10 continuous points of the reference trajectory above the threshold, and all points thereafter were excluded. To attempt to remove points associated with relatively long-lived low intensity events such as photoblinking and persistent out-of-plane orientation, points in which less than a tenth of the reference trajectory's points were above threshold over the next 250 frames were also excluded. For smoothed intensity thresholding (lower panels in Figure 3b), the same procedure was followed, but a running average of the total intensity served as the reference trajectory. This approach led to more points before the photobleach being retained.

Data Treatment: Constructing and Fitting the Autocorrelation. In cases in which thresholding was performed, following this step, the intensity trajectories were converted to LDs. For those trajectories in which background was subtracted, LD values outside the expected range of $-1 \leq \text{LD} \leq 1$ may appear (Figure 4). In low signal to noise situations, as are commonly encountered in experiment, the proportion of LD values outside the expected range may be substantial, affecting the quality of the constructed ACF, to the extent that including all LD points may prevent construction of a fittable ACF. Thus, these values were set to +1 or -1 in experimental data analysis and in simulations. For experimental and simulated trajectories without explicit background subtraction, no points outside the expected range were present and no constriction of the LD trajectory was performed.

Following this, LD trajectories were autocorrelated via

$$\text{ACF} = C(t) = \frac{\sum_{t'} a(t')a(t' + t)}{\sum_{t'} a(t')^2}$$

where $a(t) = \text{LD}(t) - \langle \text{LD}(t) \rangle$ (7)

and then fit to stretched exponential decays given by

$$C(t) = Ae^{(-t/\tau_{\text{rot}})^\beta} \quad (8)$$

using a least-squares fitting routine with the constraints $0.3 \leq \beta \leq 2.0$ and $0.3 \leq A \leq 1.2$. The fit was performed on the portion of the ACF from $t = 0$ until $C(t)$ decayed to 0.1. If the ACF did not have at least five points in which it was greater than its uncertainty (propagated from the standard deviation of LD values for that time lag), the ACF was deemed unfittable. For all fit ACFs, the average rotational correlation time, τ_c , was extracted via $\tau_c = (\tau_{\text{rot}}/\beta)\Gamma(1/\beta)$. Although it was shown previously that for very short trajectories linear fitting of the ACFs returned more accurate τ_c values,³⁹ for simplicity, all

fitting in this paper was done to stretched or compressed exponential decays. Although in these simulations all dynamics were homogeneous and thus would be expected to be purely exponential, the systems in which these analyses have been performed are suspected of dynamic heterogeneity, as can be captured with stretched exponential decays ($\beta < 1$), with the deviation of β below one indicating degree of heterogeneity. Short trajectories such as those analyzed here and measured in experiments may be best fit by nonexponential decays even when homogeneous dynamics are present for statistical reasons.⁴⁰

Rotational Correlation Time and Stretching Exponent Accuracy and Precision. As mentioned above, short trajectories of molecules exhibiting homogeneous rotational diffusion with a given rotational correlation time will not necessarily return the known correlation time and a β value of 1. This is in accordance with the expectation that the zero filling effect will dominate the estimate of the correlation function for short trajectories.³⁵ Fluctuations in the correlation function that consequently arise at long lag times affect the stretched exponential fitting, leading to deviation of τ_c and β values obtained from the true values.³⁵

Distributions of obtained τ_c and β values for simulations of noise- and background-free trajectories are shown as gray histograms in Figures 5 and S2 (Supporting Information). All

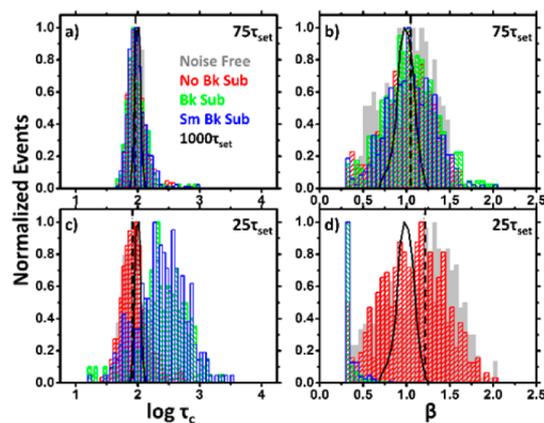


Figure 5. Log τ_c and β distributions of a simulated data set obtained with no thresholding. Distributions for noise-free trajectories are shown in solid gray, with median values of those distributions depicted by dashed lines. Histograms obtained from long trajectories ($1000 \tau_{\text{set}}$) are shown in black. Red, green, and blue correspond to nonbackground subtracted, background subtracted, and smoothed background subtracted trajectories, respectively. Upper (lower) panels show the distributions for $75 \tau_{\text{set}}$ ($25 \tau_{\text{set}}$) trajectories. The same data is shown in separate panels in Figures S2 and S3 (Supporting Information) for clarity.

results presented in this section are for simulations with 20 points/ τ_{set} and all results discussed were also found to hold at 100 points/ τ_{set} . Consistent with previous results, the distributions were quite broad, with a standard deviation of 0.11 and 0.18 for $\log(\tau_c)$ and 0.29 and 0.38 for β for the 75 and 25 τ_{set} trajectory lengths, respectively. The distributions for a 1000 τ_{set} trajectory, typically outside the range accessible by experimental measurements, are also shown as black lines in Figure 5. These distributions are much narrower than for the shorter trajectories but still show noticeable finite trajectory length effects.

We next considered how $\log(\tau_c)$ and β distributions change for 75 and 25 τ_{set} trajectories for systems with signal and background characteristics that were similar to those exhibited by experimental data ($B = 180$, $SB = 1.5$) as a function of data treatment approaches shown in Figure 3. The discussion here particularly focuses on $SB = 1.5$ because it was found that data processing often profoundly affected extracted data for $SB \leq 1.5$. In addition to the noise- and background-free results described above, Figure 5 also shows histograms for the 75 and 25 τ_{set} trajectories without thresholding and with either no, regular, or smoothed background subtraction. For the 75 τ_{set} trajectories, the distributions of $\log(\tau_c)$ were all quite similar to each other, as well as to the noise-free limit. The β distributions were also quite similar regardless of the choice of background treatment; no background subtraction, standard background subtraction and smoothed background subtraction. For the 25 τ_{set} trajectories, only the case with no data treatment (no thresholding and no background subtraction) returns τ_c and β distributions that approach those of the noise- and background-free simulations. For both standard and smoothed background subtraction, the resulting ACFs appear very stretched (see example in Figure S6, Supporting Information). Examination of individual ACFs reveals that due to inclusion of the photobleached portions of the trajectory, the ACFs start at a very low value, limiting the decay range to be fit. These slowly decaying ACFs were preferentially best fit with stretched exponentials with low β values that in turn returned anomalously high τ_c values. Although it is not obvious from the normalized plots shown in Figure 5, the total number of LD ACFs that were fit also changed as a function of data treatment. For the 25 τ_{set} trajectories with no data treatment (no background subtraction and no thresholding), all 500 ACFs were fit; however, only 233 of the possible 500 ACFs were fit when background subtraction was employed.

Figure 6 shows the same set of background subtraction approaches for data in which horizontal thresholding was applied. Relative to the no thresholding approach results shown

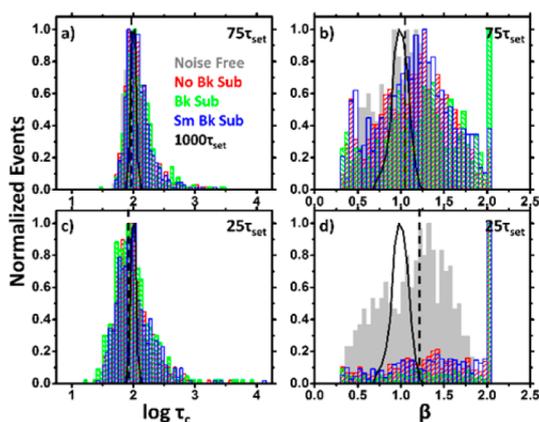


Figure 6. $\log \tau_c$ and β distributions of the same simulated data set in Figure 5 obtained with horizontal thresholding. Distributions for noise-free trajectories are shown in solid gray, with median values of these distributions depicted by dashed lines. Histograms obtained from long trajectories ($1000 \tau_{\text{set}}$) are shown in black. Red, green, and blue correspond to nonbackground subtracted, background subtracted, and smoothed background subtracted trajectories, respectively. Upper (lower) panels show the distributions for 75 τ_{set} (25 τ_{set}) trajectories. The same data is shown in separate panels in Figures S2 and S4 (Supporting Information) for clarity.

in Figure 5, the distribution of obtained $\log(\tau_c)$ and β values deviated more relative to the noise- and background-free result, with a long time tail in the $\log(\tau_c)$ distribution and a very wide range of β values, with a particular abundance of ACFs best fit with $\beta > 1.0$. This effect is more obvious in the 25 τ_{set} trajectories, though here the distributions do not vary much with particular background subtraction technique unlike in the no thresholding cases shown in Figure 5. The increased width of the distributions and propensity for high β values in the fits to ACFs of data with horizontal thresholding is related to loss of points within the trajectory that occurs upon thresholding. Although the threshold was put in place to remove points that were likely due solely to noise, the fact that excluding these points affected the measured τ_c and β distributions in a manner that pushed them from the noise-free result shows that all these points should not be excluded from the trajectories. The thresholding that removed these points resulted in gaps in the resulting trajectories, in turn yielding ACFs in which points at both early and long time lags were less well sampled than expected for a trajectory of given length. This resulted in an earlier onset of oscillations in the ACF and a tendency for the ACF decays to yield compressed exponential fits (see example in Figure S6, Supporting Information). Data analysis employing smoothed intensity thresholding did not suffer from the same problems as horizontal intensity thresholding because it retained many more points before the photobleach. As shown in Figure 7, smoothed intensity thresholding appeared robust

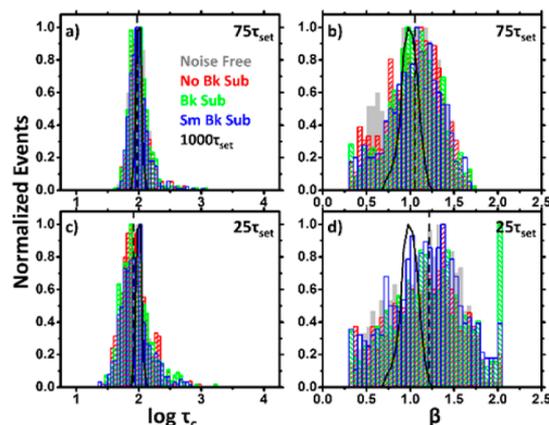


Figure 7. $\log \tau_c$ and β distributions of the same simulated data set in Figure 5 obtained with smoothed thresholding. Distributions for noise-free trajectories are shown in solid gray, with median values of these distributions depicted by dashed lines. Histograms obtained from long trajectories ($1000 \tau_{\text{set}}$) are shown in black. Red, green, and blue correspond to nonbackground subtracted, background subtracted, and smoothed background subtracted trajectories, respectively. Upper (lower) panels show the distributions for 75 τ_{set} (25 τ_{set}) trajectories. The same data is shown in separate panels in Figures S2 and S5 (Supporting Information) for clarity.

across the different background subtraction approaches, returning τ_c and β distributions that were similar to each other and relatively well matched to the noise-free results for the 75 τ_{set} trajectories. Unlike for the no thresholding case, for the 25 τ_{set} trajectories all background subtraction approaches also yielded reasonable τ_c and β distributions (Figure 7c,d).

Given that different SM experiments will have different signal to background characteristics, in Figure 8, we compared a subset including the most successful combinations of thresh-

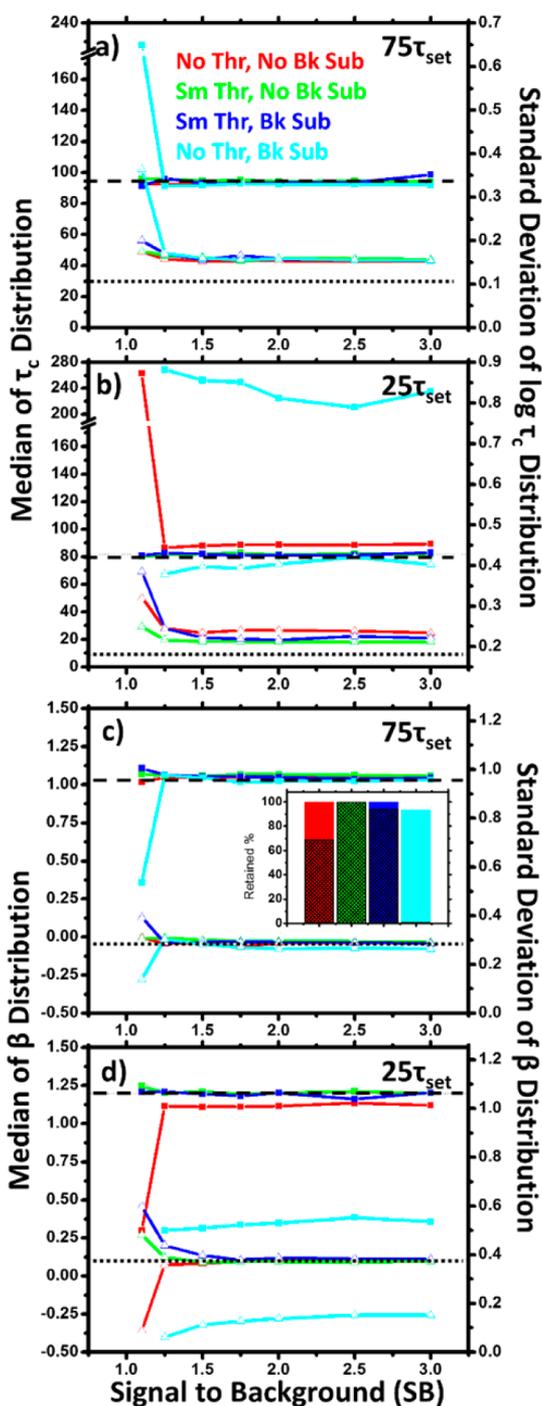


Figure 8. Median (filled squares, left axis) and standard deviations (open triangles, right axis) of (a,b) $\log(\tau_{\text{set}})$ and (c,d) β distributions as a function of SB at $B = 180$ for several data treatment approaches. Panels a and c are for trajectories that are photobleached at $75 \tau_{\text{set}}$ and panels b and d for those photobleached at $25 \tau_{\text{set}}$. In all panels, data treatment techniques are (red) no thresholding nor background subtraction, (green) smooth thresholding and no background subtraction, (blue) smoothed thresholding and background subtraction, and (cyan) no thresholding and background subtraction. In all cases, the black dashed (dotted) line represents median (standard deviation) for noise-free trajectories of that length. Inset shows the percentage of ACFs fit for all data processing techniques for both the $75 \tau_{\text{set}}$ (solid) and $25 \tau_{\text{set}}$ (mesh) trajectories at $SB = 1.1$. No ACFs were fit for the $25 \tau_{\text{set}}$ trajectories with no thresholding but background subtraction.

olding and background subtraction approaches for trajectories with a fixed B value and varying SB from 1.0 to 3.0. In the absence of any treatment of the data (no background subtraction, no thresholding), for $75 \tau_{\text{set}}$ trajectories, the obtained median and standard deviation of distributions of both $\log(\tau_c)$ and β were quite similar to those expected in the absence of noise at all values of SB. This was true of the other approaches shown as well until the lowest SB probed ($SB = 1.1$), where there was an increase in standard deviation for all approaches and a clear failure of the approach in which neither background subtraction nor thresholding was performed. For $25 \tau_{\text{set}}$ trajectories, greater deviations from the noise-free limits were evident in all approaches, particularly at low SB, though the background subtraction approach without thresholding failed at all signal to noise levels, presumably because for these short trajectories the long noise tail following the photobleach coupled with the background subtraction that introduces additional noise to the early portion of the trajectory resulted in ACFs that were dominated by noise rather than the rotational dynamics of the probe. Additionally, unlike for the longer trajectories, performing ACF analysis with neither background subtraction nor thresholding yielded poor results, both limiting the number of ACFs that were fit (inset in Figure 8c) and decreasing accuracy and precision of extracted τ_c and β values. Approaches employing smoothed thresholding did not suffer as acutely even at the lowest SB values explored.

CONCLUSIONS

Previously, it was shown that experimentally realistic trajectory length affects the ability to accurately and precisely obtain time constants and stretching exponents from autocorrelations of single molecule linear dichroism trajectories. In part because autocorrelation analysis ameliorates effects of noise, whether explicit noise and background further degrade ability to obtain τ_c and β values and judge the homogeneity of single molecule dynamics had not previously been investigated. Here, we showed through simulations that data processing methods meant to exclude noise from trajectories may have deleterious effects. In particular, we demonstrated that for trajectories of sufficient length and above a critical signal to background level, data processing in the form of background subtraction and/or thresholding does not improve the accuracy and precision of extracted τ_c and β values relative to unprocessed data. In shorter trajectories and at low SB values, where signal and noise cannot be easily distinguished, background subtraction followed by smoothed thresholding provided the optimal results. Although initial assessment of each trajectory can be used to direct data treatment choices on a trajectory to trajectory basis, using smoothed thresholding for all trajectories may be preferable to analyzing unprocessed data even for longer trajectories with higher signal to background. This approach does not show deleterious effects for any trajectories explored here and allows for automated identification of bleaching time, which, in turn, allows for facile exclusion of points after the bleach as well as identification of relevant trajectory length.

ASSOCIATED CONTENT

Supporting Information

Six supplementary figures as described in the text are presented. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*L. J. Kaufman. E-mail: kaufman@chem.columbia.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under grant numbers CHE 0744322 and CHE 1213242 as well by NSF Graduate Research Fellowships to D.T.H. and L.M.L.

■ REFERENCES

- (1) Michaelis, J.; Brauchle, C. *Chem. Soc. Rev.* **2010**, *39*, 4731–4740.
- (2) Ye, F. M.; Collinson, M. M.; Higgins, D. A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 66–82.
- (3) Kulzer, F.; Xia, T.; Orrit, M. *Angew. Chem., Int. Ed.* **2010**, *49*, 854–866.
- (4) Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. *Annu. Rev. Biochem.* **2008**, *77*, 51–76.
- (5) Woll, D.; Braeken, E.; Deres, A.; De Schryver, F. C.; Uji-i, H.; Hofkens, J. *Chem. Soc. Rev.* **2009**, *38*, 313–328.
- (6) Veigel, C.; Schmidt, C. F. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 163–176.
- (7) Kaufman, L. J. *Annu. Rev. Phys. Chem.* **2013**, *64*, 177–200.
- (8) Ediger, M. D. *Annu. Rev. Phys. Chem.* **2000**, *51*, 99–128.
- (9) Qiu, X. H.; Ediger, M. D. *J. Phys. Chem. B* **2003**, *107*, 459–464.
- (10) Reinsberg, S. A.; Qiu, X. H.; Wilhelm, M.; Spiess, H. W.; Ediger, M. D. *J. Chem. Phys.* **2001**, *114*, 7299–7302.
- (11) Wang, C. Y.; Ediger, M. D. *J. Phys. Chem. B* **1999**, *103*, 4177–4184.
- (12) Richert, R. J. *Phys.: Condens. Matter* **2002**, *14*, R703–R738.
- (13) Benninger, R. K. P.; Onfelt, B.; Neil, M. A. A.; Davis, D. M.; French, P. M. W. *Biophys. J.* **2005**, *88*, 609–622.
- (14) Quinlan, M. E.; Forkey, J. N.; Goldman, Y. E. *Biophys. J.* **2005**, *89*, 1132–1142.
- (15) Ennaceur, S. M.; Hicks, M. R.; Pridmore, C. J.; Dafforn, T. R.; Rodger, A.; Sanderson, J. M. *Biophys. J.* **2009**, *96*, 1399–1407.
- (16) Empedocles, S. A.; Neuhauser, R.; Bawendi, M. G. *Nature* **1999**, *399*, 126–130.
- (17) Mackowiak, S. A.; Herman, T. K.; Kaufman, L. J. *J. Chem. Phys.* **2009**, *131*, 244513.
- (18) Leone, L. M.; Kaufman, L. J. *J. Chem. Phys.* **2013**, *138*, 12A524.
- (19) Mackowiak, S. A.; Leone, L. M.; Kaufman, L. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1786–1799.
- (20) Zondervan, R.; Kulzer, F.; Berkhout, G. C. G.; Orrit, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 12628–12633.
- (21) Hinze, G.; Basche, T.; Vallee, R. A. L. *Phys. Chem. Chem. Phys.* **2011**, *13*, 1813–1818.
- (22) Adhikari, A. N.; Capurso, N. A.; Bingemann, D. *J. Chem. Phys.* **2007**, *127*, 114508.
- (23) Mei, E.; Tang, J. Y.; Vanderkooi, J. M.; Hochstrasser, R. M. *J. Am. Chem. Soc.* **2003**, *125*, 2730–2735.
- (24) Pramanik, R.; Ito, T.; Higgins, D. A. *J. Phys. Chem. C* **2013**, *117*, 3668–3673.
- (25) Seebacher, C.; Hellriegel, C.; Brauchle, C.; Ganschow, M.; Wohrle, D. *J. Phys. Chem. B* **2003**, *107*, 5445–5452.
- (26) Jung, C.; Schwaderer, P.; Dethlefsen, M.; Kohn, R.; Michaelis, J.; Brauchle, C. *Nat. Nanotechnol.* **2011**, *6*, 86–91.
- (27) Kirstein, J.; Platschek, B.; Jung, C.; Brown, R.; Bein, T.; Brauchle, C. *Nat. Mater.* **2007**, *6*, 303–310.
- (28) Jung, C.; Hellriegel, C.; Michaelis, J.; Brauchle, C. *Adv. Mater.* **2007**, *19*, 956–+.
- (29) Liao, Y.; Yang, S. K.; Koh, K.; Matzger, A. J.; Biteen, J. S. *Nano Lett.* **2012**, *12*, 3080–3085.
- (30) Dickson, R. M.; Norris, D. J.; Tzeng, Y. L.; Moerner, W. E. *Science* **1996**, *274*, 966–969.
- (31) Moerner, W. E.; Fromm, D. P. *Rev. Sci. Instrum.* **2003**, *74*, 3597–3619.
- (32) Chung, I. H.; Shimizu, K. T.; Bawendi, M. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 405–408.
- (33) Paeng, K.; Kaufman, L. J. *Chem. Soc. Rev.* **2014**, *43*, 977–989.
- (34) Mackowiak, S. A.; Kaufman, L. J. *J. Phys. Chem. Lett.* **2011**, *2*, 438–442.
- (35) Lu, C. Y.; Vanden Bout, D. A. *J. Chem. Phys.* **2006**, *125*, 124701.
- (36) Wei, C. Y. J.; Lu, C. Y.; Kim, Y. H.; Bout, D. A. V. *J. Fluoresc.* **2007**, *17*, 797–804.
- (37) Fourkas, J. T. *Opt. Lett.* **2001**, *26*, 211–213.
- (38) Robbins, M. S.; Hadwen, B. J. *IEEE Trans. Electron Devices* **2003**, *50*, 1227–1232.
- (39) Bingemann, D. *Chem. Phys. Lett.* **2006**, *433*, 234–238.
- (40) Wei, C. Y. J.; Kim, Y. H.; Darst, R. K.; Rossky, P. J.; Vanden Bout, D. A. *Phys. Rev. Lett.* **2005**, *95*, 173001.