A Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: Application to simulated and in vitro data

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Analytical Chemistry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>ac-2011-034375.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>12-Mar-2012</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Guo, Syuan-Ming; Massachusetts Institute of Technology, Chemistry  
                          He, Jun; Massachusetts Institute of Technology, Biological  
                          Engineering  
                          Monnier, Nilah; Massachusetts Institute of Technology, Biological  
                          Engineering  
                          Guangyu, Sun; National University of Singapore, Chemistry and  
                          Centre for Bioimaging Sciences  
                          Wohland, Thorsten; National University of Singapore, Chemistry and  
                          Centre for Bioimaging Sciences  
                          Bathe, Mark; MIT, Biological Engineering |
A Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: Application to simulated and in vitro data

Syuan-Ming Guo¹, Jun He¹, Nilah Monnier¹, Guangyu Sun², Thorsten Wohland², Mark Bathe¹*

¹Laboratory for Computational Biology & Biophysics, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
²Department of Chemistry and Centre for Bioimaging Sciences, National University of Singapore, 117543 Singapore

KEYWORDS fluorescence correlation spectroscopy, multiple hypothesis testing, Bayesian inference, statistical accuracy, diffusion

*Corresponding Author:

Mark Bathe
77 Massachusetts Avenue
Building 16, Room 255
Cambridge, MA 02139, USA
Tel. (617) 324-5685
Fax (617) 324-7554
mark.bathe@mit.edu
ABSTRACT

Fluorescence Correlation Spectroscopy (FCS) is a powerful approach to characterizing the binding and transport dynamics of macromolecules. The unbiased interpretation of FCS data relies on the evaluation of multiple competing hypotheses to describe an underlying physical process under study, which is typically unknown a priori. Bayesian inference provides a convenient framework for this evaluation based on the temporal autocorrelation function (TACF), as previously shown theoretically using model TACF curves\(^1\). Here, we apply this procedure to simulated and experimentally measured photon-count traces analyzed using a multi-tau correlator, which results in complex noise properties in TACF curves that cannot be modeled easily. As a critical component of our technique, we develop two means of estimating the noise in TACF curves based either on multiple independent TACF curves themselves or a single raw underlying intensity trace, including a general procedure to ensure that independent, uncorrelated samples are used in the latter approach. Using these noise definitions, we demonstrate that the Bayesian approach selects the simplest hypothesis that describes the FCS data based on sampling and signal limitations, naturally avoiding over-fitting. Further, we show that model probabilities computed using the Bayesian approach provide a reliability test for the downstream interpretation of model parameter values estimated from FCS data. Our procedure is generally applicable to FCS and image correlation spectroscopy, and therefore provides an important advance in the application of these methods to the quantitative biophysical investigation of complex analytical and biological systems.
INTRODUCTION

Fluorescence Correlation Spectroscopy (FCS) is an established technique to measure molecular dynamics ranging from binding to transport in living and non-living samples\textsuperscript{2-6}. As with all measurements, interpretation of FCS data relies on an inference process in which multiple competing hypotheses are tested for the ability to describe the physical process under study, which is generally assumed to be unknown \textit{a priori}, followed by interpretation of parameter values of the inferred model. Controlled molecular or other perturbations to the system then provide insight into the mechanistic origin of the inferred model, as well as its parameter values.

Broad application of FCS to biological and other complex systems has in part been hindered by sample heterogeneity\textsuperscript{7-9} and measurement limitations that include low signal-to-noise ratios\textsuperscript{10}, non-stationarity of the physical process under study, and photobleaching\textsuperscript{11}, which render objective interpretation of FCS data challenging\textsuperscript{12-14}. As a first step towards overcoming these obstacles in order to apply FCS broadly to complex chemical systems, we recently developed a multiple hypothesis-testing framework that is based on Bayesian inference that enables the automatic, objective evaluation of multiple competing models in the interpretation of FCS data from Temporal Autocorrelation Functions (TACFs)\textsuperscript{1}. In that work, we simulated analytical TACFs from various transport processes with controlled noise features in order to test systematically the ability of the proposed procedure to detect the underlying physical process in the presence of various types of noise and data limitations that are characteristic of FCS applications. Precise and independent control of these distinct noise and sampling properties of the TACF, in addition to the physical transport process, were essential to enable the systematic evaluation of the proposed theoretical procedure in a controlled setting without the complicating features of real measurements noted above.

In conventional FCS, fluorescence fluctuations are measured in femtoliter volumes at a fixed spatial position (or series of positions) using a sensitive photon counting device such as a Photomultiplier Tube (PMT) or Avalanche Photodiode (APD)\textsuperscript{15-17}. TACFs are subsequently computed from measured photon counts using a hardware correlator. These intensity-derived TACFs have characteristic noise features that cannot be controlled independently because they depend nonlinearly both on various features of the measurement system as well as on the manner in which the correlation function itself is computed. For example, multi-tau correlators result in
correlated, inhomogeneous noise that decays with increasing lag-time\textsuperscript{18-20}, whereas uniform sliding window averaging used to compute the TACF in image correlation spectroscopy results in correlated but largely homogeneous noise\textsuperscript{21}. Existing theoretical models for noise levels in TACFs are derived assuming particular forms of the correlation function such as an exponential or single hyperbolic function and therefore cannot be applied to FCS data from physical processes containing more complex correlations\textsuperscript{18, 20, 22, 23}. Calculating standard deviations of multiple TACFs measured from the same physical process provides one model-free procedure to estimating the noise in TACFs\textsuperscript{20, 24}. However, the large number of repeated measurements required for a reliable estimate of standard deviation is time-consuming and not always feasible for a non-stationary system. Importantly, Schätzel et al.\textsuperscript{19} previously showed that ignoring noise correlations in the analysis of TACFs may result in over-fitting and therefore improper interpretation of FCS data. Notwithstanding, noise correlations are still typically ignored in the analysis and interpretation of TACFs due to the lack of suitable procedures for their estimation.

To apply the theoretical Bayesian framework\textsuperscript{1} to FCS data from fluorescence intensity traces, we investigate here its ability to infer the known physical process of multi-component diffusion based on TACFs computed from simulated and experimental fluorescence intensity traces under various measurement and sampling conditions. As a critical component of the application of our procedure, we present two approaches to estimating the noise magnitude and its correlation from either multiple independent TACF curves or a single photon-count trace. In the latter, we introduce an analysis procedure that ensures that the noise covariance matrix is properly estimated from underlying intensity traces, eliminating potential bias introduced by over-fitting TACF data. Following Wohland et al.\textsuperscript{20}, fluorescence intensity traces are simulated from single-particle diffusion and fluorescence emission and collection from a confocal volume to control and vary systematically fluorophore emission and light collection properties, which is impossible in real experiments. Importantly, TACFs calculated from simulated and real intensity traces have noise properties that cannot be modeled and therefore investigated using our previous approach that applied Bayesian inference to model TACFs with superposed simulated noise\textsuperscript{1}. Subsequently, we apply the procedure to examine the effect of triplet fraction of excited fluorescence molecules and molecular diffusion from experimental samples of aqueous Atto565 and Atto565-labeled Streptavidin solutions. Subsequent developments will focus on the application of the proposed Bayesian approach to biological samples in which additional
important complexities beyond those encountered here are present, including non-stationarity of the underlying physical process, sample heterogeneity, and photobleaching.

**THEORY**

**FCS models**

The TACF of fluctuations in the measured fluorescence intensity is calculated using,

$$G(\tau) = \frac{\langle \delta F(t) \delta F(t + \tau) \rangle}{\langle F(t) \rangle^2}$$  (1)

where $\delta F(t) = F(t) - \langle F(t) \rangle$ is the fluctuation of the measured fluorescence intensity and brackets denote ensemble average, which is equal to the time-average if the system is ergodic\(^{25}\).

For a system of $N_D$ non-interacting fluorescent species undergoing normal diffusion and a 3D Gaussian ellipsoid focal volume with transverse width $w_0$ and height $z_0$, the TACF has the closed-form solution\(^{2,17,26,27}\),

$$G_D(\tau) = \sum_{i=1}^{N_D} \alpha_i \left( 1 + \frac{\tau}{\tau_{D_i}} \right)^{-1} \left( 1 + \frac{\tau}{s^2 \tau_{D_i}} \right)^{-\frac{1}{2}}$$  (2)

where $\alpha_i = (B_i)^2 \langle N_i \rangle / (\sum_i B_i \langle N_i \rangle)^2$ and $\tau_{D_i} = w_0^2 / 4D_i$ are the correlation amplitude and the diffusion time of component $i$, respectively. $s = z_0 / w_0$ is the aspect ratio of the focal volume or the molecule detection function (MDF)\(^{28}\), and $B_i$ is the brightness of component $i$ and is given by $B_i = \sigma_{abs} q_f \kappa q_D$, where $\sigma_{abs}$ is the absorption cross-section of the fluorophore, $q_f$ is the quantum yield of fluorophores, $\kappa$ is the overall transmittance of the optics that accounts for total photon loss due to light transmission and collection of optical components, $q_D$ is the quantum efficiency of the detection device.

In the case of high illumination intensity, the high fraction of molecules being excited to the non-fluorescent triplet state results in additional fluctuations in intensity on a short time scale, which can be described by\(^{29}\),
\[ G(\tau) = \left(1 + \frac{F_{\text{trip}}}{1 - F_{\text{trip}}} \exp\left(-\frac{\tau}{\tau_{\text{trip}}}ight)\right) G_D(\tau), \]  

(3)

where \(F_{\text{trip}}\) is the mean fraction of molecules in the triplet state, and \(\tau_{\text{trip}}\) is the triplet blinking time. \(F_{\text{trip}}\) increases with increasing illumination intensity and can be approximated by \(G_D(\tau)\) when \(F_{\text{trip}} \ll 1\) or \(\tau \gg \tau_{\text{trip}}\).

**Bayesian model selection**

The detailed theory for the Bayesian inference procedure based on TACF curves is described in previous work\(^1\). In that work, various distributions of noise were assumed in order to test the approach generally, whereas here we compute the noise directly from the multi-tau correlator. This results in non-uniform and correlated noise due in part to the sampling time that increases with increasing lag-time, \(\tau\). The data probability given by the parameters \(\beta\) of the model \(f(x, \beta)\) is described by a general multivariate general multivariate normal distribution,

\[
P(y | \beta) = \frac{1}{(2\pi)^{n/2} \sqrt{\text{det}(C)}} \exp\left\{ -\frac{1}{2} [y - f(x, \beta)]^T C^{-1} [y - f(x, \beta)] \right\},
\]

(4)

where \(y = \{G(\tau_1), G(\tau_2), ..., G(\tau_n)\}\) is the data vector of the measured TACF, \(f(x, \beta)\) is the analytical TACF FCS model with the vector of sampled lag time \(x = \{\tau_1, \tau_2, ..., \tau_n\}\) and the parameter vector \(\beta\), and \(C\) is the covariance matrix of the errors in the measured TACF values. Modeling the noise in the TACF with the normal distribution is justified by the central limit theorem because each value in the TACF is the average of a large number of the underlying intensity products (Eq. 1 and Supporting Information, S1). The nonlinear regression process that seeks the maximum-likelihood estimate \(\beta_{\text{MLE}}\) to maximize Eq. 4 is called generalized least squares (GLS), in contrast to weighted least squares (WLS), in which independent errors are assumed\(^30\). The estimation of \(C\) for the TACF is described in later sections.

The probability of model \(k\) given the observed data \(P(M_k | y)\) is given by Bayes’ theorem,
\[ P(M_k|y) = \frac{P(y|M_k)P(M_k)}{P(y)} \propto P(y|M_k), \]  

(5)

where the proportionality holds if the prior probabilities of all the models are equal for all \( k \). Because the normalized model probability is typically reported, in practice we only need to know the marginal likelihood of the data, \( P(y|M_k) \).

\( P(y|M_k) \) in Eq. 5 is obtained by an integral over the parameter space

\[ P(y|M_k) = \int_{\beta} P(y|\beta,M_k)P(\beta|M_k)d\beta, \]

which can be evaluated analytically using the Laplace approximation that is asymptotically exact,

\[ P(y|M_k) = (2\pi)^{n/2} |\Sigma_{Bayes}|^{1/2} P(y|\hat{\beta}_{Bayes},M_k)P(\hat{\beta}_{Bayes}|M_k) \]  

(6)

where \( \hat{\beta}_{Bayes} \) is the maximum of the posterior distribution of parameters \( P(\beta|y,M_k) \), or the Bayesian point estimate, and \( \Sigma_{Bayes} \) is the covariance matrix of \( \beta \) estimated at \( \hat{\beta}_{Bayes} \). When a uniform prior in the parameter space is chosen, \( \hat{\beta}_{Bayes} \) and \( \Sigma_{Bayes} \) are equivalent to \( \hat{\beta}_{MLE} \) and \( \Sigma_{MLE} \).

Use of the Laplace approximation is justified when the posterior distribution of the parameters approaches a Gaussian, as was shown in our previous work to be the case for typical TACF curves\(^1\). Once \( P(y|M_k) \) is computed, model probabilities may be calculated using Eq. 5. Calculation of the probabilities of a set of competing models for describing FCS data thereby enables their direct comparison and evaluation in the presence of measurement noise, data limitations, as well as sample heterogeneity, though we explore only the first two features of FCS data in this work.

**Estimation of noise and noise correlations from multiple TACFs**

Proper estimation of the noise and the noise correlation in the TACF is essential for the calculation of model probabilities using Bayesian inference, as well as for accurate parameter estimation\(^20\). We present two independent approaches to estimating the covariance matrix \( C \) in
Figure 1 Flowchart of Bayesian model selection procedures for FCS data of multiple TACF curves or a single intensity trace. See text for details.
Eq. 4 for distinct types of FCS data (Fig. 1). First, when multiple independent TACF curves measured from the same physical process are available, $\mathbf{C}$ may be estimated by calculating the sample covariance matrix $\mathbf{S}$ using all TACF curves. The matrix elements $S_{kl}$ in the sample covariance matrix $\mathbf{S}$ are given by,

$$ S_{kl} = \frac{1}{J-1} \sum_{j=1}^{J} \left( G^{(j)}(\tau_k) - \bar{G}(\tau_k) \right) \left( G^{(j)}(\tau_l) - \bar{G}(\tau_l) \right) $$

(7)

where $J$ denotes the total number of TACF curves analyzed and $G^{(j)}(\tau)$ denotes an individual curve $j$. The covariance of the mean curve $\bar{G}(\tau)$ from the average of $J$ individual curves is then given by $\mathbf{S}_m = (1/J) \cdot \mathbf{S}$. However, this sample-based estimate of the covariance matrix is only well-conditioned when the number of observations (i.e., the number of available independent TACF curves) is much larger than the dimension of $\mathbf{C}$, which is generally not true for high dimensional data such as TACF curves (typically 128 points). Therefore, we use the regularized shrinkage estimator for covariance matrices to ensure that the estimated matrix is well-conditioned and has lower mean squared error (MSE) than the sample covariance matrix$^{31,32}$. The shrinkage estimator for the covariance is typically the linear combination of the sample covariance matrix and a low dimensional shrinkage target,

$$ \mathbf{S}^* = \lambda \cdot \mathbf{T} + (1 - \lambda)\mathbf{S} $$

(8)

where $\mathbf{T}$ is the shrinkage target, which is usually diagonal, and $\lambda$ is the optimal shrinkage weight associated with the specific shrinkage target. The shrinkage weight $\lambda$ depends on the uncertainty in $\mathbf{S}$, which is determined by the number of samples from which $\mathbf{S}$ is calculated. $\lambda$ decreases monotonically as the number of available samples increases so that $\mathbf{S}^*$ approaches $\mathbf{S}$ asymptotically. While we found that different shrinkage targets all reduce the MSE and yield well-conditioned estimates for the covariance structure of the noise in the TACF, we choose the “two-way” target$^{33}$ here because it preserves the correlation structure and yields smaller uncertainty in model probabilities at the same time (Supporting Information, S2).
**Estimation of noise and noise correlations from a single photon-count trace**

Estimating the noise from multiple TACF curves is only justified when all of the curves come from repeated measurements of the same physical process and the variation across curves is purely from intrinsic statistical error in the correlation function estimator due to the finite sample size. However, in applications of FCS to *in vivo* biological data it may be preferable to evaluate model probabilities directly from the same, single sample. In this case, an alternative approach is to estimate the noise in a single TACF by calculating the noise from the underlying photon-count products\(^{20}\), assuming that the raw photon-count trace or photon arrival time information is available. For the multi-tau correlator, the TACF for a photon-count trace with an acquisition time \(T_{aq}\) is calculated from

\[
G(\tau_k) = G(k\Delta \tau) = \frac{1}{M} \sum_{m=1}^{M} \delta n_m \delta n_{m+k} \tag{9}
\]

with

\[
\bar{n}_j = \frac{1}{M} \sum_{m=1}^{M} n_{m+j} \tag{10}
\]

where \(\Delta \tau\) is the sampling time (or channel width) used to calculate \(G(\tau_k)\), \(n_m\) is the photon count at \(m\Delta \tau\), \(\delta n_m = n_m - \bar{n}_0\), \(\delta n_{m+k} = n_{m+k} - \bar{n}_k\), and \(M = T_{aq} / \Delta \tau - k\) is the number of possible products \(\delta n_m \delta n_{m+k}\). According to Eq. 9, because \(G(\tau)\) is the mean of the underlying photon-count products, the covariance of \(G(\tau_k)\) and \(G(\tau_l)\) may be calculated from these underlying samples \(\{\delta n_m \delta n_{m+k} \delta n_m \delta n_{m+l}\}_{m=1}^{M}\).

Note that calculating the variance or covariance in the photon-count products between time lags requires either independent samples of these products at each \(\tau_k\) or knowledge of the correlations between the samples of these products at each \(\tau_k\), otherwise the noise level will be underestimated. For FCS data, the photon-count products \(\{\delta n_m \delta n_{m+k}\}_{m=1}^{M}\) at any \(\tau_k\) are not
independent and their correlations are given by a fourth-order correlation function\textsuperscript{23, 24}. Here we introduce a procedure to circumvent this problem by performing a block-transformation on the photon-count products to obtain independent samples\textsuperscript{34}. This approach is used extensively in the molecular simulation community to obtain independent estimates from highly correlated molecular trajectories\textsuperscript{35}. The covariance matrix of the mean can then be calculated directly from these transformed, independent samples without the evaluation of the higher order correlation function.

The blocking method seeks the minimal averaging time (or block-time) beyond which the transformed samples (or the block-averages) of the photon-count products are no longer correlated. Defining $p_m(k) = \frac{\delta n_m \delta n_{m+k}}{\langle n \rangle \langle n_k \rangle}$, Eq. 9 becomes,

$$G(k\Delta \tau_i) = \frac{1}{M} \sum_{m=1}^{M} p_m(k).$$

(11)

The block-transformed sample $p'_m(k)$ with the block-time $t_b$ is given by the block-average of all $p_m(k)$ within $t_b$, namely,

$$p'_m(k) = \frac{1}{\nu} \left[ p_{(m-1)v+1}(k) + p_{(m-1)v+2}(k) + \ldots + p_{mv}(k) \right]$$

(12)

where $\nu = t_b / \Delta \tau_i$ is the number of samples in the block. The covariance between $p'_m(k)$ at different $m$ decrease as $t_b$ increases because correlations between the underlying photon count decay over time. Note that Eq. 11 still holds for the transformed samples $p'_m(k)$, so the covariance of $G(\tau_k)$ and $G(\tau_l)$ can be calculated from $p'_m(k)$ as well. When $t_b$ is beyond the minimal block-time needed to obtain independent blocked samples, the matrix elements $S_{kl}$ in the sample covariance matrix $S$ of $G(\tau)$ can be calculated directly from these independent transformed samples:
\[ S_{kl} = \frac{1}{M'(M'-1)} \sum_{m=1}^{M'} (p'_m(k) - \bar{p}'(k))(p'_m(l) - \bar{p}'(l)) \] (13)

where \( M' \) is the number of transformed samples. The minimal block-time depends on the correlation time of the measured physical process and the functional form of its decay. In practice, the minimal block-time can be determined by plotting \( S_{kk} \) at \( \tau_k \) as a function of the block-time \( t_b \), and finding the minimal time after which \( S_{kk} \) becomes invariant to \( t_b \) \(^{34, 35}\).

**RESULTS AND DISCUSSION**

*Effects of acquisition parameters and physical processes on model probabilities when multiple TACF measurements are available*

To evaluate the performance of the proposed Bayesian analysis procedures we first apply Approach 1 (Fig. 1) to TACF curves computed from simulated fluorescence intensity traces of two species with the same brightness undergoing normal diffusion and test their ability to detect the two species for varying conditions of acquisition and physical processes typically encountered experimentally (Supporting Information, S3 and S4). The noise correlation calculated from multiple TACF curves has a structure consistent with the expected noise correlation structure for a TACF curve calculated from a multi-tau correlator (Fig. 2B) \(^{19, 24}\). At small \( \tau_i \), the noise is dominated by uncorrelated photon counting noise due to the short sampling time \( (\Delta \tau_i) \) there. As \( \Delta \tau_i \) increases with increasing \( \tau_i \), the photon count increase and the noise becomes dominated by intensity fluctuations from particle motions, which are highly correlated on the time-scale of diffusion in the focal volume. When \( \Delta \tau_i \) is much larger than the correlation time \( \tau_D \), the noise in this regime becomes uncorrelated again because intensity fluctuations are uncorrelated on this time-scale.

To investigate systematically the effect of variable experimental noise, which can result from low count rate of the particles or insufficient acquisition time, on the ability of the Bayesian approach to identify the number of components present, we varied the optical transmittance \( \kappa \) or the acquisition time \( T \) of the photon-count trace and calculated model probabilities from the resulting TACF curves. Again, the noise levels and correlations in TACFs are estimated using multiple curves as in Approach 1 (Fig. 3A and B). As observed in our previous work\(^1\), the
Figure 2 (A) Generalized least-squares fits of the one- and two-component models to the mean, $\overline{G}(\tau)$, of eight individual TACFs, $G^{(j)}(\tau)$, where each $G^{(j)}(\tau)$ is computed using the multi-tau algorithm from a simulated photon-count trace of two distinct, point-like fluorescent species with the same brightness undergoing normal diffusion. 34 particles with $D_1 = 63.1 \mu m^2/sec$ and 35 particles with $D_2 = 121.8 \mu m^2/sec$ are simulated in a sphere of 3 $\mu m$ radius, corresponding to a mean concentration of 1 nM, with $P = 100 \mu W$, $\lambda = 514.5$ nm, $\sigma_{abs} = 2.2 \times 10^{-12}$ m$^2$, $q_\phi = 0.98$, $\kappa = 0.01$, pinhole diameter 50 $\mu m$, and acquisition time 3.3 sec, corresponding to count rates of $\sim 35$ kHz for each species. Errors bars denote noise in the mean TACF estimated using Approach 1 (see text for details). Normalized residuals from the two fits are shown in the lower panel. (B) Shrinkage estimate of the noise correlation matrix in $\overline{G}(\tau)$ estimated from 64 individual TACF curves calculated from simulated photon-count traces with the same simulation parameters as in (A) except $D_1 = 270 \mu m^2/sec$ and $D_2 = 27 \mu m^2/sec$. The simulation parameters in (B) are used for all simulations in future figures unless otherwise noted.
Figure 3  Bayesian analysis of simulated two-component diffusion using Approach 1. Eight mean TACF curves are fit to obtain each data point, where each mean TACF is the average of eight individual TACF curves generated using the same simulation parameters as in Fig. 2B. The noise correlation matrix is estimated using 64 individual TACFs and regularized using the shrinkage estimator as described in the main text. (A-E) Inferred model probabilities as a function of (A) $\kappa$, (B) acquisition time, (C) minimum sampling time, (D) $\alpha_2/\alpha_1$, and (E) $D_2/D_1$. (F) Diffusion coefficients obtained from the one- and two-component model fits in (E). Black lines represent the prescribed values of $D_1$ and $D_2$ for the two-component model. (G) $\langle N \rangle$ values obtained from the one- and two-component model fits in (E). The black lines represent the prescribed $\langle N \rangle$ values for the two-component model. In all cases, medians and upper and lower quartiles are shown for the eight mean curves analyzed for each data point. In (A) and (B), the lower panels show the relative noise level $\sigma_0/G_0$ in the TACF curves, which is calculated by averaging over the first 16 values in $\sigma(\tau)/|G(\tau)|$. The count rate per particle corresponding to the range of $\kappa$ scanned in (A) is 0.35-350 kHz. A background with count rate of 5 kHz was added to the signal before detection in (A) to avoid zero correlation at low $\kappa$. $\sigma_0/G_0$ is not monotonically decreasing in (A) because the signal at low $\kappa$ is almost from the uncorrelated background and results in noisy TACFs with $D_2/D_1$ in (E) is varied by changing $D_2$ while holding $D_1$ fixed at 63.1 $\mu$m$^2$/sec.
Bayesian procedure resolves the two-component model at low noise levels, but prefers the one-component model when the noise level increases because the additional complexity of the two-component model is not justified there given the uncertainty in the TACF data. Importantly, the three-component model is never preferred, demonstrating that the Bayesian approach does not over-fit the data. While the maximum relative noise level ($\sigma_i/G_0$) that can be tolerated for the Bayesian approach to resolve the two-component model is found here to be $\sim 10\%$, this maximum relative noise level depends on the separation of the diffusion times and the relative amplitudes of the two components.

The minimum sampling time is also expected to eliminate the ability of the Bayesian approach to resolve the two species. When the minimum sampling time exceeds the diffusion time of the fast diffusing component (55 $\mu$s), the fast component is no longer apparent in the TACF curve and the Bayesian approach correspondingly prefers the one-component model (Fig. 3C). While the sampling time is not generally a concern for single-point FCS measurements using APDs or PMTs, limitations in minimum sampling time are of major concern in image-based FCS methods where sampling time is on the scale of milliseconds or longer$^{21, 36-39}$.

Next we examined the effect of varying the underlying physical process by varying the relative brightness $B_i$ of the two species, which alters their relative amplitudes $\alpha_2 / \alpha_1$ (Fig. 3D). Experimentally, $\alpha$ may vary due to differences either in the brightness or the concentration of the particles examined. When the amplitudes of the two components are similar ($\alpha_2 / \alpha_1 \sim 1$), the Bayesian approach is able to resolve the two species (Fig. 3D). In contrast, when the relative amplitude of the two components becomes considerably different from unity, the contribution to the TACF from one of the species is too small to be detected above the noise so that the Bayesian procedure prefers the simpler, one-component model (Fig. 3D). Similar to the effect of varying $\alpha_2 / \alpha_1$, the ability of the Bayesian approach to resolve the two species depends on the ratio of their diffusion coefficients $D_2 / D_1$ (Fig. 3E). When $D_2$ and $D_1$ are approximately equal, the simpler one-component model is preferred because the two components cannot be sufficiently resolved due to their similar diffusion time-scales, as observed previously$^1$. Outside of this regime, the two-component model is preferred because the diffusion time-scales of the two components are well separated and can be distinguished at the given noise level. Again, the over-parameterized three-component model is never preferred. Note that the transition in model
probabilities at $D_2 / D_1 > 1$ is sharper than the transition at $D_2 / D_1 < 1$. This is because the value of $D_1$ is smaller in the latter case, and slower diffusion of particles increases the variability in observed TACF curves given a fixed observation time, increasing noise and reducing the ability to resolve complex models.

The parameter estimates from the two-component model are reliable when the probability of this model is high, and become less reliable when the model probability is low, with greater uncertainty in the latter case (Figure 3F, 3G, and S3). The estimated values of the diffusion coefficients are better predicted than the mean particle numbers when the two-component model probability is low because the system becomes effectively one-component when $D_2 / D_1 = 1$. In this regime, $D_1$ and $D_2$ in the model need to be close to the overall effective diffusion coefficient, whereas the amplitudes $\alpha_1$ and $\alpha_2$ need only to satisfy the constraint $\alpha_1 + \alpha_2 = \alpha$, where $\alpha$ is the overall amplitude. As a general rule, these results suggest that parameter estimates from a complex model should only be interpreted when this model is preferred decisively by the Bayesian inference procedure.

In contrast to accounting for noise correlations in the Bayesian procedure, we find that ignoring noise correlations as performed in conventional analysis of FCS data results in over-fitting due to the underestimation of the effective level of noise in the TACF (Fig. S4). The increase in mean model probabilities of complex models as well as uncertainties of model probabilities indicate noise correlations are over-fit by the extra component in more complex models when these correlations are not accounted for in the fitting process, whereas parameter estimates are more robust to ignoring noise correlations in TACFs (Fig. S5). Thus, for proper interpretation of FCS data, noise correlations should be considered for hypothesis testing of competing FCS models.

**Application of the Bayesian approach to single intensity traces**

When multiple TACF measurements are not available for a system of interest, the noise in a single TACF may be estimated directly from underlying photon-count products using the block-transformation procedure in Approach 2 (Fig. 1). The block-transformation results in increasing noise with increasing block-size $t_b$, indicating the presence of correlation in samples
Figure 4 (A) Estimated noise level in the value of $G(\tau_0)$ at $\tau = 10^{-5}$ sec as a function of block-time $t_b$ for two independent realizations (red and blue) of a simulated photon-count trace with simulation parameters $\kappa = 0.01$, $D_1 = 270 \mu m^2/sec$, $D_2 = 27 \mu m^2/sec$, acquisition time of 26.4 sec, and background = 5 kHz. For both curves, the fixed point is reached at a block-time of approximately 20 msec. Error bars are standard deviations in the noise estimate assuming independent samples. (B) Estimated noise level in the TACF in (A) with and without the blocking method. Block-time is 26.4 msec. (C) and (D) Bayesian analysis of simulated two-component diffusion using Approach 2. Simulation parameters are identical to those in Fig. 3D and E except that each curve has an acquisition time of 26.4 sec so that the overall acquisition times in Fig. 3 and Fig. 4 are the same. (E) Diffusion coefficients obtained from the one- and two-component model fits in (C). (F) $\langle N \rangle$ values obtained from the one- and two-component model fits in (C). Medians and upper and lower quartiles are shown in (C-F) from the results of fitting eight individual TACF curves.
used to estimate the noise level (Fig. 4A). When $t_b \ll \tau_D$, the rate of increase is small because fluctuations on this time-scale result primarily from uncorrelated photon counting noise. However, as the block-time increases to the scale $t_b \sim \tau_D$, the rate of increase is high because particle motion, which results in correlated intensity changes, is now the major source of the fluctuations. The estimated noise increases and approaches the true noise level because the covariance in the samples decreases with increasing $t_b$. Finally, when $t_b \gg \tau_D$ (beyond the correlation time-scale of the particle motions), the transformed photon-count products become nearly independent and further increases in $t_b$ no longer change the estimated noise. Note that the statistical error in the noise estimate also grows with increasing block-size due to fewer samples at larger $t_b$, therefore the choice of the optimal block-time is a trade-off between the bias and the variance in the noise estimate.

In practice, the optimal block-time (fixed point) can be determined by seeking the first point at which the noise estimate is invariant (within statistical error) to increasing $t_b$. However, when the acquisition time $T$ is not considerably longer than the correlation time of the physical process, which in this case is the diffusion time-scale of the slower species, determining the fixed point might be hindered by the large error in the noise estimate due to the insufficient number of transformed samples. The noise level along the TACF estimated before and after the block-transformation using 26.4 ms exhibits a large increase due to the lack of correlations accounted in the samples by the former (Fig. 4B). Importantly, this apparent reduced error in the absence of the blocking procedure may clearly lead to over-fitting complex models to experimental FCS data.

To evaluate the performance of the Bayesian procedure with noise estimated from a single photon-count trace using the block-transformation, we apply Approach 2 to TACFs calculated from simulated intensity traces with varying simulation parameters as was performed in Fig. 3. Approaches 1 and 2 exhibit similar behaviors in terms of model probabilities and parameter estimates when data with equal simulation times are analyzed (Fig. 4C, D and S6). These results demonstrate that proper estimation of the noise and its correlations in the TACF may be obtained using the block-transformation when multiple independent measurements from a sample are not available.
Application to experimental FCS measurements

To test the applicability of the proposed Bayesian approaches to experimental FCS measurements with varying physical process, we analyzed one-component fluorescein solutions with varying excitation laser powers using Approach 1 (Fig. 5A and Supporting Information, S5). The fraction of fluorescein molecules in the triplet state increases with the increase of the excitation laser power (Eq. 3). When the excitation laser power is low, the Bayesian approach prefers the one-component model because the fraction of fluorescein in the triplet state is low. As the laser power increases, the one-component with triplet state model is eventually preferred because the brightness of the dye increases as well as the fraction of dye molecules in the triplet state and the triplet state can be resolved. Again, the over-parameterized two-component with triplet state model is never preferred.

Finally, we tested the ability of the procedure to detect multiple diffusing species at distinct concentration ratios using mixtures of Atto565 and Atto565-labeled Streptavidin (Fig. 5B). For all models, the triplet blinking time was fixed at 1.5 $\mu$s, which is the triplet blinking time of Atto565 measured under the same experimental condition. As the concentration of Atto565-Streptavidin increases, the model probability transitions from the one-component to the two-component model, which is consistent with the simulations analyzed above (Fig. 3D). The above results demonstrate the capability of the Bayesian model procedure to successfully identify the number of components in experimental systems.

CONCLUSION

TACF curves generated from correlated photon-count signal have noise properties that are in general highly correlated, which may lead to over-fitting of complex models when independent noise is assumed. We present two practical approaches to properly estimate the noise and its correlation using either multiple TACF curves or a single photon-count trace, both of which can be implemented easily by FCS practitioners. Estimating noise and its correlation from a single photon-count trace is model-free and thus general for any form of TACF, which will prove useful for applications of the Bayesian approach as well as other hypothesis testing procedures to biological datasets in which multiple measurements from the same process may not be feasible.
Figure 5 Bayesian analysis of experimental FCS data with varying physical process using Approach 1. Five mean TACF curves are fit to obtain each data point, where each mean TACF is the average of two individual TACF curves with acquisition time >20 s calculated by the correlator. The noise correlation matrix is estimated using 10 individual TACFs and regularized using the shrinkage estimator as described in the main text. (A) Inferred model probabilities as a function of laser power from FCS data for fluorescein measured with varying laser power. Each mean TACF is fitted with models of one-component diffusion, one-component diffusion with one-triplet state, and two-component diffusion with one-triplet state. (B) Inferred model probabilities as a function of $\alpha_2/\alpha_1$ from FCS data for mixtures of Atto565 and Atto565-Streptavidin with $\alpha_2/\alpha_1 = \alpha(\text{Atto565-Streptavidin})/\alpha(\text{Atto565})$. Each mean TACF is fitted with one-component-one-triplet, two-component-one-triplet, and three-component-one-triplet models. The triplet blinking time is fixed at 1.5 $\mu$s for all models because the low fraction of molecules in the triplet state cannot be detected by the fitting algorithm. $\alpha_2/\alpha_1$ is always below approximately 0.61 due to the presence of some residual free Atto565 in the ‘pure’ Atto565-Streptavidin sample, which is confirmed by the analysis of the ‘pure’ Atto565-Streptavidin solution alone. Medians and upper and lower quartiles are shown from the results of fitting five mean TACF curves.
We demonstrate that proper estimation of noise correlations allows the Bayesian approach to infer multiple component diffusion models from TACFs under a variety of noise levels and sampling conditions. Consistent with previous work\textsuperscript{1}, complex models are only preferred over simple models when the data resolution (i.e., noise level) justifies the use of additional complexity, whereas simpler models are preferred when noise is high. Importantly, the interpretation of parameter estimates from a complex model is only justified when this model is preferred decisively by the Bayesian inference procedure. Thus, if noise and its correlations are properly estimated then the Bayesian approach provides a pre-screening filter for the downstream interpretation of model parameter values.

Ignoring noise correlations leads to over-fitting data and artificially increases the probability of overly complex models, leading to the over-interpretation of measured data. Thus, noise correlations should be considered generally for the proper analysis of FCS data whenever possible. We also illustrate the capability of the Bayesian approach to identify the correct number of diffusing components in experimental FCS data by analyzing experimental datasets for a two-component system with different component fractions. Incorporating additional models and/or physical processes into the current Bayesian framework is straightforward, including comparison with non-nested models that cannot be included in standard statistical tests, as shown previously\textsuperscript{1}. Thus, the proposed procedure provides a convenient framework in which to interpret FCS data generally in an objective manner, in particular when testing multiple competing hypotheses of biophysical mechanism in complex biological systems in which underlying physical processes are generally unknown. Our approach is available to the broader scientific community at http://fcs-bayes.org.

Automation of the blocking procedure presented will enable the broad application of the proposed hypothesis testing approach to parallel FCS as implemented for example in image correlation spectroscopy (ICS), in which a large number of correlation functions are derived from sub-images of high spatial and temporal resolution image data. Further, non-uniform model and parameter priors may be used, as evidence is accumulated in specific biological systems for the preference of specific models and their parameters in various contexts\textsuperscript{40, 41}. In this context, application of traditional statistics approaches that rely on pair-wise tests are impractical because they do not allow for the direct comparison of multiple non-nested competing models via the generation of model probabilities because they do not condition hypothesis-testing on the models.
themselves, as performed in Bayesian inference\textsuperscript{40, 42, 43}. Future application of the proposed procedure to complex analytical and biological samples will probe the effects of sample heterogeneity, non-stationarity of the underlying physical process, and photobleaching on model and parameter inference.

**ACKNOWLEDGEMENTS**

We are grateful to Korbinian Strimmer (University of Leipzig) for advice on covariance matrix regularization techniques and Jan Ellenberg for helpful discussions. This work was funded by MIT Faculty Start-up Funds and the Samuel A. Goldblith Career Development Professorship awarded to M.B. G.S. gratefully acknowledges a graduate scholarship from the National University of Singapore (NUS). T.W. was funded by a NUS – Baden-Württemberg grant (R-142-000-422-646).

**REFERENCES**


For TOC only

Molecular dynamics
Photon count trace

TACF curve

Noise covariance

Bayesian model selection

Model probabilities

\[ P(M) \]

Model parameters

\[ \beta \]
Supporting Information for

A Bayesian approach to the analysis of fluorescence correlation spectroscopy data II: Application to simulated and in vitro data

Syuan-Ming Guo¹, Jun He¹, Nilah Monnier¹, Guangyu Sun², Thorsten Wohland², Mark Bathe¹*

¹Laboratory for Computational Biology & Biophysics, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
²Department of Chemistry and Centre for Bioimaging Sciences, National University of Singapore, 117543 Singapore

*Corresponding Author:

Mark Bathe
77 Massachusetts Avenue
Building 16, Room 255
Cambridge, MA 02139, USA
Tel. (617) 324-5685
Fax (617) 324-7554
mark.bathe@mit.edu
To validate the assumption of normally distributed noise in the TACF (Eq. 4), we analyzed the distributions of the underlying block-transformed photon-count products in the TACF using increasing block-times. The distribution of the block-transformed photon-count products approaches the distribution of the TACF asymptotically as the block-time increases because the TACF corresponds to the block-transformed photon-count product with the maximum possible block-time (block number = 1). When the block-time is short, however, the distributions of the transformed photon-count products are not normally distributed due to the Poisson statistics of both the photon detection process and the occupation number of particles in the focal volume (Fig. S1A). As the block-time increases, the distributions become normal, as predicted by the central limit theorem (Fig. S1B), and assumed in Eq. 4 for the noise in the TACF. Note that the use of Eq. 13 to estimate the covariance matrix in the normal noise model (Eq. 4) only requires the samples to be independent, but not necessarily normally distributed.
Figure S1 Analysis of the distribution of noise in the TACF using the block-transformation. (A) Distributions of the block-transformed photon-count products at three different lag-times ($\tau = 0.2 \mu s$, $205 \mu s$, $49.2$ ms) are calculated from a simulated photon-count trace with block-time 26.4 ms. The photon-count trace is simulated for two-component diffusion with the same simulation parameters as in Fig 2B except for the simulation time of 26.4 sec. (B) Distributions of the block-transformed photon-count products calculated from the same simulated photon-count trace with block-time of 264 ms. As shown in Figure 4A, the distributions in S1A and S1B are both generated from independent samples.
**S2 Regularized shrinkage estimator for covariance matrices**

The shrinkage estimator is the linear combination of a high dimensional maximum likelihood estimate and a low dimensional shrinkage target with an optimal shrinkage weight associated with the specific shrinkage target. For covariance estimators, the high dimensional maximum likelihood estimate is the sample covariance, and the low dimensional shrinkage target can be a diagonal matrix. A good choice of the shrinkage target should yield the shrinkage estimate with the structure close to the expected structure. To choose the optimal shrinkage target for covariance estimates of noise in TACFs, we explore different shrinkage targets that preserve the characteristic structure of the noise covariance in the TACF\(^1,2\). For target B and D, the shrinkage estimate of the covariance is given by Eq. 8. Target B is a uniform diagonal matrix with the matrix element \(t_{ij}\) and its associated shrinkage weight \(\lambda\) given by

\[
t_{ij} = \begin{cases} v = E(s_{ii}) & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
\]

\[
\lambda = \frac{\sum_{i,j} \text{Var}(s_{ij})}{\sum_{i \neq j} s_{ij}^2 + \sum_i (s_{ii} - v)^2}
\]

where \(s_{ij}\) is the matrix element of the sample covariance matrix \(S\); \(E(\cdot)\) and \(\text{Var}(\cdot)\) denote the sample mean and variance respectively. Target D is the non-uniform diagonal matrix of the variance:

\[
t_{ij} = \begin{cases} s_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}
\]

\[
\lambda = \frac{\sum_{i \neq j} \text{Var}(s_{ij})}{\sum_{i \neq j} s_{ij}^2}
\]

Target D shrinks the covariance but preserves the variance. In contrast, the “two-way” target shrinks the variance and the correlation individually. The matrix element \(s_{ij}^*\) of the shrinkage estimate with the “two-way” target is given by
\[ s_{ij}^* = r_{ij}^* \sqrt{v_i^* v_j^*} \]  

(5)

where

\[ r_{ij}^* = (1 - \lambda_1) r_{ij} \]  

(6)

\[ v_i^* = \lambda_2 \cdot v_{\text{median}} + (1 - \lambda_2) v_i \]  

(7)

\[ \lambda_1 = \min \left( 1, \frac{\sum_{i \neq j} \text{Var}(r_{ij})}{\sum_{i \neq j} r_{ij}^2} \right) \]  

(8)

\[ \lambda_2 = \min \left( 1, \frac{\sum_i \text{Var}(v_i)}{\sum_i (v_i - v_{\text{median}})^2} \right) \]  

(9)

and \( r_{ij} \) is the matrix element of the sample correlation matrix. We examine structures of regularized correlation matrices and the performance of Approach 2 with different shrinkage targets when the underlying photon-count trace is short (Fig. S2). We choose the “two-way” target for the current procedure because it preserves the correlation structure and yields smaller uncertainty in model probabilities at the same time.
(A)

\[ n = 1005, \text{ no regularization} \]

\[ n = 11 \]

no regularization  

two-way

target D  

target B
Figure S2 Comparison of shrinkage estimators for covariance with different targets. (A) Top: sample correlation matrix of noise in the TACF for two-component diffusion with parameters \(\kappa = 0.01\), \(D_1 = 270 \mu m^2 / \text{sec}\), \(D_2 = 27 \mu m^2 / \text{sec}\), and \(T = 26.4\) sec. The covariance matrix is estimated from 1005 independent photon-count products after blocking transformation. Bottom: noise correlation matrices given by different shrinkage estimators for the TACF calculated from only the first 1/80 of the same photon-count trace, which gives 11 independent photon-count products after blocking transformation. (B-E) Model probabilities as a function of acquisition time calculated using Approach 2 with different covariance estimates: (B) “two-way” target (C) diagonal only (variance) (D) target D (E) target B. In all cases, medians and upper and lower quartiles are shown for the eight individual curves analyzed for each data point.
S3 FCS Simulation

Simulation of fluorescent particle diffusion using Brownian Dynamics and simulation of the photo-physics of fluorescence excitation, emission, and light collection from a confocal volume is implemented according to Wohland et al., 2001. Briefly, diffusion of a fixed number of particles corresponding to approximately 1 nM concentration \( N = 69 \) is simulated using a random walk in a sphere of 3 µm radius. A new particle is placed at a random position on the surface of the sphere when a particle leaves the sphere in order to maintain constant number density. The step size of the random walk in each Cartesian coordinate direction \((i = 1, 2, 3)\) is given by \( \Delta x_i = \sqrt{2D\Delta t} \cdot \xi \), where \( D \) is the particle diffusion coefficient, \( \xi \) is a normally distributed random number with unit variance and zero mean, and \( \Delta t \) is the size of the time-step (0.2 µs). Fluorescence excitation by a microscope objective of 63x magnification and 1.2 numerical aperture (NA) is simulated assuming a Gaussian Laser Beam (GLB) profile, with laser power \( P \) (100 µW), wavelength \( \lambda \) (514.5 nm), and characteristic beam width of 261 µm. Fluorescence emission is simulated accounting for the actual position of the particle in the focal volume, its absorption cross-section \( \sigma_{abs} \) \( (2.2 \times 10^{-20} \text{ m}^2) \), and its fluorescence quantum yield \( q_f \) (0.98). The emission photon count \( N_e \) for a particle at \( (r,z) \) from the center of the sphere is thus given by \( N_e = I(r,z) / e_{phot} \cdot \sigma_{abs} \cdot \Delta t \cdot q_f \), where \( I(r,z) \) is the GLB profile in cylindrical coordinates. Effects of photobleaching, saturation, and the triplet state of the fluorophore are ignored in the current simulation.

The mean number of collected photons from a particle depends on the Collection Efficiency Function (CEF) of the pinhole through \( \left\langle N_D \right\rangle = \kappa \cdot N_e \cdot \text{CEF}(r,z) \cdot q_D + N_{bg} \), where \( N_{bg} \) is a constant representing the ambient light, and \( \kappa \) and \( q_D \) were defined above. CEF\( (r,z) \) for a pinhole with diameter of 50 µm is calculated by numerical integration according to Rigler et al. 1993. For all simulations in this work, \( \kappa = 0.01 \) and \( q_D = 0.25 \) if not otherwise specified. \( N_{bg} = 0 \) is used except for the simulation for varying \( \kappa \), in which a constant \( N_{bg} \) corresponding to a count rate of 5 kHz is used to avoid zero correlation at low \( \kappa \). \( w_0 \) and \( z_0 \) of the focal volume, or MDF, are obtained by fitting a 3D-Gaussian function to the resulting MDF, which is given by \( \text{MDF}(r,z) = \text{CEF}(r,z) \cdot I(r,z) \). The simulation is implemented in MATLAB (The MathWorks,
Natick, MA) using the Poisson random number generator \textit{Poisson()} to simulate photon emission with specified mean \(\langle N_D \rangle\). With above simulation parameters, the corresponding count rate per particle is \(\sim 35\) kHz, which is within the typical range of count rates of fluorophores 20-60 kHz.

TACF curves are computed using the multi-tau algorithm\(^6\), which has 128 channels recording fluorescence intensities with a quasi-logarithmic scale time increment; namely, the time increment is doubled after the first 16 channels and then doubled every 8 channels. The counts in two adjacent channels are summed before being passed to the next channel with double time increment (see Wohland et al. 2001 for details).

**S4 Nonlinear regression and calculation of model probabilities**

The noise covariance matrix \(\mathbf{C}\) in Eq. 4 is estimated from multiple independent TACF curves or from the blocked-transformed photon-count products using the regularized shrinkage estimator with the “two-way” target\(^2\) (Fig. 1, Supporting Information S2). In the current work we employ a uniform parameter prior \(P(\boldsymbol{\beta}, M_k)\); therefore \(\hat{\boldsymbol{\beta}}_{\text{Bayes}}\) and \(\Sigma_{\text{Bayes}}\) are equivalent to \(\hat{\boldsymbol{\beta}}_{\text{MLE}}\) and \(\Sigma_{\text{MLE}}\), which can be computed using standard nonlinear regression algorithms. Generalized least squares regression with \(\mathbf{C}\) is performed using the MATLAB function \textit{nlinfit}, which employs the Levenberg–Marquardt algorithm to implement a nonlinear least squares regression assuming uniform, independent error terms, thereby minimizing \(\sum_i[y_i - f(x_i, \boldsymbol{\beta})]^2\). To adapt this approach to handle correlated noise, we note that the negative logarithm of Eq. 4 is proportional to \([y - f(x, \boldsymbol{\beta})]^T \mathbf{C}^{-1} [y - f(x, \boldsymbol{\beta})]\), which can be written as \([y - f(x, \boldsymbol{\beta})]^T (\mathbf{L}^{-1})^T \mathbf{L}^{-1} [y - f(x, \boldsymbol{\beta})]\), where \(\mathbf{L}\) is given by the Cholesky decomposition of the covariance matrix \(\mathbf{C} = \mathbf{LL}^T\). The regression problem can then be transformed to an ordinary least squares problem with the transformed data vector \(\mathbf{L}^{-1} y\) and the transformed vector of the function values \(\mathbf{L}^{-1} f(x, \boldsymbol{\beta})\), which can be solved using the MATLAB function \textit{nlinfit}. All TACFs are fit with the FCS models defined in Eq. 2 and 3 with an additional constant term to account for the random fluctuation in \(G_c\) about zero due to finite acquisition time. The best-fit estimate \(\hat{\boldsymbol{\beta}}_{\text{MLE}}\) and the covariance matrix of the parameters \(\Sigma_{\text{MLE}}\) are then used to compute the model probabilities using Eq. 6. The prior probability \(P(\boldsymbol{\beta}, M_k)\) is taken as a constant that is normalized
to one within the parameter range $2\gamma_f$ times the parameter standard deviation centered at $\hat{\beta}_{\text{Bayes}}$ in each dimension, namely, $\beta \in \bigcap_{i=1}^{k} \left[ \hat{\beta}_{\text{Bayes}}^{(i)} - \gamma_f \sigma_\beta^{(i)} \hat{\beta}_{\text{Bayes}}^{(i)} + \gamma_f \sigma_\beta^{(i)} \right]$, and parameter standard deviations are estimated as the square root of the diagonal terms of $\Sigma_{\text{Bayes}}$. $\gamma_f$ is chosen to be 200 so that the prior range is significantly larger than the range of parameters suggested by the MLE. The effect of different choices of prior on model probabilities was examined in detail in §.

Once $P(y|M_k)$ has been computed for each model, the normalized model probability is given by

$$P(M_k|y) = P(y|M_k) / \sum_{k=1}^{d} P(y|M_k)$$

assuming uniform model priors $P(M_k)$. 
Figure S3 Parameter estimates from the Bayesian analysis of simulations in Fig. 3A-D using Approach 1 (GLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight mean curves analyzed for each data point.
Figure S4 Bayesian analysis of the same simulations as in Fig. 3A-E using Approach 1 without incorporating noise correlations (WLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight mean curves analyzed for each data point.
Figure S5 Parameter estimates from the Bayesian analysis of simulations in Fig. S4 using Approach 1 without incorporating noise correlations (WLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight mean curves analyzed for each data point.
Figure S6 Bayesian analysis of simulations for two-component diffusion with varying simulation parameters using Approach 2 (GLS fitting). All the simulation parameters are identical as in Fig. 3 except that each curve has simulation time 26.4 sec; thus overall simulation times are the same as in Fig. 3. In all cases, medians and upper and lower quartiles are shown for the eight individual curves analyzed for each data point. The top panel in (B) shows the ratio of sample size and the dimension \( \frac{n}{p} \) for the covariance estimate.
Figure S7 Parameter estimates from the Bayesian analysis of simulations in Fig. 4C, 4D and S6A-C using Approach 1 (GLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight individual curves analyzed for each data point.
Figure S8 Bayesian analysis of the simulations in Fig. 4C, 4D and S6A-C using Approach 2 without incorporating noise correlations (WLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight individual curves analyzed for each data point.
Figure S9 Parameter estimates from the Bayesian analysis of simulations in Fig. S8 using Approach 2 without incorporating noise correlations (WLS fitting). In all cases, medians and upper and lower quartiles are shown for the eight individual curves analyzed for each data point.
S5 FCS experiments

The FCS setup was described previously in detail (Pan et al., 2007). Briefly, FCS was performed using a customized Olympus FV 300 confocal microscope (Olympus, Tokyo, Japan). Excitation was provided by the 488-nm laser line of an Argon ion laser that is focused into samples by a water-immersion objective (60×, NA 1.2; Olympus, Tokyo, Japan). The laser power, measured before the objective, was 30 µW. The emitted light, which passed an internal 3x magnification stage in the confocal microscope, passed through a 150-µm pinhole. A 50/50 beam splitter (Thorlabs, Newton, NJ) was used to split the emission light into two channels, to eliminate artifacts of detector afterpulsing in FCS measurement (Chen et al., 2003, Pan et al., 2007). The light then was focused through band-pass filters 595AF60 (Omega, Brattleboro, VT), onto avalanche-photodiode detectors (SPCM-AQR-14; Pacer, Berkshire, UK). Autocorrelation curves were computed online using a hardware correlator (Flex02-01D; Correlator.com, Bridgewater, NJ, USA).

SUPPORTING REFERENCES