Time-Integrated Fluorescence Cumulant Analysis in Fluorescence Fluctuation Spectroscopy

Bin Wu and Joachim D. Müller
School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota

ABSTRACT We introduce a new analysis technique for fluorescence fluctuation data. Time-integrated fluorescence cumulant analysis (TIFCA) extracts information from the cumulants of the integrated fluorescence intensity. TIFCA builds on our earlier FCA theory, but in contrast to FCA or photon counting histogram (PCH) analysis is valid for arbitrary sampling times. The motivation for long sampling times lies in the improvement of the signal/noise ratio of the data. Because FCA and PCH theory are not valid in this regime, we first derive a theoretical model of cumulant functions for arbitrary sampling times. TIFCA is the first exact theory that describes the effects of sampling time on fluorescence fluctuation experiments. We calculate factorial cumulants of the photon counts for various sampling times by rebinning of the original data. Fits of the data to models determine the brightness, the occupation number, and the diffusion time of each species. To provide the tools for a rigorous error analysis of TIFCA, expressions for the variance of cumulants are developed and tested. We demonstrate that over a limited range rebinning reduces the relative error of higher order cumulants, and therefore improves the signal/noise ratio. The first four cumulant functions are explicitly calculated and are applied to simple dye systems to test the validity of TIFCA and demonstrate its ability to resolve species.

INTRODUCTION

Biological systems are heterogeneous. Resolving such heterogeneous mixtures of molecules is an important and challenging problem. Fluorescence fluctuation spectroscopy (FFS) provides a promising tool for studying heterogeneous biological systems. The technique exploits fluorescence intensity fluctuations of molecules that pass through a very small optical observation volume. Each passage of a fluorescent molecule leads to a short burst of detected photons. Collectively, these diffusing molecules give rise to a stochastic fluorescence signal. Statistical analysis tools are required to extract information about the fluorescent molecules from the stochastic signal. The most widely used analysis technique, fluorescence correlation spectroscopy (FCS) (8), uses the autocorrelation function of the fluorescence signal. FCS resolves species based on the diffusion coefficient, which ultimately depends on molecular weight. Simulations, however, point out that FCS is limited to resolving species that differ in their diffusion time by at least a factor of five (10). Thus, many interesting cases, such as a monomer/dimer equilibrium, cannot be directly resolved by FCS.

Photon counting histogram (PCH) (3) and fluorescence intensity distribution analysis (5) are recently developed techniques that determine the molecular brightness and occupation number in the observation volume of the system from the experimental histogram of the photon counts. PCH complements FCS in that it distinguishes species by a difference in their molecular brightness instead of their molecular weight. PCH has been successfully applied to resolve binary dye mixtures (13) and probe protein interactions in living cells (2).

PCH is based on the analysis of the probability distribution function (pdf) of the photon counts. The probability distribution function and all of its moments contain, from a mathematical point of view, identical information (6). Moment analysis has been developed in the late 80s and early 90s to resolve heterogeneous bimolecular samples in FFS experiments (14,15,17,18). However, the potential of this approach has not been fully explored. A further development was the introduction of fluorescent cumulant analysis (FCA) (11). Cumulants are closely related to moments and have mathematical properties particularly suited for random variables. FCA uses simple analytical expressions that relate the factorial cumulants of the photon counts to the molecular brightness and occupation number in the observation volume. FCA successfully resolved binary mixtures according to brightness. We demonstrated that both FCA and PCH are equivalent techniques for analyzing FFS data.

PCH, FCS, and FCA theories explicitly or implicitly assume that the sampling time T is much smaller than the characteristic timescale of fluctuation in the fluorescence signal. However, few photons are collected during a short sampling period and the signal/noise ratio of the data is poor. Longer sampling times are needed to improve the signal, and the theory of PCH and FCA needs to be modified. Finding an exact solution for PCH at long sampling times has been difficult.

Cumulant analysis, in contrast to histogram analysis, allows an exact treatment of sampling times. In this article, we extend our previously developed FCA theory to arbitrary

Submitted March 28, 2005, and accepted for publication June 27, 2005.

Address reprint requests to Bin Wu, University of Minnesota, School of Physics and Astronomy, 116 Church St., SE, Minneapolis, MN 55455. Tel.: 612-624-6045; Fax: 612-624-4578; E-mail: binwu@physics.umn.edu.

0006-3495/05/10/2721/15 $2.00

doi: 10.1529/biophysj.105.063685

sampling times, which significantly increases the signal statistics of the technique. From a fluorescence intensity point of view, short sampling times capture the instantaneous fluorescence I, whereas longer sampling times measure the time-integrated fluorescence. To distinguish our previously developed FCA, which is only valid for short sampling times, we refer to our new theory as time-integrated fluorescence cumulant analysis (TIFCA).

Experimentally, we measure photon counts with a short sampling time. The photon counts are subsequently rebinned to obtain factorial cumulants of the photon counts for longer sampling times. Theoretical models are then fitted to experimental cumulants as a function of binning time. The fit determines simultaneously the molecular brightness, the occupation number, and the diffusion time of each species from a single measurement. Thus TIFCA is capable of distinguishing particles by both molecular brightness and diffusion time. In other words, the technique combines the advantages of both PCH and FCS.

In addition, we formulate and experimentally verify a mathematical model for the statistical error of factorial cumulants for arbitrary sampling times. The relative error of cumulants provides a measure of its statistical significance and provides weighing factors for data fitting. We also discuss the effect of binning time on the relative error of cumulants. Analysis of simple dye mixtures by FCA demonstrates that our theory successfully models the experimental data.

THEORY

We previously described factorial cumulants of photon counts in the limit of short sampling times (11). Here we expand the theory of cumulants to arbitrary binning times. We first introduce the binning function to capture the effect of binning time on the amplitude of cumulants, before deriving expressions of binning time-dependent cumulants. The variance of cumulants, which is needed to perform error analysis on experimental data, is discussed as well.

Binning function

Our starting point is Mandel's formula (9), which relates the probability distribution function $P(W)$ (pdf) of the integrated light intensity W to the pdf $p(k,T)$ of the photon counts k,

$$p(k,T) = \int_0^\infty \mathrm{Poi}(k, \eta W) P(W) dW, \tag{1}$$

where $\mathrm{Poi}(k, x)$ is the Poisson distribution with a mean of x. The parameter η describes the detection efficiency of the photo detector. For convenience, we set $\eta = 1$ and assume a constant spatial intensity profile across the detector area A. As is customary, we will set the area of the detector to 1. With these conventions, the integrated intensity W is expressed in units of photon counts per sampling time. For a given fluorescent light intensity I, W is the integral of I over the sampling time T,

$$W(t) = \int_{-T/2}^{T/2} I(t')dt'. \tag{2}$$

It is well known that the moment-generating function $Q_w(s)$ of W is equal to the factorial moment-generating function $Q_k^f(s)$ of the photon counts k (19). The logarithm of a generating function defines its corresponding cumulant-generating function. Thus Mandel's formula implicitly states that the cumulant-generating function of W equals the factorial cumulant-generating function of k. In other words, the r^{th} factorial cumulants of the photon counts $\hat{\kappa}_{[r]}(k)$ equals the r^{th} cumulants of the integrated intensity $\kappa_r(W)$:

$$\hat{\kappa}_{[r]}(k) = \kappa_r(W). \tag{3}$$

For deriving the binning function it is convenient to treat the special case of a single molecule in the entire sample volume. The result obtained for this special case is later generalized to the case of N particles in an open volume. Let us assume that a single molecule diffuses in a large, but closed volume V, illuminated by a laser with a normalized beam profile given by $\overline{PSF}(\vec{r})$. Without losing generality, let us assume that the point-spread function has a characteristic length w, and that the shape of the point-spread function is determined collectively by a set of parameters α. For example, for a three-dimensional Gaussian beam profile in an s-photon ($s = 1, 2\ldots$) excitation experiment,

$$\overline{PSF}_{3DG} = \exp\left(-\frac{2s(x^2+y^2)}{w^2} - \frac{2sz^2}{w_z^2}\right), \tag{4}$$

the characteristic length is the beam waist w, and the shape parameter is given by the squared beam waist ratio $\alpha = r = w_z^2/w^2$. According to definition (Eq. 2), the r^{th} moment of the integrated intensity,

$$\langle W(T)^r \rangle = \int_{-T/2}^{T/2} \cdots \int_{-T/2}^{T/2} g_r(t_1, t_2, \cdots t_r) dt_1 dt_2 \cdots dt_r, \tag{5}$$

is related to the r^{th} order correlation function g_r of the fluorescence intensity, which is defined as

$$g_r(t_1, t_2, \cdots t_r) = \langle I(t_1) I(t_2) \cdots I(t_r) \rangle. \tag{6}$$

Without losing generality, let us assume that the time is ordered: $t_1 < t_2 < \cdots < t_r$. For a stationary diffusion process, the correlation function is a function of the time differences only, $\tau_2 = t_2 - t_1$, $\tau_3 = t_3 - t_2$,$\cdots$, and $\tau_r = t_r - t_{r-1}$. In Appendix A, we show that g_r can be expressed as

$$g_r(t_1, t_2, \cdots t_r) = \gamma_r \lambda^r \frac{V_{PSF}}{V} G_r(\tau_2, \tau_3, \cdots, \tau_r; \tau_d, \alpha), \quad (7)$$

where λ is the molecular brightness measured in counts per second per molecule (cpsm) and V_{PSF} is the reference volume conventionally used in fluorescence fluctuation experiments,

$$V_{PSF} = \int_V \overline{PSF}(\vec{r}) d\vec{r}. \quad (8)$$

The coefficient γ_r is defined as

$$\gamma_r = \frac{\int_V (\overline{PSF}(\vec{r}))^r dr^3}{V_{PSF}}, \quad (9)$$

and the normalized correlation function G_r is given by

$$G_r(\tau_2, \cdots \tau_r; \tau_d, \alpha) = \frac{\left\{ \int_V \cdots \int_V \overline{PSF}(\vec{r}_1)\overline{PSF}(\vec{r}_2) \cdots \overline{PSF}(\vec{r}_r) \frac{1}{(4\pi D\tau_2)^{\frac{3}{2}}} \exp\left(-\frac{\|\vec{r}_2 - \vec{r}_1\|^2}{4D\tau_2} \right) \cdots \frac{1}{(4\pi D\tau_r)^{\frac{3}{2}}} \exp\left(-\frac{\|\vec{r}_r - \vec{r}_{r-1}\|^2}{4D\tau_r} \right) d\vec{r}_1 d\vec{r}_2 \cdots d\vec{r}_r \right\}}{\int_V \overline{PSF}(\vec{r})^r d\vec{r}}, \quad (10)$$

where D is the diffusion coefficient. It is straightforward to show that $G_r(0,0,\cdots,0;\tau_d,\alpha) = 1$ when $\tau_2 = \cdots = \tau_r = 0$, since

$$\lim_{\tau \to 0} \frac{1}{(4\pi D\tau)^{\frac{3}{2}}} \exp\left(-\frac{\|\vec{r}_2 - \vec{r}_1\|^2}{4D\tau} \right) = \delta(\vec{r}_2 - \vec{r}_1), \quad (11)$$

where $\delta(x)$ is Dirac's δ-function.

Suppose the characteristic length of the point-spread function is w, which means if we define a normalized coordinate $\vec{r}' = (\vec{r}/w)$, the point-spread function $\overline{PSF}(\vec{r}')$ only depends on the shape parameters α, but is independent of w. We change the integration variables $\vec{r} = w \times \vec{r}'$ and define a characteristic diffusion time $\tau_d = (w^2/4sD)$. By inspecting Eq. 10, we observe that G_r depends on τ only through (τ/τ_d), and on the point-spread function only through the shape parameters α. In other words, the normalized correlation function has the scaling property

$$G_r(\tau_2, \tau_3, \cdots, \tau_r; \tau_d, \alpha) = G_r\left(\frac{\tau_2}{\tau_d}, \frac{\tau_3}{\tau_d}, \cdots, \frac{\tau_r}{\tau_d}; 1, \alpha \right). \quad (12)$$

If we plug Eq. 7 into the r^{th} moment of the integrated intensity (Eq. 5), we obtain

$$\langle W(T)^r \rangle = \gamma_r \frac{V_{PSF}}{V} \lambda^r B_r(T; \tau_d, \alpha), \quad (13)$$

where B_r is called the r^{th} binning function, and is defined as

$$B_r(T, \tau_d, \alpha) = r! \int_{0 \le t_1 \le t_2 \le \cdots \le t_r \le T} \cdots \int G(t_2 - t_1, t_3 - t_2, \cdots, t_r - t_{r-1}; \tau_d, \alpha) dt_1 dt_2 \cdots dt_r. \quad (14)$$

Expressing the binning function in the form of Eq. 14 uses the fact that the hyper-cubic integration volume of Eq. 5 can be divided into $r!$ equivalent blocks, where within each region the times $t_1, t_2, \cdots, t_r = 0$ are time-ordered and the integration of the correlation function for each block is identical. We also change the integration range of t_i from $[(-T/2), (T/2)]$ to $[0,T]$ using the stationary property of our correlation function. Now we change the integration variables t_1 to $\tau_i = t_i - t_{i-1}$ for $i > 1$ and express the integration volume using the transformed coordinates. Because the integrand is independent of t_1, it is straightforward to integrate t_1 out. The binning function is now written in the form

$$B_r(T, \tau_d, \alpha) = r! \int_{\substack{\tau_2 \ge 0, \cdots \tau_r \ge 0 \\ \tau_2 + \cdots \tau_r \le T}} \cdots \int G(\tau_2, \tau_3, \cdots, \tau_r; \tau_d, \alpha) \times (T - \tau_2 - \cdots - \tau_r) d\tau_2 \cdots d\tau_r. \quad (15)$$

We change the integration variables to $x_i = \tau_i/\tau_d$ and arrive at the final form of the binning function, which is

$$B_r(T, \tau_d, \alpha) = \tau_d^r r! \int_{\substack{x_i > 0, i=2, \cdots r \\ x_2 + \cdots + x_r \le T/\tau_d}} \cdots \int G(x_2, x_3, \cdots, x_r; 1, \alpha) \times \left(\frac{T}{\tau_d} - x_2 - \cdots - x_r \right) dx_2 \cdots dx_r. \quad (16)$$

By inspecting Eq. 16, we clearly see that the binning function also obeys a scaling law,

$$B_r(T, \tau_d, \alpha) = \tau_d^r B_r\left(\frac{T}{\tau_d}, 1, \alpha \right), \quad (17)$$

which is important in practical application. The calculation of binning functions requires high-dimensional integration, which cannot be solved analytically for arbitrary point-spread functions. In practice, we calculate the binning function numerically. Because of its scaling property, we only need to calculate the binning function for the special case of $\tau_d = 1$. We construct a table of its function values for different T and α. The scaling property (Eq. 17) allows us to calculate the binning function for an arbitrary diffusion time by interpolation of the tabulated values.

However, an analytical expression of the binning function is obtained in the limit of short binning times ($T \ll \tau_d$). By inserting Eq. 11, which is valid in this limit, into Eq. 10, we obtain a normalized correlation function G_r of ~ 1. This

approximation is used to determine the analytical form of the binning function for short binning times,

$$B_r(T;\tau_d,\alpha) = T^r. \tag{18}$$

To calculate binning functions for arbitrary binning times one has to choose a point-spread function. A commonly used point-spread function in fluorescence fluctuation experiment is given by a three-dimensional Gaussian (Eq. 4). We calculate the normalized correlation function up to the fourth order according to Eq. 10,

$$G_2\left(\frac{\tau}{\tau_d};r\right) = \frac{1}{1+\frac{\tau}{\tau_d}}\frac{1}{\sqrt{1+\frac{\tau}{r\tau_d}}},$$

$$G_3\left(\frac{\tau_2}{\tau_d},\frac{\tau_3}{\tau_d};r\right) = \frac{3\sqrt{3}}{\left(4\frac{\tau_2\tau_3}{\tau_d^2}+4\frac{\tau_2+\tau_3}{\tau_d}+3\right)\sqrt{4\frac{\tau_2\tau_3}{r^2\tau_d^2}+4\frac{\tau_2+\tau_3}{r\tau_d}+3}},$$

$$G_4\left(\frac{\tau_2}{\tau_d},\frac{\tau_3}{\tau_d},\frac{\tau_4}{\tau_d};r\right) = \frac{2\sqrt{2}}{\left(4\frac{\tau_2\tau_3\tau_4}{\tau_d^3}+4\frac{\tau_2\tau_3+\tau_3\tau_4+\tau_2\tau_4}{\tau_d^2}+\frac{3\tau_2+4\tau_3+3\tau_4}{\tau_d}+2\right)} \frac{1}{\sqrt{4\frac{\tau_2\tau_3\tau_4}{r^3\tau_d^3}+4\frac{\tau_2\tau_3+\tau_3\tau_4+\tau_2\tau_4}{r^2\tau_d^2}+\frac{3\tau_2+4\tau_3+3\tau_4}{r\tau_d}+2}}. \tag{19}$$

We plug these correlation functions into Eq. 16 and integrate numerically. An analytical expression for the second binning function B_2 (T;τ_d,r) has been constructed earlier (11). In Fig. 1, we plot B_r/T^r up to the fourth order for $\tau_d = 1$ as a function of the binning time. We notice that for $T \ll \tau_d$, the function B_r/T^r goes to 1, as predicted by Eq. 18. We also find that higher order binning functions decay faster than lower order ones.

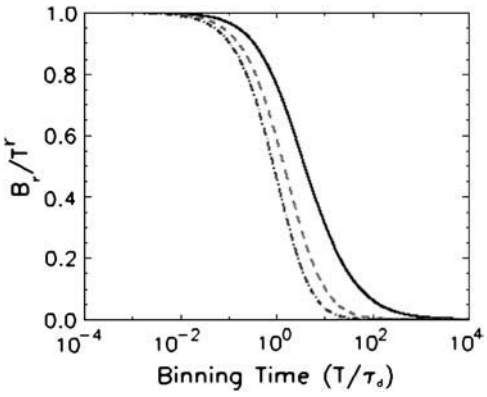

FIGURE 1 Binning function $\kappa_r(T)$ up to fourth order. The function κ_r/T^r is graphed for as a function of binning time T for second order (*solid line*), third order (*dashed line*), and fourth order (*dash-dotted line*). The binning functions are calculated for a diffusion time $\tau_d = 1$ and a squared beam waist ratio $r = 25$.

Cumulants for arbitrary binning time

The r^{th} cumulant of the integrated intensity of a single particle $\kappa_r^{(1)}(W)$ can be expressed as a function of all raw moments up to order r (11). In the thermodynamic limit $V \to \infty$, only the r^{th} moment of the integrated intensity survives,

$$\kappa_r^{(1)} = \langle W^r \rangle. \tag{20}$$

Now we treat the case of more than one molecule in the closed volume V. The r^{th} cumulant for a large number of N_{total} noninteracting, diffusing molecules is given by the sum of the r^{th} cumulant over all molecules. If the molecules are identical, we obtain by using Eq. 13,

$$\kappa_r = N_{total}\,\kappa_r^{(1)} = N_{total}\frac{V_{PSF}}{V}\gamma_r\lambda^r B_r(T;\tau_d,\alpha). \tag{21}$$

Fluctuation experiments measure fluorescence emerging from an open excitation volume, which is much smaller than the total sample volume V. We use the concentration c, which is given by N_{total}/V, to define the average occupation number N in the observation volume V_{PSF}:

$$N = V_{PSF}c = V_{PSF}N_{total}/V. \tag{22}$$

Thus, the r^{th} cumulant of the integrated fluorescence intensity simplifies to

$$\kappa_r(T) = \gamma_r\lambda^r B_r(T;\tau_d,\alpha)N. \tag{23}$$

When the binning time is short, the binning function B_r is approximated by T^r (Eq. 18). Thus the formula for the cumulant (Eq. 23) reduces to the special case for short binning times (11),

$$\kappa_r(T) = \gamma_r(\lambda T)^r N. \tag{24}$$

We rewrite Eq. 23 with the help of Eq. 17 in the following form:

$$\kappa_r(T) = N\gamma_r(\lambda\tau_d)^r B_r(T/\tau_d;1,\alpha). \tag{25}$$

Equations 23 and 25 describe the r^{th} cumulant of the integrated fluorescence intensity of diffusing molecules with a diffusion time τ_d, brightness λ, and an occupation number of N. We have written the cumulant as a function of T to stress that, experimentally, we are interested in the binning time-dependence of the cumulant. Eq. 25 is useful for calculating the cumulant function $\kappa_r(T)$, because every species is described by the same binning function, if the binning time is normalized by the diffusion time. The cumulants of a mixture of noninteracting fluorescent species are given by the sum of the cumulants of each individual species according to

the additive property of cumulants for independent random variables,

$$\kappa_r(T) = \gamma_r \sum_{i=1}^{M} \lambda_i^r B_r(T;\tau_{di},\alpha) N_i. \quad (26)$$

Experimentally we observe photon counts instead of integrated intensities. Equation 3 states that the factorial cumulants of the photon counts are identical to the cumulants of the integrated fluorescence intensity. In other words, the cumulants $\kappa_r(T)$ of the integrated fluorescence intensity are experimentally determined by calculating the factorial cumulants $\hat{\kappa}_{[r]}(T)$ of the photon counts.

Variance of the factorial cumulants

It is important to know the experimental error of factorial cumulants for judging the quality of fits to theoretical models. In statistics, the error is determined by the variance. We use a technique called moments-of-moments (7, 11) to calculate the variance of factorial cumulants. For example, the variances of the first five factorial cumulants are given in Müller (11). In this technique, to calculate the variance of r^{th} order factorial cumulant $\hat{\kappa}_{[r]}$, factorial cumulants up to the $2r^{th}$ order are needed. Here, for convenience, we rewrite the variance of the factorial cumulants $\hat{\kappa}_{[1]}$ and $\hat{\kappa}_{[2]}$ in terms of factorial cumulants (the variance of higher order factorial cumulants $\hat{\kappa}_{[r]}$ is readily calculated symbolically using the statistical software MathStatica (MathStatica, Sydney, Australia)):

$$\begin{aligned}
\mathrm{Var_M}[\hat{\kappa}_{[1]}] &= \frac{1}{n}(\hat{\kappa}_{[1]} + \hat{\kappa}_{[2]}),\\
\mathrm{Var_M}[\hat{\kappa}_{[2]}] &= \frac{1}{n}(2\hat{\kappa}_{[1]}^2 + 2\hat{\kappa}_{[2]} + 4\hat{\kappa}_{[1]}\hat{\kappa}_{[2]} + 2\hat{\kappa}_{[2]}^2 + 4\hat{\kappa}_{[3]} + \hat{\kappa}_{[4]}),
\end{aligned} \quad (27)$$

where n is the number of data points. There is a potential problem, however. It is well known that the higher order moments are notoriously difficult to estimate in experiment. The signal/noise ratio decreases rapidly with increasing order of the cumulant (11). It seems that we cannot determine experimentally the error of high order factorial cumulants accurately since it involves even higher order moments. Fortunately, in our case, the variance of the factorial cumulants is mainly determined by the lower order factorial cumulants. This can be understood as follows. In the short binning time limit, the r^{th} order factorial cumulant is proportional to T^r (11). Therefore, higher order factorial cumulants are much smaller than the lower order ones. For instance, $\mathrm{Var_M}[\hat{\kappa}_{[2]}]$ is mainly determined by $\hat{\kappa}_{[1]}$, $\hat{\kappa}_{[2]}$, and only weakly depends on $\hat{\kappa}_{[3]}$ and $\hat{\kappa}_{[4]}$. For very long binning time, the photon counts converge to a Gaussian distribution, which is solely determined by the first two cumulants. In other words, compared to the first two factorial cumulants, higher order factorial cumulants are not important in determining the error of second order factorial cumulants. The same argument applies to the variance of higher order factorial cumulants. This suggests that we may define a truncated variance $\overline{\mathrm{Var_M}}[\hat{\kappa}_{[r]}]$ for the r^{th} ($r > 1$) order factorial cumulant, which neglects contributions of all factorial cumulants higher than r. We list the truncated variance of factorial cumulant up to third order,

$$\begin{aligned}
\overline{\mathrm{Var_M}}[\hat{\kappa}_{[2]}] &= \frac{1}{n}\Big(2\hat{\kappa}_{[1]}^2 + 2\hat{\kappa}_{[2]} + 4\hat{\kappa}_{[1]}\hat{\kappa}_{[2]} + 2\hat{\kappa}_{[2]}^2\Big),\\
\overline{\mathrm{Var_M}}[\hat{\kappa}_{[3]}] &= \frac{1}{n}\Big(6\hat{\kappa}_{[1]}^3 + 18\hat{\kappa}_{[1]}^2\hat{\kappa}_{[2]} + 54\hat{\kappa}_{[2]}^2 + 6\hat{\kappa}_{[2]}^3 + 6\hat{\kappa}_{[3]}\\
&\quad + 9\hat{\kappa}_{[3]}^2 + 18\hat{\kappa}_{[1]}\hat{\kappa}_{[2]} + 18\hat{\kappa}_{[1]}\hat{\kappa}_{[2]}^2\\
&\quad + 36\hat{\kappa}_{[1]}\hat{\kappa}_{[3]} + 72\hat{\kappa}_{[2]}\hat{\kappa}_{[3]}\Big).
\end{aligned} \quad (28)$$

Another assumption of the moments-of-moments technique is that the experimental data, which in our case are photon counts, are independent of each other. This assumption is not valid for fluctuation experiments. In fact, FCS uses the correlation between photon counts to determine transport properties of particles. We know that the correlation between photon counts decays to zero for very long binning times. However, it is still necessary to investigate its effect on the variance of the factorial cumulants for short and intermediate binning times. As we will show, the effect of correlation is negligible for estimating the variance of the higher order factorial cumulants, but has a strong influence on the first order.

It is difficult to exactly calculate the variance of factorial cumulants of arbitrary order for correlated data. Here we illustrate the calculation of variance for the first two factorial cumulants in the context of fluorescence fluctuation experiments, following the procedure outlined by Qian (17). The first factorial cumulant is just the intensity, which is defined as

$$\hat{\kappa}_{[1]} = \frac{1}{n}\sum_{i=1}^{n} k_i, \quad (29)$$

where k_i is the i^{th} photon counts and n is the total number of data points. The true variance $\mathrm{Var_T}$ of $\hat{\kappa}_{[1]}$ for correlated data is given by

$$\begin{aligned}
\mathrm{Var_T}[\hat{\kappa}_{[1]}] &= \frac{1}{n}\langle (k_i - \langle k_i\rangle)^2\rangle + \frac{2}{n^2}\sum_{i>j}^{n}\langle \Delta k_i \Delta k_j\rangle\\
&= \mathrm{Var_M} + \mathrm{Var_{corr}},
\end{aligned} \quad (30)$$

where $\langle\cdots\rangle$ represents the ensemble average. The first term $\mathrm{Var_M}$ is the one determined by moments-of-moments. The second term $\mathrm{Var_{Corr}}$ contains the contribution due to correlations between photon counts. In Appendix B, we calculate the covariance or correlation in the photon counts $\langle \Delta k(t_i)\Delta k(t_j)\rangle$ for arbitrary binning times (Eq. 45). If we plug Eq. 45 into $\mathrm{Var_{corr}}$, it is written as

$$\begin{aligned}
\mathrm{Var_{corr}} &= \frac{N\gamma_2\lambda^2}{n^2}\sum_{i>j}^{n}[B_2((i-j+1)T) + B_2((i-j-1)T)\\
&\quad - 2B_2((i-j)T)].
\end{aligned} \quad (31)$$

The summation in Eq. 31 can be carried out analytically by rearranging the order of summation. The final result is very simple,

$$\mathrm{Var}_{\mathrm{corr}} = \frac{N\gamma_2\lambda^2}{n^2}[B_2(nT) - nB_2(T)] = \frac{\kappa_2(nT,\tau_d,\alpha)}{n^2} - \frac{\kappa_2(T,\tau_d,\alpha)}{n}. \quad (32)$$

Eq. 32 is exact for all binning and data acquisition times. When the binning time is short, the summation in Eq. 30 can be approximated by transforming it into an integral (17). The correction derived by this approximation contains only the first term of Eq. 32.

The variance of the second factorial cumulant can be calculated similarly. In Appendix C, we work out the correction for the case of short binning times. Since higher order correlation functions decay much faster than the lower order ones, the leading correction comes from Eq. 53. However, it involves fourth order factorial cumulants, and as we argued above, the correction should be small.

MATERIALS AND METHODS

Instrumentation

The experiments are performed on our homebuilt two-photon microscope. A mode-locked Ti:sapphire laser (Tsunami, Spectra-Physics, Mountain View, CA) pumped by an intracavity doubled Nd:YVO4 laser (Millennia Vs, Spectra-Physics, Mountain View, CA) serves as the two-photon excitation source. An excitation wavelength of 780 nm was used for all dye experiments. A 63X Plan-Apochromat oil immersion objective ($NA = 1.4$) is used to focus the laser and collect the fluorescence. The light passes through an optical filter and is detected with an avalanche photodiode (APD) (SPCM-AQ-14, Perkin-Elmer, Dumberry, Québec). The output of the APD, which produces TTL pulses, was directly connected to a data acquisition card (FLEX02, Correlator.com, Bridgewater, NJ). The recorded photon counts were stored and later analyzed with programs written for IDL version 5.4 (RSI, Boulder, CO). A program written in Fortran with a nonlinear least-squares optimization routine from the Port Library (available at http://www.netlib.org) is used to fit the theoretical model to the experimental cumulants.

Sample preparation

Alexa488 (Molecular Probes, Eugene, OR) and Rhodamine 6G (Acros Organics, Morris Plains, NJ) were dissolved in water with 0.02% (by volume) of NP-40 (Sigma, St. Louis, MO). The small amount of detergent was added to prevent Rhodamine 6G from absorbing to the surface of our sample holder.

Data analysis

We use the software MathStatica to derive formulas of factorial cumulants up to the 20th order and the variance of the factorial cumulants up to the 10th order by the technique of moments-of-moments (7). These formulas are implemented into an analysis program written in IDL to calculate the experimental factorial cumulants and their errors.

We rebin the data to determine the factorial cumulants for different sampling times. The procedure is performed as follows: We feed the recorded sequence of photon counts into software to calculate the experimental factorial cumulants of photon counts of sampling time T. To get cumulant for a sampling or binning time of $2T$, we add neighboring photon counts together to get a new sequence of photon counts with binning time $2T$. We apply the same software on the rebinned data to get the cumulants for a binning time $4T$. This process is repeated to calculate the cumulants for binning times of specific integer multiples of T. By rebinning, we calculate the factorial cumulants over binning times that cover three orders of magnitude.

We fit the experimentally determined factorial cumulants $k_{[r]}$ to theoretical cumulants $\kappa_{[r]}$ determined by Eq. 26 with a nonlinear least-squares fitting program. The reduced χ^2 of the fit is given by

$$\chi^2 = \frac{1}{(K-p)}\sum_T \sum_r^{r_0} \frac{(\hat{\kappa}_{[r]}(T) - \kappa_r(T))^2}{\mathrm{Var}[\hat{\kappa}_{[r]}(T)]}. \quad (33)$$

The value of K is the total number of cumulants used in the fit and p is the number of free fitting parameters of the model.

RESULTS

Single dye experiment

To test the TIFCA theory for arbitrary binning times, we perform experiments on simple fluorescent dye solutions. Each species is characterized by three parameters: its molecular brightness λ, the diffusion time τ_d, and the average number of molecules N in V_{PSF}. This information can be obtained by fitting the first two cumulants (11). Here we fit the first four cumulants simultaneously, while allowing γ_3 and γ_4 to be determined experimentally as free parameters. Fig. 2 shows the first four factorial cumulants as a function of binning time for Rhodamine 6G. The data was taken with a sampling time of 5 μs with a total measurement time of 600 s. The reduced χ^2 of the fit is 0.67 with a recovered brightness of $\lambda = 23{,}000$ cpsm, a diffusion time of $\tau_d = 41$ μs, and with $N = 0.72$ molecules in the observation volume. The fitted γ factors are $\gamma_3 = 0.246 \pm 0.004$ and $\gamma_4 = 0.208 \pm 0.012$. Conventional FCS analysis of the fluctuation data yields a diffusion time of $\tau_d = 42$ μs and $N = 0.73$ molecules. These parameters are, within experimental error, identical to the results of TIFCA.

Brightness is a molecular property, which has to be independent of the concentration of the solution. To check this, we perform dilution experiments on a Rhodamine 6G and an Alexa488 sample. A concentrated stock solution of each dye was repeatedly diluted by a factor of 2. We fit the factorial cumulants of all data sets globally by linking their γ factors. The fitted parameters are listed in Table 1. The γ factors returned by the fit with a reduced χ^2 of 2.1 are $\gamma_3 = 0.251$ and $\gamma_4 = 0.223$. We will use the experimentally determined values of γ_3 and γ_4 for all subsequent analysis presented in this article. To determine the concentration ratio of both stock solutions, we graph the number of molecules of Rhodamine 6G as a function of the number of molecules of Alexa488 and fit the data to a straight line (data not shown). The fitted slope of 1:3.3 determines the concentration ratio of both dye solutions. Later, we mix both stock solutions

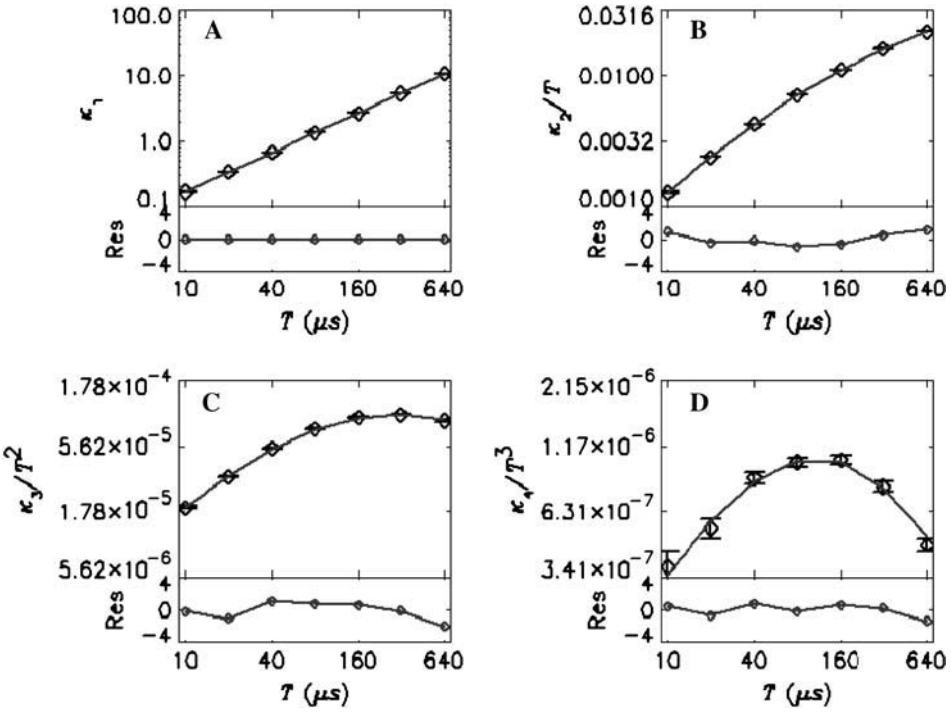

FIGURE 2 FCA analysis of a Rhodamine 6G solution. The data are taken with a sampling time of 5 μs and a total data acquisition time of 10 min. The factorial cumulants are calculated for binning times from 10 μs to 640 μs. The first four cumulants divided by $T^{\tau-1}$ are shown in A–D as diamonds. The best fit to a single species model is shown as a solid line together with the normalized residuals (*diamonds*) for each cumulant in the lower panel with a dashed line to guide the eye. The fit determined a brightness of $\lambda = 23{,}000$ cpsm, a diffusion time of $\tau_d = 41$ μs and an average number of molecules of $N = 0.72$. In the fit γ_2 was fixed to $(2\sqrt{2})^{-1}$, which corresponds to a three-dimensional Gaussian PSF, whereas γ_3 and γ_4 are free parameters. The fit determined $\gamma_3 = 0.245$ and $\gamma_4 = 0.208$.

together and perform a dilution experiment on the binary mixture. We will compare the concentration ratio recovered from the mixture with the one determined from the single species experiments.

Variance of factorial cumulants

We now investigate the variance of the factorial cumulants as a function of binning time. The variance characterizes the experimental uncertainty of measuring factorial cumulants. We first concentrate on the moments-of-moments technique and thus ignore the influence of correlations on the variance for the moment. The relative error $\delta\hat{\kappa}_{[r]}$ of the factorial cumulant $\hat{\kappa}_{[r]}$ is given by $\delta\hat{\kappa}_{[r]} = \sqrt{\mathrm{Var}[\hat{\kappa}_{[r]}]}/\hat{\kappa}_{[r]}$ and is a measure of the noise/signal ratio. We now examine the relative error $\delta\hat{\kappa}_{[r]}$ of the dye data presented in Fig. 2 as a function of binning time T. We first calculate the factorial cumulants up to the eighth order from the data set for binning times from 5 μs to 5.12 ms. Next, we calculate the variance $\mathrm{Var}[\hat{\kappa}_{[r]}]$ up to the fourth order from the factorial cumulants using the moments-of-moments technique (11). The explicit formulas for the first two factorial cumulants are shown in Eq. 27. The relative error $\delta\hat{\kappa}_{[r]}$ of the first four cumulants based on the moments-of-moments technique is shown in Fig. 3 as a function of binning time.

TABLE 1 TIFCA analysis of single dye samples

Rhodamine 6G			Alexa 488		
λ_1 (10^4 cpsm)	τ_{d1} (μs)	N_1	λ_2 (10^4 cpsm)	τ_{d2} (μs)	N_2
2.3	41	0.72	0.91	38	2.3
2.4	40	0.35	0.94	37	1.1
2.4	40	0.17	0.94	37	0.50
2.3	40	0.090	0.94	36	0.24
2.2	40	0.048	0.88	37	0.12

The table lists the fitted parameters of samples containing either Rhodamine 6G or Alexa 488. The solution of each dye is successively diluted by factors of 2 between measurements. The brightness λ, the diffusion time τ_d, and the number of molecules N of each species are shown. A linear fit of N_2 as a function of N_1 reveals a concentration ratio of 1:3.3 between Rhodamine 6G and Alexa 488.

Now, we compare the relative error $\delta\hat{\kappa}_{[r]}$ of the experimental data with the one predicted by theory. We use the fitted molecular brightness λ, diffusion time τ_d, and the number of molecules N to calculate the cumulants according to Eq. 26. The variance of these theoretical cumulants is calculated based on the moments-of-moments technique, and their relative errors are shown in Fig. 3 as solid lines. As expected, the theory agrees well with the experiments for the first two factorial cumulants, because our experimental data accurately determine the first four cumulants. However, the agreement among the relative error between experiment and theory of the third and fourth factorial cumulants is nontrivial, since it involves fifth-to-eighth order cumulants. The signal/noise of higher order cumulants deteriorates rapidly (11), and it is not obvious that we are able to extract meaningful cumulants of eighth order from the experimental data. The agreement between theory and experiment points to our earlier discussion, where we reasoned that the error is not sensitive to higher order cumulants.

To confirm this, we also calculate the truncated variances $\overline{\mathrm{Var}}[\hat{\kappa}_{[r]}]$ of cumulants $\hat{\kappa}_{[r]}$ according to Eq. 28, which ignores

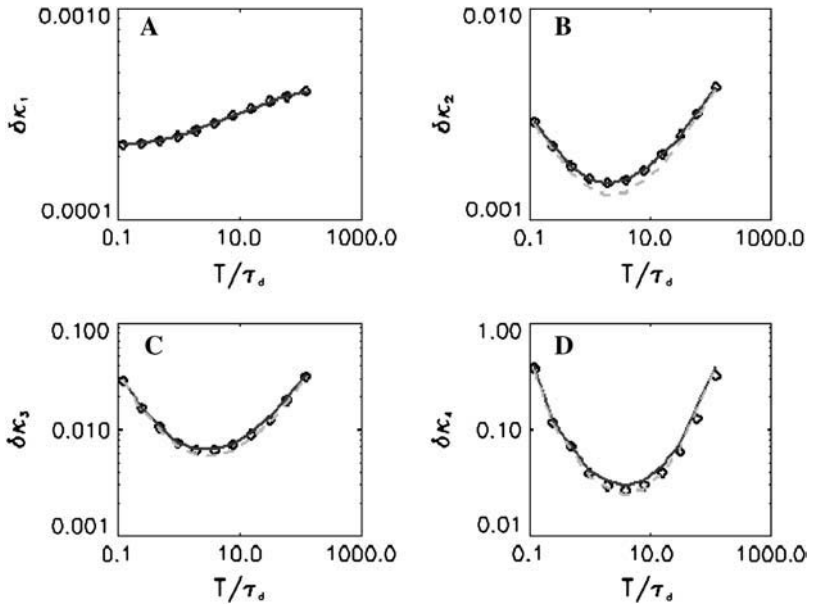

FIGURE 3 The relative error $\delta\hat{\kappa}_{[r]}$ of the factorial cumulant by moments-of-moments analysis as a function of the reduced binning time T/τ_d. The relative error is shown up to order four in $A-D$ for the Rhodamine 6G sample shown in Fig. 2. We determined the error in three different ways. The diamonds represent the relative error calculated from the experimental data. We also calculated theoretical cumulants from the fit parameters according to Eq. 26 and determined their relative error using the moments-of-moments variance (*solid line*). In addition, we also determined the relative error using the truncated moments-of-moments variance $\overline{\mathrm{Var}}_M[\hat{\kappa}_{[r]}]$ (*dashed line*). The data were taken with a sampling time of 5 μs and a total data acquisition time of 131 s.

all cumulants of order higher than r. The relative error of the factorial cumulants $\hat{\kappa}_{[r]}$ based on the truncated variance is shown in Fig. 3 for $r > 1$ as dashed lines, which is almost indistinguishable from the exact theoretical ones (*solid lines*). Both errors are indistinguishable at very short and very long binning times as discussed in Theory, above. The differences at intermediate binning times are small and are negligible for all practical purposes.

Fig. 3 also demonstrates that the relative error of the second, third, and fourth factorial cumulant has a minimum with respect to the binning time. In other words, rebinning allows us to maximize the signal/noise ratio of factorial cumulants. This increase in the signal/noise ratio was the main motivation to introduce rebinned factorial cumulants. The increase in signal/noise is especially important for higher order cumulants. For example, binning allows us to reduce the relative error of the fourth cumulant by a factor of 16 compared to short binning times.

The relative error of factorial cumulants as a function of binning time is shaped by two competing factors. The factorial cumulants for longer binning times are obtained by rebinning the original data. For each rebinning step, the sampling time increases, and the number of data points decreases. On the one hand, fewer data points lead to an increase in the variance and the relative error. On the other hand, the number of photons collected from a single molecule during time T increases for each rebinning step, which increases the signal/noise ratio. Which of the two factors dominates depends on the reduced binning time T/τ_d. For short binning times ($T \ll \tau_d$), $\hat{\kappa}_{[r]}$ is, according to Eq. 24, proportional to T_r. The variance $\mathrm{Var}[\hat{\kappa}_{[r]}]$ is dominated by the lowest power of T, which, according to the moments-of-moments technique, is given by T^r. In addition, the variance also increases, because of the reduction in data points, which is proportional to T. These two effects lead to a variance $\mathrm{Var}[\hat{\kappa}_{[r]}]$ that is proportional to T^{r+1}. Thus, the relative error $\delta\hat{\kappa}_{[r]}$ is of order $\sqrt{T^{1-r}}$. As long as $r > 1$, rebinning reduces the relative error of $\hat{\kappa}_{[r]}$ for short binning times.

In the other extreme, for very long binning times ($T \gg \tau_d$) we can assume that there is a characteristic time T_c, above which the photon counts are statistically independent. Without losing generality, we set the binning time $T = mT_c$. According to the additive property of cumulants for independent random variables, $\hat{\kappa}_{[r]}(T) = m\hat{\kappa}_{[r]}(T_c) = T/T_c\hat{\kappa}_{[r]}(T_c)$. In other words, the cumulant $\hat{\kappa}_{[r]}$ is proportional to the binning time, $\hat{\kappa}_{[r]}(T) \sim T$. The variance of $\hat{\kappa}_{[r]}$ is proportional to the highest power of T, which is again T^{r+1}. So the relative error $\delta\hat{\kappa}_{[r]}$ scales as $\sqrt{T^{r-1}}$. Thus, for $r > 1$, the relative error will increase as a function of binning time. The experimentally observed dependence of the relative error of cumulants with $r > 1$ in Fig. 3 is exactly as predicted by theory. For short binning times, the relative error decreases, reaches a minimum, and then increases at long binning times.

The moments-of-moments technique assumes statistically independent variables and ignores correlations between data. We now consider the effect of the correlation between photon counts on the variance of factorial cumulants. To do this, we need an experimental estimate of the true variance for data with embedded correlations. We calculate this experimental variance Var_E by dividing a long data set into a number of smaller segments. For each segment, we calculate the factorial cumulants, and determine the average and the variance of these factorial cumulants. For comparison, we also calculate the variance of the factorial cumulants for each segment by the moments-of-moments technique. We average all these variances to arrive at the estimated variance Var_M based on the moments-of-moments technique, which ignores correlations. Fig. 4 shows the ratio between these

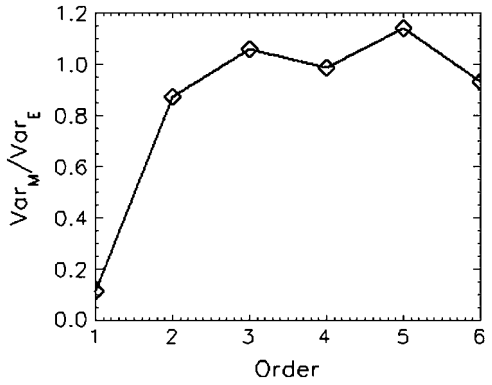

FIGURE 4 Ratio of moments-of-moments variance and the experimental variance of the factorial cumulants obtained from the Rhodamine 6G sample. The experimental variance includes correlations between the data points, which are ignored by the moments-of-moments method. Only the variance of the first cumulant is clearly influenced by the presence of correlations, whereas the variance of the higher order cumulant is accurately described by the moments-of-moments technique.

two variances. A ratio 1 shows that the two variances agree with one another, whereas a ratio different from 1 indicates the importance to include correlations in the determination of the variance. Fig. 4 shows that $\mathrm{Var_M}$ is considerably smaller than $\mathrm{Var_E}$ for the first factorial cumulant, but is very close to $\mathrm{Var_E}$ for higher order factorial cumulants.

This discrepancy between both variances of the first factorial cumulant is due to correlations in the experimental photon counts. In Fig. 5 *A*, we plot the moments-of-moments variance $\mathrm{Var_M}$ ($\diamond$) and the experimental variance $\mathrm{Var_E}$ (Δ) of $\hat{\kappa}_{[1]}$ as a function of the total data acquisition time T_{DAQ} for a fixed binning time of $T = 20$ ms. The number of data points n increases linearly with T_{DAQ}. Thus, we expect according to Eq. 27 that the variance is proportional to $1/T_{\mathrm{DAQ}}$ as was indeed observed in Fig. 5 *A*. However, $\mathrm{Var_M}$ differs from $\mathrm{Var_E}$ because of correlations in the experimental data. We now make use of Eq. 32, which describes the contribution of correlations to the variance. By adding the correction $\mathrm{Var_{Corr}}$ to the moments-of-moments variance $\mathrm{Var_M}$, we obtain the corrected true variance $\mathrm{Var_T}$, which is shown as a solid line in Fig. 5 *A*. The corrected variance $\mathrm{Var_T}$ successfully describes the experimentally determined variance $\mathrm{Var_E}$ (Δ).

The correction term described by Eq. 32 is exact for arbitrary binning times. We demonstrate this in Fig. 5 *B* where we plot the variances $\mathrm{Var_M}$ ($\diamond$),$\mathrm{Var_E}$ (Δ), and the corrected variance $\mathrm{Var_T}$ (*solid line*) as a function of binning time. $\mathrm{Var_T}$ faithfully models the experimental variance at all binning times. Note that the moments-of-moments technique underestimates the true variance for short binning times, but accurately describes the experimental variance for very long binning times. This result confirms our earlier argument that when T is large, correlations between photon counts vanish. The statistical independence of photon counts for large T allows us to use the additive property of cumulants. In other words, $\hat{\kappa}_{[2]}(nT) = n\hat{\kappa}_{[2]}(T)$, and the correction in Eq. 32 due to correlations vanishes.

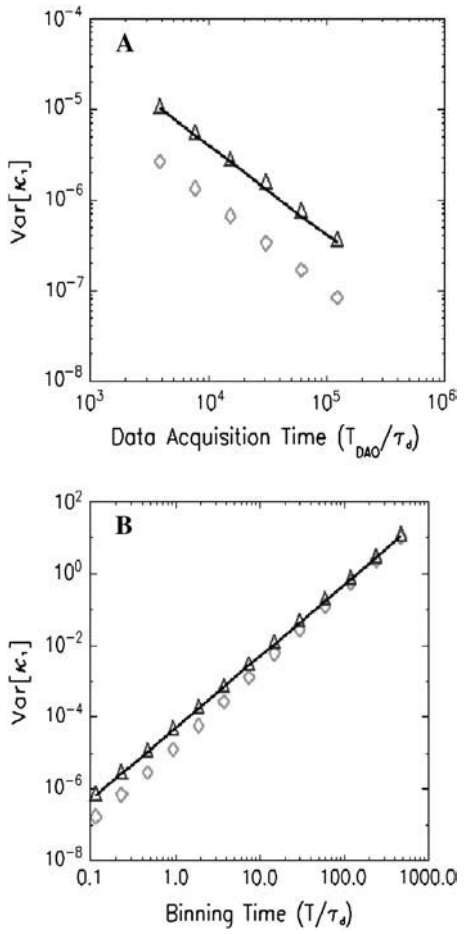

FIGURE 5 Variance of $\hat{\kappa}_{[1]}$ of a Rodamine 6G sample. Diamonds represent the variance determined by moments-of-moments. Triangles are experimentally determined variances, which are obtained by dividing the long data set into small segments and calculating the variance from the statistics of the cumulant $\hat{\kappa}_{[1]}$ of each segment. The solid line is the variance calculated according to Eq. 32, which takes correlations between photon counts into account. (*A*) Variance of $\hat{\kappa}_{[1]}$ as a function of total data acquisition time. The variance decreases inversely proportional to T_{DAQ} as expected. The moments-of-moments variance underestimates the experimental variance in $\hat{\kappa}_{[1]}$. The corrected variance describes the experimental variance accurately. (*B*) Variance of $\hat{\kappa}_{[1]}$ as a function of reduced binning time T/τ_{d}. At short binning times, the moments-of-moments technique underestimates the variance of $\hat{\kappa}_{[1]}$, but approaches the experimental variance for long binning times, because correlations between photon counts in adjacent bins vanish. The corrected variance describes the experimental variance accurately for all binning times.

Binary dye mixture

Resolving a binary dye mixture requires the determination of six parameters: the molecular brightness λ_{i}, the diffusion time τ_{di}, and the number of molecules N_{i} of each species i. The diffusion time of small organic dyes is generally very similar. In other words, we mainly rely on the molecular brightness to resolve species. Therefore, at least four factorial

cumulants are needed to resolve a binary dye mixture. We mixed stock solutions of Rhodamine 6G and Alexa 488 together and performed a fluorescence fluctuation measurement. Fig. 6 shows the first four experimental cumulants of the binary dye mixture. The fit of the experimental cumulants to a single-species model is shown as solid lines. The model and experimental data are not in agreement. The misfit is especially apparent for the third and fourth factorial cumulant. The reduced χ^2 of the fit is 153, which is not surprising, because we expect that the single-species model fails to describe experimental data of a binary dye mixture. In the same figure, we also show the fit of the experimental cumulant to a two-species model, which is shown as dashed lines. The reduced χ^2 for the two-species model is 2.35 and describes the experimental data within experimental error.

Next, we perform a dilution experiment of the binary dye mixture to judge the robustness of the technique. The concentrated stock solutions of Rhodamine 6G and Alexa 488 from the single-species experiment were mixed and diluted by factors of 2 between successive measurements. Each data set was fitted to a two-species model. The reduced χ^2 of the fits varied between 0.68 and 2.35 for the five measurements. The fitted molecular brightness and the number of molecules are shown in Fig. 7. The brightness is concentration-independent as expected. The average photon count rate of Fig. 7 *A* is $\lambda_R = 24{,}100$ cpsm for Rhodamine and $\lambda_A = 9000$ cpsm for Alexa, compared with $\lambda_R = 23{,}400$ cpsm for Rhodamine and $\lambda_A = 9200$ cpsm for Alexa determined by single-species experiments. The diffusion time of Rhodamine varies from 36 μs to 40 μs with an average of 38 μs, whereas the diffusion time of Alexa varies from 37 μs to 49 μs with an average of 43 μs. The average diffusion times of each species are within 20% with the values determined from analysis of single dye solution. A fit of the number of molecules of both species shown in Fig. 7 *B* to a straight line yields a slope of 3.6. In other words, we recover a composition of 21.7% of Rhodamine 6G and 78.3% of Alexa 488. This compares well with the composition of the stock solutions, which we determined earlier from the single-species experiments. We predicted that a mixture of the stock solutions would yield a solution with 23.3% of Rhodamine and 76.7% of Alexa.

DISCUSSION

The signal statistics of fluorescence fluctuation experiments depends strongly on the sampling time. Existing theories, such as PCH, FCS, and our previous FCA model (11), explicitly or implicitly assume that the binning time is shorter than the characteristic timescale of fluctuations. However, the shorter the sampling time is, the fewer photons are received per sampling period, and the signal/noise ratio deteriorates. Thus, we are compelled to increase the sampling or binning time to increase the signal/noise of the experiment. To analyze such data requires a theory that is not restricted to short sampling times. Cumulant provides an ideal framework to discuss the signal statistics of fluctuation experiments. Cumulants of statistically independent variables are additive, and each cumulant provides independent information about the fluctuating signal. In addition, cumulants offer an exact approach for taking arbitrary binning times into account. We developed the theory and constructed

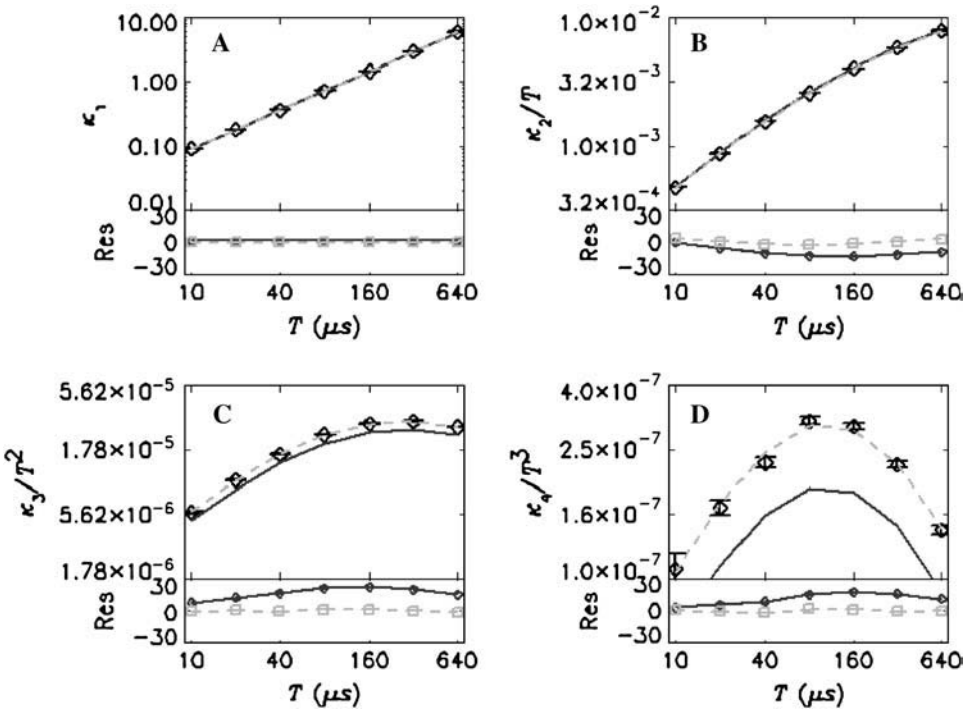

FIGURE 6 FCA analysis of a binary mixture of Rhodamine 6G and Alexa 488. The sample was measured for 30 min with a sampling time of 5 μs. The first four cumulants are plotted as $\hat{\kappa}_{[r]}/T^{r-1}$ (*diamonds*) in *A*–*D* with respect to binning time *T*. The fit to a single-species model is shown as solid line. The lower panel shows the normalized residuals of the fit as a solid line. The best fit to a two-species model is shown as a dashed line. The normalized residuals of the two-species fit are shown as a dashed line in the lower panel. The recovered parameters are $\lambda_R = 24{,}000$ cpsm, $\tau_{dR} = 40$ μs, $N_R = 0.18$ for Rhodamine, and $\lambda_A = 9300$ cpsm, $\tau_{dA} = 37$ μs, $N_A = 0.53$ for Alexa.

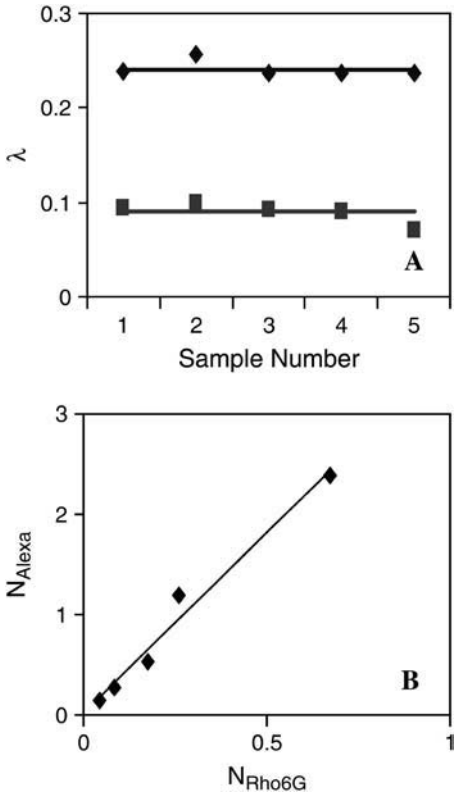

FIGURE 7 FCA analysis of a dilution experiment of a binary mixture of Rhodamine 6G and Alexa 488. Each data set was taken for 30 min with binning time of 5 μs. (*A*) The brightness of Rhodamine 6G (*diamond*) and Alexa 488 (*square*) are shown together with their average value (*solid lines*). (*B*) The number of molecules of both species recovered from the fit. The solid line is a fit of the data to a straight line. Its slope determines the concentration ratio of the mixture. We recover a composition of 22% of Rhodamine 6G and 78% of Alexa 488.

binning functions that capture the influence of sampling time on each cumulant. We decided to collect data with a short sampling time and create data of longer sampling times by rebinning of the original data with software. This approach is flexible, because it preserves the temporal information of the original data, and at the same time allows us to increase our signal/noise by rebinning. TIFCA calculates the experimental cumulants $\hat{\kappa}_{[r]}(T)$ of the photon counts as a function of binning time T and compares them to theoretical models of the cumulants. We successfully implemented the technique and verified it experimentally with simple dye studies.

Photon counting histogram and fluorescence intensity distribution analysis analyze the pdf of the photon counts. Both techniques are only applicable for short sampling times ($T \gg \tau_d$). Extending PCH theory to arbitrary binning times requires a path integral approach, which does not appear attractive for practical use. Cumulants provide an alternative approach to analyze fluctuation data. We previously demonstrated that PCH and FCA are equivalent techniques in the limit of short binning times (11). However, FCA has an advantage over PCH, because an exact extension of the technique to arbitrary binning times is feasible. We developed in this article the exact theory of time-integrated fluorescence cumulants, which experimentally describe the factorial cumulants of photon counts at arbitrary binning times. PCH analysis and FCA are sensitive to molecular brightness; FCS is sensitive to the diffusion time. TIFCA combines aspects of both PCH and FCS, because molecular brightness and diffusion time are determined simultaneously.

There has been an approach described in the literature that extends histogram analysis to longer sampling times (16). Fluorescence intensity multiple distribution analysis studies the histogram of photon counts at different sampling times, and extracts diffusion and brightness information. However, the model is based on an approximation. It utilizes the first two cumulants to correct the brightness and number of molecules for different binning times, but ignores all higher order cumulants. Resolution of two species requires information derived from the third- and fourth-order cumulants. The effect of the approximation introduced by fluorescence intensity multiple distribution analysis on the resolution of species has not been investigated yet. We stress that TIFCA is currently the only theory that is exact for arbitrary binning times. Thus, TIFCA provides an excellent framework to study the effect of binning on PCH and related techniques.

Error analysis is needed to judge the quality of models in describing the experimental data. We therefore studied the variance of the factorial cumulants. We apply moments-of-moments to calculate the variance. We assume a large sample size in our formulation of the theory. This condition is easily met in FFS experiment since we acquire routinely millions of data points. The moments-of-moments technique also assumes statistical independence of the sampled data. This requirement is problematic since FFS data are correlated. For all experimental conditions we tested, we found that the moments-of-moments variance of $\mathrm{Var}_M[\hat{\kappa}_{[r]}]$ is sufficiently accurate for $r > 2$. For $r = 1$, $\mathrm{Var}_M[\hat{\kappa}_{[1]}]$ significantly underestimates the variance and correlation must be explicitly taken into account. We derived an expression to calculate the variance of $\hat{\kappa}_{[1]}$ for correlated data, which successfully describes our experiments. In practice, $\hat{\kappa}_{[1]}$ can be measured with extreme accuracy. In fact, the accuracy is high enough to detect the presence of other noise sources, such as long-term fluctuations in the excitation power of the laser. Such additional noise sources are not accounted for in our theory, which leads to an underestimate of the true variance by Eq. 32. For $r = 2$, there is a small bias in $\mathrm{Var}_M[\hat{\kappa}_{[2]}]$, which usually can be safely ignored. We determined the leading-term correction due to correlations for short binning times. Since the bias will be maximal for short binning times, Eq. 53 allows us to calculate the largest bias introduced by the moments-of-moments technique. The leading-term correction contains fourth-order factorial cumulants and should be small, as we argued in Results, above. For very long binning times, the correlation between different bins dies away and $\mathrm{Var}_M[\hat{\kappa}_{[2]}]$ becomes asymptotically

unbiased. In practice, we found that the bias is <20% in all cases studied. To summarize, the moments-of-moments technique is suitable to calculate $\mathrm{Var}[\hat{\kappa}_{[r]}]$ for $r>1$, but correction must be applied for $\mathrm{Var}[\hat{\kappa}_{[1]}]$.

We investigated the behavior of the relative error of factorial cumulant as a function of binning time. The relative error is an ideal indicator for the importance of a factorial cumulant for data analysis. For example, $\delta\hat{\kappa}_{[r]}>1$ means $\hat{\kappa}_{[r]}$ is statistically insignificant and not useful for data analysis. Since each cumulant contains independent information for a random variable, the number of statistically significant cumulants specifies whether the data is sufficient to resolve species. The theoretical model of the variance of factorial cumulants allows us to calculate the number of significant cumulants for various experimental conditions and is useful for judging the feasibility of resolving species by experiment (11). Rebinning reduces the number of data points, but more photons are collected from a single molecule during the bin time. These two competing factors shape the dependence of the relative error on binning time T. Theory predicts that the relative error $\delta\hat{\kappa}_{[r]}$ scales as $\sqrt{T^{1-r}}$ for short binning times and as $\sqrt{T^{r-1}}$ for long binning times. Thus, for $r > 1$, rebinning reduces the relative error for short binning times, but leads to an increase at long binning times. We verified this behavior of the relative error in Fig. 3 experimentally, where the relative error $\delta\hat{\kappa}_{[r]}$ decreases for short binning times, reaches a minimum, and then increases for long binning times.

Let us briefly compare the signal/noise characteristics of PCH and TIFCA. It is straightforward to calculate factorial cumulants from the histogram of photon counts. However, only a finite number of statistically significant cumulants are contained within an experimental histogram. To resolve two species by brightness requires four statistically significant cumulants (11). Let us assume we want to improve the signal/noise ratio of the fourth cumulant by a factor of 10. Because the relative error of the fourth cumulant scales as $\sqrt{T_{\mathrm{DAQ}}^{-1}}$, we require a 100-fold longer data acquisition time T_{DAQ} to achieve a 10-fold signal/noise improvement by PCH. TIFCA, in contrast to PCH, offers two ways to improve the signal/noise ratio of the experiment. First, just as in the case of PCH, longer data acquisition times reduce the relative error of cumulants. Second, rebinning of data can reduce the relative error, without need for longer data acquisition times. In the short sampling time limit ($T \ll \tau_{\mathrm{d}}$), the relative error of the r^{th} order cumulant scales as $\sqrt{T^{1-r}}$. Thus rebinning by a factor of 5 reduces, in the short sampling limit, the relative error of the fourth cumulant by a factor of 10, whereas PCH requires a 100-times longer data acquisition time to achieve the same improvement. This example serves to illustrate the advantage of TIFCA over PCH.

However, practically, there is a limit to the improvement in signal/noise achievable by binning. Each binning step moves us away from the short sampling time regime, and the increase in signal/noise slows until the relative error of the cumulant reaches a minimum. Further rebinning decreases the signal/noise ratio. Despite this limitation, we demonstrated in Fig. 3 *D* that the relative error of the fourth cumulant improved by more than an order-of-magnitude by rebinning of the experimental data. The reduction of relative error by binning increases with the order r of the cumulant, because the relative error scales as $\delta\hat{\kappa}_{[r]}\sim\sqrt{T^{1-r}}$ for $T \ll \tau_{\mathrm{d}}$. This is confirmed by the data in Fig. 3. The improvement in signal/noise achieved for each cumulant increases with its order.

As we mentioned earlier, the signal/noise ratio of cumulants decreases rapidly with its order and experimental data only contain a finite number of cumulants with $\delta\hat{\kappa}_{[r]}<1$. For example, the relative error of cumulant in Fig. 3 increases by approximately one order-of-magnitude for each successive cumulant. Thus, we only have four factorial cumulants with relative error <1 when the binning time is small. But the relative error of the fifth-order cumulant, $\hat{\kappa}_{[5]}$ is reduced to <1 by rebinning and reaches a minimum value of 0.1. We have not made use of the additional information provided by $\hat{\kappa}_{[5]}$ in this article, where we focus on introducing the technique. But it is straightforward to extend the technique to include higher order cumulants.

To simplify the discussion regarding our new technique, we acquired data for relatively long times to have four significant cumulants at all sampling times. However, much shorter sampling times are sufficient for resolving the binary mixture presented in Fig. 6 by TIFCA. If we only take the first 2 min of acquired data and perform TIFCA analysis, we arrive at a reduced χ^2 of 13 for a single-species fit. A two-species fit determines $\lambda_{\mathrm{R}} = 23{,}000$ cpsm, $\tau_{\mathrm{dR}} = 39\ \mu$s, $N_{\mathrm{R}} = 0.22$ for the Rhodamine, and $\lambda_{\mathrm{A}} = 7600$ cpsm, $\tau_{\mathrm{dA}} = 37\ \mu$s, $N_{\mathrm{A}} = 0.53$ for Alexa, which is in excellent agreement with the result obtained for the complete data set. A fit of the PCH of the same data sampled at 10 μs to a single-species fit yields a reduced χ^2 of 1.8. Thus, in contrast to TIFCA, PCH is unable to resolve this binary mixture.

The point-spread function (PSF) of the laser at the focal point of the objective is complicated (4). In FFS experiments, the PSF is conventionally modeled by a two-dimensional Gaussian, three-dimensional Gaussian, or a Gaussian Lorentzian function. These serve as approximations of the physical PSF. FFS is only able to distinguish between different PSF models if the signal statistics are sufficient. For example, with a bright dye sample, FCS can easily distinguish between a two-dimensional Gaussian and a three-dimensional Gaussian PSF by least-squares fitting. If two model functions lead within experimental error to identical fitting results, then choosing a PSF is simply a matter of convention and mathematical convenience. For example, the Gaussian-Lorentzian and the three-dimensional Gaussian model fit experimental autocorrelation curves equally well (12). The situation is similar for PCH analysis (1), where both the three-dimensional Gaussian and Gaussian Lorentzian PSF describe experimental histograms. TIFCA depend

on the PSF model in two ways: the g factors and binning functions. The binning function determines the shape of the cumulant function and is determined by integration of correlation functions, which ultimately depend on the PSF. A two-dimensional Gaussian PSF is unable to fit experimental cumulants (data not shown), which is not surprising because the two-dimensional Gaussian also fails to fit the experimental autocorrelation function. In the current implementation of TIFCA, we chose a three-dimensional Gaussian model for calculation of the correlation functions. The computed binning functions based on a three-dimensional Gaussian PSF describe the shape of the experimental cumulants very well. The γ-factors are associated with the amplitude of the cumulant and characterize the shape of the PSF. According to definition (Eq. 9), γ_1 is always 1. The second g-factor, γ_2, has to be calculated from the PSF model, because TIFCA is unable to measure this parameter directly. We choose the three-dimensional Gaussian model and set γ_2 to $(2\sqrt{2})^{-1}$. However, higher-order γ factors can be determined by experiment, because a single species is completely specified by its first two cumulant functions (11). This allows us to treat γ_3 and γ_4 as free parameters, which are determined from fitting a single-species model to the first four cumulant functions. We recovered values of $\gamma_3 = 0.251$ and $\gamma_4 = 0.223$, which differ from the values predicted for a three-dimensional Gaussian model. This discrepancy highlights that the model functions only *approximately* describe the physical PSF. Instead of constructing a better PSF model function, we choose to determine the higher-order γ factors experimentally. In other words, we perform a calibration experiment that determines γ_3 and γ_4. Once the g factors are calibrated, we fix them in the analysis of subsequent experiments. We believe that this calibration procedure adds flexibility for adapting TIFCA to other instruments with slightly different optical properties.

The diffusion coefficient determines the shape, and the brightness and number of molecules controls the amplitude of the cumulant function. Thus, TIFCA not only distinguishes species by brightness, but also by their diffusion coefficients. If the diffusion coefficients of two species differ substantially, then the presence of the two species is reflected in an altered shape of the cumulant functions. However, in this article we resolved a binary mixture based on brightness alone, because the diffusion coefficient of both fluorophores is essentially identical and the shape of the cumulant functions is determined by a single diffusion time. Resolving species purely by brightness is experimentally the more challenging case.

CONCLUSION

This article introduces TIFCA, a new analysis technique that extracts information from the time-integrated fluorescence cumulants. A central concept of the theory is the binning function, which characterizes the influence of sampling time on cumulants. We verified the theory of TIFCA with simple dye experiments and determined the brightness, the number of molecules, and the diffusion time of fluorophores. Note that, in contrast to existing theories, TIFCA is exact for arbitrary sampling times. Thus, TIFCA provides us with a tool to investigate the effects of sampling time on other analysis techniques, such as PCH. We also developed an error analysis method by introducing equations for the variance of cumulants. An important feature of TIFCA is the reduction of the relative error of experimental cumulants by rebinning of data to longer sampling times. The improvement of the signal/noise ratio by rebinning increases the sensitivity of TIFCA in resolving species considerably.

APPENDIX A

The fluorescence intensity of a single particle located at position $\vec{r}$ at time t is given by

$$I(\vec{r}(t)) = \lambda \overline{PSF}(\vec{r}(t)). \tag{34}$$

Plugging Eq. 34 into the definition of the r^{th} order correlation function Eq. 6, we obtain the correlation function as

$$g_{\mathrm{r}}(t_1, t_2, \cdots t_{\mathrm{r}}) = \lambda^{\mathrm{r}} \int_{\mathrm{V}} \cdots \int_{\mathrm{V}} \overline{PSF}(\vec{r}_1)\overline{PSF}(\vec{r}_2)\cdots\overline{PSF}(\vec{r}_{\mathrm{r}}) \Pr(\vec{r}_1, t_1, \vec{r}_2, t_2 \cdots \vec{r}_{\mathrm{r}}, t_{\mathrm{r}}) d\vec{r}_1 d\vec{r}_2 \cdots d\vec{r}_{\mathrm{r}}, \tag{35}$$

where $\Pr(\vec{r}_1, t_1, \vec{r}_2, t_2 \cdots \vec{r}_{\mathrm{r}}, t_{\mathrm{r}})$ is the probability distribution function of the particle located at the position $\vec{r}_{\mathrm{i}}$ at time t_{i} ($i = 1, \ldots r$).

Using the Markov property of the diffusion process, we write the correlation function for $t_1 < t_2 \ldots < t_{\mathrm{r}}$ as

$$g_{\mathrm{r}}(t_1, t_2, \cdots t_{\mathrm{r}}) = \lambda^{\mathrm{r}} \int_{\mathrm{V}} \cdots \int_{\mathrm{V}} \overline{PSF}(\vec{r}_1)\overline{PSF}(\vec{r}_2)\cdots\overline{PSF}(\vec{r}_{\mathrm{r}}) \Pr(\vec{r}_1, t_1)\Pr(\vec{r}_2, t_2|\vec{r}_1, t_1)\cdots\Pr(\vec{r}_{\mathrm{r}}, t_{\mathrm{r}}|\vec{r}_{\mathrm{r}-1}, t_{\mathrm{r}-1}) \times d\vec{r}_1 d\vec{r}_2 \cdots d\vec{r}_{\mathrm{r}}. \tag{36}$$

The probability distribution of finding the molecule at position $\vec{r}_1$ at time t_1 is

$$\Pr(\vec{r}_1, t_1) = \frac{1}{V}, \tag{37}$$

and the transition probability that the particle is at position $\vec{r}_2$ at time t_2, if it was at position $\vec{r}_1$ at time t_1, is given by

$$\Pr(\vec{r}_2, t_2|\vec{r}_1, t_1) = \frac{1}{(4\pi D(t_2 - t_1))^{\frac{3}{2}}} \exp\left(-\frac{\|\vec{r}_2 - \vec{r}_1\|^2}{4D(t_2 - t_1)}\right). \tag{38}$$

Since we have assumed that the diffusing system is in equilibrium, the stochastic process $I(t)$ is a stationary process, and the correlation function $g_{\mathrm{r}}(t_1, t_2, \ldots t_{\mathrm{r}})$ is invariant with respect to time translation, that is, g_{r} is a function of $t_2 - t_2, t_3 - t_2, \cdots, t_{\mathrm{r}} - t_{\mathrm{r}-1}$ only. Let us define

$$\tau_1 = t_1, \tau_2 = t_2 - t_1, \tau_3 = t_3 - t_2, \cdots, \tau_{\mathrm{r}} = t_{\mathrm{r}} - t_{\mathrm{r}-1}. \tag{39}$$

Using the definition of the γ-factor in Eq. 9 and with the help of Eqs. 37 and 39, the correlation function g_{r} can be put into closed form

$$g_r(t_1, t_2, \cdots t_r) = \gamma_r \lambda^r \frac{V_{PSF}}{V} G_r(\tau_2, \tau_3, \cdots, \tau_r; \tau_d, \alpha), \quad (40)$$

where G_r is defined in Eq. 10.

APPENDIX B

We consider the correlation in photon counts for an arbitrary binning time T. Let us assume that $t_2 > t_1 + T$ and $t_1 > T$. The correlation of photon counts is determined by the correlation of the integrated intensity,

$$\langle \Delta k(t_1)\, \Delta k(t_2)\rangle = N_{total} \langle W(t_1)\, W(t_2)\rangle, \quad (41)$$

where $k(t_i)$ denotes the total photon counts detected in the time interval $(t_i, t_i + T)$, W_i is the integrated intensity for a single molecule during the same time interval, and N_{total} is the total number of molecules.

The correlation in W is related to the intensity correlation function by

$$\langle W(t_1)W(t_2)\rangle = \int_{t_1}^{t_1+T} dt_1' \int_{t_2}^{t_2+T} dt_2' \langle I(t_1')I(t_2')\rangle. \quad (42)$$

Plugging Eq. 7 into the above equation, we obtain

$$\langle \Delta k(t_1)\Delta k(t_2)\rangle = N\gamma_2\lambda^2 \int_{t_1}^{t_1+T} dt_1' \int_{t_2}^{t_2+T} dt_2' G_2(t_2' - t_1'). \quad (43)$$

The following procedure is similar to the derivation of the binning function. We define the time difference as $\tau = t_2 - t_1$, and transform the above equation into

$$\langle \Delta k(t_1)\Delta k(t_1+\tau)\rangle = N\gamma_2\lambda^2 \left(\int_{\tau}^{\tau+T} dx\, G_2(x)(T + \tau - x) + \int_{\tau-T}^{\tau} dx\, G_2(x)(T - \tau + x) \right). \quad (44)$$

Using the definition of the second binning function B_2, we simplify the above equation to

$$\langle \Delta k(t_1)\Delta k(t_1+\tau)\rangle = N\gamma_2\lambda^2 \frac{(B_2(\tau+T) + B_2(\tau-T) - 2B_2(\tau))}{2}. \quad (45)$$

When the binning time is small ($T \ll \tau_d$), the correlation function reduces to the special case of

$$\langle \Delta k(t_1)\Delta k(t_1+\tau)\rangle = N\gamma_2(\lambda T)^2 \frac{1}{2}\frac{d^2 B_2(\tau)}{d\tau^2} = N\gamma_2(\lambda T)^2 G_2(\tau). \quad (46)$$

APPENDIX C

The second factorial cumulant is calculated from the photon counts by

$$\hat{\kappa}_{[2]} = \frac{1}{n}\sum_{i=1}^{n} \left(k_i - \frac{1}{n}\sum_{j=1}^{n} k_j \right)^2 - \frac{1}{n}\sum_{i=1}^{n} k_i. \quad (47)$$

The variance of the above equation is given by

$$\mathrm{Var}[\hat{\kappa}_{[2]}] = \left\langle \left\{ \frac{1}{n}\sum_{i=1}^{n} [(\Delta k_i)^2 - \langle (\Delta k_i)^2\rangle] - \left[\left(\frac{1}{n}\sum_{i=1}^{n} \Delta k_i \right)^2 - \left\langle \frac{1}{n}\sum_{i=1}^{n} \Delta k_i \right\rangle^2 \right] - \frac{1}{n}\sum_{i=1}^{n} \Delta k_i \right\}^2 \right\rangle, \quad (48)$$

with $\Delta k_i = k_i - \langle k\rangle$.

We first identify the terms of order $(1/n)$ and denote the rest as correction

$$\mathrm{Var}[\hat{\kappa}_{[2]}] = \frac{1}{n}\left(\langle \Delta k^4\rangle - 2\langle \Delta k^3\rangle - \langle \Delta k^2\rangle^2 + \langle \Delta k^2\rangle \right) + Var_{corr}. \quad (49)$$

The first term in Eq. 49 equals the moments-of-moments variance. The second term Var_{corr} contains contribution of the correlation in the photon counts

$$\mathrm{Var}_{corr} = \frac{2}{n^2}\sum_{i<j}^{n} \left[\langle \Delta k_i^2 \Delta k_j^2\rangle - \langle \Delta k_i^2\rangle\langle \Delta k_j^2\rangle + \langle \Delta k_i \Delta k_j\rangle - \langle \Delta k_i^2 \Delta k_j\rangle - \langle \Delta k_i \Delta k_j^2\rangle \right] - \frac{2}{n^3}\sum_{i,j,l} \left[\langle \Delta k_i^2 \Delta k_j \Delta k_l\rangle - \langle \Delta k_i^2\rangle\langle \Delta k_j \Delta k_l\rangle - \langle \Delta k_i \Delta k_j \Delta k_l\rangle \right] + \frac{1}{n^4}\sum_{i,j,l,m} \left[\langle \Delta k_i \Delta k_j \Delta k_l \Delta k_m\rangle - \langle \Delta k_i \Delta k_j\rangle\langle \Delta k_l \Delta k_m\rangle \right]. \quad (50)$$

The above formula is still exact. Since the higher order correlation function decays much faster than the lower order ones, we expect the contribution of the second and third terms in Eq. 50 to be small compared with the first term. As a further approximation, we only consider the case of short binning times. In this limit, the summation in Eq. 50 can be transformed into integration. Following a similar procedure as in Qian (17), we derive the relation between fluctuation moments and the correlation functions, for $i < j < l < m$ as

$$\begin{aligned} \langle \Delta k_i \Delta k_j\rangle &= \hat{\kappa}_{[2]} G_2(t_j - t_i) \\ \langle \Delta k_i^2 \Delta k_j\rangle &= \hat{\kappa}_{[3]} G_3(0, t_j - t_i) + \hat{\kappa}_{[2]} G_2(t_j - t_i) \\ \langle \Delta k_i^2 \Delta k_j^2\rangle &= \hat{\kappa}_{[4]} G_4(0, t_j - t_i, 0) + \hat{\kappa}_{[3]}(G_3(t_j - t_i, 0) \\ &\quad + G_3(0, t_j - t_i)) + \hat{\kappa}_{[2]} G_2(t_j - t_i) \\ &\quad + 2\hat{\kappa}_2^2 G_2^2(t_j - t_i) + (\hat{\kappa}_{[2]} + \hat{\kappa}_{[1]})^2. \end{aligned} \quad (51)$$

We plug these moments into Eq. 50 and change the summation into integration. Then we obtain the lowest order correction due to correlations,

$$\mathrm{Var}_{corr} = \frac{2}{(nT)^2} \left(\int_0^{nT} \left[\hat{\kappa}_{[4]} G_4(0, \tau, 0) + 2\hat{\kappa}_{[2]}^2 G_2^2(\tau) \right] \times (nT - \tau) d\tau \right). \quad (52)$$

In the case of the three-dimensional Gaussian PSF model, $G_4(0,\tau,0;\tau_d,r) = G_2(\tau, \tau_d/2, r)$ and the second term in Eq. 52 is much smaller than the first term. So Eq. 52 can be simplified to

$$\mathrm{Var}_{\mathrm{corr}} = \frac{\hat{\kappa}_{[4]}(T)}{(nT)^2} B_2(nT; \tau_d/2, r). \quad (53)$$

This work was supported by grants from the National Institutes of Health (No. GM64589) and the National Science Foundation (No. PHY-0346782).

REFERENCES

1. Chen, Y., M. Tekmen, L. Hillesheim, J. Skinner, B. Wu, and D. J. Müller. 2005. Dual-color photon-counting histogram. *Biophys. J.* 88: 2177–2192.
2. Chen, Y., L. N. Wei, and J. D. Muller. 2003. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proc. Natl. Acad. Sci. USA.* 100:15492–15497 [erratum appears in *Proc. Natl. Acad. Sci. USA*, 101:1792].
3. Chen, Y., J. D. Müller, P. T. So, and E. Gratton. 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophys. J.* 77:553–567.
4. Hess, S. T., and W. W. Webb. 2002. Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophys. J.* 83:2300–2317.
5. Kask, P., K. Palo, D. Ullmann, and K. Gall. 1999. Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. *Proc. Natl. Acad. Sci. USA.* 96:13756–13761.
6. Kendall, M. G., and A. Stuart. 1977. Chapt. 3: Moments and cumulants. *In* The Advanced Theory of Statistics, Vol. 1. MacMillan Publishing, New York. 57–96.
7. Kendall, M. G., and A. Stuart. 1977. Chapt. 12: Cumulants of sampling distributions. *In* The Advanced Theory of Statistics, Vol. 1. MacMillan Publishing, New York. 293–328.
8. Magde, D., E. Elson, and W. W. Webb. 1972. Thermodynamics fluctuations in a reacting system: measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.* 29:705–708.
9. Mandel, L. 1958. Fluctuations of photon beams and their correlations. *Proc. Phys. Soc.* 72:1037–1048.
10. Meseth, U., T. Wohland, R. Rigler, and H. Vogel. 1999. Resolution of fluorescence correlation measurements. *Biophys. J.* 76:1619–1631.
11. Müller, J. D. 2004. Cumulant analysis in fluorescence fluctuation spectroscopy. *Biophys. J.* 86:3981–3992.
12. Müller, J. D., Y. Chen, and E. Gratton. 2003. Fluorescence correlation spectroscopy. *In* Methods of Enzymology, Vol. 361. G. Marriott and I. Parker, editors. Academic Press, San Diego, CA. 69–92.
13. Müller, J. D., Y. Chen, and E. Gratton. 2000. Resolving heterogeneity on the single molecular level with the photon-counting histogram. *Biophys. J.* 78:474–486.
14. Palmer, A. G., and N. L. Thompson. 1989. High-order fluorescence fluctuation analysis of model protein clusters. *Proc. Natl. Acad. Sci. USA.* 86:6148–6152.
15. Palmer, A. G., and N. L. Thompson. 1987. Molecular aggregation characterized by high order autocorrelation in fluorescence correlation spectroscopy. *Biophys. J.* 52:257–270.
16. Palo, K., U. Mets, S. Jager, P. Kask, and K. Gall. 2000. Fluorescence intensity multiple distributions analysis: concurrent determination of diffusion times and molecular brightness. *Biophys. J.* 79: 2858–2866.
17. Qian, H. 1990. On the statistics of fluorescence correlation spectroscopy. *Biophys. Chem.* 38:49–57.
18. Qian, H., and E. L. Elson. 1990. On the analysis of high order moments of fluorescence fluctuations. *Biophys. J.* 57:375–380.
19. Saleh, B. 1978. Chapt. 3: Point processes. *In* Photoelectron Statistics With Applications to Spectroscopy and Optical Communication. Springer-Verlag, Berlin. 57–84.