Probability Distribution Analysis of Single-Molecule Fluorescence Anisotropy and Resonance Energy Transfer

Stanislav Kalinin,* Suren Felekyan, Matthew Antonik,† and Claus A. M. Seidel*

Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine Universität, Universitätsstrasse 1, Geb 26.32, 40225 Düsseldorf, Germany

Received: March 22, 2007; In Final Form: May 25, 2007

Analysis of anisotropy in single-molecule fluorescence experiments using the probability distribution analysis (PDA) method is presented. The theory of anisotropy-PDA is an extension of the PDA theory recently developed for the analysis of Förster resonance energy transfer (FRET) signals [Antonik, M.; et al. *J. Phys. Chem. B* **2006**, *110*, 6970]. The PDA method predicts the shape of anisotropy histograms for any given expected ensemble anisotropy, signal intensity distribution, and background. Further improvements of the PDA theory allow one to work with very low photon numbers, i.e., starting from the level of background signal. Analysis of experimental and simulated data shows that PDA has the major advantage to unambiguously distinguish between shot noise broadening and broadening caused by heterogeneities in the sample. Fitting of experimental histograms yields anisotropy values of individual species, which can be directly compared with those measured in ensemble experiments. Excellent agreement between the ensemble data and the results of PDA demonstrates a good absolute accuracy of the PDA method. The precision in determination of mean values depends mainly on the total number of photons, whereas the ability of PDA to detect the presence of heterogeneities strongly depends on the time window length. In its present form PDA can be also applied to computed fluorescence parameters such as FRET efficiency and scatter-corrected fluorescence anisotropy. Extension of the PDA theory to low photon numbers makes it possible to apply PDA to dynamic systems, for which high time resolution is required. In this way PDA is developed as a sensitive tool to detect biomolecular heterogeneities in space and time.

1. Introduction

Recently, single-molecule detection (SMD) fluorescence techniques have been applied to distance measurements by Förster resonance energy transfer (FRET).[1−9] Combining FRET and SMD allows one to monitor distance changes at the subpopulation level, with millisecond time resolution (see, e.g., refs 6,9). Advanced SMD techniques such as *m*ultiparameter *f*luorescence *d*etection (MFD)[10−12] and alternating-laser excitation (ALEX)[13−15] help to avoid many experimental artefacts, in particular, incomplete and/or unspecific labeling and local dye quenching. Simultaneous measurements of FRET and polarization anisotropy by MFD can also to some extent reduce uncertainties in the angular coupling factor $\langle \kappa^2 \rangle$.[5]

In single-molecule FRET (SM-FRET) experiments the fluorescence signal is detected by two or more detectors in two wavelength ranges (e.g., "green" and "red"), which cover the fluorescence spectra of the donor (D) and the acceptor (A) dyes, respectively. From these data, a distribution of a FRET-related parameter (usually FRET efficiency) is generated. However, due to the presence of a background signal and the very limited numbers of photons that are collected from a single molecule, shot-noise-broadened distributions are always observed. It is clear that the obtained distributions cannot be directly interpreted as being due to an actual distribution of donor−acceptor distances.

Most often, the experimental FRET efficiency histograms are fitted with a weighted sum of Gaussian distributions, for which mean values are calculated (see, e.g., ref 8). The mean values of these Gaussian distributions can then be used to calculate donor−acceptor distances according to well-known expressions.[16,17] Unfortunately, this approach to the analysis of SM-FRET data extracts very limited information. Although it may (or may not) provide the correct mean FRET efficiencies, the information on the width and the shape of the distributions is completely ignored. As a result, important information about the actual distribution of distances and/or number of different molecular states is lost. In other words, one main advantage of single-molecule experiments—the ability to analyze the heterogeneity of a system—is not utilized.

Only recently have quantitative methods such as the probability distribution analysis (PDA)[18] been developed for the analysis of FRET data allowing for Ångström resolution in optical spectroscopy. In contrast to the usual fit using a Gaussian function which requires two variables, center and width, it is the fundamental idea of the PDA approach that a single parameter, which corresponds to the mean, determines automatically the maximum width and asymmetry of the distribution. This sensitive approach is essential for the shape analysis of single-molecule FRET histograms and allows one to separate shot-noise broadening from heterogeneity of relevant physical parameters (e.g., donor−acceptor distances).[18−20] In addition to PDA, a second method which can be directly applied to experimental single-molecule data measured on freely diffusing molecules, is the proximity ratio histogram (PRH) analysis,

* To whom correspondence should be addressed. E-mail: stanislav.kalinin@uni-duesseldorf.de; cseidel@gwdg.de.

† Present address: Department of Physics & Astronomy, The University of Kansas, 1082 Malott Hall, 1251 Wescoe Hall Drive, Lawrence, KS 66045-7582, U.S.A.

10.1021/jp072293p CCC: $37.00
Published on Web 08/03/2007

proposed by Nir et al.[19] The main difference between PDA and PRH is that PRH analysis can be applied to selected bursts of varying duration, whereas PDA requires that the burst data are cut into time windows of equal length (or equal time bins). Moreover, PDA has been originally developed for the analysis of green to red signal ratio (rather than proximity ratio) histograms. Here we show that PDA can be applied to histograms of the FRET efficiency itself or any measure of it, such as the proximity ratio or the green/red signal ratio, whereby no effect on the accuracy is noticeable.

In addition to PDA and PRH analysis, there are several other papers that deal with quantitative analysis of shot-noise distributions but are less applicable to real SM-FRET experiments on freely diffusing molecules. A complementary approach, which is intended for the analysis of fluorescence traces of immobilized molecules, has been developed by the group of Yang.[20,21] Theoretical treatment of photon statistics in SMD-FRET experiments is presented by Gopich and Szabo.[22] Finally, we should mention that the theory of two-dimensional fluorescence intensity distribution analysis (2D-FIDA)[23] is also closely related to PDA. The key idea of PDA is to use experimental signal intensity distribution (or burst size distribution in the method of Nir et al.[19]) rather than predict this distribution by assuming a certain shape of the observation volume. An important advantage of PDA and PRH analysis is the possibility to apply a burst-selection procedure prior to further analysis.

Specific labeling of a macromolecule with two different fluorophores for a FRET experiment can be rather difficult.[17] It is therefore helpful to use fluorescence anisotropy to distinguish and characterize separate species.[24−28] Single-molecule anisotropy measurements are complementary to FRET experiments or can even be performed simultaneously with an MFD setup.[6,11,12] To our knowledge, there is no quantitative tool for precise calculation of average fluorescence anisotropies from single-molecule data or detection of broadening due to inhomogeneities. The main purpose of this work is to extend the theory of PDA to single-molecule fluorescence anisotropy data and to demonstrate the absolute accuracy (i.e. the ability to extract correct mean values with low systematic errors) of the method. With a bulk setup, the anisotropy can be easily measured with ±0.001 accuracy, whereas the accuracy of FRET efficiency measurements depends on several correction factors (e.g., detection efficiencies), which can be difficult to measure with comparable accuracy.

In this work we show that not only the precision (i.e., reproducibility) but also the absolute accuracy of PDA is comparable with that of ensemble measurements. We also present error analysis of PDA and, in particular, investigate the influence of time window length. For time windows with small numbers of photons we introduce a modification of the PDA theory, which removes the current requirement that the background contribution to the total signal be small. Because the fluorescence intensity measured from freely diffusing molecules is heavily weighted toward low photon counts, taking into account all time windows improves the accuracy of shape determination and consequently leads to a better understanding of underlying processes.

2. Theory

The theory of anisotropy-PDA (as briefly outlined below) is closely related to the previously published theory of FRET-PDA.[18] "Green" and "red" detection channels are naturally replaced by the two polarization channels, which collect photons with parallel (||) and perpendicular (⊥) polarization with respect to polarization of the excitation light. Here we focus mainly on significant improvements of the PDA theory and its applications to the analysis of single-molecule anisotropy data.

The measured fluorescence signal S consists of fluorescence (F) and background (B) photons. In this work the values of S, F, and B are expressed in photon count numbers per time window of a fixed length. The signal is divided into two components (e.g., "green" and "red" or "parallel" and "perpendicular") and measured by two or more single-photon counting detectors. In general, the probability of observing a certain combination of photon counts in two detection channels 1 and 2, $P(S_1,S_2)$, is given by a product of independent probabilities[18]

$$P(S_1,S_2) = \sum_{F_1+B_1=S_1;F_2+B_2=S_2} P(F)P(F_1,F_2|F)P(B_1)P(B_2) \tag{1}$$

In eq 1, $P(F)$ describes the fluorescence intensity distribution, i.e., the probability of observing exactly F fluorescence photons per time window. The second term in eq 1, $P(F_1,F_2|F)$, is the conditional probability of observing a particular combination of F_1 and F_2, provided the total number of fluorescence photons is F. $P(B_1)$ and $P(B_2)$ represent the background intensity distributions.

2.1. Theory of Anisotropy-PDA. In anisotropy experiments the signal is divided into parallel and perpendicular components, denoted by the subscripts || and ⊥, respectively. Each polarization is detected by an individual detector. PDA starts with a calculation of $P(S_{||},S_{\perp})$, which is given by eq 1. For a time window of a given length, $P(B_{||})$ and $P(B_{\perp})$ can be described by a Poisson distribution[29]

$$P(B) = \frac{\langle B\rangle^B \exp(-\langle B\rangle)}{B!} \tag{2}$$

where $\langle B\rangle$ is the average number of background photons per time window. $\langle B_{||}\rangle$ and $\langle B_{\perp}\rangle$ can be measured independently with a clean buffer sample.

Rotations of macromolecules in solution are usually fast on the time scale of milliseconds. In the following we assume that for any integration time (or a time window) the molecules sample all possible orientations with respect to polarization of the excitation light. Therefore, the probabilities of a detected photon being "parallel" or "perpendicular" ($p_{||}$ and $p_{\perp}$, respectively) are related to the fluorescence anisotropy $\langle r\rangle$ according to well-known expressions for ensemble anisotropy:[30,31]

$$p_{||} = \left\langle \frac{F_{||}}{F_{||}+F_{\perp}} \right\rangle = \frac{1+2\langle r\rangle}{2+\langle r\rangle}; \quad p_{\perp} = \frac{1-\langle r\rangle}{2+\langle r\rangle} \tag{3}$$

In a real experiment the fluorescence signal can be additionally depolarized by the microscope objective, which is taken into account by the use of two correction factors, l_1 and l_2:[32]

$$p_{||} = \frac{1+\langle r\rangle(2-3l_1)}{1+\langle r\rangle(2-3l_1)+G-G\langle r\rangle(1-3l_2)}; \quad p_{\perp} = 1-p_{||} \tag{4}$$

In eq 4, G stands for the ratio of the detection efficiencies, $g_{||}$ and $g_{\perp}$, of the two detection channels ($G = g_{\perp}/g_{||}$). The parameters l_1 and l_2 describe mixing of polarizations in the microscope objective, according to eq 5:[32]

$$\begin{aligned} F_{||} &\propto (1-l_1)F_{||,\mathrm{true}} + l_1F_{\perp,\mathrm{true}} \\ F_{\perp} &\propto l_2F_{||,\mathrm{true}} + (1-l_2)F_{\perp,\mathrm{true}} \end{aligned} \tag{5}$$

In eq 5 $F_{\parallel,\mathrm{true}}$ and $F_{\perp,\mathrm{true}}$ denote the "true" polarized signals that would be observed with a perfect setup.

Analogous to a FRET experiment, the conditional probability $P(F_\parallel,F_\perp|F)$ can be expressed as a binomial distribution[18]

$$P(F_\parallel,F_\perp|F) = \frac{F!}{F_\parallel!F_\perp!} p_\parallel^{F_\parallel} p_\perp^{F_\perp} = \frac{F!}{F_\parallel!(F-F_\parallel)!} p_\parallel^{F_\parallel}(1-p_\parallel)^{F-F_\parallel} \quad (6)$$

Substitution of eqs 2 and 6 into eq 1 yields an expression for $P(S_\parallel,S_\perp)$, where the only fitting parameter is $p_\parallel$.

2.2. Deconvolution of $P(F)$. The fluorescence intensity distribution $P(F)$ in eq 1 is not measurable directly. A practical approximation used in the previous work[18] reads

$$P(S_\parallel,S_\perp) = \sum_{F_\parallel+B_\parallel=S_\parallel;F_\perp+B_\perp=S_\perp} P(S)P(F_\parallel,F_\perp|S-B_\parallel-B_\perp)\cdot P(B_\parallel)P(B_\perp) \quad (7)$$

where $P(S)$ is the total experimental intensity distribution ($S = S_\parallel + S_\perp = F + B$). The use of eq 7 implies that the probabilities $P(S)$ and $P(B)$ are independent, which is reasonable only for $B \ll S$. In order to apply PDA also for data with small photon numbers ($S \approx B$), the fluorescence intensity distribution $P(F)$ is needed. In this work we calculate $P(F)$ by deconvolution. The model for $P_\mathrm{M}(F)$ is found by using a χ^2 maximum likelihood test (eq 8)

$$\chi^2 = \sum_{S=0}^{S_\mathrm{max}} W(S)[P(S) - P_\mathrm{M}(F) \otimes P(B)]^2 \rightarrow \min \quad (8)$$

In eq 8 "⊗" denotes the convolution operator, and S_max is the maximum number of photons per time window. The relationship for the weights $W(S) = 1/P(S)$ is valid because each point of the experimental intensity distribution $P(S)$ represents a number of events and therefore obeys Poisson statistics. For a single species $P_\mathrm{M}(F)$ can be expressed as a weighted sum of Poisson distributions, representing contributions from different spatial elements of the illumination volume[23,29,33]

$$P_\mathrm{M}(F) = \sum_i P(q_i) \frac{q_i^F \exp(-q_i)}{F!} \quad (9)$$

Here q_i denote the mean (expected) numbers of photons emitted by a molecule situated in the ith volume element, and $\{P(q_i)\}$ stand for the relative probabilities of finding a molecule in the corresponding volume elements. In practice a predefined set $\{q_i\}$ is used, consisting of 100−200 elements and ranging from 10^{-3}...10^{-2} to S_max photons per time window. As we analyze data measured under single-molecule conditions, the probability of finding several molecules and/or species in the focus at the same time is not considered here, but can be accounted for.[23] We do not make any *a priori* assumptions regarding the shape of $\{P(q_i)\}$, which is one of the most notable differences between the PDA and FIDA methods.[23,34] The distribution $\{P(q_i)\}$ is reconstructed by using the maximum entropy method (see, e.g.,-refs 35,36). This method yields a distribution that maximizes the function

$$\Theta = \nu s - \chi^2 \quad (10)$$

where χ^2 is given by eq 8, ν is a constant, and s is an entropy-like function for the coefficients $\{P(q_i)\}$,[36] here given by

$$s = \sum_i - P(q_i) \log P(q_i) + \mathrm{const} \quad (11)$$

Maximization of Θ (eq 10) was performed as proposed by Vinogradov and Wilson.[37] Our model-free approach allows us to use any empirical burst selection algorithm and apply the analysis to selected bursts[18] because the coefficients $\{P(q_i)\}$ are optimized for each measurement. It also becomes possible to use relatively long time windows (Δt), which can be comparable or even longer than the FCS-diffusion time t_D. In the following we discuss advantages and disadvantages of using the experimental intensity distribution, $P(S)$, in the calculations.

Following eq 9, the representation of the fluorescence intensity distribution $P_\mathrm{M}(F)$ also appears to be convenient for the model function of the signal intensity distribution $P_\mathrm{M}(S)$. The convolution of fluorescence and background in eq 8 is calculated according to

$$P_\mathrm{M}(S) \equiv P_\mathrm{M}(F) \otimes P(B) = \sum_i P(q_i) \frac{(q_i + \langle B\rangle)^S \exp\{-(q_i + \langle B\rangle)\}}{S!} \quad (12)$$

Moreover, if the distribution $P_\mathrm{M}(F)$ is known for one species with brightness Q, the appropriate fluorescence distribution $P'_\mathrm{M}(F)$ can be calculated for another species with the same diffusion properties and a different brightness Q'

$$P'_\mathrm{M}(F) = \sum_i P(q_i) \frac{(Q'q_i/Q)^F \exp(-Q'q_i/Q)}{F!} \quad (13)$$

Equation 13 is valid if the saturation effects (e.g., triplet formation) are unchanged or negligible.

2.3. PDA Histograms of Fluorescence Anisotropy and FRET Efficiency. In the previous work the PDA method has been applied to $S_\mathrm{G}/S_\mathrm{R}$ signal ratio histograms,[18] where the subscripts G and R refer to the "green" and "red" detection channels, respectively. Here we obtain the two-dimensional probability distribution $P(S_\parallel,S_\perp)$ or $P(S_\mathrm{G},S_\mathrm{R})$ for FRET-PDA as an intermediate result (eq 1), from which one-dimensional probability histograms can be generated for any parameter of interest as defined below. The fluorescence anisotropy and the energy transfer efficiency[30] are the most commonly used parameters in polarization and FRET studies, respectively. For anisotropy-PDA we have chosen the scatter-corrected fluorescence anisotropy r_S:[25]

$$r_\mathrm{S} = \frac{G(S_\parallel - \langle B_\parallel\rangle) - (S_\perp - \langle B_\perp\rangle)}{G(S_\parallel - \langle B_\parallel\rangle)(1-3l_2) + (S_\perp - \langle B_\perp\rangle)(2-3l_1)} \quad (14)$$

Provided that the correction factors l_1, l_2, and G (cf. eq 4) and mean background contributions ($\langle B_\parallel\rangle$ and $\langle B_\perp\rangle$) are known with sufficient accuracy, the value of r_S should be directly comparable with the bulk anisotropy measured under similar conditions. Assuming zero background in eq 14 leads to an expression for the "experimental" fluorescence anisotropy r_exp:[32]

$$r_\mathrm{exp} = \frac{GS_\parallel - S_\perp}{GS_\parallel(1-3l_2) + S_\perp(2-3l_1)} \quad (15)$$

Unlike signal ratios (S_G/S_R and $S_{||}/S_\perp$), r_S and r_{exp} are "normalized" by the total signal. Therefore, these parameters have a limited range of values and are defined for any total signal $S > 0$.

In analyses of FRET data, the FRET efficiency E is frequently used:

$$E = \frac{F_A/\Phi_{F(A)}}{F_D/\Phi_{F(D)} + F_A/\Phi_{F(A)}} \quad (16)$$

In eq 16, F_D and F_A denote the detected intensity of the donor and the acceptor fluorescence emissions, respectively, and $\Phi_{F(D)}$ and $\Phi_{F(A)}$ are the corresponding fluorescence quantum yields. F_D and F_A are related to measurable S_G and S_R signals according to

$$S_G - \langle B_G \rangle = g_G F_D$$
$$S_R - \langle B_R \rangle = g_R F_A + \alpha g_G F_D$$

where g_G and g_R denote the detection efficiencies of the "green" and "red" detection channels, respectively, and α is the crosstalk from green donor signal into the red detection channel of acceptor.

The corresponding "uncorrected" parameter is the proximity ratio E_{exp}, given by

$$E_{exp} = \frac{S_R}{S_G + S_R} \quad (17)$$

The parameters E_{exp} and E are also normalized in the same sense as the scatter-corrected and experimental fluorescence anisotropies.

To obtain a theoretical 1D histogram of FRET efficiency, one has to calculate the sums of probabilities $P(S_G,S_R)$, for which the corresponding FRET efficiency (eq 16) falls into the same (ith) histogram bin, i.e., $E_i \le E(S_G,S_R) < E_{i+1}$:

$$P(E_i \le E < E_{i+1}) = \sum_{E_i \le E(S_G,S_R) < E_{i+1}} P(S_G,S_R) \quad (18)$$

PDA histograms of other parameters of interest can be generated in a similar way. For instance, for the fluorescence anisotropy (eq 14) one obtains

$$P(r_{S,i} \le r_S < r_{S,i+1}) = \sum_{r_{S,i} \le r(S_{||},S_\perp) < r_{S,i+1}} P(S_{||},S_\perp) \quad (19)$$

3. Materials and Methods

3.1. Chemicals. Rhodamine 110 (Rh110) was purchased from Radiant Dyes (Germany), enhanced green fluorescent protein (EGFP) was obtained from BioVision. All other chemicals were of analytical grade and were used as received.

3.2. Fluorescence Measurements. The experiments were performed using a multiparameter fluorescence detection (MFD) technique, as described elsewhere.[11] Briefly, single-molecule fluorescence detection was performed using a confocal epi-illuminated microscope with excitation by an active mode-locked Ar^+ laser (73.5 MHz, 150 ps) (Inova Sabre, Coherent, Palo Alto, CA) at 476.5 nm. The linearly polarized beam was focused into solution with a 60 × 1.2 water immersion objective (UPlan Apo, Olympus, Hamburg, Germany). The laser power at the objective was 372 μW. The diameter of the focus in the xy-plane was about 1.2 μm, as estimated by fluorescence correlation spectroscopy (FCS). Fluorescence detection was performed with the same objective, but with a confocal pinhole ($\varnothing = 100$ μm) that results in a detection volume of ~3 fL.

Sample molecules diffusing freely through the solution occasionally pass through the detection volume, resulting in brief (~0.5−3 ms) bursts of fluorescence. Dilute solutions of molecules (~50 pM) ensure that only single molecules are detected, with each burst corresponding to a single molecule. The fluorescence signal was separated from the laser light by a dichroic beamsplitter (Q485DCLP, AHF, Analysentechnik, Tübingen, Germany) and further divided into its parallel and perpendicular components by a polarizing beamsplitter cube (VISHT11, Gsänger, Planegg, Germany). The color range of detection channels was selected by interference filters (HQ533/46, AHF Analysentechnik, Tübingen, Germany). The photons were detected by two single-photon avalanche diodes (Micro Photon Devices (MPD), Bolzano, Italy) coupled to PC-based time-correlated single-photon counting (TCSPC) modules (modified SPC 132, Becker and Hickel GmbH, Berlin, Germany).

3.3. Simulations. Simulations of single-molecule measurements were performed by using a Brownian dynamics approach.[23,38−40] The spatial intensity distribution of the observation volume was assumed to be a 3D Gaussian:

$$I(x,y,z) = \sqrt{8}\exp\left(-\frac{2x^2+2y^2}{\omega_0^2} - \frac{2z^2}{z_0^2}\right) \quad (20)$$

Note that the approximation (eq 20) is not required by the theory of PDA. The "molecules" were allowed to diffuse freely in an open volume, as will be described elsewhere [Felekyan, S.; et al. Manuscript in preparation]. The size of the simulation volume was defined by a surface $I(x,y,z) = I(\mathbf{0})\exp(-50)$.

At each time step Δt, normally distributed random numbers Δx, Δy, and Δz were added to x, y, and z coordinates of each molecule, respectively, where $\langle \Delta x^2 \rangle = \langle \Delta y^2 \rangle = \langle \Delta z^2 \rangle = 2D\Delta t$, and D stands for the diffusion coefficient. The average number of photons emitted by the molecule and registered by the ith detector at each time step was $\langle F_i \rangle = Q_i I(x,y,z)\Delta t$, where Q_i stands for the brightness of the molecule as measured in ith detection channel. The total brightness $Q = \sum Q_i$ of each species was taken to be 100 kHz, which was similar to typical experimental values measured with our setup. The diffusion time ($t_D = \omega_0^2/4D$) of all species was 0.5 ms, and the time step was 0.005 ms. Poisson-distributed background was added when necessary. The algorithm was thoroughly tested by analyzing simulated data by FCS, FIDA, and MFD methods.

3.4. PDA Analysis. The recorded sequence of photon events was divided into equal noninterleaving time windows; for each of them, $S_{||}$, $S_\perp$ and $S = S_{||} + S_\perp$ were calculated. The duration of the time window was 1 ms unless stated otherwise. The experimental distribution $P(S_{||},S_\perp)$ was obtained by counting the number of time windows with particular $S_{||}$ and $S_\perp$. A threshold for total photon number $S_{min} \le S \le S_{max}$ was then applied. The choice of a minimum threshold S_{min} depends on whether the approximation, eq 7, is used, as discussed below in section 4.1. A maximum photon number threshold S_{max} was applied to reduce the computation time, which is generally proportional to S_{max}^3. The value of S_{max} was chosen so that no more than 0.5% of the time windows was discarded.

The experimental data are presented as a 1D histogram of any ratiometric or normalized parameter of interest X, such as the signal ratio $S_{||}/S_\perp$, scatter-corrected fluorescence anisotropy r_S (eq 14), signal ratio S_G/S_R, or proximity ratio E_{exp} (eq 17). The experimental histogram $P(X)$ consists of n bins X_1, ..., X_n of the parameter X. Fitting of experimental histograms involves

minimization of reduced χ^2-values (χ_r^2) by using a Levenberg−Marquardt optimization algorithm. The model used to fit the experimental data is given by eqs 1−6. Depending on whether polarization or FRET data are fitted, the only variable parameter is $p_{||}$ or p_G, respectively. As an example, for each value of $p_{||}$ suggested by the optimization algorithm, a 2D model probability distribution $P(S_{||},S_\perp)$ is first calculated (eq 1). From this distribution a 1D model histogram $P_M(X)$ is generated for $X = X_1..., X_n$ (cf. eqs 18 and 19). The model histogram is then normalized so that its sum is set to be equal to the experimental number of time windows, for which $S_{min} \le S \le S_{max}$ is fulfilled. The model is fitted to the data histogram by minimizing the reduced χ^2-value (χ_r^2), which is given by

$$\chi_r^2 = \frac{1}{n-m+1} \sum_{i=1..n} \frac{[P(X_i) - P_M(X_i)]^2}{P(X_i)} \qquad (21)$$

In eq 21, n is the number of histogram bins, and m is the number of model parameters, which here is equal to one.

4. Results and Discussion

4.1. Deconvolution of *P*(*F*). As discussed in section 2.2, the fluorescence intensity distribution $P(F)$ is not directly measurable in a real experiment. However, a realistic fluorescence distribution $P(F)$ and the corresponding signal distribution $P(S)$ for a given mean background $\langle B \rangle$ can be easily simulated. In this case the distribution $P(F)$ is easily obtained by "switching off" the background signal in the simulation. We have used this approach to check if the deconvolution procedure (eqs 8−11) recovers the correct $P(F)$.

The simulated intensity distributions $P(F)$ and $P(S)$ are shown in Figure 1A as empty circles and filled squares, respectively. The deconvoluted $P_M(F)$ (eqs 8−11) shown in Figure 1A appears to be essentially identical to the "true" (i.e., simulated) distribution $P(F)$.

Figure 1B shows the $S_{||}/S_\perp$ histogram (gray) obtained from simulated data and the theoretical $S_{||}/S_\perp$ histogram obtained by least-squares fits (eqs 1−6). Due to very low threshold $S_{min} = 5$, comparable with the mean background $\langle B \rangle = 5$, the histograms exhibit distinct "spikes" due to the digital photon numbers, which appear at ratios of small integer numbers.[18] Fitting of the histogram recovers the ratio of $\langle F_{||}/F_\perp \rangle = 1$ set in the simulation with the accuracy of better than 0.5%, even for a quite low signal-to-background ratio ($\langle B \rangle = 5$, $\langle S \rangle / \langle B \rangle \cong 3$, $\langle B_{||}/B_\perp \rangle = 1.5$; see Figure 1B).

The theory based on deconvolution of $P(F)$ already works well for S_{min} of as low as 3 photons ($\chi_r^2 = 0.96$), whereas the approximation, eq 7[18] requires $\langle B \rangle \ll S_{min}$. In the particular case shown in Figure 1, eq 7 fails already at signal numbers $S_{min} \le 10$, which results in high χ_r^2-values ($\chi_r^2 = 2.57$) and systematic deviations visible in the weighted residuals (Figure 1B); i.e. for low-signal levels the convolution (eqs 1 and 8) is mandatory to recover correct results. We have empirically found that for the use of approximation, eq 7, $S_{min} > 3\langle B \rangle$ should be usually sufficient.

4.2. What Are the Specific Advantages of Ratiometric and Normalized Fluorescence Parameters? In single-molecule fluorescence spectroscopy, both ratiometric parameters (such as $S_{||}/S_\perp$ and S_G/S_R) and normalized parameters (such as anisotropy and FRET efficiency) are commonly used to characterize a signal recorded simultaneously in two detection channels. Thus, there is a question as to which parameter is the most appropriate. We will show in the following that there is no unique solution for this problem, because each parameter has its specific advantages and disadvantages in data analysis and presentation of results. However, it is emphasized here that, in this context, PDA calculates a two-dimensional (2D) probability distribution $P(S_1,S_2)$ as an intermediate general result (eq 1), from which a 1D probability histogram of any interesting parameter can be generated. The PDA method is therefore suitable for the analysis of both types of parameters.

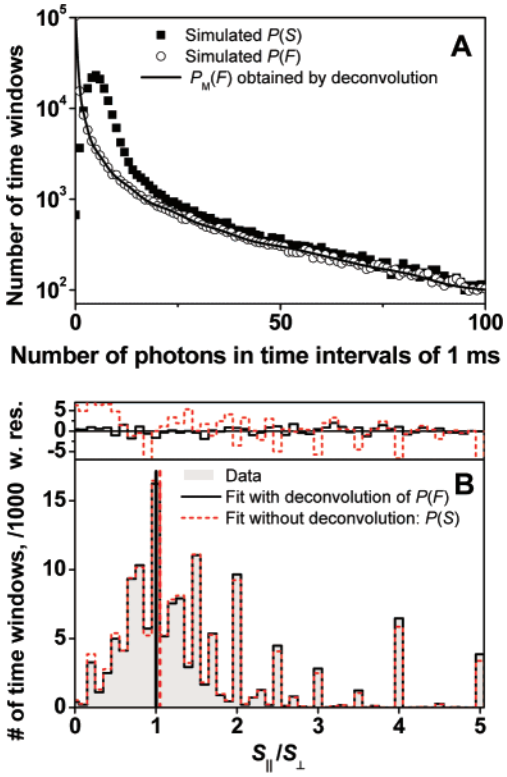

Figure 1. (A) Simulated intensity distribution $P(S)$ (squares) and the corresponding $P(F)$ (open circles). Bold line shows the distribution $P_M(F)$ calculated by deconvolution (eqs 8−11). (B) $S_{||}/S_\perp$ ratio histograms obtained from the simulated data (gray area; $S_{min} = 5$, $S_{max} = 200$), and theoretical $S_{||}/S_\perp$ histograms calculated by using eqs 1−6 (black solid line; $\langle F_{||}/F_\perp \rangle = 1.003 \pm 0.008$; $\chi_r^2 = 0.81$). The red dotted line shows the best fit to the simulated data without deconvolution (eq 7; $\langle F_{||}/F_\perp \rangle = 1.051 \pm 0.008$; $\chi_r^2 = 13.2$). Weighted residuals are also plotted (above). For the simulations we assumed equal brightness in the two detection channels ($Q_{||} = Q_\perp = 50$ kHz, i.e., $\langle F_{||}/F_\perp \rangle = 1$; $\langle B_{||} \rangle = 3$, $\langle B_\perp \rangle = 2$).

The ratiometric parameters $S_{||}/S_\perp$ and S_G/S_R are directly related to the experiment and have a good discrimination between distinct species using a logarithmic scale. In our experience, this representation of S_G/S_R has the specific advantage that the width of the shot noise limited distributions does not vary drastically over the range of interest (e.g., $0 \le E \le 1$), so that complex mixtures with a wide range of parameters can be directly visually analyzed. However, the ratiometric parameters have the disadvantage that they have no defined border cases: i) infinity, if the denominator is zero; ii) no real value if the nominator is zero and logarithmic x-scale is used. Because of these border cases, some data points cannot be displayed on a signal ratio histogram with a logarithmic x-axis (e.g., all points with $S_G = 0$ or $S_R = 0$), which biases the data analysis for small photon numbers.

Commonly used spectroscopic parameters, such as anisotropy and FRET efficiency cannot be directly measured in single-molecule experiments, because the contribution of background signal and partly unknown calibration factors influence the

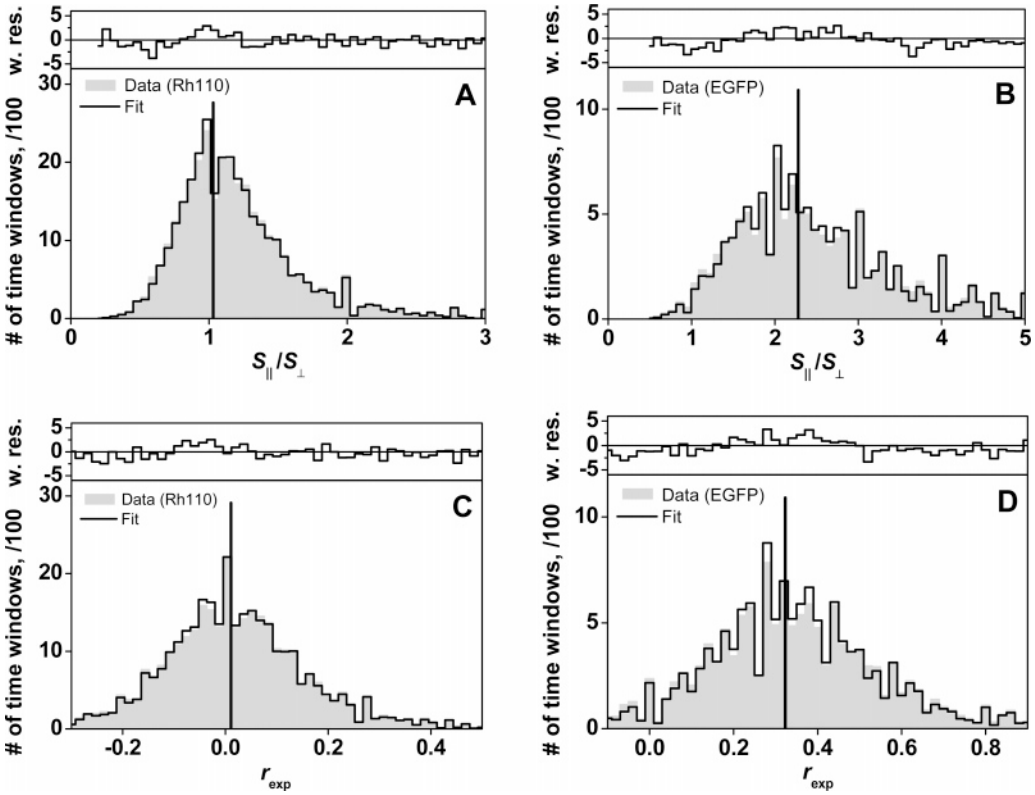

Figure 2. (A, B) Experimental histograms of the number of time windows with particular $S_{||}/S_{\perp}$ (gray areas) are displayed together with best PDA fits (solid lines). Weighted residuals plots are also shown. (A) Histograms are obtained from Rh110 data. The steady-state anisotropy calculated by PDA is 0.010 ± 0.003 (χ_r^2= 1.57). (B) $S_{||}/S_{\perp}$ ratio histogram obtained from EGFP data. The corresponding fluorescence anisotropy is 0.320 ± 0.006 (χ_r^2= 2.11). For both cases $S_{min} = 15$, $S_{max} = 150$, $\langle B_{||}\rangle = 2.7$, $\langle B_{\perp}\rangle = 0.9$. (C,D) Experimental histograms of the number of time windows with particular experimental fluorescence anisotropy (eq 15) obtained from (C) Rh110 and (D) EGFP data. Fitting of the anisotropy histograms yields $r = 0.011 \pm 0.003$ (χ_r^2= 1.47) for Rh110 and $r = 0.323 \pm 0.006$ (χ_r^2= 2.69) for EGFP.

parameter values significantly. The corresponding direct measurable quantities are the experimental anisotropy r_{exp} (eq 15) and the proximity ratio E_{exp} (eq 17), which are primarily useful for qualitative analysis. Only if background intensities and additional calibration parameters (see eqs 14 and 16) are known, can spectroscopically more meaningful background-corrected parameters such as scatter-corrected fluorescence anisotropy r_S and FRET efficiency E be calculated. They have the advantage that their mean values are directly related to molecular properties of interest. However, in these calculations, mean background intensities are usually subtracted, so that in some cases E distributions with negative values are obtained, which have no spectroscopic meaning. The reason why background-corrected parameter distributions are partly distorted is due to the fact that the probability to measure a certain total signal value is a product of the independent probabilities to register fluorescence and background signal values (eq 1), which corresponds to a convolution of their intensity distributions (eq 7). This result is clearly not equivalent to the subtraction of a mean background value from a parameter distribution. To conclude, if one wants go beyond the direct classic analysis of mean values by analyzing the width and the shape of distributions, one has to calculate the shot-noise limit of a distribution and to apply a proper background correction. This is achieved through PDA, which directly analyses the 2D S_1-S_2 signal distribution generating shot-noise limited parameter distributions, which include a background correction obtained by deconvolution. Background subtraction is not always sufficient even when high threshold values are applied. For instance, the data shown in Figure 1B corrected by simple background subtraction cannot be satisfactorily fitted, yielding $\chi_r^2 = 31.8$ for a low threshold $S_{min} = 5$ (see Supporting Information, Figure S1A) and $\chi_r^2 = 3.34$ for a high threshold $S_{min} = 20$ (see Supporting Information, Figure S1B). In this way only background and shot-noise corrected distributions with a spectroscopic meaning are obtained. For display purposes these underlying distributions can be used to generate models (eqs 18 and 19) for parameter distributions, which have been obtained corrected by mean background signal.

Moreover, it is important to note that normalized functions of r_S, r_{exp}, E_{exp}, and E have the advantage of being defined for any total signal $S > 0$. Thus, no data are excluded from the PDA histograms, which is the prerequisite for the quantitative analysis of molecular fractions at low photon numbers. The use of normalized functions also has disadvantages, because they are usually displayed using limited linear scales with a small dynamic range. For example, on a logarithmic ratio scale one can easily distinguish between two species with signal ratios of 100 and 10. However, on a linear scale of the proximity ratio, for example, these two species fall into a few bins of the PDA histogram, i.e., the shot-noise limited width of normalized parameters being proportional to $\langle E\rangle(1 - \langle E\rangle)$[22] changes drastically on the linear scale of the histogram. This makes a visual interpretation of the histograms very difficult (see, for example that the E-distribution of "species 2" is severely distorted, Supporting Information, Figure S2).

4.3. Measurements on Rh110 and EGFP. To prove that PDA also works well for experimental data, we analyzed $S_{||}/S_{\perp}$ ratio histograms obtained from measurements of Rh110 and EGFP.[25] The corresponding experimental histograms and the best fits are shown in Figure 2. As expected, shot-noise limited distributions were observed (Figure 2). Fitting of $S_{||}/S_{\perp}$ histo-

TABLE 1: Values of the Fluorescence Anisotropy *r* as Calculated by Fitting Histograms of Various Parameters by Using PDA[a]

		r calculated by fitting PDA histograms of			
sample	*r* (bulk)	$S_{\|\|}/S_{\perp}$(log)	r_{exp}	r_S	Gaussian fit of r_S
simulation	0.4	0.399 ± 0.005	0.398 ± 0.005	0.400 ± 0.006	0.429 ± 0.004
	0.25	0.251 ± 0.005	0.252 ± 0.005	0.252 ± 0.005	0.268 ± 0.005
	0	−0.002 ± 0.005	−0.002 ± 0.005	−0.001 ± 0.005	− 0.004 ± 0.005
	−0.2	−0.201 ± 0.002	−0.201 ± 0.002	− 0.200 ± 0.002	−0.216 ± 0.002
Rh110	0.008 ± 0.002	0.010 ± 0.003	0.010 ± 0.003	0.011 ± 0.003	0.008 ± 0.002
EGFP	0.315 ± 0.002	0.320 ± 0.006	0.323 ± 0.006	0.323 ± 0.006	0.325 ± 0.008

[a] Error bars represent standard deviations as determined by fitting 10 independent sets of data. For simulations the following parameters were taken: $\langle B_{\|}\rangle = 3$, $\langle B_{\perp}\rangle = 1.5$.

TABLE 2: Values of the FRET Efficiency *E* as Calculated by Fitting Histograms of Various Parameters by Using PDA[a]

		E calculated by fitting PDA histograms of			
sample	*E* (bulk)	S_G/S_R (log)	E_{exp}	E	Gaussian fit of *E*
simulation	0.1	0.100 ± 0.002	0.100 ± 0.002	0.100 ± 0.002	0.095 ± 0.002
	0.25	0.251 ± 0.002	0.250 ± 0.002	0.251 ± 0.002	0.249 ± 0.002
	0.333	0.333 ± 0.003	0.333 ± 0.003	0.333 ± 0.003	0.330 ± 0.002
	0.5	0.501 ± 0.004	0.501 ± 0.004	0.501 ± 0.003	0.498 ± 0.003
	0.667	0.666 ± 0.002	0.666 ± 0.002	0.666 ± 0.002	0.667 ± 0.002
	0.9	0.900 ± 0.002	0.900 ± 0.002	0.900 ± 0.002	0.902 ± 0.002

[a] Error bars represent standard deviations as determined by fitting 10 independent sets of data. For simulations the following parameters were taken: $\langle B_G\rangle = 3$, $\langle B_R\rangle = 1.5$.

grams of Rh110 and EGFP yields anisotropy values of 0.010 ± 0.002 and 0.320 ± 0.006, respectively. These values can be directly compared with the ensemble steady-state anisotropies of Rh110 and EGFP, which are 0.008 ± 0.002 and 0.315 ± 0.002, respectively. Very similar results can be obtained by fitting PDA histograms of experimental fluorescence anisotropy, which are respectively presented in graphs C and D of Figure 2. These results demonstrate good absolute precision of the PDA method as applied to the analysis of single-molecule fluorescence polarization data.

4.4. Estimation of Mean Values. In this section we focus on the question of how to characterize the parameters of a single species, i.e., to determine the mean value of the parameter of interest, e.g., the FRET efficiency or fluorescence anisotropy. Without the use of the PDA approach (eqs 1−6) even this problem is not trivial, because shot-noise broadened distributions often exhibit strongly asymmetrical shapes.[22] For this reason, fitting a Gaussian peak does not always provide the correct mean value. In the next section we show how to decide if one or more species contribute to the observed distribution when visual analysis fails. Furthermore, we investigate the influence of the total number of photons and the time window length on the accuracy of PDA in determining mean values as well as in detecting heterogeneities.

4.4.1. Errors in Estimating Parameters from PDA Histograms. In section 2.3 we extended the PDA method to the analysis of fluorescence anisotropy and FRET efficiency histograms. In order to compare the effectiveness of the PDA method when fitting histograms of various experimental parameters (e.g., r_{exp} and r_S), we subdivided the simulated data and the experimental data presented in Figure 2 into 10 equal parts. Each subset was then fitted separately, allowing us to calculate the standard deviations, as summarized in Table 1. In the case of simulated data the "true" values were exactly known, so it was also possible to estimate the systematic errors. Table 1 makes clear that the choice of experimental parameter values to be shown in a histogram has only a minor effect on the accuracy of the analysis. Thus, the PDA approach is not limited to fitting of signal ratio histograms as originally demonstrated but can also analyze 1D histograms of many other parameters of interest such as r_{exp} and r_S.

This conclusion holds for FRET efficiency histograms as well. Table 2 reports the calculated average values and random deviations, obtained from the above data by fitting histograms of S_G/S_R signal ratio, proximity ratio E_{exp} (eq 17), and FRET efficiency (eq 16). As for the anisotropy analysis, the choice of histogram parameter has a negligible effect on the accuracy of PDA (Table 2). Examples of signal ratio, proximity ratio, and FRET efficiency histograms are presented in Figure 3.

It is also possible to fit distributions of "corrected" parameters (i.e., scatter-corrected anisotropy, r_S, and FRET efficiency, *E*) with Gaussian distributions, from which mean values can be calculated (Tables 1 and 2). For the case of FRET efficiency this method is widely used.[1,15] One can see from Table 2 that for moderate FRET efficiencies Gaussian fits provide reasonable mean values of *E*. However, for extreme values of mean *E* the apparent shot-noise broadened distributions become asymmetric,[22] and systematic errors become noticeable. For instance, for "true" $E = 0.1$ (see Table 2) PDA is ca. 20 times more accurate than Gaussian fitting in terms of corresponding errors in donor−acceptor distances. Furthermore, fitting of anisotropy histograms by Gaussian distributions always leads to high systematic errors of at least 5% (see Table 1). By taking into account that the PDA method additionally provides a meaningful distribution width, we demonstrate that it has clear advantages over simple Gaussian fitting of shot-noise broadened distributions of *E* and r_S.

4.4.2. Influence of the Number of Photons. Given that the total number of photon events ($N_{photons}$) is constant, the length of time window has no systematic effect on the calculated mean values of $\langle F_{\|}/F_{\perp}\rangle$ and the observed standard deviation (σ), as shown in Figure 4 (open circles). However, σ decreases with the number of photons (full squares), where a $\sigma \propto 1/\sqrt{N_{photons}}$ dependence is expected (dashed line). If histograms are compared, this case is observed when values are obtained from the same number of time windows but differ in the window duration. Thus, the total photon number, $N_{photons}$, decreases if smaller windows are used. To avoid this effect one has to increase the number of time windows. Figure 4 shows that this strategy works well for a homogeneous population with a common mean value of $\langle F_{\|}/F_{\perp}\rangle$, where only the first moment

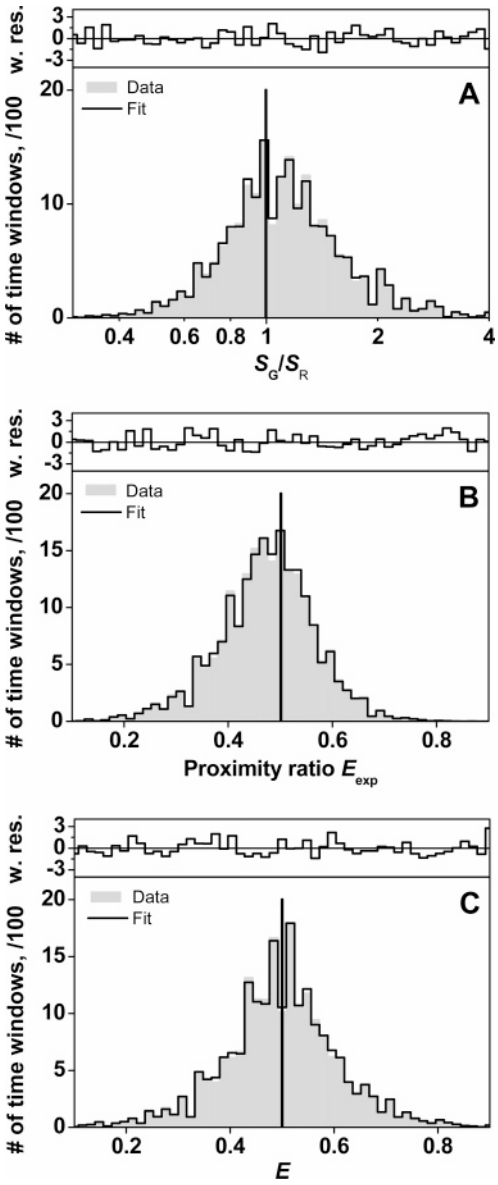

Figure 3. Simulated PDA histograms (gray areas) of (A) the signal ratio S_G/S_R, (B) proximity ratio $E_{exp} = S_R/(S_G + S_R)$, (C) FRET efficiency (calculated by eq 16 using $\Phi_{F(D)} = \Phi_{F(A)} = 1$, $g_G = g_R = 1$ and $\alpha = 0$), generated from the same data set. For the simulations we assumed $\langle F_G/F_R \rangle = 1$, $\langle B_G \rangle = 2$, and $\langle B_R \rangle = 1.2$. The best PDA fits are also shown (solid lines). The calculated underlying mean $\langle F_G/F_R \rangle$ values are 0.997, 0.997 ($E_{exp} = 0.501$), and 0.998 ($E = 0.500$), respectively. Weighted residuals are shown in upper subplots.

of the distribution is relevant. As we will show in section 4.5, this observation does not hold for the analysis of heterogeneous populations, where higher moments become relevant, which describe the shape of the underlying probability density function.

4.5. Analysis of Heterogeneities by PDA. High accuracy in the determination of mean values does not immediately imply the ability to detect heterogeneities. In this section we demonstrate the advantages of using PDA to characterize systems where more than one species are present.

4.5.1. Application of PDA to Heterogeneous Mixtures. In order to show that broadening caused by heterogeneities can be separated from shot-noise broadening, we simulated a single-molecule anisotropy experiment for a mixture of two species at equal concentrations, with $r_{(1)} = 0.1$ and $r_{(2)} = 0.2$. The brightness of the two species was assumed to be equal, which is required by the theory of PDA in its present form (cf. eq 1).

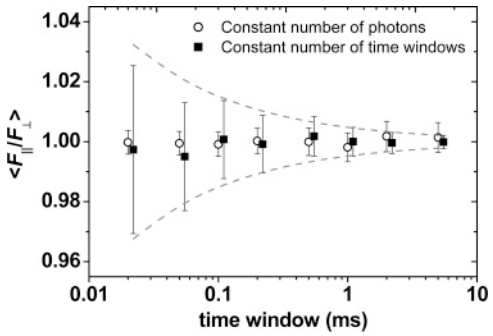

Figure 4. Dependence of calculated $\langle F_G/F_R \rangle$ ratio on the duration of the time window. The error bars represent standard deviations (σ) as determined according to χ_r^2 ($\langle F_G/F_R \rangle \pm \sigma$) $= \chi_{min}^2 + \sqrt{2/n}$, where n is the number of histogram bins. $\langle F_G/F_R \rangle$ values are obtained by fitting PDA histograms of the proximity ratio E_{exp} (eq 17); $S_{min} = 1$. The data contain a constant number of photons (open circles, $N_{photons} = 3 \times 10^6$) or constant number of time windows (full squares; in this case the total number of photons is proportional to the time window length increasing linearly from 6×10^4 to 1.5×10^7). The dotted lines represent error limits estimated according to $\sigma \propto N_{photons}^{-1/2}$.

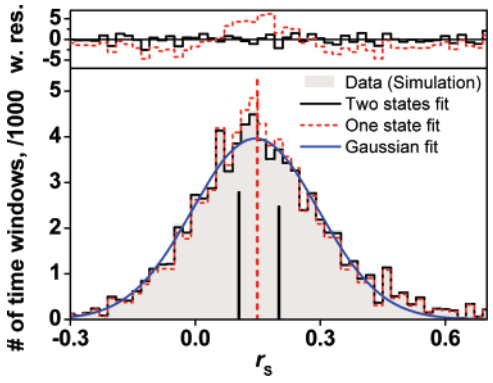

Figure 5. Fluorescence anisotropy histograms obtained from simulated data (mixture of two species at equal concentrations with $r_{(1)} = 0.1$ and $r_{(2)} = 0.2$; $\langle B_{||} \rangle = 2$, $\langle B_{\perp} \rangle = 1.2$). The best possible PDA fit (red dotted line) yields $\chi_r^2 = 6.90$ ($r = 0.149$). Weighted residuals are shown in the upper plot. The solid black line represents a fit by a sum of two PDA histograms calculated for $r_{(1)} = 0.105$ (53%) and $r_{(2)} = 0.201$ (47%); $\chi_r^2 = 1.02$. In both cases $S_{min} = 15$, $S_{max} = 150$. The blue line shows a Gaussian fit ($\langle r_S \rangle = 0.144$, $\sigma = 0.30$) to the simulated data histogram.

The simulated anisotropy histogram and the best shot-noise limited PDA fit with $\langle r_S \rangle = 0.149$ (dotted red line) are shown in Figure 5. The theory of PDA cannot explain the observed broadening by shot noise only, which is seen from the weighted residuals plot (upper plot in Figure 5), and the reduced χ^2-value $\chi_r^2 = 6.90$. Note that the anisotropy histogram exhibits only one broad peak, and it is not directly obvious that several species are present. The histogram can also be very well fitted with one Gaussian peak with $\langle r_S \rangle = 0.144$ and a standard deviation $\sigma = 0.30$ (see Figure 5), providing no direct evidence for the presence of multiple states. Please note the systematic error (0.006) in the estimation of the mean value compared to that in the single-component PDA analysis.

Only the use of PDA can unambiguously show that multiple species contribute to the observed distribution. It appears to be possible to fit the histogram in Figure 5 by a weighted sum of two shot-noise limited PDA histograms (black line). The fit yields $\chi_r^2 = 1.02$, $r_{(1)} = 0.105$ (53%), and $r_{(2)} = 0.201$ (47%), which are in excellent agreement with the stated values. These results show no pronounced dependence on initial guess parameters and can be obtained starting from, e.g., $r_{(1)} = 0$ (33%) and $r_{(2)} = 0.4$ (67%). However, we realize that fitting of

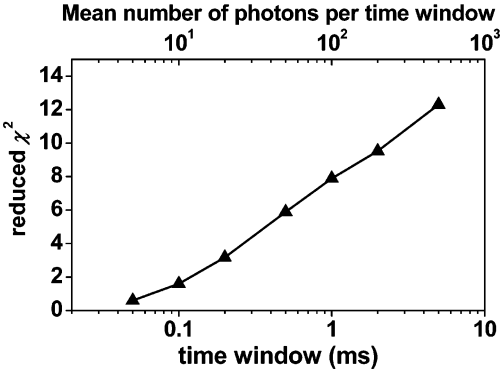

Figure 6. Reduced χ^2-values obtained by fitting of simulated $S_{\parallel}/S_{\perp}$ ratio histograms for heterogeneous mixtures by assuming a shot-noise limited distribution of $S_{\parallel}/S_{\perp}$, plotted as a function of the time window length. The simulated data contain 3×10^6 photons emitted by two different species at equal concentrations, $\langle F_{\parallel}/F_{\perp}\rangle_1 = 1$, $\langle F_{\parallel}/F_{\perp}\rangle_2 = 1.5$. The data are fitted assuming a single $\langle F_{\parallel}/F_{\perp}\rangle$ ratio, as described in section 4.5. The maximum number of photons (S_{max}) is chosen to include 99.5% of all time windows and $S_{min} = 3$.

TABLE 3: Reduced χ^2-Values Obtained for Fit Assuming a Single Species Instead of Two by 2D-FIDA and PDA

$\langle F_{\parallel}/F_{\perp}\rangle_2$	r	χ_r^2, PDA	χ_r^2, 2D-FIDA
1.1	0.032	0.89	0.73
1.2	0.063	1.15	0.70
1.3	0.091	2.35	0.91
1.4	0.118	4.35	1.07
1.5	0.143	7.89	1.44

experimental PDA histograms obtained for real mixtures of different species can be more complicated because the individual fluorescence intensity distributions $P_j(F)$ are required. Individual anisotropy values can differ due to local quenching processes, rather than because of different rotational freedoms of the dyes. Therefore, in many real cases a strong correlation between anisotropy and brightness has to be expected. Work on extension of the PDA theory to the case of multiple species with different brightnesses is in progress.

4.5.2. Sensitivity of PDA To Determine Heterogeneities. We have demonstrated that the PDA method is able to detect the presence of multiple species with different anisotropies. One may then ask what minimal difference in anisotropies can be unambiguously distinguished from shot-noise broadening, and how does PDA compare with other methods in this respect. Another method that can detect this kind of heterogeneity is 2D-FIDA.[34] Unlike PDA, this method predicts the shape of the experimental distribution $P(S)$. Here we compared the sensitivities (i.e., the ability to detect the presence of heterogeneities) of PDA and 2D-FIDA as applied to resolution of multiple species.

We simulated several MFD experiments on a mixture of two different species at equal concentrations: "species 1" has equal brightness in parallel and perpendicular channels $\langle F_{\parallel}/F_{\perp}\rangle_1 = 1$, whereas "species 2" has somewhat higher brightnesses in the parallel channel (i.e., $\langle F_{\parallel}/F_{\perp}\rangle_2 > 1$). As shown in Table 3 we investigated five cases for $\langle F_{\parallel}/F_{\perp}\rangle_2$: 1.1, 1.2, 1.3, 1.4, and 1.5.

In the simulation the average number of molecules in the "laser focus" was 0.01 and 1 for PDA and 2D-FIDA, respectively. The latter value was reported to be optimal for FIDA methods.[23,34] To guarantee the same statistics, 3×10^6 fluorescence photon events were generated for each case. The data were fitted, assuming that only one species was present, and the quality of fit was judged by reduced χ^2-values (χ_r^2) and weighted residuals plots. Two-dimensional FIDA histograms contained typically about 300 bins. The same number of bins was also used for generating PDA histograms so that χ_r^2-values obtained for 2D-FIDA and PDA fits could be directly compared. This procedure was repeated three times for each value of $\langle F_{\parallel}/F_{\perp}\rangle_2$. One can see that already at $\langle F_{\parallel}/F_{\perp}\rangle_2 \geq 1.3$ the PDA fits become unacceptable with $\chi_r^2 > 2$ (Table 3). One of the reasons could be that PDA can work with very long time windows ($> t_D$) as compared to FIDA (usually less than 1/10 of the diffusion time,[23,34] here 40 μs), which means that PDA time windows contain much higher numbers of photons. As stated in section 4.4 for the case of mixtures, the simple relationship of a constant standard deviation with a constant photon number is not applicable anymore.

To investigate this, we varied the length of the time window and re-analyzed the simulated data set for $\langle F_{\parallel}/F_{\perp}\rangle_2 = 1.5$. The total number of photons was always the same, irrespective of the number of time windows. The dependence of χ_r^2 on the duration of the time window is illustrated in Figure 6. Shorter time windows contain too few photons, and broadening due to heterogeneities is difficult to detect on top of shot-noise broadening. These results suggest that obtaining good photon statistics due to relatively long (comparable with t_D) time windows is at least one of the reasons for the good sensitivity of the PDA method. It is important to emphasize that the main advantage of PDA is not the accuracy of mean-value determination, which depends mainly on the total number of photons (see section 4.4) and can be achieved by many methods, including fitting properly corrected data histograms by Gaussian distributions. In addition, the PDA approach allows one to access the width and the shape of the distributions, providing information on the number of molecular species and/or distribution of physical properties. For the same reason we can expect that the burstwise PRH analysis proposed by Nir et al.[19] should also be very sensitive in resolving multiple fluorescent species.

Conclusions

In this work we present several important steps toward quantitative analysis of shot-noise broadened distributions, obtained in single-molecule fluorescence experiments. The theory of probability distribution analysis (PDA),[18] originally applied to FRET data, has been extended to single-molecule fluorescence anisotropy. Calculation of $P(S_{\parallel},S_{\perp})$ or $P(S_G,S_R)$ allows one to predict the shape of 1D distributions of any parameter that can be derived from photon count numbers in two detection channels. In particular, PDA can be applied to FRET efficiency and fluorescence anisotropy histograms. By means of simulations and analysis of experimental data we have shown that the PDA method provides high accuracy in the determination of relevant mean values. The use of PDA can reveal the presence of multiple fluorescent species, which is difficult or impossible to detect by other techniques. Extension of the PDA theory to low photon numbers makes it possible to apply PDA to dynamic systems, for which high time resolution is required.

Acknowledgment. S.K. acknowledges the Alexander von Humboldt Foundation for a Humboldt Research Fellowship. This study was supported by the Deutsche Forschungsgemeinschaft (DFG) Grant SE 1195/10-2 and the Volkswagen Foundation Grant I/78 837. We thank Dr. Steven Magennis for valuable comments on the manuscript.

Supporting Information Available: Experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Deniz, A. A.; Dahan, M.; Grunwell, J. R.; Ha, T. J.; Faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3670−3675.

(2) Ha, T. *Methods* **2001**, *25*, 78−86.

(3) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743−747.

(4) Lipman, E. A.; Schuler, B.; Bakajin, O.; Eaton, W. A. *Science* **2003**, *301*, 1233−1235.

(5) Rothwell, P. J.; Berger, S.; Kensch, O.; Felekyan, S.; Antonik, M.; Wöhrl, B. M.; Restle, T.; Goody, R. S.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1655−1660.

(6) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15516−15521.

(7) Diez, M.; Zimmermann, B.; Börsch, M.; König, M.; Schweinberger, E.; Steigmiller, S.; Reuter, R.; Felekyan, S.; Kudryavtsev, V.; Seidel, C. A. M.; Gräber, P. *Nat. Struct. Mol. Biol.* **2004**, *11*, 135−141.

(8) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754−2759.

(9) Joo, C.; McKinney, S. A.; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T. *Cell* **2006**, *126*, 515−527.

(10) Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A. M. *J. Biotechnol.* **2001**, *86*, 163−180.

(11) Kühnemuth, R.; Seidel, C. A. M. *Single Mol.* **2001**, *2*, 251−254.

(12) Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A. M. *Anal. Chem.* **2006**, *78*, 2039−2050.

(13) Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8936−8941.

(14) Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S. *Biophys. J.* **2005**, *88*, 2939−2953.

(15) Lee, N. K.; Kapanidis, A. N.; Koh, H. R.; Korlann, Y.; Ho, S. O.; Kim, Y.; Gassman, N.; Kim, S. K.; Weiss, S. *Biophys. J.* **2007**, *92*, 303−312.

(16) Förster, T. *Ann. Phys.* **1948**, *2*, 55−75.

(17) Haas, E. *Chem. Phys. Chem.* **2005**, *6*, 858−870.

(18) Antonik, M.; Felekyan, S.; Gaiduk, A.; Seidel, C. A. M. *J. Phys. Chem. B* **2006**, *110*, 6970−6978.

(19) Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S. *J. Phys. Chem. B* **2006**, *110*, 22103−22124.

(20) Watkins, L. P.; Chang, H. Y.; Yang, H. *J. Phys. Chem. A* **2006**, *110*, 5191−5203.

(21) Watkins, L. P.; Yang, H. *Biophys. J.* **2004**, *86*, 4015−4029.

(22) Gopich, I.; Szabo, A. *J. Chem. Phys.* **2005**, 014707.

(23) Kask, P.; Palo, K.; Ullmann, D.; Gall, K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13756−13761.

(24) Forkey, J. N.; Quinlan, M. E.; Goldman, Y. E. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 1−35.

(25) Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M. *J. Phys. Chem. A* **1999**, *103*, 331−336.

(26) Ha, T.; Laurence, T. A.; Chemla, D. S.; Weiss, S. *J. Phys. Chem. B* **1999**, *103*, 6839−6850.

(27) Rosenberg, S. A.; Quinlan, M. E.; Forkey, J. N.; Goldman, Y. E. *Acc. Chem. Res.* **2005**, *38*, 583−593.

(28) Karolin, J.; Fa, M.; Wilczynska, M.; Ny, T.; Johansson, L. B. A. *Biophys. J.* **1998**, *74*, 11−21.

(29) Fries, J. R.; Brand, L.; Eggeling, C.; Köllner, M.; Seidel, C. A. M. *J. Phys. Chem. A* **1998**, *102*, 6601−6613.

(30) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 2nd ed.; Kluwer Academic/Plenum Publishers: New York, 1999.

(31) Valeur, B. *Molecular Fluorescence: Principles and Applications*; Wiley-VCH Verlag: Weinheim, 2002.

(32) Koshioka, M.; Saski, K.; Masuhara, H. *Appl. Spectrosc.* **1995**, *49*, 224−228.

(33) Rigler, R.; Mets, Ü. *SPIE* **1992**, *1921*, 239−248.

(34) Kask, P.; Palo, K.; Fay, N.; Brand, L.; Mets, Ü.; Ullmann, D.; Jungmann, J.; Pschorr, J.; Gall, K. *Biophys. J.* **2000**, *78*, 1703−1713.

(35) Brochon J. C. *Methods of Enzymology*; Academic Press: New York, 1994; Vol. 240; pp 262−311.

(36) Livesey, A. K.; Skilling, J. *Acta Crystallogr., Sect. A* **1985**, *41*, 113−122.

(37) Vinogradov, S. A.; Wilson, D. F. *Appl. Spectrosc.* **2000**, *54*, 849−855.

(38) Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Goodwin, P. M.; Keller, R. A. *J. Phys. Chem. B* **1997**, *101*, 3626−3632.

(39) Laurence, T. A.; Kapanidis, A. N.; Kong, X. X.; Chemla, D. S.; Weiss, S. *J. Phys. Chem. B* **2004**, *108*, 3051−3067.

(40) Dix, J. A.; Hom, E. F. Y.; Verkman, A. S. *J. Phys. Chem. B* **2006**, *110*, 1896−1906.