Accuracy of maximum likelihood estimates of a two-state model in single-molecule FRET

Irina V. Gopich

*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases,
National Institutes of Health, Bethesda, Maryland 20892, USA*

(Received 30 September 2014; accepted 29 November 2014; published online 16 January 2015)

Photon sequences from single-molecule Förster resonance energy transfer (FRET) experiments can be analyzed using a maximum likelihood method. Parameters of the underlying kinetic model (FRET efficiencies of the states and transition rates between conformational states) are obtained by maximizing the appropriate likelihood function. In addition, the errors (uncertainties) of the extracted parameters can be obtained from the curvature of the likelihood function at the maximum. We study the standard deviations of the parameters of a two-state model obtained from photon sequences with recorded colors and arrival times. The standard deviations can be obtained analytically in a special case when the FRET efficiencies of the states are 0 and 1 and in the limiting cases of fast and slow conformational dynamics. These results are compared with the results of numerical simulations. The accuracy and, therefore, the ability to predict model parameters depend on how fast the transition rates are compared to the photon count rate. In the limit of slow transitions, the key parameters that determine the accuracy are the number of transitions between the states and the number of independent photon sequences. In the fast transition limit, the accuracy is determined by the small fraction of photons that are correlated with their neighbors. The relative standard deviation of the relaxation rate has a “chevron” shape as a function of the transition rate in the log-log scale. The location of the minimum of this function dramatically depends on how well the FRET efficiencies of the states are separated. [<http://dx.doi.org/10.1063/1.4904381>]

I. INTRODUCTION

In single-molecule Förster resonance energy transfer (FRET) spectroscopy, likelihood-based methods can be used to get the maximum information about a molecule's conformational dynamics from available experimental data.^{1–6} Fluctuations of the distance between donor and acceptor labels attached to the molecule lead to the fluctuation of the donor and acceptor photon intensity. The sequence of photons can be analyzed using the likelihood function of a model that describes the interdyer distance changes. The parameters of the model are estimated by maximizing the likelihood function.

When the processes being studied are slow, photon sequences can be divided into time bins and studied using various methods, including Hidden Markov Models.^{7,8} The observables are the numbers of donor and acceptor photons detected during bin time. This analysis is applicable when the bin time is much shorter than the time scale of conformational changes.⁹ To study fast conformational dynamics, photon sequences must be analyzed without binning.^{10–16} The observables in this case are the photon colors (i.e., donor, green, or acceptor, red) and photon arrival times (or the interphoton times) so that all information in the measurements are utilized. Such analysis of single-molecule photon sequences has been used to find transition rates of fast folding proteins^{17,18} (in which case there can be several transitions during smallest bin time) and transition path time, which is comparable to the interphoton times.^{5,19,20}

In this paper, we study the likelihood function that is used to analyze photon sequences without binning with recorded photon colors and arrival times.¹² We are interested in the accuracy of the parameters obtained by the Maximum Likelihood (ML) method. The problem of the accuracy of ML estimates has been a focus of numerous studies, including those in single-molecule FRET.^{21,22} There can be two types of errors. First, the model used to construct the likelihood function may not be accurate. Second, even when the model is perfect, there are uncertainties due to the fact that the amount of data is limited. This is the type of errors that is the subject of the present paper.

We consider the simplest two-state model of conformational dynamics. For this model, we study the uncertainties of the parameters obtained from the arrival times and colors of photons. Statistical properties of the ML estimates from photon arrival trajectories for the two-state model have been studied numerically by Waligórska and Molski.¹⁴ The results of simulations for this model have also been reported in Ref. 23. Here, we present an extensive study of the two-state ML estimates and derive analytical expressions for the standard deviations of the extracted parameters in some special and limiting cases.

Our aim is to understand the behavior of the errors in the extracted parameters. We would like to know how the errors depend on various parameters such as the number of photons (photon intensity), transition rates, and duration of the photon sequences. Intuitively, the parameters should be more accurately estimated when the transitions between the states are slow and the separation between the two FRET

efficiencies is large. In this limit, many photons are detected before the interdy distance changes; the states can be readily identified and the sequence of photons can be transformed into a state trajectory. The error in the transition rate decreases when the number of transitions between the two states increases so that faster transitions result in more accurate estimates of the transition rates given photon sequences with the same total duration. In the case when the transition rates are very large and the states of the molecule cannot be directly determined, the accuracy should decrease. In this paper, we make these intuitive ideas more precise and provide simple expressions for the variance of the ML estimates obtained from analytical theory and validated by simulations.

We start the paper by formulating the model and presenting our main results for the errors of the ML estimates in Sec. II. The variances of the ML estimates were obtained analytically in a special case (when the FRET efficiencies in two states are 0 and 1) and in two limiting cases of fast and slow conformational dynamics. The results are discussed and compared with simulations in Sec. III. The last three sections contain the details of the derivation.

II. MODEL AND RESULTS

A. Model

Consider a molecule at equilibrium undergoing transitions between two conformational states with a small and a large interdy distance corresponding to high and low FRET efficiencies. The transitions between the two states are described by the rate constants k_1 ($1 \rightarrow 2$) and k_2 ($2 \rightarrow 1$). The equilibrium populations of the states are $p_1 = k_2/(k_1 + k_2)$ and $p_2 = 1 - p_1$. The statistics of the acceptor and donor photon counts in each state are Poissonian and determined by the acceptor, n_{Ai} ($i = 1, 2$), and donor, n_{Di} , photon count rates (or fluorescence intensities). The photon count rates are the mean numbers of photon counts detected per unit time. These parameters depend on the distance between donor and acceptor fluorophores, laser intensity, background noise, etc.²⁴ The assumption of Poisson statistics implies that all these parameters are fixed when the molecule is in a given conformational state.

The ratio of the acceptor count rate to the total count rate is the apparent FRET efficiency, so that each state can be characterized by the FRET efficiencies $\mathcal{E}_1 = n_{A1}/(n_{A1} + n_{D1})$ and $\mathcal{E}_2 = n_{A2}/(n_{A2} + n_{D2})$, which are the probabilities of observing an acceptor rather than a donor photon when the molecule is in state $i = 1, 2$. Unlike the photon count rates, the apparent FRET efficiencies usually do not depend on the laser intensity.

A sequence of photons with recorded photon colors and arrival times is analyzed by constructing a likelihood function and optimizing it with respect to the model parameters. In general, the likelihood function involves all four photon count rates.^{12,14} Here, we assume that the total (acceptor and donor) photon count rate in two states is the same, $n_{A1} + n_{D1} = n_{A2} + n_{D2} \equiv n$. In this case, the likelihood function can be presented in a form that does not involve photon count rates.¹² The likelihood function of j th sequence of photons (or photon trajectory) with N_j photons of color c_i (acceptor or donor) and interphoton times τ_i is

$$L_j = \mathbf{1}^\top \left(\prod_{i=2}^{N_j} \mathbf{F}(c_i) e^{\mathbf{K}\tau_i} \right) \mathbf{F}(c_1) \mathbf{p}_{eq}, \quad (1)$$

where $\mathbf{F}(\text{acceptor}) = \mathbf{E}$, $\mathbf{F}(\text{donor}) = \mathbf{I} - \mathbf{E}$, $\mathbf{E}$ is the diagonal matrix of FRET efficiencies, $\mathbf{I}$ is the 2×2 unity matrix, $\mathbf{K}$ is the rate matrix, $\mathbf{1}^\top$ is a row vector with unit elements, and $\mathbf{p}_{eq}$ is the column vector of equilibrium populations

$$\mathbf{E} = \begin{pmatrix} \mathcal{E}_1 & 0 \\ 0 & \mathcal{E}_2 \end{pmatrix}, \quad \mathbf{p}_{eq} = \begin{pmatrix} p_1 \\ 1 - p_1 \end{pmatrix}, \quad (2)$$

$$\mathbf{K} = \begin{pmatrix} -k_1 & k_2 \\ k_1 & -k_2 \end{pmatrix} = k \begin{pmatrix} -(1 - p_1) & p_1 \\ 1 - p_1 & -p_1 \end{pmatrix}.$$

Here, $k = k_1 + k_2$ is the relaxation rate. The likelihood function in Eq. (1) reads from the right to the left. Note that the states of the molecule are not observed, so the first photon can be emitted from the first or second state with probabilities p_1 and $p_2 = 1 - p_1$. Therefore, the first term on the right in Eq. (1) is the vector of equilibrium probabilities $\mathbf{p}_{eq}$. The next term $\mathbf{F}(c_1)$ corresponds to the first detected photon and depends on the color of this photon (donor or acceptor). The evolution of the conformational states until the second photon is detected at time τ_2 is described by the matrix exponential, $\exp(\mathbf{K}\tau_2)$. The next term $\mathbf{F}(c_2)$ depends on the color of the second detected photon, and so on. The final multiplication by $\mathbf{1}^\top$ sums over all conformational states.

The above likelihood function can be used to analyze the experiments with both immobilized and diffusing molecules¹⁷ without need to infer photon intensities. In the experiments with freely diffusing molecules, many short photon sequences (bursts of photons) emitted by different molecules are observed.^{25–27} The total likelihood of M independent photon sequences is the product of the likelihoods of individual sequences, $L = \prod_{j=1}^M L_j$. The total duration of all photon sequences is T and the total number of all photons is $N = \sum_{j=1}^M N_j = nT$. It is assumed that there are many photons in each photon sequence (or burst) so that $N \gg M$.

The resulting likelihood function depends on four model parameters, FRET efficiencies of the states $\mathcal{E}_1$ and $\mathcal{E}_2$, the relaxation rate $k = k_1 + k_2$, and the population $p_1 = k_2/k = 1 - p_2$. These parameters are found by numerical optimization of the log-likelihood. We assume that all four parameters are unknown, except for the special case considered in Subsection II B, in which the FRET efficiencies of the states are known.

The likelihood function in Eq. (1) is exact for the model described above. This means that when photon sequences corresponding to the model are analyzed with the likelihood function in Eq. (1), the ML estimates converge to the true values in the limit of large number of photons. When the number of photons is infinite, the likelihood function is a delta-function at the true parameter values. When the number of photons is large, but finite, the likelihood function at the maximum can be approximated by a multidimensional Gaussian. The errors of the parameters can be estimated from the curvature of the likelihood function at the maximum, which is related to the Fisher information.^{21,28,29} This approximation of normal distribution is appropriate when the number of photons is large, so the error of the parameter estimation is small. In addition,

the estimates should not be close to the boundaries (which are 0 and 1 for the FRET efficiencies and population and 0 for the relaxation rate).

Below, we first present the main results (and the key ideas behind them) for the variances of ML estimates, which are found from the curvature of the likelihood function at the maximum. Detailed derivations are given in the subsequent sections. In general, the above likelihood function is a complicated function of parameters and can be studied only numerically. To study this likelihood analytically, we first consider a special case in which the interdyne distance in the two states is much smaller and much larger than the Förster radius, so the FRET efficiencies of the states are 0 and 1. Then, we present the analytic results for arbitrary FRET efficiencies when conformational changes are fast and slow compared to the time between photons. These analytic results are validated by simulations.

B. Exactly solvable model

Assume that the FRET efficiencies in two states are known and equal to 0 and 1. This means that the state of the molecule is known at the moment when a photon is detected, i.e., the donor photons are emitted only when the molecule is in state 1 and the acceptor photons are from state 2 (see Fig. 1(a)). Then, the likelihood function in Eq. (1), which is a product of matrices, simplifies and becomes a product of scalar functions (see Eq. (11)) and can be treated analytically. There are only two parameters that are found by maximizing the log-likelihood, i.e., the relaxation rate k and the population p_1 .

The relative variance of the relaxation rate can be found in terms of special functions (see Sec. IV) and becomes relatively

simple for equally populated states ($p_1 = p_2 = 1/2$)

$$\frac{\langle \delta k^2 \rangle}{k^2} = \frac{1}{N} \frac{4k/n}{\zeta(3, 1+n/(2k))}, \quad (3)$$

where N is the total number of all photons and $\zeta(z, a) \equiv \sum_{i=0}^{\infty} 1/(i+a)^z$ is the Hurwitz zeta function (a generalization of the Riemann zeta function $\zeta(z) \equiv \sum_{i=1}^{\infty} 1/i^z$).

The square root of Eq. (3) is the relative standard deviation of the relaxation rate, which is shown in Fig. 1(b). This function has a “chevron” (or “V”) shape as a function of k/n in the log-log scale, where n is the photon count rate. At small relaxation rates, $k \ll n$, the relative quadratic error is $\langle \delta k^2 \rangle/k^2 \rightarrow 2n/kN = 2/kT$, where $T = N/n$ is the duration of all photon sequences. $kT/2$ is the mean number of transitions during T . Thus, when the relaxation rate is small, the error is large because there are few transitions in all photon sequences. When the relaxation rate is large, $k \gg n$, the error is $\langle \delta k^2 \rangle/k^2 \rightarrow 4k/(Nn\zeta(3))$. The transition between the fast and slow regions is at the minimum of the above function, $k/n \approx 0.756$. Thus, the relative standard deviation of the relaxation rate is large when k is small (small number of transitions) and when k is large (many transitions between two photons). The optimal standard deviation is when the numbers of photons and transitions are comparable.

The relative variance of the population for equally populated states is

$$\frac{\langle \delta p_1^2 \rangle}{p_1^2} = \frac{k/n}{N (\psi(1 + \frac{n}{2k}) - \psi(\frac{n}{2k}) - \frac{k}{n}) + Mk/n}, \quad (4)$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ is the digamma function (the logarithmic derivative of the gamma function). The relative standard deviation of the population (the square root of Eq. (4)) is shown in Fig. 1(c). When the relaxation rate is small, $k \ll n$, then

FIG. 1. Transitions between two states with FRET efficiencies 0 and 1. (a) Acceptor (red) and donor (green) photon trajectory and state (blue) trajectory. Photon arrival times correspond to the centers of circles. The states are equally populated, total photon count rate $n = 100$ in arbitrary time units, two-state relaxation rate $k = 2k_1 = 2k_2 = 100$. (b) and (c) Relative standard deviation of the ML estimate of the relaxation rate (b) and population (c) as a function of the ratio of the relaxation rate, k , and photon count rate, n (the square roots of Eqs. (3) and (4)). The number of photon sequences is $M = 1, 10^2, 10^3, 10^4$; the number of photons in all sequences is $N = 200\,000$.

$\langle \delta p^2 \rangle / p^2 \sim 1/(Nk/(2n) + M) = 1/(kT/2 + M)$. Even when there are few transitions in all photon sequences, the population can be accurately determined if there are enough independent photon sequences ($M \gg 1$). When the relaxation rate is large, $k \gg n$, the quadratic error is $\propto 1/N$ (assuming that $N \gg M$). The population is very accurately determined even when the relaxation rate is much faster than the interphoton times (see Fig. 1(c)). This is because the FRET efficiencies in this special case are assumed to be known, and the population is related to the average FRET efficiency ($\bar{\mathcal{E}} = p_1\mathcal{E}_1 + p_2\mathcal{E}_2 = 1 - p_1$), which can be determined with high accuracy from the total number of photons. This is not the case when all parameters are unknown.

C. Slow transitions

The above results can be extended to arbitrary FRET efficiencies by considering the limiting cases of fast and slow transitions separately. Here, we present the results when the transitions between the states are slow. In this case, there are many photons detected while the molecule stays in one state and so one can readily determine a state trajectory from the sequence of photons (by comparing color patterns or by binning and comparing FRET efficiencies in bins). One can also count the number of transitions between the states and the number of photons in each state. The likelihood function in Eq. (1) can be replaced by a simpler likelihood with new observables which are the number of photons in each state and the times between transitions. The details of the analysis of this likelihood function are given in Sec. V.

The error of the FRET efficiency in this limit does not depend on the relaxation rate

$$\langle \delta \mathcal{E}_i^2 \rangle = \frac{\mathcal{E}_i(1 - \mathcal{E}_i)}{p_i N}, \quad i = 1, 2, \quad (5)$$

where $p_i N$ is the mean number of photons emitted from state i ($i = 1, 2$). The above variance of the FRET efficiency ML estimate is the same as the variance due to shot noise.^{24,30}

The variances of the relaxation rate and of population are

$$\frac{\langle \delta k^2 \rangle}{k^2} = \frac{p_1^2 + p_2^2 + M/kT}{p_1 p_2 (kT + 2M)}, \quad (6a)$$

$$\langle \delta p_1^2 \rangle = \frac{2p_1 p_2}{kT + 2M}, \quad (6b)$$

$$\frac{\langle \delta k \delta p_1 \rangle}{k} = \frac{p_1 - p_2}{kT + 2M}. \quad (6c)$$

The above equations reduce to those for the exactly solvable model in Eqs. (3) and (4) in the limit $k/n \ll 1$ when $p_1 = p_2 = 1/2$.

The errors of the relaxation rate and of the population in Eq. (6) are independent of the FRET efficiencies of the states and photon intensity (n). They are determined by the number of transitions between the states, $kT/2$, and the number of independent photon sequences, M . The relaxation rate and the population are uncoupled when the populations of the states are equal.

When there are more transitions than photon sequences, $kT \gg 2M$, the errors of both k and p_1 can be estimated from

the following simple arguments. The mean number of transitions $1 \rightarrow 2$ (the same as those of $2 \rightarrow 1$ on average) in the photon trajectory of duration $T = N/n$ is $N_{1 \rightarrow 2} = p_1 k_1 T = p_1 p_2 k T$. In the limit of slow transitions, it can be shown (see Sec. V) that parameters k_1 and k_2 are independent and have variance $\langle \delta k_1^2 \rangle / k_1^2 = 1/N_{1 \rightarrow 2} = \langle \delta k_2^2 \rangle / k_2^2$. The error of $k = k_1 + k_2$ can be obtained using $\langle \delta k^2 \rangle = \langle \delta k_1^2 \rangle + \langle \delta k_2^2 \rangle$. The error of $p_1 = k_2/(k_1 + k_2)$ can be obtained by calculating $\langle \delta p_1^2 \rangle = \langle (k_1 \delta k_2 - k_2 \delta k_1)^2 / k^4 \rangle$ and using $\langle \delta k_1 \delta k_2 \rangle = 0$. Similarly, the covariance of k and p_1 can be obtained using $\langle \delta k \delta p_1 \rangle = \langle (\delta k_1 + \delta k_2)(k_1 \delta k_2 - k_2 \delta k_1) / k^2 \rangle$. In this way, we get the results in Eq. (6) with M/kT set to zero.

In the opposite limit when transitions during a single photon sequence are infrequent and there are more sequences than transitions ($kT \ll 2M$), the error of k is still determined by the number of transitions in all photon sequences. However, the error of p_1 is determined by the number of photon trajectories M

$$\begin{aligned} \frac{\langle \delta k^2 \rangle}{k^2} &= \frac{1}{2p_1 p_2 k T}, \\ \langle \delta p_1^2 \rangle &= \frac{p_1 p_2}{M}. \end{aligned} \quad (7)$$

Thus, the populations (but not the rates) can be accurately estimated even when there are no transitions during photon trajectories. Since it has been assumed that each photon trajectory starts randomly, the color of the first photon in the trajectory has information about the equilibrium populations.

D. Fast transitions

The case where the transition rates are fast is more challenging. In the limit of very fast transitions, all photons are uncorrelated and the only parameter that can be accurately determined is the mean FRET efficiency, $\bar{\mathcal{E}} = p_1\mathcal{E}_1 + p_2\mathcal{E}_2$. This is why the errors of all four parameters increase as k/n increases. When k/n is large, but finite, only photons separated by short interphoton times are correlated. The pairs of consecutive photons that are separated by times $\sim 1/k$ provide additional information about those parameters that determine the interphoton time distributions. The number of such pairs is $\propto Nn/k$. The interphoton time distributions involve only the mean FRET efficiency, $\bar{\mathcal{E}}$, the relaxation rate, k , and the FRET efficiency variance due to two-state transitions, $\bar{\mathcal{E}}^2 - \bar{\mathcal{E}}^2 = p_1 p_2 (\mathcal{E}_2 - \mathcal{E}_1)^2$. Consequently, the relative quadratic error of the relaxation rate is $\propto k/n$ (i.e., inversely proportional to the number of the correlated pairs of photons). To get the remaining parameters, $\mathcal{E}_1$, $\mathcal{E}_2$, and p_1 , a significant number of three consecutive photons must be correlated (see Sec. VI). The probability of detecting such photons is proportional to $(n/k)^2$. As a result, the quadratic errors of the FRET efficiencies and of the population increase as $\propto (k/n)^2$.

The details of the derivation are considered in Sec. VI. It will be shown that the quadratic errors (variances) in the limit of large relaxation rate, $k/n \gg 1$, and small separation between FRET efficiencies of the states, $p_1 p_2 (\mathcal{E}_2 - \mathcal{E}_1)^2 \ll \text{Min}(\bar{\mathcal{E}}^2, (1 - \bar{\mathcal{E}})^2)$, are

$$\langle \delta \mathcal{E}_1^2 \rangle = \frac{1}{N} \frac{k^2}{n^2} \frac{4\bar{\mathcal{E}}^3(1-\bar{\mathcal{E}})^3}{p_1^2(\mathcal{E}_2 - \mathcal{E}_1)^4}, \quad (8a)$$

$$\langle \delta \mathcal{E}_2^2 \rangle = \frac{1}{N} \frac{k^2}{n^2} \frac{4\bar{\mathcal{E}}^3(1-\bar{\mathcal{E}})^3}{p_2^2(\mathcal{E}_2 - \mathcal{E}_1)^4}, \quad (8b)$$

$$\langle \delta p_1^2 \rangle = \frac{1}{N} \frac{k^2}{n^2} \frac{16\bar{\mathcal{E}}^3(1-\bar{\mathcal{E}})^3}{(\mathcal{E}_2 - \mathcal{E}_1)^6}, \quad (8c)$$

$$\frac{\langle \delta k^2 \rangle}{k^2} = \frac{1}{N} \frac{k}{n} \frac{8\bar{\mathcal{E}}^2(1-\bar{\mathcal{E}})^2}{p_1^2 p_2^2 (\mathcal{E}_2 - \mathcal{E}_1)^4}. \quad (8d)$$

These equations are applicable only when the variances are small ($\ll 1$). When the error is large, the assumptions used to derive these equations do not hold. The above variances confirm the dependence on k/n mentioned above. They are very sensitive to $\mathcal{E}_2 - \mathcal{E}_1$.

The restriction of small separation between the states can be removed for the case when the populations are equal and $\mathcal{E}_1 + \mathcal{E}_2 = 1$. In this case, it can be shown that

$$\langle \delta \mathcal{E}_1^2 \rangle = \langle \delta \mathcal{E}_2^2 \rangle = \frac{1}{N} \frac{k^2}{n^2} \frac{1}{4\Delta^2 F(\Delta^{-2} - 1)}, \quad (9a)$$

$$\langle \delta p_1^2 \rangle = \frac{1}{N} \frac{k^2}{n^2} \frac{1}{4\Delta^4 F(\Delta^{-2} - 1)}, \quad (9b)$$

$$\frac{\langle \delta k^2 \rangle}{k^2} = \frac{1}{N} \frac{k}{n} \frac{4}{Li_3(\Delta^4)} \left(1 + \frac{Li_2^2(\Delta^4)}{2\ln(1 - \Delta^4) Li_3(\Delta^4)} \right)^{-1}, \quad (9c)$$

where $\Delta = \mathcal{E}_2 - \mathcal{E}_1$, $Li_m(z) = \sum_{k=1}^{\infty} z^k / k^m$ is the polylogarithm function, and $F(z) = 3z \ln z - 8(1+z) \ln(1+z) + 6(2+z) \times \ln(2+z) - (4+z) \ln(4+z) + (z-1)[Li_2(-4/z) - 4Li_2(-2/z) + 4Li_2(-1/z)]$. When the separation between the two FRET efficiencies is small, $\Delta \ll 1$, $Li_m(\Delta^4) \rightarrow \Delta^4$, $F(\Delta^{-2} - 1) \rightarrow \Delta^2$, and Eqs. (9) reduce to Eqs. (8) (with $\bar{\mathcal{E}} = 1/2$, $p_1 = p_2 = 1/2$). When $\mathcal{E}_2 = 1$ and $\mathcal{E}_1 = 0$, $\Delta = 1$, $Li_3(1) = \zeta(3)$, and Eq. (9c) become $\langle \delta k^2 \rangle / k^2 = 4k / (nN\zeta(3))$. This coincides with the exactly solvable model in Eq. (3) in the limit $k/n \gg 1$. However, the population variances in Eqs. (9b) and (4) differ. This is because in the exactly solvable model the FRET efficiencies are assumed to be known, whereas here all four parameters are unknown.

The errors in Eq. (9) depend on the separation of the FRET efficiencies, $\Delta = \mathcal{E}_2 - \mathcal{E}_1$, as Δ^2 and Δ^4 . Since $\Delta \leq 1$, the approximation of small Δ in Eq. (8) turns out to be pretty accurate. Even when $\mathcal{E}_1 = 0.1$ and $\mathcal{E}_2 = 0.9$, the difference between Eq. (9c) and its approximate version Eq. (8d) is only 11%.

III. COMPARISON WITH SIMULATIONS AND DISCUSSION

Simulations to find the errors of the ML estimates were performed for various relaxation rates and FRET efficiencies. We generated photon sequences and analyzed them to find FRET efficiencies $\mathcal{E}_1$ and $\mathcal{E}_2$, the relaxation rate k , and the population of the first state, p_1 . These parameters were found by maximizing the log-likelihood $\sum_{j=1}^M \ln L_j$, with L_j in Eq. (1). The standard deviations were obtained by evaluating numerically the covariance matrix at the maximum from the Hessian matrix. The parameters of the simulations were chosen to keep the errors small (less than 30%), so that the likelihood function

can be well approximated by a multidimensional Gaussian near the maximum.¹⁴ For each data set, we simulated 200 photon sequences of 10 ms duration and photon count rate $n = 100 \text{ ms}^{-1}$. These parameters are similar to those used in single-molecule FRET experiments on fast protein folding.²³ Each estimation experiment was repeated 10 times and the resulting standard deviations were averaged. All sequences used to extract parameters have the same number of photons (200 000) and same total duration (2 s). The states in the model were equally populated; the relaxation rate k varied from 0.01 ms^{-1} to 1000 ms^{-1} . There were four different sets of FRET efficiencies, including ones that are very close (0.4 and 0.6) or well separated (0.1 and 0.9). The values of FRET efficiencies were chosen so that $\mathcal{E}_1 + \mathcal{E}_2 = 1$. This allows us to find simple expressions for the standard deviation analytically, Eq. (9), and to test the applicability of the results in Eq. (8). Simulations with the non-symmetric FRET efficiencies and various state populations show similar trends.²³

The results of the simulations for the standard deviations of the parameters are shown in Fig. 2. It can be seen that there are two qualitatively different regions, which we refer to as slow and fast transition regions. When the transitions between the states are fast, the standard deviations dramatically depend on how different the FRET efficiencies of the states are. When the transitions are slow, the errors are insensitive to the values of FRET efficiencies. The results of the simulations agree with the standard deviations obtained in the limiting cases of fast (Eqs. (8) and (9)) and slow (Eqs. (5) and (6)) transitions, confirming that these equations capture the dependence of the standard deviation on the model parameters.

Let us consider the plots in more detail. The relative standard deviation of the relaxation rate in Fig. 2(a) has a “chevron” (“V”) shape, similar to the analytically solvable model in Fig. 1(b). The shape of the curves in the slow transitions regime is independent of the FRET efficiencies and is described by Eq. (6a) (magenta dashed line). As follows from Eq. (6a), the accuracy of the relaxation rate is mainly determined by the mean number of transitions in all sequences, $kT/2$. It does not matter whether there are many short sequences or few long sequences as long as they have the same total duration.

Unlike the relaxation rate, the accuracy of the population in the slow transition region is determined by the sum of the number of transitions, $kT/2$, and the number of independent sequences, M , as follows from Eq. (6b). In the limit of very slow transitions, the standard deviation of the population does not depend on the relaxation rate (see Fig. 2(b) and Eq. (7)). For the slowest transition rate used in the simulations, $k = 0.01 \text{ ms}^{-1}$, there are only 10 transitions during all 200 sequences on average so that the error of the relaxation rate is large ($>30\%$). However, the error of the population in this case is 7%. This is because the initial state of the molecule in each photon sequence is random, so that additional information about populations comes from the first photon of the sequence.

The FRET efficiencies in the slow transition region are estimated very accurately, independent of the transition rates (see Eq. (5)). When normalized to $\sqrt{\mathcal{E}_i(1-\mathcal{E}_i)}$, $i = 1, 2$, the standard deviations of the FRET efficiency in the first and

FIG. 2. Accuracy of the maximum likelihood estimates for a two-state model as a function of the ratio of the relaxation rate and the photon count rate, k/n , at various FRET efficiencies. Relative standard deviation of the relaxation rate (a) and of the population (b) (circles). (c) Standard deviation of the FRET efficiencies of the first (reverse triangles) and second (triangles) states normalized to $\sqrt{\mathcal{E}_i(1-\mathcal{E}_i)}$, $i = 1, 2$. The states are equally populated, the FRET efficiencies $\mathcal{E}_1$ and $\mathcal{E}_2$ are 0.4 and 0.6 (red), 0.3 and 0.7 (orange), 0.2 and 0.8 (green), and 0.1 and 0.9 (blue). Purple lines are the standard deviations for $\mathcal{E}_1 = 0$ and $\mathcal{E}_2 = 1$ (calculated using Eqs. (3) and (4)). Dashed lines are the errors in the fast transition limit (the square roots of Eqs. (8), black, and (9), cyan) and in the slow transition limit (the square roots of Eqs. (5) and (6), magenta).

second states are the same and equal to $\sqrt{2/N}$ for all FRET efficiencies (see Fig. 2(c)).

Now consider the behavior of the plots in the fast transition region. The standard deviations of all parameters increase as the transition rate increases. As follows from the theory in Sec. VI, the estimates of the FRET efficiencies and the population are coupled in the fast transition region. The mean

FRET efficiency, which is a combination of these parameters ($\bar{\mathcal{E}} = p_1\mathcal{E}_1 + p_2\mathcal{E}_2$), can be determined very accurately even when the transitions are extremely fast. To estimate all three parameters, photon sequences must have a sufficient number of three consecutive photons separated by short times $\sim 1/k$. The probability of such triads is $\propto n^2/k^2$, so the standard deviation of the FRET efficiencies and the population is $\propto k/n$ (see Eqs. (8)). On the other hand, the estimation of the relaxation rate requires only pairs of consecutive photons to be correlated and hence the relative standard deviation of the relaxation rate is $\propto \sqrt{k/n}$ (see Eq. (8d)).

As follows from Eqs. (8), the standard deviations of the FRET efficiencies and the relaxation rate depend on the difference between the FRET efficiencies as $(\mathcal{E}_2 - \mathcal{E}_1)^{-2}$, while the population is even more sensitive to the difference, $(\mathcal{E}_2 - \mathcal{E}_1)^{-3}$. As can be seen in Fig. 2, the simulation results agree with these asymptotics (black dashed lines) when the relaxation rate is large and all standard deviations are small. When any of the standard deviations is large, the Gaussian approximation cannot be used to estimate the variance of the parameters. This is why the simulated data in Fig. 2(c) (red and orange triangles for $k/n = 2, 5, 10$) deviate from the asymptotics.

The results in Eqs. (8) were derived in Sec. VI assuming that the difference between the FRET efficiencies of the states is small. To check how accurate Eqs. (8) are, we also derived the variances in the fast transition region for the special parameter values that were used in the plots in Fig. 2, Eqs. (9), which are applicable for arbitrary separation of the FRET efficiencies (cyan lines in Fig. 2). It turns out that the approximation in Eqs. (8) is pretty accurate and can be used even when the separation between FRET efficiencies is not small.

In the above analysis, all four parameters are unknown. In the exactly solvable model, which we also plot in Figs. 2(a) and 2(b) for the reference, the FRET efficiencies are assumed to be known (0 and 1). In the slow transition region, the estimates of the FRET efficiencies are independent from the estimates of the rates and population, so the plots agree with the exactly solvable model. In the fast transition region, the error of the population in the exactly solvable model is small (see Fig. 2(b), purple line). This is because the FRET efficiencies are known and the population in the exactly solvable model is the same as the average FRET efficiency, which can be estimated very accurately.

Note that in the slow transition region, the standard deviations of both the relaxation rate and the population are independent of photon intensity n . The key parameter here is the number of transitions during all photon sequences rather than the number of photons. In the slow transition region, increasing photon intensity does not improve the accuracy of the transition rates. In contrast, in the fast transition region, the standard deviations depend on the photon intensity as $1/n$ (the relaxation rate) and $1/n\sqrt{n}$ (the population and FRET efficiencies). The relaxation rate where the slow transition region changes to the fast transition region (the minimum in the plot in Fig. 2) can be estimated by equating the fast and slow asymptotics in Eqs. (8d) and (6a) (with $M \ll kN/n$)

$$k/n \sim \frac{p_1 p_2 (\mathcal{E}_2 - \mathcal{E}_1)^2}{2\bar{\mathcal{E}}(1-\bar{\mathcal{E}})} \sqrt{\frac{p_1^2 + p_2^2}{2p_1 p_2}}. \quad (10)$$

This border between the two regions is intensity-dependent. It moves towards larger relaxation rates as the photon count rate increases.

The above results for the standard deviations allow one to determine the amount of data required to estimate parameters to a given accuracy. For example, when the transitions between the two states are slow, one needs the trajectories with at least 100 transitions between the states in order to estimate the transition rates with the accuracy 10%. When the transition rate is fast and comparable to the photon count rate ($k \sim n$), one needs at least $\sim 1\,560\,000$ photons to estimate all four parameters with the accuracy 10% for the molecule with equally populated states and $\mathcal{E}_1 = 0.4$ and $\mathcal{E}_2 = 0.6$. However, when $\mathcal{E}_1 = 0.3$ and $\mathcal{E}_2 = 0.7$, the amount of required photons reduces to $\sim 25\,000$. It should be noted that these estimates of the required data are for the ideal case when the model is known and accurate. In reality, there are other factors, such as fluorophore blinking or submicrosecond chain dynamics, that lead to the deviation from the assumed two-state kinetic model and result in additional errors. The effect of these factors on the accuracy of the ML estimates is beyond the scope of this paper.

In summary, the analysis of the variances of the ML estimates shows that there are two regions of parameter space where the accuracy of these estimates is qualitatively different. In the slow transition region, there are two key parameters, i.e., the number of transitions that occur during all photon sequences and the number of independent photon sequences. The error of the transition rate is determined by the number of transitions (see Eq. (6a)), whereas the error of the population, Eq. (6b), is determined by both parameters. The error of the FRET efficiencies in the slow transition region is small and depends on the number of photons emitted by the corresponding states in all photon sequences (see Eq. (5)). In the fast transition region, the accuracy of the relaxation rate is determined by the number of correlated pairs of photons, Eq. (8d), whereas the accuracy of the population and FRET efficiencies are determined by the number of correlated triplets of photons, Eqs. (8a)–(8c).

The rest of the paper deals with the technical aspects of the derivation of the results in the paper.

IV. EXACTLY SOLVABLE MODEL

In this section, we consider the special case when the FRET efficiencies in the two states are 0 and 1 and provide the details of the derivation of the parameter variances. The estimated parameters are just the relaxation rate and the population, since the FRET efficiencies are assumed to be known. In this case, the state of the molecule is known at the time the photon is detected, i.e., if the photon is green (i.e., from the donor), it has been emitted by the molecule in the first state with $\mathcal{E}_1 = 0$; red (acceptor) photons correspond to the second state with $\mathcal{E}_2 = 1$ (see Fig. 1(a)). The color of i th photon, c_i , is directly connected to the state of the molecule (donor $\rightarrow$ state 1, acceptor $\rightarrow$ state 2), therefore, in this particular case, c_i denotes both the color of the photon and the state of the molecule. The product of matrices in the likelihood function, Eq. (1), becomes

the product of scalar functions (because matrices $\mathbf{F}(c_i)$ have only one non-zero element)

$$L_j = \prod_{i=1}^{N_j-1} G(c_{i+1}, \tau_{i+1}; c_i) p_{in}. \quad (11)$$

Here, $G(c_{i+1}, \tau_{i+1}; c_i)$ depends on the colors of the consecutive photons, c_i and c_{i+1} , and the time between i th and $(i+1)$ th photon, τ_{i+1} ; the product is over all photons; p_{in} is the population of the state corresponding to the first photon in the trajectory (i.e., p_1 when the first photon is green and p_2 when it is red). The conditional transition probability $G(\alpha, \tau; \beta)$ ($\alpha, \beta = 1, 2$) is the probability to be in state α at time τ provided that initially the molecule was in state β

$$G(\alpha, \tau; \beta) \equiv G_{\alpha\beta}(\tau) = [\exp(\mathbf{K}\tau)]_{\alpha\beta} = p_\alpha + (\delta_{\alpha\beta} - p_\alpha) e^{-k\tau}, \quad (12)$$

where $k = k_1 + k_2$ is the two-state relaxation rate, $p_1 = k_2/(k_1 + k_2) = 1 - p_2$, $\delta_{\alpha\beta}$ is the Kronecker delta ($\delta_{\alpha\beta}$ is 1 when α and β are equal and 0 otherwise). The last equality is obtained by evaluating the matrix exponential, $\exp(\mathbf{K}\tau)$, with $\mathbf{K}$ from Eq. (2).

The likelihoods L_j of various photon sequences are multiplied together, so that the log-likelihood of M sequences is

$$\ln L = \ln \prod_{j=1}^M L_j = \sum_i \ln G(c_{i+1}, \tau_{i+1}; c_i) + M_1 \ln p_1 + M_2 \ln p_2, \quad (13)$$

where M_1 and $M_2 = M - M_1$ are the numbers of the trajectories that start with donor and acceptor photon, respectively. The sum over i is performed over photons in all sequences. This log-likelihood depends on the observed interphoton times τ_i and photon colors c_i . The log-likelihood is a function of two variables, the population p_1 and the relaxation rate k .

The parameters and their errors are found by optimizing numerically the log-likelihood and evaluating the curvature of the likelihood function at the maximum. To obtain the errors analytically in the limit of large photon counts, we regroup the terms in Eq. (13) by color and replace the summation by an integration over the distribution of interphoton times. For example, the summation with respect to the times between two acceptor consecutive photons is replaced by

$$\begin{aligned} \sum_i \ln G(A, \tau_{i+1}; A) &= \sum_i \ln G_{22}(\tau_{i+1}) \\ &\rightarrow N \int_0^\infty [\ln G_{22}(t)] G_{22}(t) p_2 n e^{-nt} dt, \end{aligned} \quad (14)$$

where N is the total number of photons in all photon trajectories and n is the total count rate, which is the same in both states. Here, $n \exp(-nt)$ is the probability density that two consecutive photons are separated by time t and $G_{22}(t) p_2$ is the probability that both these photons are red (acceptor) and thus were emitted by the molecule in state 2. (More accurately, the factor in the above equation is $N - M$ instead of N , which accounts for the last photon in the trajectory, however, since we assume $N \gg M$, this correction can be neglected.)

Similar replacements are performed for the sums with other colors, so that, for any function $\phi(c_{i+1}, \tau_{i+1}; c_i) \equiv \phi_{c_{i+1}, c_i}(\tau_{i+1})$ that depends on colors c_i

$$\sum_i \phi(c_{i+1}, \tau_{i+1}; c_i) \rightarrow N \sum_{\alpha\beta} \int_0^\infty \phi_{\alpha\beta}(t) G_{\alpha\beta}(t) p_\beta n e^{-nt} dt. \quad (15)$$

Using this in Eq. (13), we get the log-likelihood function

$$\ln L = N \sum_{\alpha, \beta=1,2} \int_0^\infty [\ln G'_{\alpha\beta}(t)] G_{\alpha\beta}(t) p_\beta n e^{-nt} dt + M[p_1 \ln p'_1 + (1-p_1) \ln(1-p'_1)]. \quad (16)$$

Here, we used the fact that, in the limit of large number of photons, the number of photon sequences starting with the donor (acceptor) photon is $M_1 = p_1 M$ ($M_2 = p_2 M$). The primes in $\ln G'_{\alpha\beta} \equiv \ln G_{\alpha\beta}(t; k', p'_1)$ indicate the dependence on the parameters that are being varied, k' and p'_1 . These differ from the true parameters k and p_1 .

The estimates for the population and relaxation rate are found by maximizing the log-likelihood with respect to $v_1 = k'$ and $v_2 = p'_1$. For large numbers of photons, these estimates converge to the true values. The variances are found from the curvature of the log-likelihood at the maximum. The curvature is determined by the Hessian matrix $\mathbf{H}$ (the Fisher information matrix) with the elements $[\mathbf{H}]_{ij} = -\partial_{v_i} \partial_{v_j} \ln L$. The inverse of the Hessian matrix results in the covariance matrix in the limit

of large number of photons, which determines the errors of the extracted parameters

$$\begin{pmatrix} \langle \delta k^2 \rangle & \langle \delta k \delta p_1 \rangle \\ \langle \delta k \delta p_1 \rangle & \langle \delta p_1^2 \rangle \end{pmatrix} = \mathbf{H}^{-1}. \quad (17)$$

The first derivatives of the log-likelihood in Eq. (16) are zero when $k' = k$ and $p'_1 = p_1$, as expected. This can be verified by differentiating Eq. (16) and using $d \ln G'_{\alpha\beta} / dv_i = (dG'_{\alpha\beta} / dv_i) / G'_{\alpha\beta}$ and $\sum_\alpha G_{\alpha\beta} = 1$.

Calculation of the second derivative can be simplified using

$$G_{\alpha\beta} \frac{d^2 \ln G'_{\alpha\beta}}{dv_i dv_j} = \frac{G_{\alpha\beta}}{G'_{\alpha\beta}} \frac{d^2 G'_{\alpha\beta}}{dv_i dv_j} - \frac{G_{\alpha\beta}}{G'^2_{\alpha\beta}} \frac{dG'_{\alpha\beta}}{dv_i} \frac{dG'_{\alpha\beta}}{dv_j}.$$

The first term in the right-hand side vanishes after removing the primes (i.e., setting $k' = k$ and $p'_1 = p_1$) and summing using $\sum_\alpha G_{\alpha\beta} = 1$. The second term contributes to the second derivative of Eq. (16) at the maximum

$$-\frac{d^2 \ln L}{dv_i dv_j} = N \sum_{\alpha, \beta=1}^2 \int_0^\infty \frac{(dG_{\alpha\beta}/dv_i)(dG_{\alpha\beta}/dv_j)}{G_{\alpha\beta}(t)} p_\beta n e^{-nt} dt + \delta_{i2} \delta_{j2} \frac{M}{p_1 p_2}. \quad (18)$$

Using the expressions for $G_{\alpha\beta}$ in Eq. (12) and replacing $z = kt$, we find that the elements of the Hessian matrix are

$$-k^2 \frac{d^2 \ln L}{dk^2} = \mathbf{H}_{11} = N \frac{n}{k} p_1 p_2 \int_0^\infty z^2 \left(\frac{p_2}{p_1 + p_2 e^{-z}} + \frac{p_1}{p_2 + p_1 e^{-z}} + \frac{2}{1 - e^{-z}} \right) e^{-(2+n/k)z} dz, \quad (19a)$$

$$-\frac{d^2 \ln L}{dp_1^2} = \mathbf{H}_{22} = N \frac{n}{k} \int_0^\infty (1 - e^{-z})^2 \left(\frac{p_1}{p_1 + p_2 e^{-z}} + \frac{p_2}{p_2 + p_1 e^{-z}} \right) e^{-nz/k} dz + N \frac{k(p_1^2 + p_2^2)}{(k+n)p_1 p_2} + \frac{M}{p_1 p_2}, \quad (19b)$$

$$-k \frac{d^2 \ln L}{dk dp_1} = \mathbf{H}_{12} = N \frac{n}{k} p_1 p_2 \int_0^\infty z(1 - e^{-z}) \left(\frac{1}{p_2 + p_1 e^{-z}} - \frac{1}{p_1 + p_2 e^{-z}} \right) e^{-(1+n/k)z} dz + N \frac{kn(p_2 - p_1)}{(k+n)^2}. \quad (19c)$$

The above integrals can be expressed in terms of special functions (Lerch's transcendent, the Hurwitz zeta function, and the hypergeometric function). The covariance matrix is the inverse of the Hessian matrix (see Eq. (17)). When the two states are equally populated, the off-diagonal terms vanish. The variances of the relaxation rate and populations in this case are the reciprocals of Eqs. (19a) and (19b), which can be shown to be given by Eqs. (3) and (4).

V. SLOW TRANSITIONS

In this section, we consider photon sequences in the limit when the transitions between the states are slow. All four parameters (the relaxation rate, population, and FRET efficiencies) are assumed to be unknown. An observed sequence of photons may result from different state trajectories. However, when transitions between the states are rare and there are enough photons so that the identity of a state can be determined before the state changes, there is a single (most probable) state trajectory defined from the observed photons, and all other state trajectories are very unlikely. In this section, we

study such a case assuming that the state trajectory can be readily obtained from the observed photon trajectory. Once the state trajectory has been found, then the new observables are the numbers of photons emitted by each state, the number of transitions between the states, the time spent in each state, and the initial state for each trajectory (when there are many independent photon sequences).

The corresponding likelihood function of M photon trajectories is

$$\prod_{j=1}^M L_j \sim \mathcal{E}_1^{N_{A1}} \mathcal{E}_2^{N_{A2}} (1 - \mathcal{E}_1)^{N_{D1}} (1 - \mathcal{E}_2)^{N_{D2}} e^{-k_1 T_1 - k_2 T_2} \times k_1^{N_{1 \rightarrow 2}} k_2^{N_{2 \rightarrow 1}} p_1^{M_1} (1 - p_1)^{M_2}, \quad (20)$$

where N_{Ai} , N_{Di} are the total numbers of the acceptor and donor photons in state i ($i = 1, 2$), T_i is the total time spent in state i in all trajectories, $N_{1 \rightarrow 2}$ and $N_{2 \rightarrow 1}$ are the numbers of $1 \rightarrow 2$ and $2 \rightarrow 1$ transitions, respectively; M_i is the number of photon trajectories that start from state i . In this equation, $p_1^{M_1} (1 - p_1)^{M_2}$ is the probability that M_1 photon sequences start with the molecule in the first state and $M_2 = M - M_1$ sequences

start with the molecule is in second state. $\exp(-k_1T_1 - k_2T_2)$ is the probability that there are no transitions during times T_1 and T_2 . Since the probability of one $1 \rightarrow 2$ transition during short time Δt is $k_1\Delta t$, $N_{1 \rightarrow 2}$ transitions result in the factor $k_1^{N_{1 \rightarrow 2}}$ in the above likelihood. Similarly, $k_2^{N_{2 \rightarrow 1}}$ corresponds to the $2 \rightarrow 1$ transitions. Finally, we note that $\mathcal{E}_i$ and $1 - \mathcal{E}_i$ are the probabilities that a photon emitted by the molecule in state i ($i = 1, 2$) is red (acceptor) and green (donor), respectively. Therefore, N_{Ai} acceptor and N_{Di} donor photons result in the factor $\mathcal{E}_i^{N_{Ai}}(1 - \mathcal{E}_i)^{N_{Di}}$.

The log of the above likelihood is

$$\begin{aligned} \ln L = \ln \prod_{j=1}^M L_j \sim & N_{A1} \ln \mathcal{E}_1 + N_{D1} \ln(1 - \mathcal{E}_1) \\ & + N_{A2} \ln \mathcal{E}_2 + N_{D2} \ln(1 - \mathcal{E}_2) + N_{1 \rightarrow 2} \ln k_1 - k_1 T_1 \\ & + N_{2 \rightarrow 1} \ln k_2 - k_2 T_2 + M_1 \ln p_1 + M_2 \ln(1 - p_1). \end{aligned} \quad (21)$$

When there are just few sequences, the last two terms in the above equation can be neglected. In this case, the parameters $\mathcal{E}_1$, $\mathcal{E}_2$, k_1 , and k_2 are independent. The most likely FRET efficiencies are found by setting the first derivatives with respect to $\mathcal{E}_i$ to zero, resulting in the estimates $\mathcal{E}_i = N_{Ai}/(N_{Ai} + N_{Di})$, $i = 1, 2$. In the limit of large photon counts, $N_{Ai} \rightarrow \mathcal{E}_i p_i N$, $N_{Di} \rightarrow (1 - \mathcal{E}_i) p_i N$, where N is the total number of photons, so that the maximum likelihood estimates converge to the true parameters. The errors of the FRET efficiencies are obtained from the second derivatives, $\langle \delta \mathcal{E}_i^2 \rangle = (d^2 \ln L / d \mathcal{E}_i^2)^{-1} = N_{Ai} N_{Di} / (N_{Ai} + N_{Di})^3$. In the limit of large photons, this becomes $\langle \delta \mathcal{E}_i^2 \rangle = \mathcal{E}_i(1 - \mathcal{E}_i)/N$.

Similarly, most likely rates are estimated as $k_i = N_{i \rightarrow j} / T_i$ and the corresponding variances are $\langle \delta k_i^2 \rangle / k_i^2 = (d^2 \ln L / d k_i^2)^{-1} / k_i^2 = 1 / N_{i \rightarrow j}$. In the limit of large number of photons, $T_i = p_i T$, $N_{1 \rightarrow 2} = N_{2 \rightarrow 1} = p_1 p_2 k T$, so that the ML estimates of k_i reduce to the true values and $\langle \delta k_i^2 \rangle / k_i^2 = 1 / (p_1 p_2 k T)$.

When the number of photon sequences is large so the last terms in Eq. (21) become significant, $\mathcal{E}_i$ are still independent, but k_1 and k_2 are coupled. We consider the estimates of the relaxation rate k and the population p_1 in this case. Replacing $k_1 = p_2 k$ and $k_2 = p_1 k$ and rearranging the terms, we get from Eq. (21)

$$\begin{aligned} \ln L = \ln \prod_{j=1}^M L_j \sim & N_{A1} \ln \mathcal{E}_1 + N_{D1} \ln(1 - \mathcal{E}_1) \\ & + N_{A2} \ln \mathcal{E}_2 + N_{D2} \ln(1 - \mathcal{E}_2) \\ & + (N_{1 \rightarrow 2} + N_{2 \rightarrow 1}) \ln k - k(p_1 T_2 + (1 - p_1) T_1) \\ & + (M_1 + N_{2 \rightarrow 1}) \ln p_1 + (M_2 + N_{1 \rightarrow 2}) \ln(1 - p_1). \end{aligned} \quad (22)$$

The estimation of the FRET efficiencies is the same as before, the variances of the FRET efficiencies result in Eq. (5). To find the ML estimates of the relaxation rate and the population, we set the first derivatives of the above equation with respect to k and p_1 to zero. One can verify that these estimates coincide with the true values in the limit of large number of photons. The variances of the parameters $v_1 = k$ and $v_2 = p_1$ are obtained by evaluating the second derivatives at the maximum. In this way, we find that the elements of the Hessian matrix

$[\mathbf{H}]_{ij} = -\partial_{v_i} \partial_{v_j} \ln L$ are

$$\begin{aligned} -\frac{\partial^2}{\partial k^2} \ln L &= \frac{N_{1 \rightarrow 2} + N_{2 \rightarrow 1}}{k^2}, \\ -\frac{\partial^2}{\partial p_1^2} \ln L &= \frac{M_1 + N_{2 \rightarrow 1}}{p_1^2} + \frac{M_2 + N_{1 \rightarrow 2}}{(1 - p_1)^2}, \\ -\frac{\partial^2}{\partial p_1 \partial k} \ln L &= T_2 - T_1. \end{aligned} \quad (23)$$

Inverting the Hessian matrix and replacing $T_i = p_i T$, $M_i = p_i M$, ($i = 1, 2$), and $N_{1 \rightarrow 2} = N_{2 \rightarrow 1} = p_1 p_2 k T$, we get the variance and covariance in the limit of large number of photons given in Eq. (6).

VI. FAST TRANSITIONS

In this section, we consider the case when the conformational relaxation rate is much larger than the photon count rate. The number of independent photon trajectories does not matter in this limit, so we consider just one trajectory. When the transitions between the states are infinitely fast, the correlation between any two photons is lost, so that the likelihood function in Eq. (1) converges to

$$L = \prod_i g(c_i) = \bar{\mathcal{E}}^{N_A} (1 - \bar{\mathcal{E}})^{N_D}, \quad (24)$$

where

$$g(c_i) = \mathbf{1}^T \mathbf{F}(c_i) \mathbf{p}_{eq}. \quad (25)$$

Here, c_i is the color of i th photon (donor or acceptor), $g(\text{acceptor}) = \bar{\mathcal{E}}$, $g(\text{donor}) = 1 - \bar{\mathcal{E}}$, $\bar{\mathcal{E}} = \mathcal{E}_1 p_1 + \mathcal{E}_2 p_2$ is the mean FRET efficiency. The product in Eq. (24) is over all photons, N_A and $N_D = N - N_A$ are the total numbers of acceptor and donor photons, respectively. In this likelihood function, there is no information about the relaxation rate. Only the mean FRET efficiency $\bar{\mathcal{E}}$ can be accurately determined with the variance $\langle \delta \bar{\mathcal{E}}^2 \rangle = \bar{\mathcal{E}}(1 - \bar{\mathcal{E}})/N$.

When the transition rates are large but finite, only photons separated by short interphoton times ($\sim 1/k$) are correlated. They determine the dependence of the likelihood on the rates. Assume that no more than two consecutive photons are correlated. Then, the likelihood function in Eq. (1) can be approximated as

$$L = \prod_i \frac{g(c_{i+1}, \tau_{i+1}; c_i)}{g(c_{i+1})}, \quad (26)$$

where the product is over all photons and

$$g(c_{i+1}, \tau_{i+1}; c_i) = \mathbf{1}^T \mathbf{F}(c_{i+1}) e^{K \tau_{i+1}} \mathbf{F}(c_i) \mathbf{p}_{eq}. \quad (27)$$

For example, for two consecutive acceptor photons separated by time τ , $c_i = c_{i+1} = A$ and $g(A, \tau; A) \equiv g_{AA}(\tau) = \bar{\mathcal{E}}^2 + \sigma_{\bar{\mathcal{E}}}^2 \exp(-k\tau)$, where $\sigma_{\bar{\mathcal{E}}}^2 = p_1 p_2 (\mathcal{E}_2 - \mathcal{E}_1)^2$. This function is equal to the auto-correlation function of the FRET efficiency $\langle \mathcal{E}(\tau) \mathcal{E}(0) \rangle$ when the system jumps between $\mathcal{E}_1$ and $\mathcal{E}_2$.

When the interphoton times are longer than the times between transitions, $g(c_{i+1}, \tau_{i+1}; c_i) / g(c_{i+1}) \rightarrow g(c_i)$, and thus Eq. (26) reduces to Eq. (24). In the special case, when the FRET efficiency of the first and second states are 0 and 1, respectively, $g(D) = p_1$, $g(A) = p_2$, and $g(c_{i+1}, \tau_{i+1}; c_i) / g(c_{i+1})$

$= G(c_{i+1}, \tau_{i+1}; c_i)$, where $G(\alpha, \tau; \beta)$ is the Green's function defined in Eq. (12). The likelihood function in Eq. (26) in this case coincides with Eq. (11) and is, therefore, exact.

The above approximation scheme can be extended to include correlation between three consecutive photons

$$L = \prod_i \frac{g(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i)}{g(c_{i+2}, \tau_{i+2}; c_{i+1})}, \quad (28)$$

where $g(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i)$ involves three consecutive photons

$$g(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i) = \mathbf{1}^\top \mathbf{F}(c_{i+2}) e^{K\tau_{i+2}} \mathbf{F}(c_{i+1}) e^{K\tau_{i+1}} \mathbf{F}(c_i) \mathbf{p}_{eq}. \quad (29)$$

If any one of the τ 's is long compared to the time between transitions ($k\tau \gg 1$), the corresponding factor in the product in Eq. (28) becomes equal to that in Eqs. (24) and (26). For example, when $\tau_{i+2} \rightarrow \infty$, $g(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i)/g(c_{i+2}, \tau_{i+2}; c_{i+1}) \rightarrow g(c_{i+1}, \tau_{i+1}; c_i)/g(c_{i+1})$. Only when all three photons are bunched together is there a difference between Eqs. (28) and (26).

We now take the log of Eq. (28), regroup the terms in the sums by color (e.g., the sum of $\ln g(c_{i+1}, \tau_{i+1}; c_i)$ has four terms, corresponding to the acceptor-acceptor, donor-donor, acceptor-donor, and donor-acceptor photons), and replace the sums by the averages. For any functions, $\phi(c_{i+1}, \tau_{i+1}; c_i) \equiv \phi_{c_{i+1}, c_i}(\tau_{i+1})$ and $\phi(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i) \equiv \phi_{c_{i+2}, c_{i+1}, c_i}(\tau_{i+2}, \tau_{i+1})$ that depend on colors

$$\begin{aligned} & \sum_i \phi(c_{i+1}, \tau_{i+1}; c_i) \\ & \rightarrow N \sum_{\alpha\beta} \int_0^\infty \phi_{\alpha\beta}(t) g_{\alpha\beta}(t) n e^{-nt} dt, \\ & \sum_i \phi(c_{i+2}, \tau_{i+2}; c_{i+1}, \tau_{i+1}; c_i) \\ & \rightarrow N \sum_{\alpha\beta\gamma} \int_0^\infty \int_0^\infty \phi_{\alpha\beta\gamma}(t_1, t_2) g_{\alpha\beta\gamma}(t_1, t_2) n^2 e^{-n(t_1+t_2)} dt_1 dt_2, \end{aligned} \quad (30)$$

where the indexes α, β, γ indicate color (donor, D , or acceptor, A) $g_{\alpha\beta}(t) = g(\alpha, t; \beta)$ and $g_{\alpha\beta\gamma}(t_1, t_2) = g(\alpha, t_1; \beta, t_2; \gamma)$. The first equation here is similar to Eq. (15), except that $G_{\alpha\beta}(t)p_\beta$ is replaced by a more general probability $g_{\alpha\beta}(t)$ that two photons separated by time t have colors α and β . In the second equation, three consecutive photons are involved in the averaging; $g_{\alpha\beta\gamma}(t_1, t_2)$ is the probability that these photons have colors α, β , and γ and $n^2 \exp(-n(t_1+t_2))$ is the probability density that the times between the photons are t_1 and t_2 .

Applying the averaging in Eq. (30) to the log of Eq. (28), we have

$$\begin{aligned} \ln L &= N \sum_{\alpha\beta\gamma} \int_0^\infty \int_0^\infty [\ln g'_{\alpha\beta\gamma}(t_1, t_2)] g_{\alpha\beta\gamma}(t_1, t_2) n^2 e^{-n(t_1+t_2)} dt_1 dt_2 \\ & - N \sum_{\alpha\beta} \int_0^\infty [\ln g'_{\alpha\beta}(t)] g_{\alpha\beta}(t) n e^{-nt} dt. \end{aligned} \quad (31)$$

Here, the primes in $g'_{\alpha\beta\gamma}(t_1, t_2)$ and $g'_{\alpha\beta}(t)$ indicate that these functions depend on the parameters that will be found by

optimizing the log-likelihood (the variables of the likelihood function) as opposed to the parameters that are given (or used to simulate photon sequences). The sums in the above expressions are over all color combinations. For example, the second sum includes four terms. Explicitly, the functions $g_{\alpha\beta}(t)$ and $g_{\alpha\beta\gamma}(t)$ in the above equations can be written as

$$\begin{aligned} g_{\alpha\beta}(t) &= g_\alpha g_\beta + S_\alpha S_\beta \sigma_\varepsilon^2 e^{-kt}, \\ g_{\alpha\beta\gamma}(t_1, t_2) &= g_\alpha g_\beta g_\gamma + S_\alpha S_\beta g_\gamma \sigma_\varepsilon^2 e^{-kt_1} + g_\alpha S_\beta S_\gamma \sigma_\varepsilon^2 e^{-kt_2} \\ & + S_\alpha x_\beta S_\gamma \sigma_\varepsilon^2 e^{-k(t_1+t_2)}, \end{aligned} \quad (32)$$

where $g_\alpha = g(\alpha)$, which is equal to $\bar{\mathcal{E}}$, when α is acceptor, or $1 - \bar{\mathcal{E}}$, when α is donor, S_α is the sign function ($S_A = 1$ and $S_D = -1$), $\sigma_\varepsilon^2 = p_1 p_2 (\mathcal{E}_2 - \mathcal{E}_1)^2$ is the FRET efficiency variance, and $x_A = p_1 \mathcal{E}_2 + p_2 \mathcal{E}_1 = \mathcal{E}_1 + \mathcal{E}_2 - \bar{\mathcal{E}}$, $x_D = 1 - x_A$.

The log-likelihood in Eq. (31) is our approximation in the case of fast transitions. In the limit of very fast transitions, $\ln L \rightarrow N \sum_\alpha g_\alpha \ln g'_\alpha$, which corresponds to the likelihood in Eq. (24) where all photons are uncorrelated. It involves only the mean FRET efficiency, which can be determined very accurately. Other parameters cannot be determined from the likelihood in Eq. (24). The approximation in Eq. (26) is the correction to the limit of uncorrelated photons. It involves the FRET efficiency variance σ_ε^2 and the relaxation rate k . To determine all four parameters, we have to consider the approximation in Eq. (28) (and (31)), which involves three bunched photons.

The parameters we are interested in are $u_1 = \mathcal{E}_1$, $u_2 = \mathcal{E}_2$, $u_3 = p_1$, and $u_4 = k$. Explicit expressions in Eq. (32) show that more convenient parameters are $v_1 = \bar{\mathcal{E}} = p_1 \mathcal{E}_1 + (1 - p_1) \mathcal{E}_2$, $v_2 = k$, $v_3 = \sigma_\varepsilon^2 = p_1(1 - p_1)(\mathcal{E}_1 - \mathcal{E}_2)^2$, and $v_4 = x_A = p_1 \mathcal{E}_2 + (1 - p_1) \mathcal{E}_1$. To get the variances of the ML estimates of u_i , we evaluate the second derivatives with respect to v_i , convert them into the derivatives with respect to u_i , and invert the resulting Hessian matrix. The derivatives with respect to u_i are $\partial^2 \ln L / \partial u_i \partial u_j = \sum_{k,n} (\partial^2 \ln L / \partial v_k \partial v_n) (\partial v_k / \partial u_i) (\partial v_n / \partial u_j)$ (the first derivatives of $\ln L$ are zero at the maximum). In matrix notation, the Hessian matrixes $\mathbf{H}_v$ (with the elements $[\mathbf{H}_v]_{ij} = -\partial_{v_i} \partial_{v_j} \ln L$) and $\mathbf{H}_u$ (with the elements $[\mathbf{H}_u]_{ij} = -\partial_{u_i} \partial_{u_j} \ln L$) are related by

$$\mathbf{H}_u = (\partial_u \mathbf{v})^\top \cdot \mathbf{H}_v \cdot \partial_u \mathbf{v}, \quad (33)$$

where $\partial_u \mathbf{v}$ is the matrix with elements $\partial v_i / \partial u_j$

$$\partial_u \mathbf{v} = \begin{pmatrix} p_1 & 1-p_1 & -\Delta & 0 \\ 0 & 0 & 0 & 1 \\ -2p_1(1-p_1)\Delta & 2p_1(1-p_1)\Delta & (1-2p_1)\Delta^2 & 0 \\ 1-p_1 & p_1 & \Delta & 0 \end{pmatrix}, \quad (34)$$

where $\Delta = \mathcal{E}_2 - \mathcal{E}_1$. The variances are the diagonal terms of $\mathbf{H}_u^{-1}$.

The first derivative of Eq. (31) with respect to a parameter v leads to the expression that contains $\sum_{\alpha\beta\gamma} (dg'_{\alpha\beta\gamma}/dv) \times (g_{\alpha\beta\gamma}/g'_{\alpha\beta\gamma})$ in the first term and a similar expression in the second term. When the parameters coincide with the true ones, $g' = g$ and the derivative is zero as it should be since the sum of $g'_{\alpha\beta\gamma}$ over all colors is 1. In the second derivative, we use

$$g \frac{d^2 \ln g'}{dv^2} = \frac{g}{g'} \frac{d^2 g'}{dv^2} - \frac{g}{g'^2} \left(\frac{dg'}{dv} \right)^2$$

and the fact that in the maximum $g' = g$, so that the term with the second derivative disappears after summation over the

colors. Therefore, the derivative with respect to parameters v_i and v_j is

$$\begin{aligned} -\frac{d^2 \ln L}{dv_i dv_j} &= N \sum_{\alpha\beta\gamma} \int_0^\infty \frac{(dg_{\alpha\beta\gamma}/dv_i)(dg_{\alpha\beta\gamma}/dv_j)}{g_{\alpha\beta\gamma}(t_1, t_2)} n^2 e^{-n(t_1+t_2)} dt_1 dt_2 \\ &\quad - N \sum_{\alpha\beta} \int_0^\infty \frac{(dg_{\alpha\beta}/dv_i)(dg_{\alpha\beta}/dv_j)}{g_{\alpha\beta}(t)} n e^{-nt} dt. \end{aligned} \quad (35)$$

We now will evaluate these derivatives to lowest order in n/k . Consider the second derivative with respect to $v_1 = \bar{\epsilon}$. The terms in $g_{\alpha\beta\gamma}$ in Eq. (32) that contain $\sigma_{\bar{\epsilon}}^2 \exp(-kt)$ can be neglected because they result in the terms $\propto n/k$, which are small. Therefore,

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_1^2} &= N \sum_{\alpha\beta\gamma} \frac{(d(g_{\alpha\beta}g_{\beta\gamma})/dv_1)^2}{g_{\alpha\beta}g_{\beta\gamma}} - N \sum_{\alpha\beta} \frac{(d(g_{\alpha\beta})/dv_1)^2}{g_{\alpha\beta}} \\ &= N \sum_{\alpha} \frac{(dg_{\alpha}/dv_1)^2}{g_{\alpha}} \\ &= \frac{N}{\bar{\epsilon}(1-\bar{\epsilon})}, \end{aligned} \quad (36)$$

where we used $\sum_{\alpha} g_{\alpha} = 1$ and $(dg_{\alpha}/dv_1)^2 = 1$. This term does not depend on n/k .

The derivatives with respect to v_2 and v_3 are found by collecting the leading terms $\propto n/k$; therefore, the terms containing $\exp(-k(t_1+t_2))$ can be neglected since they result in the next order terms. When these terms are omitted, it follows from Eq. (32) that

$$\begin{aligned} g_{\alpha\beta\gamma}(t_1, t_2) &\approx g_{\alpha\beta}(t_1)g_{\beta\gamma}(t_2)/g_{\beta}, \\ \frac{dg_{\alpha\beta\gamma}(t_1, t_2)}{dv_{2,3}} &\approx \frac{dg_{\alpha\beta}(t_1)}{dv_{2,3}}g_{\gamma} + g_{\alpha}\frac{dg_{\beta\gamma}(t_2)}{dv_{2,3}}. \end{aligned} \quad (37)$$

For example, using the above approximations in Eq. (35), we get the second derivative of $\ln L$ with respect to v_2

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_2^2} &\approx N \sum_{\alpha\beta\gamma} \int_0^\infty \frac{\left(\frac{dg_{\alpha\beta}(t_1)}{dv_2}g_{\gamma} + g_{\alpha}\frac{dg_{\beta\gamma}(t_2)}{dv_2}\right)^2}{g_{\alpha\beta}(t_1)g_{\beta\gamma}(t_2)} g_{\beta} n^2 e^{-n(t_1+t_2)} dt_1 dt_2 \\ &\quad - N \sum_{\alpha\beta} \int_0^\infty \frac{(dg_{\alpha\beta}/dv_2)^2}{g_{\alpha\beta}(t)} n e^{-nt} dt \\ &\approx N \sum_{\alpha\beta} \int_0^\infty \frac{(dg_{\alpha\beta}/dv_2)^2}{g_{\alpha\beta}(t)} n e^{-nt} dt. \end{aligned} \quad (38)$$

Here, the cross term in the first integrand is $\propto (dg_{\alpha\beta}(t_1)/dv_2)(dg_{\beta\gamma}(t_2)/dv_2) \propto \exp(-k(t_1+t_2))$, and therefore is neglected. In the remaining terms, we performed the integration with respect to one of the t 's by replacing $\int_0^\infty n \exp(-nt)/g_{\mu\nu}(t) dt$ by the leading term $1/g_{\mu\nu}$.

Performing similar calculations, we get the leading terms (with respect to n/k) of the derivatives

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_2^2} &= [\mathbf{H}_v]_{22} \approx N \int_0^\infty \sum_{\alpha\beta} \frac{(dg_{\alpha\beta}/dv_2)^2}{g_{\alpha\beta}(t)} n e^{-nt} dt \\ &= \frac{n}{k} \frac{N\sigma_{\bar{\epsilon}}^4}{k^2} \int_0^\infty z^2 \mathcal{F}_2(z) e^{-(2+n/k)z} dz, \end{aligned} \quad (39a)$$

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_3^2} &= [\mathbf{H}_v]_{33} \approx N \int_0^\infty \sum_{\alpha\beta} \frac{(dg_{\alpha\beta}/dv_3)^2}{g_{\alpha\beta}(t)} n e^{-nt} dt \\ &= N \frac{n}{k} \int_0^\infty \mathcal{F}_2(z) e^{-(2+n/k)z} dz, \end{aligned} \quad (39b)$$

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_2 \partial v_3} &= [\mathbf{H}_v]_{23} \\ &\approx N \int_0^\infty \sum_{\alpha\beta} \frac{(dg_{\alpha\beta}/dv_2)(dg_{\alpha\beta}/dv_3)}{g_{\alpha\beta}(t)} n e^{-nt} dt \\ &= -\frac{n}{k} \frac{N\sigma_{\bar{\epsilon}}^2}{k} \int_0^\infty z \mathcal{F}_2(z) e^{-(2+n/k)z} dz, \end{aligned} \quad (39c)$$

$$\begin{aligned} -\frac{\partial^2 \ln L}{\partial v_4^2} &= [\mathbf{H}_v]_{44} \\ &= N \int_0^\infty \int_0^\infty \sum_{\alpha\beta\gamma} \frac{(dg_{\alpha\beta\gamma}(t_1, t_2)/dv_4)^2}{g_{\alpha\beta\gamma}(t_1, t_2)} n^2 e^{-n(t_1+t_2)} dt_1 dt_2 \\ &= \frac{n^2}{k^2} N\sigma_{\bar{\epsilon}}^4 \int_0^\infty \int_0^\infty \mathcal{F}_3(z_1, z_2) e^{-(2+n/k)(z_1+z_2)} dz_1 dz_2, \end{aligned} \quad (39d)$$

where $g_{\alpha\beta}$ and $g_{\alpha\beta\gamma}$ are given in Eq. (32), $\mathcal{F}_2(z) \equiv \sum_{\alpha\beta} 1/g_{\alpha\beta}$ with the replacement $kt \rightarrow z$

$$\begin{aligned} \mathcal{F}_2(z) &= \frac{1}{\bar{\epsilon}^2 + \sigma_{\bar{\epsilon}}^2 e^{-z}} + \frac{1}{(1-\bar{\epsilon})^2 + \sigma_{\bar{\epsilon}}^2 e^{-z}} \\ &\quad + \frac{2}{\bar{\epsilon}(1-\bar{\epsilon}) - \sigma_{\bar{\epsilon}}^2 e^{-z}}, \end{aligned} \quad (40)$$

and $\mathcal{F}_3(z_1, z_2) \equiv \sum_{\alpha\beta\gamma} 1/g_{\alpha\beta\gamma}$ with the replacements $kt_1 \rightarrow z_1$, $kt_2 \rightarrow z_2$.

These calculations show that $[\mathbf{H}_v]_{22}$, $[\mathbf{H}_v]_{33}$, and $[\mathbf{H}_v]_{23}$ are expressed in terms of the functions for two consecutive photons, $g_{\alpha\beta}$, and are of the order of n/k which is considered here as a small parameter. $[\mathbf{H}_v]_{44}$ is expressed in terms of the function for three consecutive photons, $g_{\alpha\beta\gamma}$, and is $O(n^2/k^2)$. Similar reasonings show that all $[\mathbf{H}_v]_{i4}$ ($i = 1, 2, 3, 4$) are quadratic in n/k , $O(n^2/k^2)$. All other terms are linear in n/k , except $[\mathbf{H}_v]_{11}$, which is not small ($O(1)$). Using this in Eq. (33) and collecting the leading terms in n/k , we get

$$\begin{aligned} \langle \delta \mathcal{E}_1^2 \rangle &= [\mathbf{H}_u^{-1}]_{11} = \frac{p_2^2}{[\mathbf{H}_v]_{44}}, \\ \langle \delta \mathcal{E}_2^2 \rangle &= [\mathbf{H}_u^{-1}]_{22} = \frac{p_1^2}{[\mathbf{H}_v]_{44}}, \\ \langle \delta p_1^2 \rangle &= [\mathbf{H}_u^{-1}]_{33} = \frac{4p_1^3 p_2^3}{\sigma_{\bar{\epsilon}}^2 [\mathbf{H}_v]_{44}}, \\ \langle \delta k^2 \rangle &= [\mathbf{H}_u^{-1}]_{44} = \frac{[\mathbf{H}_v]_{33}}{[\mathbf{H}_v]_{22}[\mathbf{H}_v]_{33} - [\mathbf{H}_v]_{23}^2}. \end{aligned} \quad (41)$$

Thus, the variance of k is $\propto k/n$ and is given in terms of $[\mathbf{H}_v]_{22}$, $[\mathbf{H}_v]_{33}$, and $[\mathbf{H}_v]_{23}$. These elements are the derivatives of the likelihood function with respect to k and $\sigma_{\bar{\epsilon}}^2$, which

characterize the two-photon function $g_{\alpha\beta}(t)$. The variances of $\mathcal{E}_1$, $\mathcal{E}_2$, and p_1 are $\propto k^2/n^2$ and all are expressed in terms of a single element, $[\mathbf{H}_v]_{44}$. This element is the derivative of the likelihood function with respect to x_A , which only enters into the three-photon function $g_{\alpha\beta\gamma}(t_1, t_2)$.

A. Small variance limit

The above errors can be simply expressed when the FRET efficiency variance $\sigma_{\mathcal{E}}^2 = p_1 p_2 (\mathcal{E}_1 - \mathcal{E}_2)^2$ is small. In this case, $\mathcal{F}_2(z) = \bar{\mathcal{E}}^{-2}(1 - \bar{\mathcal{E}})^{-2}$ and $\mathcal{F}_3(z) = \bar{\mathcal{E}}^{-3}(1 - \bar{\mathcal{E}})^{-3}$. Using these in Eq. (39) (with n/k in the exponent set to 0), we get

$$\begin{aligned} [\mathbf{H}_v]_{22} &= \frac{n}{4k^3} \frac{N\sigma_{\mathcal{E}}^4}{\bar{\mathcal{E}}^2(1 - \bar{\mathcal{E}})^2}, \\ [\mathbf{H}_v]_{33} &= \frac{n}{2k} \frac{N}{\bar{\mathcal{E}}^2(1 - \bar{\mathcal{E}})^2}, \\ [\mathbf{H}_v]_{23} &= -\frac{n}{4k^2} \frac{N\sigma_{\mathcal{E}}^2}{\bar{\mathcal{E}}^2(1 - \bar{\mathcal{E}})^2}, \\ [\mathbf{H}_v]_{44} &= \frac{n^2}{4k^2} \frac{N\sigma_{\mathcal{E}}^4}{\bar{\mathcal{E}}^3(1 - \bar{\mathcal{E}})^3}. \end{aligned} \quad (42)$$

Using these in Eq. (41), we get the equations given above in Eq. (8).

B. Special case of equal populations

In the special case used in the simulations in Fig. 2, $p_1 = p_2 = 1/2$ and $\mathcal{E}_1 + \mathcal{E}_2 = 1$, so that $\bar{\mathcal{E}} = 1/2$ and $\sigma_{\mathcal{E}}^2 = (\mathcal{E}_2 - \mathcal{E}_1)^2/4 = \Delta^2/4$. In this case, $\mathcal{F}_2(z)$ in Eq. (40) becomes $\mathcal{F}_2(z) = 16/(1 - \Delta^4 \exp(-2z))$. The integrals in Eqs. (39a), (39b), and (39c) can be evaluated using

$$\int_0^\infty z^m \mathcal{F}_2(z) e^{-2z} dz = 2^{3-m} m! \Delta^{-4} Li_{m+1}(\Delta^4), \quad (43)$$

where $Li_m(z) = \sum_{k=1}^\infty z^k/k^m$ is the polylogarithm function. In this way, we get the elements $[\mathbf{H}_v]_{22}$, $[\mathbf{H}_v]_{33}$, and $[\mathbf{H}_v]_{23}$, which, after using in Eq. (41), result in the variance of the relaxation rate in Eq. (9c).

The calculation of $[\mathbf{H}_v]_{44}$ is more complex. Using Mathematica (Wolfram Research, Inc.) in evaluating the integral in Eq. (39d), we can get

$$\begin{aligned} [\mathbf{H}_v]_{44} &= \frac{n^2}{k^2} N \Delta^2 F(\Delta^{-2} - 1), \\ F(z) &= 3z \ln z - 8(1+z) \ln(1+z) \\ &+ 6(2+z) \ln(2+z) - (4+z) \ln(4+z) \\ &+ (z-1) [Li_2(-4/z) - 4Li_2(-2/z) + 4Li_2(-1/z)]. \end{aligned} \quad (44)$$

The variances of FRET efficiencies and population are $\langle \delta \mathcal{E}_1^2 \rangle = \langle \delta \mathcal{E}_2^2 \rangle = 1/(4[\mathbf{H}_v]_{44})$ and $\langle \delta p_1^2 \rangle = 1/(4\Delta^2[\mathbf{H}_v]_{44})$, according to Eq. (41). This results in Eqs. (9a) and (9b) given in Sec. II.

ACKNOWLEDGMENTS

I thank A. Szabo, H. S. Chung, and S. Pressé for helpful discussions and very useful comments. This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institutes of Health.

- ¹G. F. Schröder and H. Grubmüller, *J. Chem. Phys.* **119**, 9920 (2003).
- ²M. Andrec, R. M. Levy, and D. S. Talaga, *J. Phys. Chem. A* **107**, 7454 (2003).
- ³S. A. McKinney, C. Joo, and T. Ha, *Biophys. J.* **91**, 1941 (2006).
- ⁴L. P. Watkins, H. Chang, and H. Yang, *J. Phys. Chem. A* **110**, 5191 (2006).
- ⁵H. S. Chung, J. M. Louis, and W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11837 (2009).
- ⁶M. Pirchi, G. Ziv, I. Riven, S. S. Cohen, N. Zohar, Y. Barak, and G. Haran, *Nat. Commun.* **2**, 493 (2011).
- ⁷Y. Liu, J. Park, K. Dahmen, Y. Chemla, and T. Ha, *J. Phys. Chem. B* **114**, 5386 (2010).
- ⁸B. G. Keller, A. Kobitski, A. Jašchke, G. U. Nienhaus, and F. Noé, *J. Amer. Chem. Soc.* **136**, 4534 (2014).
- ⁹I. V. Gopich, *Chem. Phys.* **396**, 53 (2012).
- ¹⁰S. Kou, X. Xie, and J. Liu, *J. R. Stat. Soc.: Ser. C (Appl. Stat.)* **54**, 469 (2005).
- ¹¹M. Jäger, A. Kiel, D. Herten, and F. Hamprecht, *ChemPhysChem* **10**, 2486 (2009).
- ¹²I. V. Gopich and A. Szabo, *J. Phys. Chem. B* **113**, 10965 (2009).
- ¹³M. Hajdziona and A. Molski, *J. Chem. Phys.* **134**, 054112 (2011).
- ¹⁴M. Waligórska and A. Molski, *Chem. Phys.* **403**, 52 (2012).
- ¹⁵L. P. Watkins and H. Yang, *J. Phys. Chem. B* **109**, 617 (2005).
- ¹⁶K. R. Haas, H. Yang, and J. W. Chu, *J. Phys. Chem. B* **117**, 15591 (2013).
- ¹⁷H. S. Chung, I. V. Gopich, K. McHale, T. Cellmer, J. M. Louis, and W. A. Eaton, *J. Phys. Chem. A* **115**, 3642 (2011).
- ¹⁸H. S. Chung, T. Cellmer, J. M. Louis, and W. A. Eaton, *Chem. Phys.* **422**, 229 (2013).
- ¹⁹H. S. Chung, K. McHale, J. M. Louis, and W. A. Eaton, *Science* **335**, 981 (2012).
- ²⁰H. S. Chung and W. A. Eaton, *Nature* **502**, 685 (2013).
- ²¹L. P. Watkins and H. Yang, *Biophys. J.* **86**, 4015 (2004).
- ²²S. Kawai, D. Cooper, C. Landes, H. D. Mootz, H. Yang, and T. Komatsuzaki, *J. Phys. Chem. B* **117**, 8061 (2013).
- ²³H. S. Chung and I. V. Gopich, *Phys. Chem. Chem. Phys.* **16**, 18644 (2014).
- ²⁴I. V. Gopich and A. Szabo, in *Single-Molecule Biophysics: Experiment and Theory*, *Advances in Chemical Physics*, edited by T. Komatsuzaki, M. Kawakami, S. Takahashi, H. Yang, and R. J. Silbey (Wiley, New York, 2012), Vol. 146, pp. 245–297.
- ²⁵A. Deniz, T. Laurence, M. Dahan, D. Chemla, P. Schultz, and S. Weiss, *Annu. Rev. Phys. Chem.* **52**, 233 (2001).
- ²⁶B. Schuler, E. A. Lipman, and W. A. Eaton, *Nature* **419**, 743 (2002).
- ²⁷E. Sisamakias, A. Valeri, S. Kalinin, P. J. Rothwell, and C. A. M. Seidel, *Methods Enzymol.* **475**, 455 (2010).
- ²⁸A. Van den Bos, *Parameter Estimation for Scientists and Engineers* (Wiley, 2007).
- ²⁹K. R. Haas, H. Yang, and J. -W. Chu, *J. Chem. Phys.* **139**, 121931 (2013).
- ³⁰M. Dahan, A. A. Deniz, T. Ha, D. S. Chemla, P. G. Schultz, and S. Weiss, *Chem. Phys.* **247**, 85 (1999).