Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics 284 (2002) 423–437
Chemical Physics
www.elsevier.com/locate/chemphys

Statistical approaches for probing single-molecule dynamics photon-by-photon

Haw Yang[1], X. Sunney Xie*

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

Received 19 December 2001

Abstract

The recently developed photon-by-photon approach [H. Yang, X.S. Xie, J. Chem. Phys., 2002 (in press)] for single-molecule fluorescence experiments allows measurements of conformational fluctuation with time resolution on a vast range of time scales. In that method, each photon represents a data point, thereby affording better statistics. Here, we utilize the information carried by each detected photon to better differentiate theoretical models for the underlying dynamical processes – including two- and three-state models, and a diffusive model. We introduce a three-time correlation analysis, which is based on time series analyses, and the Kullback–Liebler distance, which is based on information theory principles [Elements of Information Theory, Wiley, New York, 1991]. The feasibility of and general procedures for applying these methods to single-molecule experiments are examined via computer simulations.

1. Background

In recent years, optical single-molecule spectroscopy has been proving instrumental in accessing information that is otherwise concealed in ensemble-averaged experiments [4–7]. Since molecules are studied one at a time, the distribution of molecular properties can be directly measured. In addition, by following the molecular process in real time and observing the reaction intermediates, the mechanism of complex biochemical reactions can be elucidated [8–10]. A less obvious, yet very important advantage of single-molecule spectroscopy is its capability of extracting dynamic information through statistical analyses of trajectories [11–15].

For example, Lu et al. [12] reported the observation of enzymatic turnovers of a single cholesterol oxidase (CO*x*) molecule, which catalyzes the oxidation of cholesterol by oxygen. It was found that the enzymatic reaction rate of CO*x* fluctuates with time. This observation is pertinent to the idea of ''dynamic disorder'' of a chemical reaction [16], which arises when the time scale of molecular structure change is comparable to or slower than that of the reaction. As such, the reaction rate is regulated by molecular conformational changes. Indeed, slow conformational fluctuations of single biomolecules have been observed for CO*x* [12] and

* Corresponding author.
[1] Present address: Department of Chemistry, University of California at Berkeley, CA 94720, USA.
E-mail address: xie@chemistry.harvard.eduw (X.S. Xie).

0301-0104/02/$ - see front matter
PII: S0301-0104(02)00672-9

other enzymes [15], $tRNA^{Phe}$ [17], and DNA [18–21]. In those experiments, spectroscopic observables that reflect the microscopic environment of a fluorescent probe are used to study the structural fluctuations of a single biomolecule.

Typically, an optical experimental observable of a single molecule is recorded as a function of time. Such a time trajectory often exhibits stochastic fluctuations. To gain insight about the underlying dynamics, statistical treatment and theoretical modeling of experimental data are indispensable [22–33]. One common approach is the two-time correlation method (autocorrelation), which computes the correlation in the fluctuations of observables that have been recorded at two separate times. Such a correlation function for a stationary process depends only on one variable t, the temporal separation of the two points in time. The autocorrelation function from experimental data is then compared to those obtained from theoretical models.

Two kinds of theoretical models have been used to describe dynamic disorder or conformational changes. The first kind assumes that the molecule can switch among discrete conformational states that yield different molecular observables [23,25,28], whereas the second kind assumes a continuous distribution of molecular conformation such that the observables also vary continuously [23,26,28]. Due to limited trajectory length, however, it is often difficult to ascribe a suitable theoretical model, even qualitatively, to single-molecule data based on the one-dimensional (1D) autocorrelation alone. For instance, Schenter et al. [23] analyzed the single-molecule enzymatic reaction experiment by Lu et al. [12] via the correlation-function approach and compared the experimental data to various theoretical models. It was found that both a continuous-diffusive model and a discrete two-state model fit equally well the single-molecule enzymatic turnover trajectories. While higher-order correlation functions in principle can improve the discrimination power of different models, such analyses often demand a great number of data points of observables. Evidently, experimental and theoretical strategies that afford better discriminating power over different theoretical models are needed.

A new strategy: probing single-molecule conformational changes photon-by-photon. Among the observables that have been successfully applied to single-molecule studies – including the efficiency of Förster resonance energy transfer (FRET), the lifetime of excited-state, or the emission spectrum – of particular interest is the excited-state lifetime of a fluorescent molecule. Many of the ensemble-averaged measurements for natural chromophores have shown fluorescence lifetime decays that are multi-exponential. Ensemble-averaged experiments, however, could not distinguish whether the said multi-exponential fluorescence lifetime decay is static or dynamic in nature. The dynamical aspect can be studied by single-molecule spectroscopy. Considering a fluorescent probe embedded in a biomolecule, the chromophore is subjected to fluorescence quenching by quenchers Q, which, for instance, can be a peptide residue such as tryptophan or tyrosine. The fluorescence decay rate of the chromophore can be represented by $\gamma = \gamma_0 + \gamma_q$, where γ_0 is the fluorescence decay rate in the absence of the quencher and γ_q is the quenching rate. In the case where the quenching rate γ_q is a constant, the fluorescence intensity decay is a single-exponential. Due to structural fluctuations, however, the quenching rate becomes a function of time $\gamma_q(t)$ and reflects the changes in the immediate chemical environment of the fluorescent probe. As a result, single-molecule fluorescence decay can no longer be described by a single-exponential function. The fluctuations of the single-molecule lifetime contain detailed dynamic information. In this light, single-molecule fluorescence lifetime holds the promise of serving as a sensitive measure for conformational dynamics.

Photon statistics, however, demands that a number of photons be obtained, or "binned", in order to reach a statistically reasonable estimate of the lifetime [34]. Therefore, single-molecule lifetime measurement using binned time-correlated single-photon counting (TCSPC) is oblivious to events that occur on a time scale faster than or comparable to the bin time ($\sim$ms). To improve the time resolution and statistics for single-molecule studies, we have recently introduced a new strategy for studying biomolecular conformational dynamics [1,35]. Instead of solely building a

histogram of the delay time τ of each detected photon relative to its excitation pulse (Fig. 1) as in the earlier single-molecule lifetime measurements [36–38], we record in real time both τ and the chronological time t_p of each detected photon. The sequence of such "time stamped" photons, $\{\tau_1(t_1), \ldots, \tau_n(t_n)\}$, is then subjected to statistical analysis. Such a scheme allows for measurements of fast dynamics or for measurements of dynamics over a broad range of time scales. Within the context of the time-stamped photon sequence detection scheme, the remainder of this paper discusses two approaches that offer improved discriminating power for various theoretical models.

The first approach analyzes experimental data with *higher-order correlation functions* [39]. In the case of three-time correlation functions, one calculates the correlation in the fluctuation of observables that are collected at three different times. The correlation function depends on two variables: t_1, the time separation between the first and the second event; and t_2 the time separation between the second and the third event. Since there are two variables in this three-time correlation function, one can construct a two-dimensional (2D) correlation surface based on the three-time correlation analysis [2]. The 2D correlation surface potentially provides more information over the 1D autocorrelation function. As will be discussed in Section 3, the 1D correlation function cannot determine whether a simulated trajectory is based upon a continuous model or a discrete model, whereas the three-time correlation method demonstrates a superior discrimination power over the 1D approach. In particular, we show that the three-time correlation analysis allows distinct differentiation of a two-state model from a degenerate three-state model, where the system switches among three discrete states but two of the states exhibit a nearly identical value in the experimental observable.

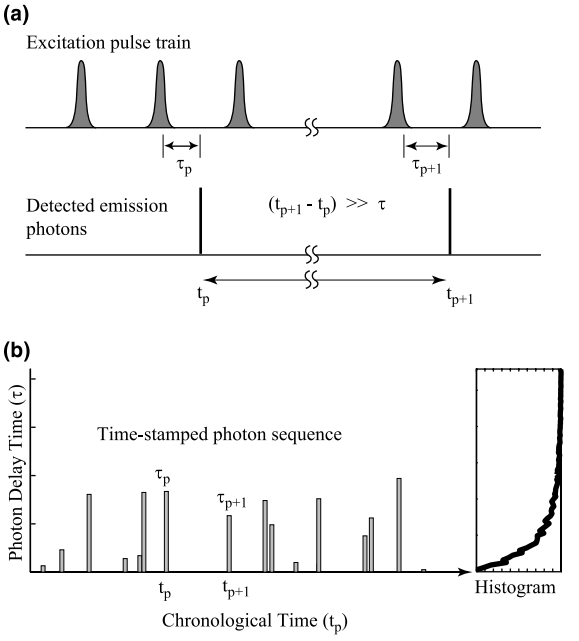


Fig. 1. (a) A schematic representation of a train of excitation pulses and detected emission photons. The delay time τ_p and chronological time t_p for each detected photon are sequentially recorded in a computer for data analysis. (b) A schematic representation of the time-stamped photon sequence. Note that the two τ and t axes are not on the same scale. To the right is the histogram of τ_p, corresponding to the conventional TCSPC result.

The second approach analyzes the *probability structure* of the observable sequence directly. Lu et al. [12] used 2D joint probability histograms to visualize the subtle differences between correlated and uncorrelated reaction steps (events) in an enzymatic cycle. To quantify the dissimilarity, Cao [27,28] defined difference distribution function as $\delta(1, m+1) = p(1, m+1) - p(1)p(m+1)$, where $p(1, m+1)$ is the joint probability of two events separated by m events and $p(m)$ is the probability of the mth event. In a new effort to quantify the 2D joint probability histogram employed by Lu et al., we introduce the use of the Kullback–Leibler entropy deficiency functional to compute the mutual information content between a correlated and uncorrelated probability map [3,40]. It is important to note that our approach, based on information theory principles, can potentially provide more information about the underlying molecular process. In particular, we explore the mutual information approach from information theory and compare it to the correlation methods.

The merits of the two statistical approaches are discussed by examples of three theoretical models: (1) discrete two-state model; (2) discrete three-state

model; and (3) continuously diffusive model [41]. Section 2 defines the functional form of various correlation functions and average mutual information, followed by analytical expressions of the three models. As an illustration of applying the theoretical analyses to time-stamped observables, Section 3 compares the analytical solutions to numerical simulations.

2. Theoretical considerations

Consider a biomolecule that undergoes slow interconversion among different conformational states with different fluorescence lifetimes. Given a kinetic scheme of interconversion, all the dynamic information is contained in the time-dependent state propagator $\mathbf{G}(f,t_f|i,t_i)$, which describes the conditional probability of the system arriving at state f at time t_f given the system being at state i at time t_i. The structure of $\mathbf{G}(f,t_f|i,t_i)$ is determined by the rate constants that characterize the kinetic scheme. One goal of statistical analyses of single-molecule trajectories is therefore to determine the interconversion rates from experimental observables – which, in the present case, are the time-stamped photon sequences $\{\tau_p(t_p)\}$ from TCSPC. This section discusses two approaches that allow classification of the underlying reaction scheme and determination of the interconversion rate constants.

Correlation methods. The first approach utilizes correlation analyses of the observables. Specifically, we calculate the two-time fluctuation-correlation function, or autocorrelation, of two events separated by period t

$$C_2(t)=\left\langle \delta\gamma(t)^{-1}\delta\gamma(0)^{-1}\right\rangle=\left\langle \delta\gamma^{-2}\right\rangle\tilde{C}_2(t), \tag{1a}$$

where $\delta\gamma(t)^{-1}=\gamma(t)^{-1}-\langle\gamma^{-1}\rangle$, and $\tilde{C}_2(t)$ is the normalized autocorrelation function, defined as

$$\tilde{C}_2(t)=\frac{\left\langle \gamma(t)^{-1}\gamma(0)^{-1}\right\rangle-\langle\gamma^{-1}\rangle^2}{\lim_{t\to 0}\left\langle \gamma(t)^{-1}\gamma(0)^{-1}\right\rangle-\langle\gamma^{-1}\rangle^2}. \tag{1b}$$

Similarly, we define a three-time fluctuation-correlation function of three events separated by periods t_1 and t_2

$$\begin{aligned}C_3(t_1,t_2)&=\left\langle \delta\gamma(t_1+t_2)^{-1}\delta\gamma(t_1)^{-1}\delta\gamma(0)^{-1}\right\rangle\\&=\left\langle \delta\gamma^{-3}\right\rangle\tilde{C}_3(t),\end{aligned} \tag{2a}$$

where the normalized three-time correlation function is

$$\tilde{C}_3(t_2,t_1)=\frac{\left\langle \delta\gamma(t_1+t_2)^{-1}\delta\gamma(t_1)^{-1}\delta\gamma(0)^{-1}\right\rangle}{\lim_{t_1\to 0}\lim_{t_2\to 0}\left\langle \delta\gamma(t_1+t_2)^{-1}\delta\gamma(t_1)^{-1}\delta\gamma(0)^{-1}\right\rangle}, \tag{2b}$$

where

$$\begin{aligned}&\left\langle \delta\gamma(t_1+t_2)^{-1}\delta\gamma(t_1)^{-1}\delta\gamma(0)^{-1}\right\rangle\\&\quad=\left\langle \gamma(t_1+t_2)^{-1}\gamma(t_1)^{-1}\gamma(0)^{-1}\right\rangle\\&\qquad-\left(\left\langle \gamma(t_2)^{-1}\gamma(0)^{-1}\right\rangle+\left\langle \gamma(t_1)^{-1}\gamma(0)^{-1}\right\rangle\right.\\&\qquad\left.+\left\langle \gamma(t_1+t_2)^{-1}\gamma(0)^{-1}\right\rangle\right)\langle\gamma^{-1}\rangle+2\langle\gamma^{-1}\rangle^3.\end{aligned} \tag{2c}$$

The temporal correlation comes from either the $\langle\gamma(t)^{-1}\gamma(0)^{-1}\rangle$ or the $\langle\gamma(t_2+t_1)^{-1}\gamma(t_1)^{-1}\gamma(0)^{-1}\rangle$ terms. They can be expressed in terms of the propagator $G(f,t_f|i,t_i)$ as

$$\left\langle \gamma(t)^{-1}\gamma(0)^{-1}\right\rangle=\sum_{f,i}\gamma_f^{-1}G_{fi}(t)\gamma_i^{-1}\rho_i^{\mathrm{eq}} \tag{3}$$

and

$$\begin{aligned}&\left\langle \gamma(t_2+t_1)^{-1}\gamma(t_1)^{-1}\gamma(0)^{-1}\right\rangle\\&\quad=\sum_{f,k,i}\gamma_f^{-1}G_{fk}(t_2)\gamma_k^{-1}G_{ki}(t_1)\gamma_i^{-1}\rho_i^{\mathrm{eq}},\end{aligned} \tag{4}$$

where $G_{fi}(t)=G(f,t_f|i,t_i)$ at $t=t_f-t_i$, and γ_i^{-1} and ρ_i^{eq} are the fluorescence lifetime and equilibrium probability for finding the system at state i, respectively. In practice, the correlation function is calculated from experimental data utilizing the general expressions for higher lifetime moments [35]:

$$\langle\gamma(t)^{-z}\rangle=\frac{1}{TA_0}\sum_{\{p\}}\frac{\tau_p^{z-1}}{\Gamma(z)}, \tag{5}$$

$$\begin{aligned}\langle\gamma(t)^{-z_1}\gamma(0)^{-z_2}\rangle&=\frac{1}{(T-t)A_0^2\Delta}\sum_{\{q>p\},\{p\}}\frac{\tau_q^{z_1-1}}{\Gamma(z_1)}\\&\quad\times\hat{u}(t_q-t_p-t)\frac{\tau_q^{z_2-1}}{\Gamma(z_2)},\end{aligned} \tag{6}$$

$$\langle \gamma(t_1+t_2)^{-z_1}\gamma(t_1)^{-z_2}\gamma(0)^{-z_3}\rangle = \frac{1}{(T-t)A_0^3\Delta^2}\sum_{\{s>q>p\},\{q>p\},\{p\}}\frac{\tau_q^{z_1-1}}{\Gamma(z_1)}\hat{u}(t_s-t_q-t_2)\times\frac{\tau_q^{z_2-1}}{\Gamma(z_2)}\hat{u}(t_q-t_p-t_1)\frac{\tau_q^{z_3-1}}{\Gamma(z_3)}, \quad (7)$$

where T is the total length of the single-molecule trajectory; $A_0=\bar{I}/\overline{\gamma^{-1}}$ is the proportionality constant that compensates for lifetime-dependent fluorescence intensity I, or sampling rate; $\Gamma(z)=\int_0^\infty y^{z-1}e^{-y}\,dy$; Δ is the chronological time increment along a single-molecule trajectory such that there is no more than one photon within each Δ interval; and $\hat{u}$ is defined as: $\hat{u}(t_p-t_s)=1$ for $(-\frac{1}{2}\Delta)\leqslant t_p-t_s\leqslant(\frac{1}{2}\Delta)$ and $\hat{u}(t_p-t_s)=0$ otherwise. Eqs. (5)–(7) allow direct calculation of the dynamical quantities on the left-hand of the equation via the experimentally realizable expressions on the right-hand side.

Information theory methods. Qualitatively, the correlation-function approach describes how fast the molecule loses its conformational memory. Higher-order correlations would be required in order to determine the dynamic variables for kinetic schemes involving more than two conformational states. On the other hand, statistical analyses based on information theory allow direct examination of the probability structure of an observable sequence. The information theory approach computes the difference in the information content between correlated and uncorrelated events along a time series [3]. Specifically, if $p[\tau_p(t)]$ and $p[\tau_q(0)]$ are probabilities for stochastic delay time stamps being τ_q and τ_p at time t and 0, respectively, and $p[\tau_p(t),\tau_q(0)]$ is the joint probability. The mutual information is defined as

$$\mathbf{Mi}[\tau_q(t);\tau_p(0)] = p[\tau_q(t),\tau_p(0)]\times\log\frac{p[\tau_q(t),\tau_p(0)]}{p[\tau_q(t)]p[\tau_p(0)]}, \quad (8a)$$

where $p[\tau_p(0)]$ (or $p[\tau_q(t)]$) can be calculated by summing over all conformational states i (or f):

$$p[\tau_p(0)] = \sum_i P_i[\tau_p(0)]P(1|I_i,\Delta)\rho_i^{\mathrm{eq}} \quad (8b)$$

and

$$p[\tau_q(t),\tau_p(0)] = \sum_{f,i}P_f[\tau_q(t)]P(1|I_f,\Delta)G_{fi}(t)\times P_i[\tau_p(0)]P(1|I_i,\Delta)\rho_i^{\mathrm{eq}}, \quad (8c)$$

where $P_i(\tau)=\gamma_i\exp(-\gamma_i\tau)\,d\tau$ (or P_f) is the probability of obtaining a τ event given the system is at state i (or f), and $P(n|I_i,\Delta)=(I_i\Delta)^n\exp[-I_i\Delta]/n!$ (or $P(n|I_f,\Delta)$) is the probability of obtaining n photons within a time interval of Δ from the i (or f) state the emission intensity of which is I_i (or I_f). Eqs. (8a) and (8c) is essentially the reduction in the uncertainty of event $\tau_q(t)$ due to the knowledge of event $\tau_p(0)$. In the case of the continuously diffusive model (Section 2.2), the summations in Eq. (8c) are replaced by integrations.

If a summation is carried out over the p and q indices in Eq. (8a), we have the relative entropy between the joint distribution $p[\tau(t),\tau(0)]$ and the product distribution $p[\tau(t)]p[\tau(0)]$, or Kullback–Leibler distance, $\mathscr{D}$ for two delay times $\tau(t)$ and $\tau(0)$ separated by chronological time

$$\mathscr{D}(p[\tau(t),\tau(0)]|p[\tau(t)]p[\tau(0)]) = \sum_{\tau_q}\sum_{\tau_p}p[\tau_q(t),\tau_p(0)]\log\frac{p[\tau_q(t),\tau_p(0)]}{p[\tau_q(t)]p[\tau_p(0)]}. \quad (9)$$

The relative entropy is a measure of the ineffectiveness of assuming the distribution being $p[\tau(t)]p[\tau(0)]$ while the true distribution is $p[\tau(t),\tau(0)]$. Not withstanding $\mathscr{D}$ is not a true measure of the distance between two distributions, it is often useful to think of $\mathscr{D}$ as a "distance" between distributions [3]. For example, $\mathscr{D}=0$ means that there is no loss of effectiveness (or information) when assuming the two conditions $\tau(t)$ and $\tau(0)$ being independent. This is seen in the cases where there are no correlations in fluorescence lifetime fluctuations.

2.1. Discrete state models

For these models, the molecule in question is considered to interconvert between discrete conformational states X_i in its electronic ground state. The molecule can be brought to the excited-state manifold via pulsed laser excitation. The observable in the experiment is the fluorescence lifetime,

obtainable from a TCSPC setup in the form of photon delay time $\{\tau_p(t_p)\}$, the delay time between the excitation pulse and the arrival of the detected photon. The time scale of structural fluctuation is assumed to be much longer than the fluorescence lifetime that is associated with each discrete conformational state. The following derivation assumes ergodic conditions, i.e., the trajectory length of a single-molecule observation is much longer than the time scales of relevant dynamic processes.

Formal treatment. Within the context of discrete-state models, the molecular dynamics are considered to follow the master equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_i(t) = \sum_{f,i} k_{if}p_f(t) - k_{fi}p_i(t), \tag{10}$$

where p_i and p_f denote the probability densities of the molecule being in the electronic-ground conformational state i and f, respectively, and k_{fi} is the macroscopic rate constant for interconversion from state i to state f. $\mathbf{P}(t)$ can be solved readily by recasting Eq. (10) into a matrix form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}(t) = \mathbf{K}\cdot\mathbf{P}(t), \tag{11}$$

where $K_{ij} = k_{ij}$ and $K_{ii} = -\sum_{i\neq j} k_{ji}$.

In order to quantify the time-dependent fluctuation in fluorescence lifetime, we need to know the conditional probability $p(f,t_1|i,t_0)$ given that the probability density of the molecule being in state i at time t_0 is $p_i(t_0)$. To calculate the probability density $\mathbf{P}(t_1)$ at time t_1 when the probability density is $\mathbf{P}(t_0)$ at time t_0, we define the conditional probability density matrix

$$\mathbf{G}(t_1,t_0) = \mathbf{G}(t_1-t_0,0) \equiv \mathbf{G}(t), \tag{12}$$

which satisfies the condition

$$\mathbf{G}(t_2+t_1) = \mathbf{G}(t_2)\cdot\mathbf{G}(t_1). \tag{13}$$

Then $\mathbf{P}(t)$ can be expressed as

$$\mathbf{P}(t) = \mathbf{G}(t)\cdot\mathbf{P}(0). \tag{14}$$

$\mathbf{G}(t)$ is therefore governed by the same master equations that govern $\mathbf{P}(t)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{G}(t) = \mathbf{K}\cdot\mathbf{G}(t) \tag{15}$$

Eq. (15) can be solved formally by Laplace transform methods

$$\tilde{\mathbf{G}}(z) = (z\mathbf{I}-\mathbf{K})^{-1}, \tag{16}$$

and

$$\mathbf{G}(t) = \mathscr{L}^{-1}\left[\tilde{\mathbf{G}}(z)\right]. \tag{17}$$

The above-outlined formal treatment is general to multiple discrete state models. Two- and three-state models will be discussed below.

Two conformational states in the electronic ground state. In this case, the single molecule fluctuates between X_1 and X_2 states:

$$X_1 \underset{k_{12}}{\overset{k_{21}}{\rightleftharpoons}} X_2 \tag{18}$$

in which k_{12} and k_{21} are the interconversion rates between the two conformers, and γ_1^{-1} and γ_2^{-1} are the effective lifetimes of the excited states, respectively. Under ergodic conditions, the time-dependent probability density $\mathbf{P}(t)$ of the single-molecule is

$$\mathbf{P}(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix}, \tag{19}$$

where p_1 and p_2 denote the probability density of the molecule being in states X_1 and X_2, respectively. The master equations that describe the dynamics of the kinetic scheme are:

$$\frac{\mathrm{d}}{\mathrm{d}t}p_1(t) = -k_{12}p_1(t) + k_{21}p_2(t), \tag{20}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}p_2(t) = k_{12}p_1(t) - k_{21}p_2(t). \tag{21}$$

The master equations can be expressed in a matrix form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}(t) = \mathbf{K}\cdot\mathbf{P}(t), \tag{22}$$

where

$$\mathbf{K} = \begin{pmatrix} -k_{21} & k_{12} \\ k_{21} & -k_{12} \end{pmatrix}. \tag{23}$$

The time-domain $\mathbf{G}(t)$ matrix (cf. Eq. (15)) is recovered by inverse Laplace transformation of $\tilde{\mathbf{G}}(z)$ and reads

$$\mathbf{G}(t) = \frac{1}{k_{12}+k_{21}}\left[\mathrm{e}^{-(k_{12}+k_{21})t}\mathbf{G}^{\mathrm{neq}} + \mathbf{G}^{\mathrm{eq}}\right], \tag{24}$$

where the conditional probability density matrix elements are

$$\mathbf{G}^{\mathrm{neq}} = \begin{pmatrix} k_{21} & -k_{12} \\ -k_{21} & k_{12} \end{pmatrix}, \qquad \mathbf{G}^{\mathrm{eq}} = \begin{pmatrix} k_{12} & k_{12} \\ k_{21} & k_{21} \end{pmatrix}. \tag{25}$$

They conform to condition Eq. (13) and conservation of probability density, $\sum_i \mathbf{G}^{\mathrm{eq}}_{ii} = 1$.

The correlation functions Eqs. (3) and (4) can then be readily calculated to be

$$C_2(t) = \frac{(\gamma_1^{-1} - \gamma_2^{-1})^2 k_{12} k_{21}}{(k_{12} + k_{21})^2} \mathrm{e}^{-(k_{12}+k_{21})t} \tag{26}$$

and

$$C_3(t_1, t_2) = \frac{(\gamma_1^{-1} - \gamma_2^{-1})^3 (k_{12} - k_{21}) k_{12} k_{21}}{(k_{12} + k_{21})^3} \mathrm{e}^{-(k_{12}+k_{21})(t_1+t_2)}. \tag{27}$$

Eqs. (26) and (27) suggest that both the autocorrelation function and the three-time correlation function regress to zero eventually with a time constant of $(k_{12} + k_{21})$.

We next consider the signal strength of $C_2(t)$ for the two-state model. It is clear from Eqs. (1a) and (1b) that the signal strength is determined by the magnitude of the lifetime variance $\langle \delta\gamma^{-2} \rangle$. In Fig. 2, the lifetime variance $\langle \delta\gamma^{-2} \rangle$ is plotted as a function of the equilibrium constant K_{eq} $(= k_{12}/k_{21})$ and the ratio of two lifetimes $\gamma_1^{-1}/\gamma_2^{-1}$. The plot suggests that, given a two-state system with a lifetime ratio $\gamma_1^{-1}/\gamma_2^{-1}$, the greatest signal amplitude is given by a $K_{\mathrm{eq}} = 1$ condition.

Alternatively, one could directly analyze the structure of the propagator $\mathbf{G}_{fi}(t)$ using Eqs. (8a) and (8c). A mutual information map of events separated by $T_c = (\text{mean count rate})^{-1}$ is shown in Fig. 3. For a stationary process, the map is symmetric with respect to the diagonal axis. When the events are correlated – that is, there is dynamic disorder – the map shows that a short delay time is *less* likely to be followed by another short delay time, as indicated by the dark area around the origin. On the other hand, a long delay time is *more* likely to be followed by another long delay time as indicated by the lighter area along the diagonal axis. Indeed, the diagonal line $\mathbf{Mi}[\tau(T_c); \tau(0)]$ carries the information of the occurrence of same-length delay times separated by T_c. Fig. 4(A) displays plots of $\mathbf{Mi}[\tau(T_c); \tau(0)]$ against delay time length τ at various K_{eq} values. For a given $\phi = \gamma_1^{-1}/\gamma_2^{-1}$, $\mathbf{Mi}[\tau(T_c); \tau(0)]$ rises sharply from negative values at short delay times, then crosses 0 at an equivocal delay time τ_e, reaches a local maximum $\tau_{\max}$, and finally declines asymptotically to 0. At the equivocal delay time τ_e, the joint

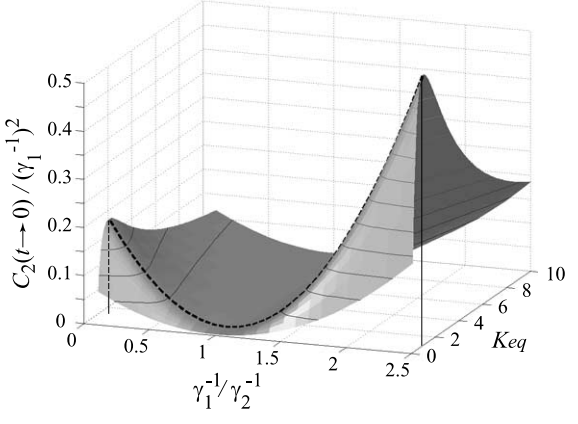


Fig. 2. Dependence of the amplitude of the autocorrelation function, $C_2(t \to 0)$, in the $\phi = \gamma_1^{-1}/\gamma_2^{-1}$ and $K_{\mathrm{eq}} = k_{21}/k_{12}$ parameter space for the two-state model. The dark dashed line denotes the $K_{\mathrm{eq}} = 1$ cross-section, where the signal strength is maximum at a given ϕ.

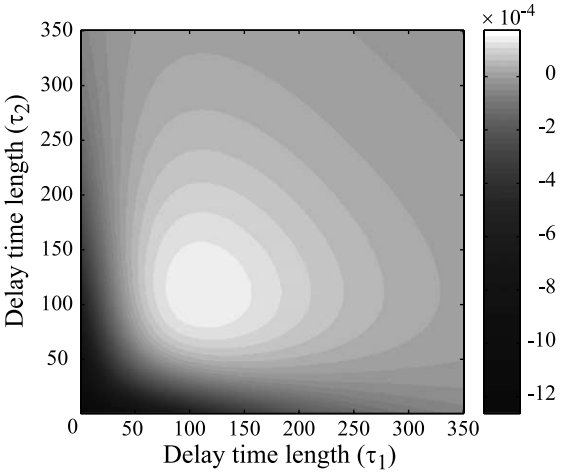


Fig. 3. Mutual information map $\mathbf{Mi}[\tau_2(T_c); \tau_1(0)]$ of delay times that are separated by $T_c = (\text{mean count rate}) - 1$ for a two-state model. The parameters used in this plot are: $k_{12} = 2\ \mathrm{s}^{-1}$, $k_{21} = 5\ \mathrm{s}^{-1}$, $T_c = 1.4$ ms, $\gamma_1^{-1} = 100$ ps, and $\gamma_1^{-1} = 40$ ps.

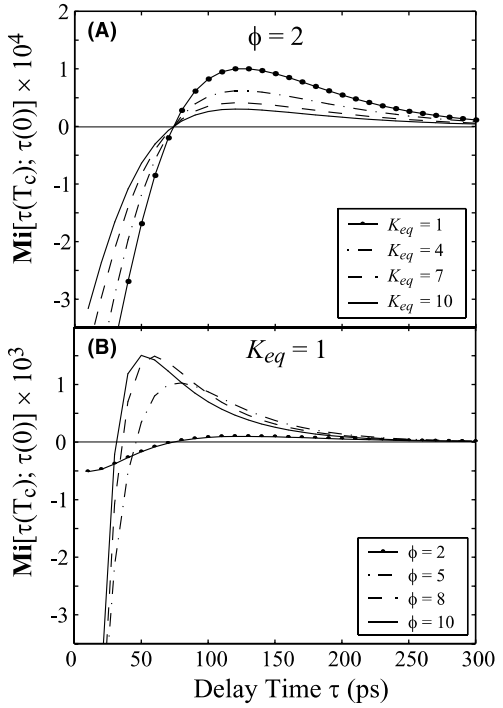


Fig. 4. Diagonal plots $\mathbf{Mi}[\tau(T_c);\tau(0)]$ for mutual information of same-length delay time stamps separated by T_c for the two-state model with $T_c = 5$ ms, $k_{12} = 2$, $\gamma_1^{-1} = 100$ ps, and $\gamma_1^{-1} = 40$ ps: (A) at various $K_{eq} = k_{21}/k_{12}$ values with $\phi = 2$; (B) at various $\phi = \gamma_1^{-1}/\gamma_2^{-1}$ values with $K_{eq} = 1$.

probability of two photons separated by T_c with arrival time length τ_e is identical to the probability of independently obtaining two photons of arrival time τ_e. The magnitude of $\mathbf{Mi}[\tau_{max}(T_c);\tau_{max}(0)]$ is dependent of K_{eq}: a greater $\mathbf{Mi}[\tau_{max}(T_c);\tau_{max}(0)]$ is favored by a two-state system with equal population in either state. This trend is also observed in Fig. 4(B), which shows different ϕ ratios for $K_{eq} = 1$.

Three conformational states in the electronic ground state. In this case, there are three conformational states X_1, X_2, and X_3 with respective fluorescence lifetimes γ_1^{-1}, γ_2^{-1}, and γ_3^{-1} (cf. Fig. 5). The rate constant matrix for this system is

$$\mathbf{K} = \begin{pmatrix} -(k_{21}+k_{31}) & k_{12} & k_{13} \\ k_{21} & -(k_{12}+k_{32}) & k_{23} \\ k_{31} & k_{32} & -(k_{13}+k_{23}) \end{pmatrix}. \tag{28}$$

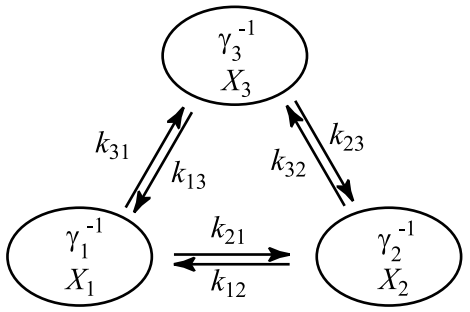
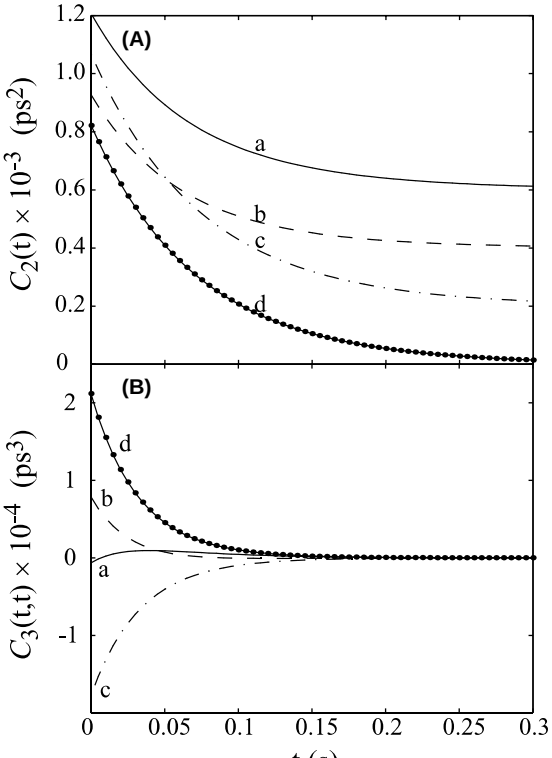


Fig. 5. A discrete three-state model. The molecule is assumed to thermally fluctuate among three conformational states in its electronic ground state. Each of the conformational states X_i has a fluorescent lifetime of γ_i^{-1}. The rate constants k_{fi} are associated with conformational changes from state i to state f.

In equilibrium, the six rate constants, k_{12}, k_{21}, k_{23}, k_{32}, k_{13}, k_{31} are related by

$$\frac{k_{12}}{k_{21}} \times \frac{k_{13}}{k_{31}} \times \frac{k_{32}}{k_{23}} = 1. \tag{29}$$

The solution to the master equations Eq. (15), $\mathbf{G}_{fi}(t)$, can be obtained as outlined earlier. With $\mathbf{G}_{fi}(t)$, the correlation functions $C_2(t)$ and $C_3(t_2,t_1)$ can be calculated accordingly. Due to the complexity of the solution (not shown here), however, it is not apparent from the expressions of $C_2(t)$ and $C_3(t_2,t_1)$ how the interconversion rate constants are related to the correlation functions. Therefore, the following discussion will proceed with numerical examples illustrating the characteristics of the three-state model. Consider, among the three conformational states, the X_1 state being either the most stable or the least stable one. In addition, the effective lifetime of X_1 can be either the longest or the shortest. With these considerations, four sets of control variables are chosen and the corresponding autocorrelation functions $C_2(t)$ are shown in Fig. 6(A). All of the four $C_2(t)$ traces decay to zero monotonically in a none-single-exponential manner. Without the added offset, it would be difficult to tell apart the four $C_2(t)$ traces in Fig. 6(A). Further, they bear close resemblance to the $C_2(t)$ functions in the two-state case. This observation indicates that it would be very difficult, if not impossible, to distinguish a two-state model from a two-state one based on the



	γ_1^{-1}	γ_2^{-1}	γ_3^{-1}	k_{12}	k_{21}	k_{13}	k_{31}	k_{23}
a	30	60	100	2	7	3	8	4.5
b	100	60	30	2	7	3	8	4.5
c	30	60	100	7	2	8	3	4.5
d	100	60	30	7	2	8	3	4.5

Fig. 6. Plots of (A) autocorrelation $C_2(t)$ and (B) three-time correlation $C_3(t,t)$ for the discrete three-state model. The units for the interconversion rate constants k_{fi} and fluorescence lifetime τ_i are s^{-1} and ps, respectively. For clarity, the (a), (b), and (c) traces in (A) are offset by 0.15, 0.1, and 0.05 units, respectively.

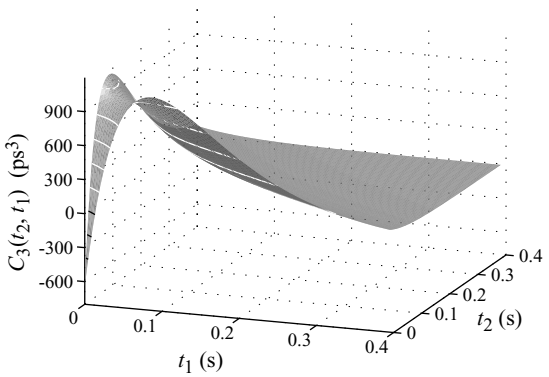


Fig. 7. Two-dimensional three-time correlation function $C_3(t_1, t_2)$ for a discrete three-state model. The dynamical control parameters are: $k_{12} = 2\ s^{-1}$, $k_{21} = 7\ s^{-1}$, $k_{13} = 3\ s^{-1}$, $k_{31} = 8\ s^{-1}$, $k_{23} = 4.5\ s^{-1}$, $\gamma_1^{-1} = 30$ ps, $\gamma_2^{-1} = 60$ ps, and $\gamma_3^{-1} = 100$ ps.

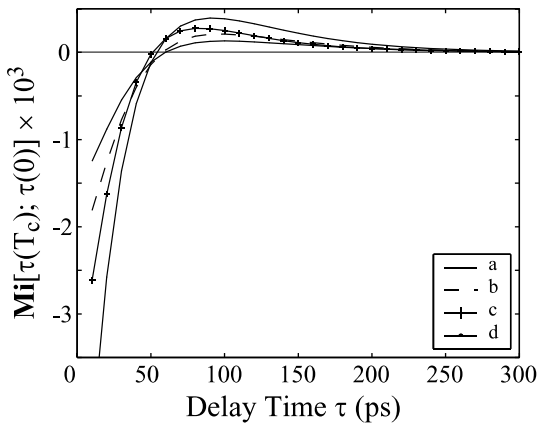


Fig. 8. Diagonal plots of $\mathbf{Mi}[\tau(T_c); \tau(0)]$ for the discrete three-state model. The dynamical control parameters of plots (a–d) are the same as those of the corresponding plots in Fig. 7.

autocorrelation function alone. The discrimination of the correlation-function approach is greatly improved as higher-order terms are considered. Fig. 6(B) shows plots for the diagonal terms (for $t_1 = t_2$) of the three-time correlation function $C_3(t_1, t_2)$. The $C_3(t,t)$ plots clearly show that the four (a)–(d) curves belong to four different sets of control variables. In fact, a 3D plot of $C_3(t_1, t_2)$ reveals subtle structures (Fig. 7) that would allow quantitative determination of the control variables of a three-state model. The above four cases can also be distinguished from their mutual information plots (Fig. 8).

2.2. Continuous distribution: a diffusive model

Oftentimes the fluorescence lifetime of a chromophore in biomolecules is dominated by quenching through excited-state electron transfer. The electron transfer rate, in turn, depends exponentially on the fluorophore-quencher distance r. Therefore, we have for the effective radiative decay rate $\gamma \to \gamma_0 e^{-\beta r(t)}$, where γ_0 is the rate constant

when the fluorophore-quencher distance is 0, and β is a constant characterizing the efficiency of electron transfer and its value lies between 1.1 and 1.4 Å^{-1} for proteins [42]. In the discussion that follows, we assume that the fluorophore-quencher distance r is a stochastic variable, which can be described by a Langevin equation. As the molecular conformation fluctuates, the effective radiative decay rate γ fluctuates continuously through the dependence $r(t)$. For the sake of discussion, we define a dimensionless variable $x(t) = \beta[r(t) - r_0]$, where r_0 is the equilibrium fluorophore-quencher distance. We then rewrite the effective radiative decay rate as

$$\gamma(t) = \gamma_0 e^{-x(t)}. \tag{30}$$

The dynamics of $x(t)$, which is also a stochastic variable, then follows the Langevin equation of Brownian motion [43]

$$\frac{d}{dt}x(t) = -\lambda x(t) + f_x(t), \tag{31}$$

where λ is the drift coefficient that characterizes the correlation time of the stochastic variable $x(t)$. In Eq. (31), $f_x(t)$ is a Gaussian random variable with mean $\langle f_x(t)\rangle = 0$ and variance $\langle f_x(t)f_x(t_0)\rangle = 2\lambda\theta\delta(t - t_0)$ where θ characterizes the magnitude of the fluctuation. The conditional probability density $G(x_1, t_1|x_0, t_0)$ that gives the probability of finding $x(t_1) = x_1$ given $x(t_0) = x_0$, can be shown to follow the well-studied Fokker–Planck equation for an Ornstein–Uhlenbeck process

$$\frac{\partial}{\partial t}G(x,t|x_0,t_0) = \frac{\partial}{\partial x}[\lambda x G(x,t|x_0,t_0)] + \lambda\theta\frac{\partial^2}{\partial x^2}G(x,t|x_0,t_0). \tag{32}$$

The solution to Eq. (32) is [44],

$$G(x,t|x_0,t_0) = \left[2\pi\theta\left(1 - e^{-2\lambda(t-t_0)}\right)\right]^{-1/2} \times \exp\left[-\frac{\left(x - x_0 e^{-\lambda(t-t_0)}\right)^2}{2\theta(1 - e^{-2\lambda(t-t_0)})}\right] \tag{33}$$

and the equilibrium distribution of x is

$$\rho_x^{eq} = \frac{1}{\sqrt{2\pi\theta}}e^{-x^2/2\theta}. \tag{34}$$

With the equilibrium distribution of Eq. (34) and the conditional probability of Eq. (33), the autocorrelation function of Eqs. (1a) and (1b) and the three-time correlation function of Eqs. (2a), (2b) and (2c) for the continuous-diffusive model are, respectively

$$C_2(t) = \gamma_0^{-2}\exp(\theta)\left[-1 + \exp(\theta e^{-\lambda t})\right] \tag{35}$$

and

$$C_3(t) = \gamma_0^{-3}\exp(3\theta/2)\left[2 + \exp\left(\theta(1 + e^{\lambda t_1} + e^{\lambda t_2}) \times e^{-\lambda(t_1+t_2)}\right) - \exp\left(\theta e^{-\lambda t_1}\right) - \exp\left(\theta e^{-\lambda t_2}\right) - \exp\left(\theta e^{-\lambda(t_1+t_2)}\right)\right] \tag{36}$$

By analogy to the two-state case, it is also possible to completely determine θ and λ in the continuous-diffusive model from $C_2(t)$ alone. However, we stress that, given a certain single-molecule trajectory, higher-order correlation functions are necessary for discriminating a diffusive model from a two-state model, as will be discussed in the next section.

The mutual information content of the diffusive model can also be calculated using Eq. (33). Fig. 9 shows plots of $\mathbf{Mi}[\tau(T_c);\tau(0)]$ at various θ values. As shown in the figure, the $\mathbf{Mi}[\tau(T_c);\tau(0)]$ function persists longer at greater θ values. This can be understood by the fact that θ is a measure of the heterogeneity of the system such that a greater θ translates to increased probability of detecting longer delay time stamps τ.

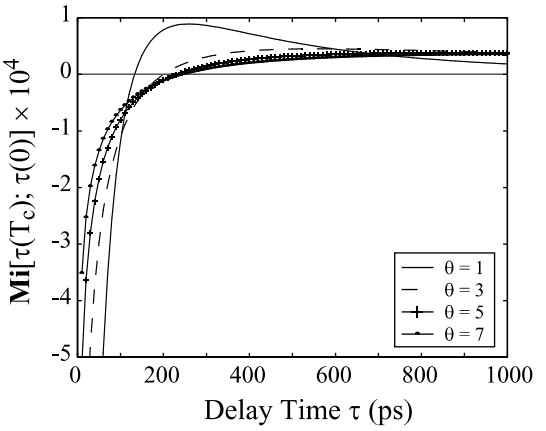


Fig. 9. Diagonal plots of $\mathbf{Mi}[\tau(T_c);\tau(0)]$ for the diffusive model with the dynamic parameter θ at 1, 3, 5, and 7 when the drifting coefficient λ is held at 2. These results are obtained by numerical integration of Eqs. (8a) and (8c).

3. A case discussion

It is important to assess the feasibility of statistical approaches for treating single-molecule data. In the following discussion, we show how the two new approaches are utilized by applying them to computer-simulated data. General procedures are as follows: given a single-molecule trajectory, we first construct a delay time histogram and calculate its autocorrelation function using Eqs. (1a), (1b), (5), and (6). Then we fit various models (two-state, three-state, and diffusive) to those two plots simultaneously. The resultant parameters are used to calculate three-time correlation surfaces and mutual information maps using analytical expressions of Eqs. (2a), (2b), (2c), (8a) and (8c). The analytical curves are compared with those computed from data to assess the viability of assigning a certain model to the single-molecule trajectory. We discuss the relative merits of different methods by example of a degenerate system, in which two conformational states exhibit a commensurable fluorescence lifetime in the three-state model.

Computer simulations for discrete state models. The computer simulations for single-molecule trajectories have been carried out in a similar way to a recent paper by Schenter et al. [23]. Briefly, the probability for the molecule to "jump" out of state i during time interval $\mathrm{d}t$ was $\Gamma_i \mathrm{d}t$, where $\Gamma_i = \sum_{f\neq i} k_{fi}$. The various states are molecular conformational states in the electronic ground state, or the excited states. The interconversion time scale is assumed to be much greater than that of fluorescence lifetime. In the two-state model simulation, the ground-state conformational kinetics follows Eq. (18). The excitation rate is 4×10^4 s^{-1}. The radiative decay rate γ_0 is set to 2.5×10^8 s^{-1}. The quenching rates for the excited states X_1 and X_2 are 9.75×10^9 and 2.475×10^{10} s^{-1}, respectively. The interconversion rates k_{21} and k_{12} are 4.9 and 4.3 s^{-1}, respectively. By virtue of separation of time scale for conformational interconversion and fluorescence lifetime, we accelerate the simulation using a multiple-time step method; the time step $\mathrm{d}t$ for propagating conformational changes was much greater than that for simulating excited-state lifetime. A random number $y \in [0,\ldots,1]$ was drawn from a uniform distribution. The molecule was considered to escape successfully if the random number y was less than $\Gamma_i \mathrm{d}t$. Once the molecule had escaped, it could arrive at one of the states $f \neq i$. Another random number $z \in [0,\ldots,1]$ was drawn to determine which state the molecule would reach by comparing z with the probability of landing on state f was k_{fi}/Γ_i. If a transition occurred from a ground conformational state to its excited-state, the chronological time t_p is recorded. At this time, the simulation time step $\mathrm{d}t$ was switched from 1 μs per step for conformational propagation to 1 ps per step for excited-state lifetime simulation. If a transition occurred from an excited-state to a ground state, the time that the molecule spent in the excited-state was recorded as a delay time stamp τ_p, and the (t_p, τ_p) pair was written to a storage device for analysis. This process was repeated until the desired number of photons was collected.

Comparison of the effectiveness of the correlation method and mutual information approach. Given an experimental trajectory of observables, in the current case the photon delay time stamp, one typically constructs a histogram to obtain the static distribution of events. The scrambled histogram, however, does not offer information about the *dynamics*, such as the interconversion rates among states. The correlation-function approach provides a means to obtain such information. Shown in Fig. 10(A) is a histogram of 10^6 simulated delay time stamps from a two-state model. The corresponding $C_2(t)$ plot is shown in Fig. 10(B). In both figures and hereafter, plots that are generated from simulated data are represented by open circles, and those that are generated from analytical expressions are represented by smooth curves. Both Figs. 11(A) and (B) are overlaid with fits to the previously discussed discrete two-state, three-state, and the continuously diffusive models. The fitted curves are obtained by simultaneously fitting the three models to lifetime histogram in Fig. 10(A) and the autocorrelation curve in Fig. 10(B). Evidently, based on the autocorrelation function alone, it is difficult to ascribe an appropriate model to the data. In Fig. 10(C), we extend the correlation analysis to the three-time correlation method. Again, the three-time correlation

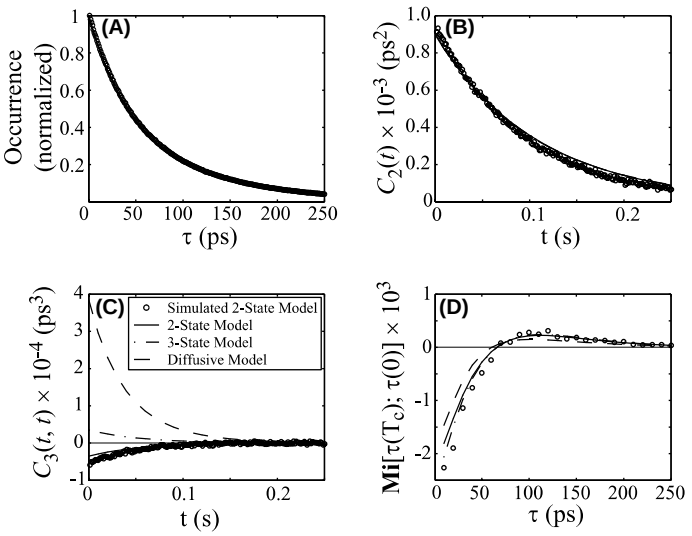


Fig. 10. Application of the statistical data analyses to a computer simulated two-state single-molecule trajectory (106 photons). (A) Photon delay time histogram overlaid by fits to the two-state, and three-state models as well as the diffusive model. (B) Autocorrelation function and fits to the theoretical models. (C) Diagonal plots of the three-time correlation curve from simulated data and analytical curves for the theoretical models using the fitted parameters from simultaneous fits to (A) and (B). (D) Diagonal plots of the mutual information function. The resulting parameters are: $k_{21} = 4.9$ s^{-1}, $k_{12} = 4.3$ s^{-1}, $\gamma_1^{-1} = 100$ ps, and $\gamma_2^{-1} = 40$ ps for the two-state model; $k_{12} = 7.56$ s^{-1}, $k_{21} = 2.97$ s^{-1}, $k_{13} = 3.09$ s^{-1}, $k_{31} = 4.90$ s^{-1}, $k_{23} = 1.21$ s^{-1}, $\gamma_1^{-1} = 100.01$ ps, $\gamma_2^{-1} = 99.97$ ps, and $\gamma_3^{-1} = 40$ ps for the three-state model; and $\gamma_0 = 1.62 \times 10^{10}$ s^{-1} (61.7 ps), $\theta = 0.1798$, and $\lambda = 8.818$ s^{-1} for the continuously diffusive model. The three-state model in this example served as a degenerate system in which two of the three fitted lifetimes are nearly identical. The smooth, analytical curves are also calculated using fitted parameters from simultaneous fits to (A) and (B).

function calculated from simulated data is shown as open circles. The parameters obtained from fitting of Figs. 11(A) and (B) are used to generate the smooth curves for the three models from their respective analytical expression. It is evident that both the three-state and the diffusive models display significant deviation from the correct two-state model. On the other hand, the diagonal $\mathbf{Mi}[\tau(T_c); \tau(0)]$ plot on Fig. 10(D) does not show improved discrimination power over the autocorrelation method. This is because the state observable τ-based $\mathbf{Mi}[\tau(T_c); \tau(0)]$ measures the mutual information of two probability maps separated by only one chronological time T_c. When two or more states exhibit a similar state observable (degeneracy), $\mathbf{Mi}[\tau(T_c); \tau(0)]$ could not distinguish it from a model that involves fewer states. Its differentiation power, nevertheless, can be improved by calculating mutual information maps that utilize multiple chronological time separations as in the case of a three-time correlation function. Therefore, analyses utilizing multiple chronological times will be necessary to extract the molecular dynamics embedded in a single-molecule trajectory. In practice, however, the statistics degrades from the lower-order correlation analysis $C_2(t)$ to the higher-order ones such as $C_3(t_1, t_2)$. A great number of photon counting events will be necessary in order to discriminate with statistical confidence different theoretical models. This point is illustrated in the following discussion using the same simulation parameters.

Number of data points and effectiveness. While 1000 data points is a sufficient number to provide a glimpse of the autocorrelation decay rate (data not shown), the three-time correlation function constructed from a 1000-event trajectory is swamped by statistical noise. As more data points are taken into account, the statistics gradually improve. Shown in Figs. 11(A)–(C) are various analyses from 10^4 photons, whereas in Figs. 11(D)–(F) are those from 10^5 photons. Although the mutual information analysis of one chronological time is not as discriminative as the

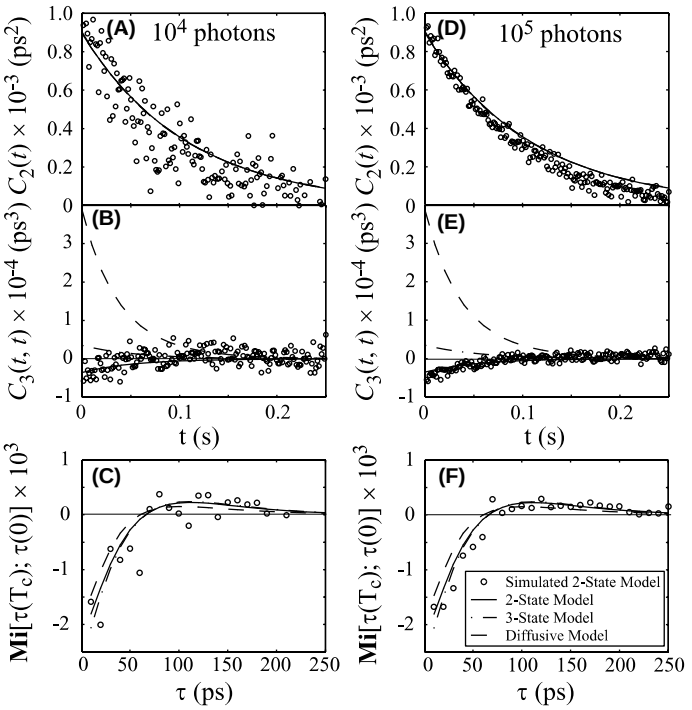


Fig. 11. Comparison of (A) autocorrelation function, (B) diagonal three-time correlation curve, and (C) diagonal mutual information plot of 10^4 simulated photons with different models in the inset of (F). (D)–(F): Same as (A)–(C) except for 10^5 simulated photons. The smooth, analytical curves are calculated using parameters from simultaneously fitting various theoretical models to delay time histogram and autocorrelation function of the 10^6-photon data.

three-time correlation method, for completeness, the $\mathbf{Mi}[\tau(T_c);\tau(0)]$ plots are also shown in Figs. 11(C) and (F). We see in Fig. 11(B) that 10^4 photons are sufficient to differentiate a continuously diffusive model from discrete switching models, despite the fact that it is still difficult to distinguish a two-state model from a three-state one. With 10^5 photons, as in Fig. 11(E), it becomes evident that the simulated data can be ascribed to a two-state model.

The above discussion suggests that it is possible to establish, in a coarse-grained sense, a theoretical model that describes the dynamics of single-molecule conformational fluctuations. This can be achieved by progressively applying higher-order correlation methods to experimental trajectories. Only a moderate number of events are necessary to distinguish a discrete model from a continuously diffusive one. For example, only 10^5 photons are needed in the present case, and it is within the limit of current experimental capability to obtain that many photons in a single-molecule trajectory.

4. Summary

In this paper, we introduce new statistical methods that utilize the recently developed photon-by-photon strategy in single-molecule spectroscopy. In this detection scheme, the delay time τ of each detected photon relative to its excitation pulse in TCSPC is recorded in real time to give a delay time sequence $\{\tau_p(t_p)\}$. The delay time sequence is then subjected to statistical analyses without further averaging, or binning. This method offers higher time resolution and improved statistics over other means that are based on binned single-molecule trajectory. The feasibility of this approach to resolving conformational dynamics models has been examined via statistical

analyses and corroborated by computer simulations. In particular, we have derived analytical solutions to various theoretical models – including the discrete two-state and three-state models as well as a continuous-diffusive model. The analytical solutions and statistical methods presented in this paper are general and are expected to be applicable to other observables, such as fluorescence wavelength, to examine spectral diffusion accompanying chemical or physical processes. They provide a basis for comparison of various statistical methods used in data processing. In some cases, the 1D two-point autocorrelation approach is unable to distinguish one kinetic model from another. To improve the discriminative power of the correlation approach, a 2D, three-time correlation method has been introduced. The three-time correlation analysis shows superior performance over the autocorrelation approach. We have also shown that three-time correlation analysis is capable of handling models of degenerate state observables. Furthermore, for such an analysis, only a moderate number of photons ($\sim 10^5$) is required, which is experimentally feasible.

An alternative method based on information theory principles has also been proposed. An implementation of the qualitative joint-probability histograms used by Lu et al. has been quantified by measuring the mutual information, or the Kullback–Leibler distance, between the correlated and uncorrelated histograms. This novel approach allows a *direct and quantitative* analysis of the probability structure of event series from a single-molecule trajectory. Since this paper is intended to introduce improved yet data-demanding statistical approaches that afford better discrimination power over various kinetic models, in-depth discussion of extension of the mutual information method is omitted here. We note, however, that this approach can potentially offer a wealth of information regarding the underlying molecular processes. For example, it can be readily extended to include two chronological times in the data analysis to improve its discriminative power, as in the case of the three-time correlation analysis. We expect that such powerful information theory-based methods when further developed will facilitate studies of single-molecule chemical physics and biophysics.

Acknowledgements

We acknowledge helpful discussions with Pallop Karnchanaphanurach, Guobin Luo, and Antoine M. van Oijen. This work was supported by the National Institutes of Health.

References

[1] H. Yang, X.S. Xie, J. Chem. Phys., 2002 (in press).
[2] N. Gershenfeld, The Nature of Mathematical Modeling, Cambridge University Press, New York, 1999.
[3] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[4] S.M. Nie, R.N. Zare, Ann. Rev. Biophys. Biomol. Struct. 26 (1997) 567.
[5] X.S. Xie, J.K. Trautman, Ann. Rev. Phys. Chem. 49 (1998) 441.
[6] W.E. Moerner, M. Orrit, Science 283 (1999) 1670.
[7] S. Weiss, Science 283 (1999) 1676.
[8] R.D. Vale, T. Funatsu, D.W. Pierce, L. Romberg, Y. Harada, T. Yanagida, Nature 380 (1996) 451.
[9] H. Noji, R. Yasuda, M. Yoshida, K. Kinosita, Nature 386 (1997) 299.
[10] A. Ishijima, H. Kojima, T. Funatsu, M. Tokunaga, H. Higuchi, H. Tanaka, T. Yanagida, Cell 92 (1998) 161.
[11] H.P. Lu, X.S. Xie, Nature 385 (1997) 143.
[12] H.P. Lu, L.-Y. Xun, X.S. Xie, Science 282 (1998) 1877.
[13] L. Edman, Z. Foldes-Papp, S. Wennmalm, R. Rigler, Chem. Phys. 247 (1999) 11.
[14] D.S. Talaga, W.L. Lau, H. Roder, J.Y. Tang, Y.W. Jia, W.F. DeGrado, R.M. Hochstrasser, Proc. Natl. Acad. Sci. USA 97 (2000) 13021.
[15] T. Ha, A.Y. Ting, J. Liang, W.B. Caldwell, A.A. Deniz, D.S. Chemla, P.G. Schultz, S. Weiss, Proc. Natl. Acad. Sci. USA 96 (1999) 893.
[16] R. Zwanzig, Acc. Chem. Res. 23 (1990) 148.
[17] Y.-W. Jia, A. Sytnik, L. Li, S. Vladimirov, B.S. Cooperman, R.M. Hochstrasser, Proc. Natl. Acad. Sci. USA 94 (1997) 7932.
[18] L. Edman, U. Mets, R. Rigler, PNAS 93 (1996) 6710.
[19] S. Wennmalm, L. Edman, R. Rigler, PNAS 94 (1997) 8271.
[20] M.I. Wallace, L. Ying, S. Balasubramanian, D. Klenerman, J. Phys. Chem. B 104 (2000) 11551.
[21] C. Eggeling, J.R. Fries, L. Brand, R. Günter, C.A.M. Seidel, Proc. Natl. Acad. Sci. USA 95 (1998) 1556.
[22] E. Geva, J.L. Skinner, Chem. Phys. Lett. 288 (1998) 225.
[23] G.K. Schenter, H.P. Lu, X.S. Xie, J. Phys. Chem. A 103 (1999) 10477.
[24] A.M. Berezhkovskii, A. Szabo, G.H. Weiss, J. Chem. Phys. 110 (1999) 9145.
[25] A.M. Berezhkovskii, A. Szabo, G.H. Weiss, J. Phys. Chem. B 104 (2000) 3776.
[26] N. Agmon, J. Phys. Chem. B 104 (2000) 7830.

[27] J.S. Cao, Chem. Phys. Lett. 327 (2000) 38.
[28] S.L. Yang, J.S. Cao, J. Phys. Chem. B 105 (2001) 6536.
[29] J.N. Onuchic, J. Wang, P.G. Wolynes, Chem. Phys. 247 (1999) 175.
[30] J.J. Portman, P.G. Wolynes, J. Phys. Chem. A 103 (1999) 10602.
[31] A. Molski, J. Hofkens, T. Gensch, N. Boens, F. De Schryver, Chem. Phys. Lett. 318 (2000) 325.
[32] E. Barkai, R. Silbey, Chem. Phys. Lett. 310 (1999) 287.
[33] E. Barkai, Y. Jung, R. Silbey, Phys. Rev. Lett. 8720 (2001) 7403.
[34] M. Köllner, J. Wolfrum, Chem. Phys. Lett. 200 (1992) 199.
[35] H. Yang, L. Cai, P. Karnchanaphanurach, X.S. Xie, Phys. Rev. Lett., 2002 (in press).
[36] S.A. Soper, L.M. Davis, E.B. Shera, J. Opt. Soc. Am. B 9 (1992) 1761.
[37] W.P. Ambrose, P.M. Goodwin, J.C. Martin, R.A. Keller, Science 265 (1994) 364.
[38] X.S. Xie, R.C. Dunn, Science 265 (1994) 361.
[39] J. Wang, P. Wolynes, Phys. Rev. Lett. 74 (1995) 4317.
[40] S. Kullback, R.A. Leibler, Ann. Math. Stat. 22 (1951) 79.
[41] N. Agmon, J.J. Hopfield, J. Chem. Phys. 78 (1983) 6947.
[42] C.C. Moser, J.M. Keske, K. Warncke, R.S. Farid, P.L. Dutton, Nature 355 (1992) 796.
[43] N.G. van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier Science Publishers, Amsterdam, 1992.
[44] H. Risken, The Fokker–Planck Equation, second ed., Springer, New York, 1989.