Single molecule kinetics. I. Theoretical analysis of indicators

James B. Witkoskie and Jianshu Cao
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 11 February 2004; accepted 2 July 2004)

Single molecule experiments reveal intriguing phenomenon in chemical and biological systems. Several indicators of complex dynamics, including ''intensity'' correlations, ''event'' correlations, and characteristic functions have been proposed, but extraction of information from these indicators can be difficult since these indicators only observe certain characteristics of the system. Generally, for systems that follow Poisson kinetics, all of these indicators contain similar information about the relaxation times of the system and the connections between different relaxation times, but the information is convoluted in different ways so the strength of various indicators is system specific. The paper discusses the theoretical implications and information content of various data analysis methods for single molecule experiments and demonstrates the relationships between indicators. Under certain conditions, common indicators contain all available information about systems with Poisson kinetics between degenerate states, but extraction of this information is generally not numerically feasible. The paper also discusses practical issues associated with these analyses, which motivates a numerical study based on Bayes' formula in the companion paper [J. Witkoskie and J. S. Cao, J. Chem. Phys. **121**, 6373 (2004), following paper]. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1785783]

I. INTRODUCTION

Over the last few years, many scientists used single molecule experiments to reveal the nature of dynamic systems. The first elegant single molecule experiment by Moerner and Orrit explored low temperature glasses, where the spectral diffusion of a single chromophore probes the environment. Recent experiments extend the single molecule technique to probe complex biological systems at room temperature. The systems studied with single molecule techniques have become much more complex, which makes interpretation of data more difficult.[1,2] Experiments by Chu, by Xie, and by other groups reveal the mechanisms of chemical reactions in biomolecular systems and the associated time scales of these reactions.[3–14] Other experiments demonstrate single molecule spectroscopy's ability to distinguish between heterogeneous and homogeneous relaxation in glassy systems.[15–17] All of these analyses require the determination of characteristic times and pertinent configurations or states from the frequency and count of the photons emitted during the experiment. The switching between configurations and the photon statistics are stochastic processes that create large uncertainties in the data retrieved from experiments on these systems. Considering these uncertainties, analysis of these experiments requires the use of robust statistical methods.

The stochastic fluctuations in single molecule systems stimulated interest in the statistical mechanics community.[18–30] These single molecule experiments give exciting insight into microscopic systems including the role of system bath-interactions and fluctuations. For bulk experiments one is generally only able to measure ''intensity'' correlations in the system, but these single molecule experiments allowed theorists to propose several new indicators of various dynamics in these systems, including the event-averaged quantities, ''two-event echo,'' and number density, which we discuss in recent papers.[26–28] A more recent proposal suggests the use of a generating function to examine blinking sequences.[29] Most of these analyses depend on large amounts of data to remove uncertainties and accurately measure indicators. Since the typical experiment is on the order of seconds to minutes, these large data requirements may be experimentally impossible to obtain.

The paper is an elaboration of previous work and a demonstration of the role of recently proposed indicators in the framework of single molecule kinetics.[21,22,26–29,31,32] The results in this paper are important in order for us to properly assess how to combine indicators to reveal the physics of the system. In this paper we introduce a standard single molecule model using Poisson kinetics in Sec. II. We simulate data with a specific kinetic scheme in order to give a numerical example of the application of these indicators to single molecule experiments. Through both analytical work and applications to the simulated data, we demonstrate the information content of each indicator and discuss both strengths and possible difficulties related to these indicators. We show that the indicators contain similar information about connections between relaxation times in complementary forms, and this information may be easier to extract from one indicator or another. Understanding the relationships between the signatures of various indicators is important in unambiguously determining possible kinetic schemes.

Under common conditions the indicators can theoretically contain all the available information of the system, but extraction of this information may not be numerically feasible. We also discuss that nonuniqueness of the underlying kinetics prevent determination of the exact kinetic scheme, but the degeneracy of the different schemes come from linear transformations. In the companion paper[33] we introduce a computational approach based on Bayesian statistics to analyze data from single molecule experiments, which can over-

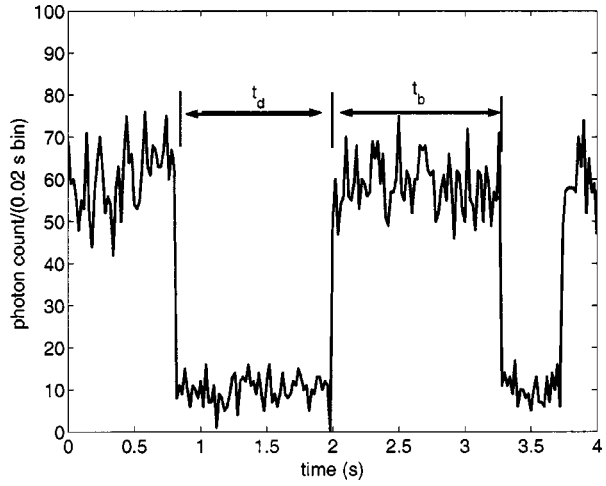

FIG. 1. The telegraph signal with 20% Poisson shot noise for the standard single molecule experiment that we examine in this paper. The molecule is either considered bright with high intensity of photon emission, or dark with low photon intensity. The event correlation function measures the duration of bright events and dark events, labeled t_b and t_d in the figure.

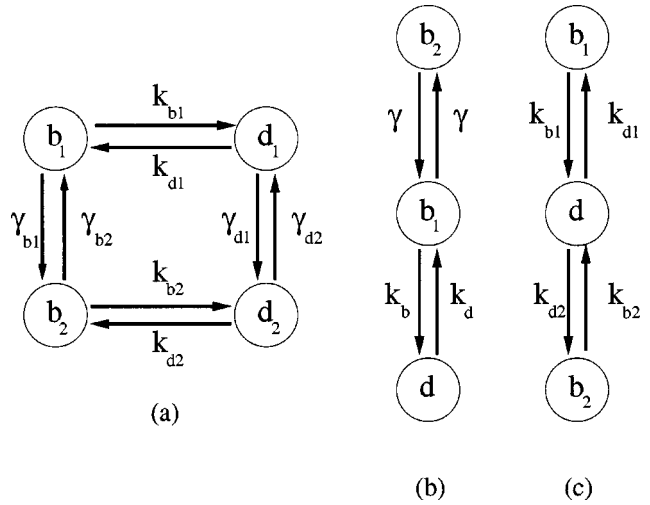

FIG. 2. Possible kinetic schemes. (a) A diagrammatic depiction of the kinetic scheme that we analyze in this paper and the companion paper. (b) and (c) show two indistinguishable schemes. For example, the probability distribution for a sequence produced from the model in (b) with parameters $k_{b1}=3\ s^{-1}$, $k_{d1}=1\ s^{-1}$ and $\gamma=2\ s^{-1}$ is indistinguishable from the model in (c) with $k_{b1}=6\ s^{-1}$, $k_{d1}=(2/5)\ s^{-1}$, $k_{b2}=1\ s^{-1}$, $k_{d2}=(3/5)\ s^{-1}$. These two sets of parameters correspond to cutting different connections in the model in (a), which one would expect to behave qualitatively differently, but this intuition is false.

come the deficiencies in the indicators. This paper along with the companion paper concentrates on combining the insight from indicators with numerical methods to approach single molecule problems.[33]

After discussing the standard four state Poisson model in Sec. II, including the important role of initial conditions, we will examine the three common indicators, intensity correlation, ''event'' correlation, and characteristic function, in Secs. III, IV, and V, respectively. Our analysis will include discussions of the information about relaxation times and connectivity contained in each indicator. For each indicator, we will also discuss situations where a limited number of moments of each indicator can theoretically contain all available information, but this information does not uniquely determine the kinetic schemes. The paper also examines numerical examples of each indicator, which demonstrates several difficulties with the use of indicators.

II. SYSTEM OF INTEREST

Our analysis will primarily focus on the blinking model that appears extensively in the literature.[20–29,31,32] The model views the single molecule experiment as a stochastic process that switches between two different set of states. A number of states, labeled ''bright,'' emit photons in a laser field, but we cannot distinguish these states from each other. The other states, labeled ''dark,'' do not emit photons and also cannot be distinguished from each other. The resulting signal resembles the telegraph signal demonstrated in Fig. 1, where we are able to see the molecule switch from a bright state to a dark state and vice versa, but we do not know which bright or dark state the system is in.

This simple model does not consider uncertainty in the state of the system, bright or dark, and also neglects restrictions on temporal resolution. One can compensate for these simplifying assumptions, but for many systems the error introduced by these assumptions is minimal. The photon shot noise is instrumental in limiting both our ability to determine the state and the temporal resolution. In many of these experiments, the photon emission rate for the bright state is much larger than the dark state background photon emission rate. Our temporal resolution is limited by our ability to collect enough photons in a bin to determine the state of the system. If both the bright state photon emission is much larger than the blinking rates and the difference between the bright and dark state photon emission rates is also much larger than the blinking rates, the binning time can be chosen so that there is little ambiguity in the state or times of transitions. The unambiguous sharp transitions are apparent in the simulation in Fig. 1, where we show a single-molecule trajectory where the bright and dark state lifetimes are 1 s (or other appropriate unit), the bright photon emission rate is 3000 photon/s and the dark emission rate is 500 photon/s. The bin size is 20 ms and the number of photons in each bin is given by a Poisson distribution. These relative rates are reasonable for many single molecule experiments. As can be seen from this trace, the state can be unambiguously assigned and the time resolution is adequate.

This simple kinetic model successfully explains correlations in the length of time spent in bright states in the experiments by Xie.[3] Mathematicians studied similar models related to ion channels.[34,35] Our discussion primarily addresses systems that follow Poisson kinetics,

$$\dot{\rho} = -\mathbf{K}\rho, \tag{1}$$

which is a type of hidden Markov chain (HMC). A HMC is a model where the kinetics of state transitions only depend on the current state, but we cannot directly observe the states. We need to infer the states from data, which does not specify the states uniquely or may only have a probabilistic dependence on the current state. Although most of our discussion will address discrete state Poisson kinetics, we can include

the continuous limit to get a diffusion type of equation, like the fluctuating bottle-neck model whose equation is given in Eq. (23) (also see Refs. 20–29, 31, 32, and 36).

The discrete state HMC can be used to interpret many experiments that fit multiexponential distributions. These models explain dynamic heterogeneity that results in long lived memory effects as we discussed in previous work.[26–28] Determination of this heterogeneity or the time scales of other underlying mechanisms in single molecule experiments requires detailed analysis of data acquired from experiments.

A. Simple model

As a demonstration of the philosophical approach to single molecule problems, we study a simple four-state model in both papers,[33] whose kinetic scheme is outlined in Fig. 2(a). The model has four states and the interconversion between the states is governed by Poisson kinetics. Two states are bright with labels "b1" and "b2," and two states are dark with labels "d1" and "d2." The equation for the probability density is given by the simple kinetic equation,

$$\begin{bmatrix} \dot{\rho}_{b1} \\ \dot{\rho}_{b2} \\ \dot{\rho}_{d1} \\ \dot{\rho}_{d2} \end{bmatrix} = \begin{bmatrix} -(\gamma_{b1}+k_{b1}) & \gamma_{b2} & k_{d1} & 0 \\ \gamma_{b1} & -(\gamma_{b2}+k_{b2}) & 0 & k_{d2} \\ k_{b1} & 0 & -(\gamma_{d1}+k_{d1}) & \gamma_{d2} \\ 0 & k_{b2} & \gamma_{d1} & -(\gamma_{d2}+k_{d2}) \end{bmatrix} \begin{bmatrix} \rho_{b1} \\ \rho_{b2} \\ \rho_{d1} \\ \rho_{d2} \end{bmatrix}. \tag{2}$$

For numerical calculations, $k_{b1}=0.75\,\mathrm{s}^{-1}$, $k_{d1}=0.50\,\mathrm{s}^{-1}$, $k_{b2}=0.33\,\mathrm{s}^{-1}$, $k_{d2}=0.22\,\mathrm{s}^{-1}$, and $\gamma_{b1}=\gamma_{b2}=\gamma_{d1}=\gamma_{d2}=0.1\,\mathrm{s}^{-1}$. Note that $b1$ is not connected to $d2$ and $b2$ is not connected to $d1$. We break the links between these states to avoid difficulties associated with nonuniqueness discussed in Sec. IV. Except for certain ranges of parameters, the eight parameter model can reproduce the waiting time distribution for the four-state model with more parameters due to the nonuniqueness.[34] The numerical values of these kinetic constants are chosen so that the waiting time distribution for both the bright and dark states are obviously not monoexponential, which we could determine from the indicators presented in Secs. III, IV, and V, but the modulation rate between the states does not cause apparent time separation in any of the indicators.

Since many experiments fit the waiting time distributions to a biexponential, this four-state model is a reasonable minimal model for many systems. For the particular constants we simulate, detailed balance holds, i.e., there is a vector ρ_{eq} such that $\dot{\rho}_{\mathrm{eq}}=0$ and $k_{ij}(\rho_{\mathrm{eq}})_j=k_{ji}(\rho_{\mathrm{eq}})_i$, but we do not need detailed balance or even Poisson kinetics to apply these approaches. In fact, detailed balance violations are generally difficult to determine.[37] The cholesterol oxidase experiment by Xie and the RNA hairpin experiment by Chu[3,9,10] monitored macromolecule that requires a substrate, which the reaction depletes from the environment. The replenishment of the substrate is a transport process, which is not in equilibrium although the system reaches a steady state. When treated appropriately, the lack of detailed balance will not affect any indicators discussed in this paper or the numerical routine discussed in the companion paper.[33]

From the model system with the specified parameters, we generate a sequence of bright and dark states for 25 molecules with a duration of 300 s, which allows about 150 observed turn-over events per molecule and 7500 pieces of data.[33] These data sets are much smaller than the data sets collected in the experiments by Xie.[3]

B. Matrix notation for Poisson kinetics

To discuss these HMC more generally we introduce a simple matrix notation for the kinetics of the system. We write the kinetic matrix as three contributions

$$\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d=\begin{bmatrix} \mathbf{\Gamma}^{(++)} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}^{(--)} \end{bmatrix}+\begin{bmatrix} \mathbf{0} & \mathbf{K}^{(+-)} \\ \mathbf{K}^{(-+)} & \mathbf{0} \end{bmatrix}+\begin{bmatrix} \mathbf{K}^{(++)} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}^{(--)} \end{bmatrix}. \tag{3}$$

We denote the bright state with a+sign and the dark state with a−sign. $\mathbf{K}^{(\pm\mp)}$ corresponds to transitions from a dark state to a bright state, or vice versa. $\mathbf{K}_d^{(\pm\pm)}$ is the decay caused by $\mathbf{K}^{(\pm\mp)}$. If probability is preserved then $(\mathbf{K}_d^{\pm\pm})_{jk}=-\Sigma_i\mathbf{K}_{ij}^{\mp\pm}\delta_{jk}$. The case of diagonal $\mathbf{K}^{(+-)}$ and $\mathbf{K}^{(-+)}$ corresponds to the modulated reaction models discussed extensively.[18–29,36,38] The $\mathbf{\Gamma}^{(\pm\pm)}$ matrices correspond to unseen transitions between two bright states or two dark states.

For the four state model, whose kinetic equation is outlined in Eq. (2), these matrices are

$$\mathbf{K}_d^{(++)}=-\mathbf{K}^{(-+)}=\begin{bmatrix} -k_{b1} & 0 \\ 0 & -k_{b2} \end{bmatrix},$$

$$\mathbf{K}_d^{(--)}=-\mathbf{K}^{(+-)}=\begin{bmatrix} -k_{d1} & 0 \\ 0 & -k_{d2} \end{bmatrix}, \tag{4}$$

and

$$\mathbf{\Gamma}^{(++)}=\begin{bmatrix} -\gamma_{b1} & \gamma_{b2} \\ \gamma_{b1} & -\gamma_{b2} \end{bmatrix} \quad \mathbf{\Gamma}^{(--)}=\begin{bmatrix} -\gamma_{d1} & \gamma_{d2} \\ \gamma_{d1} & -\gamma_{d2} \end{bmatrix}. \tag{5}$$

The waiting time distribution for the first visible transition into the dark state denoted by i given that we started in the bright state j at $t=0$ has the Laplace transform $Q_{ij}^{(-+)}(s)=(\mathbf{K}^{(-+)}[\mathbf{1}s-\mathbf{K}^{(++)}-\mathbf{\Gamma}^{(++)}]^{-1})_{ij}$. Note that our notation is transposed relative to standard probability notation. A similar expression exists for the dark to bright

transition, $Q_{ij}^{(+-)}(s)=(\mathbf{K}^{(+-)}[\mathbf{1}s-\mathbf{\Gamma}^{(--)}-\mathbf{K}^{(--)}]^{-1})_{ij}$. Below we will discuss a characteristic function, which requires us to define the combined matrix,

$$\mathbf{Q}=\begin{bmatrix} 0 & \mathbf{Q}^{(+-)} \\ \mathbf{Q}^{(-+)} & 0 \end{bmatrix}. \tag{6}$$

From the matrices $\mathbf{\Gamma}$ and $\mathbf{K}$ or $\mathbf{Q}$, all observables of the system can be determined, but we want to determine the properties of the matrix from the observables, which is a nontrivial inversion problem.

C. Equilibrium distribution

For various indicators the initial condition plays a pivotal role in determining the strength and clarity of various signatures. Since the system has Poisson kinetics, it has a well-defined steady state distribution, which for convenience we will call the equilibrium distribution ρ_{eq} even if detailed balance does not hold. Generally, ρ_{eq} will not be the initial condition of our calculations since we select certain configurations to start our measurement. For the intensity correlation discussed below, we start monitoring the system at an arbitrary time, but we only consider configurations that are in a specific state, bright or dark at that time. For this scenario, the initial condition is trivially $\delta^{(\pm)}\rho_{\mathrm{eq}}/|\delta^{(\pm)}\rho_{\mathrm{eq}}|$, where

$$\delta^{(+)}=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \delta^{(-)}=\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{7}$$

and $|\delta^{(\pm)}\rho_{\mathrm{eq}}|=\Sigma_i(\delta^{(\pm)}\rho_{\mathrm{eq}})_i$. For the event correlation function, we start the observation times after an observable transition. If we start from an observed transition from the dark state to the bright state and average over all observed transitions, the initial condition is $\rho_{\mathrm{ic}}=\mathbf{K}^{(+-)}\delta^{(-)}\rho_{\mathrm{eq}}/|\mathbf{K}^{(+-)}\delta^{(-)}\rho_{\mathrm{eq}}|$, which is the stationary flux introduced previously.[26–28] A similar expression exists for starting measurements on a bright to dark transition. Generally, this initial condition is not as strongly influenced by long time correlations as the intensity correlation since long bright or dark periods do not have as large of a contribution to the initial condition as they do for the intensity correlation initial condition.[39] It is important to note that we must average over all observed transitions. This stationary flux initial condition is adequate if the duration of a bright or dark period in the molecular trajectory is shorter than the experimental measurement times so that many measurements can be made on the same time sequence. If this is not the case, the initial condition must be modified accordingly.

III. INTENSITY CORRELATION

As discussed in Sec. II, the need to determine the time scales of the underlying conformational dynamics of these experiments motivated many authors, including our previous work, to propose various indicators. As a result, the recent literature contains extensive references to these binary ''bright-dark'' systems.[21,22,26–28,36] The popular indicators measure correlations in the state of the system or analyze blinking between the bright and dark states. Below we will discuss three major approaches to interpreting single molecule experiments with binary—''bright-dark''—informations, which we refer to as intensity correlations, event correlations (see Sec. IV), and characteristic functions (see Sec. V).[21,22,26–29,36]

The intensity correlation approach measures the probability of being in the bright or dark state at multiple times.[26–28] This correlation function requires the existence of two sets of distinguishable states. Unlike some quantities discussed below, this correlation function cannot be applied to single photon statistics since single photons do not have a well defined duration. The time scales of the correlation functions determine the relaxation of the system to the equilibrium distribution from the initial state, but does not contain the time scales for time spent in the bright or dark manifold. The one-time ''bright-bright'' intensity correlation function measures the joint probability of being in the bright state at $t=0$ and at t. The general expression for a Poisson kinetic process is

$$C_{\pm\cdots\pm}(t_1,\ldots,t_n)=\mathbf{1}^T\left\{\prod_n \delta^{(\pm)}e^{-(\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d)(t_i-t_{i-1})}\right\}\frac{\delta^{(\pm)}\rho_{\mathrm{eq}}}{|\delta^{(\pm)}\rho_{\mathrm{eq}}|}, \tag{8}$$

where the matrices are time ordered. The $\mathbf{1}^T$ is a unit row vector that denotes a sum over the remaining components and the resulting quantity is a scalar. The bright-bright correlation function measures the eigenvalues that correspond to the relaxation of the system from the nonequilibrium initial condition of being in the bright state at $t=0$. The matrix that governs this relaxation is $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$.

If the eigenvalues of the matrix $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$ are unique, the two time correlation functions $C_{\pm\pm\pm}(t_1,t_2)$ contain all of the available information about the other correlation functions through a decomposition of the correlation function into a sum of multiexponentials,

$$C_{\pm\cdots\pm}(t_1,\ldots,t_n)=\sum_{i_1,\ldots,i_n} a^{\pm}_{\pm i_1,\ldots,\pm i_n}\prod_{m=1}^{n} e^{\lambda_{i_m}(t_i-t_{i-1})}, \tag{9}$$

where we can then relate the coefficients $a^{\pm}_{\pm i_1,\ldots,\pm i_n}$ to each other. As a result, we can determine all of the available information about the system from accurate two time measurements.[40] For completeness, we outline the proof in Appendix A. We use the term all available information because different kinetic schemes can result in the same waiting time distributions and therefore produce the same sequence, although from a given kinetic matrix the waiting time distributions are uniquely determined.[34] We present an example of this nonuniqueness in Sec. IV.

Although one can theoretically extract information on the duration of time spent in the bright or dark manifold by measuring an infinite number of multitime correlations, it is not practically feasible. The approach also has a difficulty when one set of states is short lived and the system is almost always bright or dark since the correlations stay near the base line. One must simultaneously solve all of the eigenvalues at once, whereas the event correlation methods that we will discuss in Sec. IV below separates the bright and dark pro-

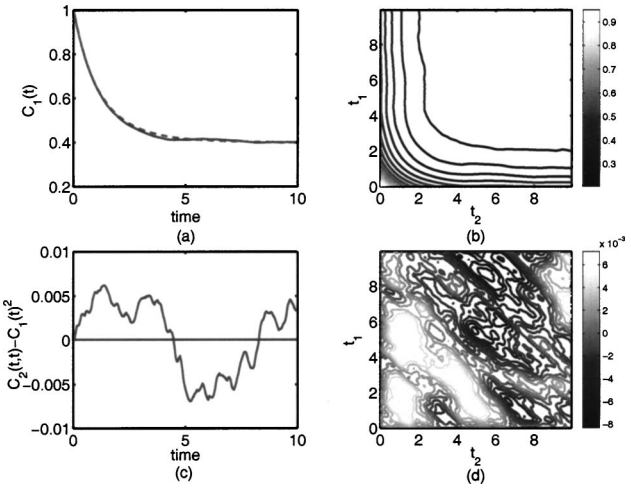

FIG. 3. Memory in the correlation function generated from the sequences of 25 molecules with kinetic scheme discussed in Sec. II. (a) shows the single-time bright-bright correlation function $C_{++}(t)$ measured from the simulation vs the exact solution for the model. As can be seen, the data is not sufficient for the determination of weak features. (b) shows a contour plot of the two-time bright-bright-bright correlation function $C_{+++}(t_1,t_2)$. (c) shows the deviation of $C_{+++}(t,t)$ from $C_{++}(t)^2$. The deviations coincide with deviations of the simulation data from the exact correlation function and cannot be considered an indication of memory. These deviations are less than 1% of the normalized correlation function and well within the range of noise. (d) shows a two dimensional plot of $C_{+++}(t_1,t_2) - C_{++}(t_1)C_{++}(t_2)$, which shows the scale of the noise in the system. Once again, the largest deviations are less than 1% of the maximum height of the correlation functions.

cesses. The intensity correlation function does not contain any readily observable signatures of conformational dynamics.[26–28] Another difficulty with this approach is the need to accurately calculate the correlation function for a wide range of times and for multiple times, since the intensity correlation relaxation time scale will generally be larger than the duration of events. The intensity correlation function does not probe individual elements of $\mathbf{\Gamma}$, $\mathbf{K}$, and $\mathbf{K}_d$, only the total matrix, so that determination of the kinetics is more difficult than the event correlation.

A. Numerical example

The deviations of the population from the equilibrium distribution by knowing that the system is in the bright or the dark state at a given time are smaller than knowing that the system made a transitions from bright to dark or vice versa at that time. As a result, the signatures of memory can be weak. The weak signatures are made apparent in Fig. 3. The figure compares the two time ''bright-bright-bright'' correlation function to the predictions for a single time correlation function, $C_{+++}(t_1,t_2)-C_{++}(t_1)C_{++}(t_2)$, for the model system discussed in Sec. II. The correlation function was calculated with a sliding time window from 25 molecular trajectories. The maximum deviations in the correlation function are around 1% relative to the steady state value of the correlation function. This deviation is of the same order as the deviations of the correlation function calculated from the data relative to the exact calculation, and because of correlations in the fluctuations and the high redundancy in the sliding window, the deviations appear to be systematic.

The deviations can be seen in Fig. 3(a), which compares the bright-bright correlation function $C_{++}(t)$ determined from the data with the model prediction. The two-time correlation function $C_{+++}(t_1,t_2)$ shown in Fig. 3(b) does not have any obvious features that represent memory. Figure 3(c) compares the values of C_{+++} for $t_1=t_2$ with the square of the single-time correlation function. The deviations between the two correlation functions coincide with deviations presented in Fig. 3(a), which indicates that it is probably a data artifact. Figure 3(d) shows a contour plot of the deviations for $0<t_1,\ t_2<10$. The maximum deviations are around 1% of the total magnitude of the function and may not indicate any memory effects considering the data's deviations from the exact correlation function. These difficulties make it almost impossible to determine the coefficients of the decomposition of the intensity correlations a_{ij} above, but suggest that other methods of determining the presence of memory are necessary. One should also note that the deviations do not give a clear signature of a specific kinetic motif since the deviations occur at short time which makes determination of specific features difficult.

IV. EVENT CORRELATION

The event correlation approach determines the statistics of events. An event is the duration of a bright or dark period. In other words, the start of a bright event is the time when the molecule makes a transition from dark to bright. The event ends when the molecule turns dark again. The definitions of time of events are demonstrated in Fig. 1. We can easily generalize these definitions to photon statistics, where the events are the arrival times of individual photons so the times of events are the times separating photons. The time scales extracted from the measurements pertain to the duration of events and not the overall relaxation measured by the intensity correlation approach. One determines the correlations in the length of duration of multiple events, such as two bright events separated by one dark event, or separated by two dark and one bright event, etc. The most important indicator of the event correlation family is a bright event versus an adjacent dark event. This quantity is the strongest, but also the most important for cases where we can determine all available information. Previous work concentrated on two bright events separated by a single dark event, but these measurements are not as strong and cannot be simply inverted even when the eigenvalues are distinct.[26–28] The approach contains possible flags to infer non-Poisson or nonrenewal behavior, but these flags also require a large amount of data to see them. One of these flags is the ''two-event echo,'' which shows a rise in the two-event bright-bright duration probability density function compared to the predictions of uncorrelated behavior for $t_1\approx t_2\gg\tau$, where t_i is the duration time of an event and τ is a characteristic time of the system. The position and height of the echo is a measure of the memory of the system.[26–28]

The *n*-event probability density function, which we will call the *n*-event correlation function, directly specifies the probability of the path of the system. The correlation or any other quantity is a sum over all possible paths so the event

correlation is in many ways the basis of all other measurements. The probability density for the 1st n transitions is

$$P(t_1,\ldots,t_n)=1^T\left[\prod_{i=1}^{n}\mathbf{K}e^{(\mathbf{K}_d+\mathbf{\Gamma})(t_i-t_{i-1})}\right]\rho_{\mathrm{ic}}^{(\pm)}. \tag{10}$$

The initial condition $\rho_{\mathrm{ic}}^{(\pm)}$ is chosen because the measurements start at a transition and one averages over all of the possible starting transitions. The superscript that appears in the initial condition is to specify the initially observed state. As has been pointed out by Verbeek and Orrit, the single event probability does not contain any information about the overall relaxation of the system. In order to extract meaningful information about conformational kinetics, multiple time event densities are required, whereas the single-time correlation function contains some information about conformational kinetics.[26–28] As discussed in Sec. II, event densities can be sensitive to photon shot noise, which can cause uncertainty in the state of the molecule, and data binning, which can cause uncertainty in the time of transitions, if the photon emission rate is not high enough.

Similar to the two-time intensity correlation function, when the matrices $\mathbf{\Gamma}^{(\pm\pm)}+\mathbf{K}_d^{\pm\pm}$ have distinct eigenvalues, the two-event correlation function contains all of the available information about the Poisson process. The derivation follows the results in Appendix A with some redefinitions of various quantities and is also outlined by Fredkin and Rice.[40] It is important to note that the eigenvalues for the decay constants will be determined by the submatrices $\mathbf{K}_d^{(++)}+\mathbf{\Gamma}^{(++)}$ and $\mathbf{K}_d^{(--)}+\mathbf{\Gamma}^{(--)}$ instead of the complete matrix $\mathbf{K}_d+\mathbf{\Gamma}+\mathbf{K}$ and the coupling between the exponential decays for the two different times will be determined by the matrices $\mathbf{K}^{(+-)}$ and $\mathbf{K}^{(+-)}$ instead of $\delta^{(\pm)}$, so one can separate the contributions from $\mathbf{K}^{(\pm\mp)}$ and $\mathbf{\Gamma}^{(\pm\pm)}+\mathbf{K}_d^{\pm\pm}$.

The result implies that there is no additional information in higher densities, but the analysis of the two-event correlation function requires binning small amounts of data into a multidimensional array. Since the probability of falling into a specific bin is a Bernoulli variable, statistical fluctuations in the number of events that fall into a bin can affect determination of these relaxation times. One method of avoiding the binning difficulty is the calculation of the covariance of the times of two events. Although this calculation can indicate a correlation between adjacent events, one loses much of the information contained in the density since the covariance is an averaged quantity.

A. Event correlations and nonuniqueness

The ''n-event'' density function can be rewritten in a revealing form

$$P(t_1,\ldots,t_n)=1^T\left[\prod_{i=1}^{n}\mathbf{K}e^{(\mathbf{K}_d+\mathbf{\Gamma})(t_i-t_{i-1})}\right]\rho_{\mathrm{ic}}$$
$$=q^T\left[\prod_{i=1}^{n}\mathbf{S}^{-1}\mathbf{K}\mathbf{S}e^{-\mathbf{\Lambda}(t_i-t_{i-1})}\right]p, \tag{11}$$

where $\mathbf{S}$ is the similarity transform that diagonalizes $\mathbf{K}_d+\mathbf{\Gamma}$ and $\mathbf{\Lambda}$ are the corresponding eigenvalues. The vectors, $p=\mathbf{S}^{-1}\rho_{\mathrm{ic}}$ and $q^T=1^T\mathbf{S}$, result from the similarity transform. We assume $\mathbf{\Lambda}$ is diagonal because having an algebraic multiplicity greater than 1 usually requires additional constraints. This form demonstrates a lack of uniqueness since any system with the same eigenvalues, the same matrix $\mathbf{S}^{-1}\mathbf{K}\mathbf{S}$, and the same vectors q^T and p give the same probability distribution for multiple time events.[34] In other words, since the density is restricted to probe along the vectors p and q^T complete information about the underlying matrix is lost even if we can determine all of the higher order correlations or event densities, but the different models are related through a linear transformation.

A simple example of two indistinguishable kinetic schemes is presented in Figs. 2(b) and 2(c). These are three-state models with two bright and one dark state. We can map the three-state model into a model with one bright state and one dark state with multiexponential waiting times for the duration of each state. Processes with one bright state and one dark state with possibly non-mono-exponential waiting times are known as alternating renewal processes. If we set $k_{b1}=3\ \mathrm{s}^{-1}$, $k_{d1}=1\ \mathrm{s}^{-1}$, $k_{b2}=k_{d2}=0\ \mathrm{s}^{-1}$, $\gamma=2\ \mathrm{s}^{-1}$, the bright state waiting time distribution with dimensionless time is $(2/5)6e^{-6t}+(3/5)e^{-t}$. The dark state waiting time is a simple Poisson process e^{-t}. The model in Fig. 2(c) can achieve the same waiting-time distribution with the parameters $k_{b1}=6\ \mathrm{s}^{-1}$, $k_{d1}=(2/5)\ \mathrm{s}^{-1}$, $k_{b2}=1\ \mathrm{s}^{-1}$, $k_{d2}=(3/5)\ \mathrm{s}^{-1}$. These models have different connections, but the ambiguity about the underlying kinetic scheme cannot be resolved.

B. Numerical example

The advantages of the event correlation over the other measurements mainly stem from its separation of the bright and dark events and the separation of $\mathbf{K}_d+\mathbf{\Gamma}$, which determines the eigenvalues, and $\mathbf{K}$ which determines the connectivity. The event correlation is the best indicator in many respects. As shown in Figs. 4(a) and 4(b), we are able to easily fit the waiting time distribution generated by 25 molecules to a biexponential form. This biexponential form gives us the number of states in both the bright and dark manifolds and also restricts the range of eigenvalues. A log plot of the binned data is not obviously biexponential, but it is obviously not monoexponential, which would justify performing a numerical fit to the biexponential, which is quite good.

The deviations of the two-event correlation versus the prediction from single waiting time distributions, $P(t_1,t_2)-P(t_1)P(t_2)$, are on the order of 2%, which is stronger than the correlation function and the echo at later times indicates conformational dynamics. It is important to note that Fig. 4, presents the adjacent ''bright-dark event'' correlations. The ''bright-bright event'' correlations has weaker memory effects and cannot be used in the analysis for system with distinct eigenvalues since one needs to know about the dark state to determine the connectivity. For the four-state model, one only needs to fit the data to a biexponential, versus a triexponential and a constant for the correlation functions.

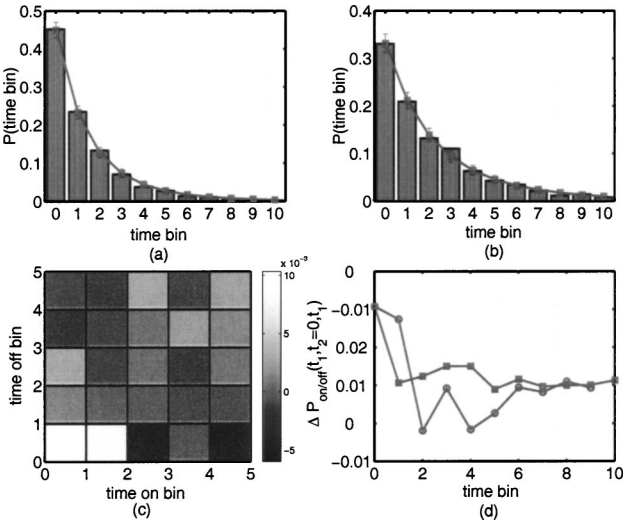

FIG. 4. Event density generated from data of 25 molecules. (a) shows the bright waiting time distribution measured from the data for 25 molecules vs the predictions. (b) shows the dark waiting time distribution measured from 25 molecules vs predictions. (c) Shows the error of the two-time bright-dark density vs the predictions of uncorrelated behavior, $\Delta P_{\mathrm{on/off}}(t_1,t_2) = P_{\mathrm{on/off}}(t_1,t_2) - P_{\mathrm{on}}(t_1)P_{\mathrm{off}}(t_2)$. (d) shows two slices of (c), $t_2=0$ (□) and $t_2=t_1$, (○).

The reduction in the number of fits also makes use of the event quantities more feasible.

V. CHARACTERISTIC FUNCTION

In this section we examine a recently proposed indicator, the characteristic function, and the related moments that we previously referred to as ''number densities''.[21,22,26–29,36] The characteristic function and related moments have an extensive history although suggestions of introducing them into the field of single molecules is fairly recent.[21,22,28,29,36] The ''characteristic function'' approach examines the probability of observing n transitions from bright to dark or vice versa in a period of time $P(n,t)$.[29] One classic indicator derived from the characteristic function is the Mandel Q parameter that measures deviations from a Poisson process,

$$\frac{\langle n^2(t)\rangle - \langle n(t)\rangle^2}{\langle n(t)\rangle} - 1, \tag{12}$$

where $\langle n^i(t)\rangle = \Sigma_n n^i P(n,t)$ is the expectation of the number of events. Wang and Wolynes and Silbey and co-workers also discuss using moments in several single molecule experiments.[30,38] One attractive property of the characteristic function indicators is that the large amounts of information can be carried in quantities measured at a single time. The binning with respect to the number of observed events is more natural than binning time. Often one measures moments or number densities, which resemble the use of correlations and avoids binning. These properties help us avoid problems associated with binning our data in multidimensional arrays.

If there is a visible transition into state j at $t=0$, the probability of observing n more transitions by time t is given by

$$P(n,t|j,t=0) = \sum_{i,k} \left[\int_t^\infty dt Q_{ik}(t) \right] * [\mathbf{Q}^{*n}(t)]_{k,j}. \tag{13}$$

In this expression, $\mathbf{Q}$ is the matrix defined in Eq. 6, $*$ is the convolution operator, and $*n$ denotes n convolutions of the same term. The $\mathbf{Q}(t)^{*n}$ term represents the probability of making n transitions before time t, and the $[\int_t^\infty dt Q_{ik}(t)]$ term represents the probability of not making the next transition. The expression in the Laplace domain is given by

$$P(n,s|j,t=0) = 1^T \left[\frac{\mathbf{1} - \mathbf{Q}(s)}{s} \right] [\mathbf{Q}^n(s)]. \tag{14}$$

The above equation is valid for any waiting time distributions, not just Poisson kinetics. The formalism can be easily applied to photons statistics, where a transition corresponds to a detected photon. Since we do not know the original or final state we must introduce the initial distribution and sum over the initial states. As will be discussed below the best initial condition starts from a switching event, switching from bright to dark or vice versa, since this initial condition allows simple analysis to determine if the process is the results of single waiting-time distributions. Supplemented with an initial condition $\rho_{\mathrm{ic}}^{(\pm)}$, the probability of observing n transitions in a HMC with bright and dark states becomes

$$P(n,s|\mathrm{ic}) = 1^T [\mathbf{1}s - \mathbf{\Gamma} - \mathbf{K}_d]^{-1} (\mathbf{K}[\mathbf{1}s - \mathbf{\Gamma} - \mathbf{K}_d]^{-1})^n \rho_{\mathrm{ic}}. \tag{15}$$

The characteristic function for the number of events, $n \to \zeta$, as a function of the Laplace variable $t \to s$ is

$$\begin{aligned} G(\zeta,s|\mathrm{ic}) &= \sum_n P(n,s|\mathrm{ic}) e^{in\zeta} \\ &= 1^T [\mathbf{1}s - \mathbf{\Gamma} - \mathbf{K}_d - e^{i\zeta}\mathbf{K}]^{-1} \rho_{\mathrm{ic}}. \end{aligned} \tag{16}$$

This characteristic function comes from the discrete Fourier transform, which can be easily related to the generating function by Brown (a Z transform) through analytic continuation.[29] For numerical applications the discrete Fourier transform has better software support and the general approach to inverting a Z transform is by evaluating it on the unit circle, in which case it becomes the discrete Fourier transform.

A. Moments of the characteristic function

As with many functions, the Fourier or Laplace transforms generally do not give good insight into the probability function. Because of this difficulty, the most convenient and informative procedure for analyzing a characteristic function will come from the moments. If there is one state with arbitrary waiting-time distribution $Q(s)$ and we see the transitions from the state back into itself, an intermittent blink discussed in Refs. 26 and 27, we have complete information by only knowing the expected number of blinks as a function of time. This scenario also applies to photon counting statistics, where the blinks are emitted photons. For photon statistics, the photon emission rates from the system will be lower than the bright-dark kinetics. As a result the background photon detection rate becomes important. For a Poisson kinetics with Poisson background, the kinetic rates in $\mathbf{K}_d$ need to be

increased by the background rate and the emission from the background will correspond to a visible transition back into the same state. For a single state, the first moment is known as the renewal function and it is related to the waiting-time distribution by the renewal theorem,

$$Q(s)=s\langle n(s)\rangle/(1+s\langle n(s)\rangle), \tag{17}$$

where $\langle n(s)\rangle$ is the Laplace transform of the expected number of events, $\langle n(t)\rangle=\Sigma_n nP(n,t)$, whose derivative is the number density that we discussed previously.[26] The relation is only valid if the process starts from a blink. If the process does not start from a blink, i.e., arbitrary start time, the expected number of blinks grows linearly in time and the relation is lost, $\langle n(t)\rangle=t/\tau$, where τ is the average time between events, and any memory effects are contained in the higher moments, which are difficult to measure. This analysis can be extended to processes with one bright and one dark state if one measures the expected number of transitions initially starting from a bright to dark transition and then measures the expected number of transitions initially starting from a dark to bright transition, one can deduce both the bright and dark waiting-time distributions, in Laplace space. Systems that have only a single bright and single dark state, can be described by two waiting-time distributions and are known as alternating renewal processes. Systems with a single waiting times for both bright and dark states have a simple memory. The system remembers how long it has been in a particular state, but does not know the path that it took to get there.

B. Deviations from simple memory

The characteristic function also contains information about the nonrenewal nature of the process. For a process with a single waiting-time distribution (renewal process),

$$\sum_n n^2P(n,s)=\langle n^2(s)\rangle=\langle n(s)\rangle+2s\langle n(s)\rangle^2, \tag{18}$$

and deviations from this relation imply more complex memory effects. Similar relationships can be derived for an alternating renewal process, supplemented with the appropriate initial conditions.

For the case of intermittent blinks discussed in Refs. 26 and 27, the inverse Laplace of transform of Eq. (18) suggests a simple indicator

$$I(t)=\frac{\langle n^2(t)\rangle-\langle n(s)\rangle-2\int d\tau\langle n(t-\tau)\rangle\partial_t\langle n(\tau)\rangle}{\langle n(t)\rangle}. \tag{19}$$

For long times $\langle n(t)\rangle\approx\mu t+b$, where μ is the average waiting time and b is an off set caused by the non-Poisson nature of the system. In the long time limit, the indicator function behaves as

$$I(t)\rightarrow\frac{\langle n^2(t)\rangle-\mu^2t^2-4\mu bt-\mu t}{\mu t}. \tag{20}$$

Note that for a Poisson process, $b=0$ and we recover the Mandel Q parameter.

C. Long lived correlations and the moments of the characteristic function

To give insight into the memory effects captured by these indicators, we consider a simple example with long lived bright states and intermittent dark states. The transition matrix $\mathbf{Q}$ is given by

$$\mathbf{Q}(s)=\begin{bmatrix}1-p & p\\ p & 1-p\end{bmatrix}\begin{bmatrix}\frac{1}{1+s} & 0\\ 0 & \frac{2}{2+s}\end{bmatrix}, \tag{21}$$

and in the limit as $t\rightarrow\infty$ the indicator approaches $\lim_{t\rightarrow\infty}\langle n(t)\rangle I(t)=(1/18p)[(1/p-2)]$. For $p=0$, the indicator is infinite since the memory is infinite. For $p=1/2$ there is no memory of the previous transition and the indicator is zero, but the process is not Poisson. This is a simple renewal process (single waiting-time distribution) with biexponential waiting time. The renewal indicator becomes negative for $p>1/2$ because the system prefers to flip-flop between states and a fast transition is followed by a slow one, i.e., anticorrelated steps. The anticorrelation is a sign of a break down of detailed balance resulting in circulation through the configurations. The important observation is that the indicator is based on first and second cumulants, which are easy to measure. Comparisons of higher order cumulants of $P(n,t)$ are not as easily measured, and it is difficult to extract meaningful information from these higher cumulants.

In these long persistence situations, large deviations from a normal distribution for intermediate times are possible.[21,22,29,36] For our example with $p\approx0$ at intermediate times, one observes a superposition of two normal distributions. Although the two-event measurements contain some of this information, the information may not be as explicit if the two exponents are comparable $k_1\approx k_2$, but p is still small.

D. Moments of a Poisson process and the extraction of kinetic schemes

From Eq. (16) for the HMC, the expected number of blinks (also know as renewals) can be written as

$$\langle n(s)\rangle=\frac{1}{s}1^T\mathbf{K}[\mathbf{1}s-\mathbf{\Gamma}-\mathbf{K}-\mathbf{K}_d]^{-1}\rho_{\mathrm{ic}}, \tag{22}$$

where $[\mathbf{1}s-\mathbf{\Gamma}-\mathbf{K}-\mathbf{K}_d]^{-1}$ is the matrix for the relaxation to the equilibrium distribution from the fluctuation that results in the initial transition at $t=0$. These are the same decay constants measured by the intensity correlation functions, but the quantities are integrated because of the $1/s$ term, which often makes extraction of this information difficult. Similar to discussions above, these eigenvalues are different than the two-event density eigenvalues, and the two measurements provide complementary information. One advantage of the characteristic function over the intensity correlation function is that the characteristic function directly probes the transition matrix $\mathbf{K}$ which is also measured in the event density so the characteristic function can be viewed as a mixture of the intensity and event indicators.

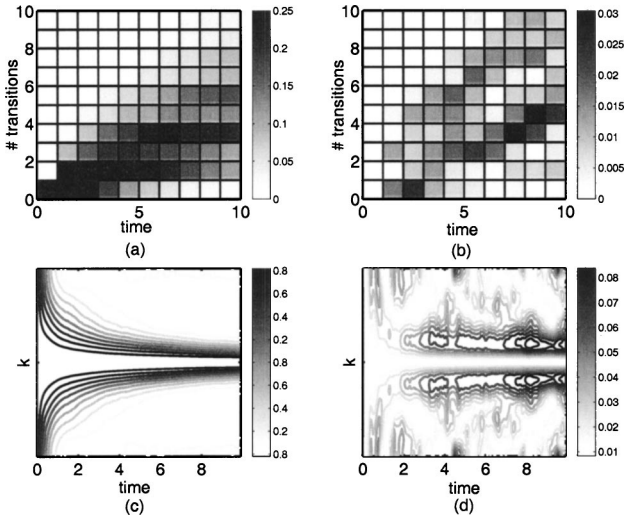

FIG. 5. $P(n|t)$ and the absolute values of its generating function $G(k|t)$. (a) shows $P(n|t)$ determined from data of 25 molecules. (b) shows the errors between $P(n|t)$ and predictions for alternating renewal process. (c) shows the generating function $G(k|t)$ determined from the data. (d) Shows the absolute errors between $G(k|t)$ and the predictions for alternating renewal processes.

For $\mathrm{Re}(s)>0$, the system has a convergent Taylor expansion in terms of the generating variables around $\zeta=0$, so we only need to know the moments to determine the functional form. If the eigenvalues of the matrix $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$ are distinct we can use a similar procedure to those applied to the intensity correlation function to relate higher moments to the first few lower moments, so we can theoretically determine the entire generating function from the first few moments. This procedure is discussed in Appendix B. Unlike the intensity correlation function, the relations are highly nonlinear and requires higher order moments so this extraction is not practical. Introducing a multiple time moment expansion will prevent the nonlinearities, but multiple times also remove the advantages of having a large amount of information contained in a two variable function.

E. Numerical examples

The above analysis demonstrates the difficultly in using the generating function to see explicit details of the system. Although the position and variance of the number of observed transitions give indications of memory effects, the underlying causes of these memory effects are not obvious. The lack of specific features makes it difficult for the generating function to distinguish features in the data. The lack of other features can be seen in Fig. 5. In this figure, we examine the data generated for the model discussed in Sec. II. The initial condition is a dark to bright transition and every dark to bright transition in the trajectory is used as an initial condition. Figure 5(a) shows a histogram of the number of renewals as a function of time $P(n,t)$ determined from the data sequence.

The histogram is compared against the expected probability for an alternating renewal process in Fig. 5(b). The alternating renewal waiting time distributions are determined from event correlations. The maximum error is 3% with typical errors around 0.5%. The characteristic function is calculated in Fig. 5(c). As can be seen, most of the details of this distribution are hidden by noise in the data. The only notable features are the central peak, whose width we can measure through the second moment and the weak peaks at $k=\pm\pi$. The peaks at $k=\pm\pi$ are the result of the alternating renewal nature of the generating function. The dark events are longer lived than the bright events so the number of renewals is more likely to be odd than even if we start from a renewal into a bright state. Figure 5(d) compares the generating function calculated from the data and the alternating renewal process predictions. The deviations between the alternating renewal prediction extracted from the data and the complete set of single molecule data are small by $<1\%$ relative to the maximum values of the function—unity—except in the vicinity of $k=0,\ \pm\pi$. The error in this vicinity will grow in height, but it will shrink in width, and results from the long time Gaussians measured from the real system and predicted for an alternating renewal process. These peaks indicate a memory effect, but it is difficult to discern a motif associated with this memory effect.

Brown used the characteristic function approach to compare the signatures of the four-state model with the fluctuating bottle-neck model.[29] The fluctuating bottle-neck model corresponds to a one dimensional diffusion process in a harmonic well with a reaction rate that depends quadratically on the coordinate.

$$\partial_t P_{\pm}(t)=D\nabla^2 P_{\pm}(t)+\nabla[kxP_{\pm}(t)]-\kappa_{\pm}x^2P_{\pm}(t)+\kappa_{\mp}x^2P_{\mp}. \tag{23}$$

Consistent with previous notation $+$ represents the bright state and $-$ represents the dark state. He demonstrates that for perfect data the characteristic functions are different for the four state model and the fluctuating bottle-neck model, as long as the temporal fluctuations of the rate constant is slow in comparison with the rates of transitions. His example shows the exact generating functions with temporal fluctuations in the rate constants that are 25 times slower than the average rate of reaction. As demonstrated in Fig. 6, when these measurements are made on stochastic data with the more interesting scenario of the fluctuations in the rate constant being the same order as the average transition rates, the signature cannot be successfully deduced. The figure compares data generated from a diffusion process with $D=k=\kappa_+=1$ and $\kappa_{-1}=K_{\mathrm{eq}}\kappa_+=2$, with a four-state model with $\gamma_{b1}=\gamma_{d1}=0.289\,706$, $\gamma_{b2}=\gamma_{d2}=1.710\,29$, $k_{b1}=1/2k_{d1}=0.417\,953$, $k_{b2}=1/2k_{d2}=4.436\,15$. The constants for the four-state model are chosen according to the procedures outlined in Ref. 29, which are close to maximizing the similarity of the two models.

The data includes the trajectories of 25 molecules run over 100 time units resulting in an average of 150 turn overs.[33] The initial condition starts from a bright to dark and the system is averaged over all possible initial transitions, which gives about 1250 total trajectories of 10 time units. Figure 6(a) shows the absolute value of the fluctuating bottleneck characteristic function determined from the data and Fig. 6(b) shows the absolute value of the four-state

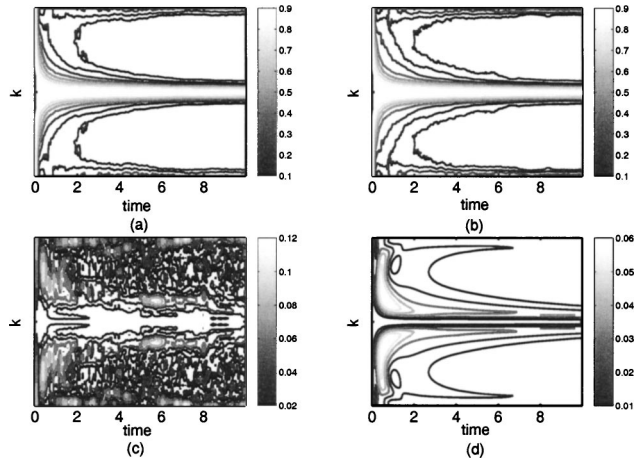

FIG. 6. Comparison the fluctuating bottle-neck model and the four-state model. (a) shows the absolute value of $G(k|t)$ for the fluctuation bottle-neck model as determined from data of 25 molecules. (b) shows the absolute value of $G(k|t)$ for the four-state model. (c) shows the absolute value of the difference in the two models. (d) Shows the theoretical results of (c).

model characteristic function determined from the data. Figure 6(c) shows the absolute value of the difference between the two models calculated from the data and Fig. 6(d), is the ideal plot of Fig. 6(c). Even the ideal signal is only about 5% of the total signal, and the noise in the characteristic function is around 10%, which prevents any strong conclusions using generating function methods. The signal would be weaker if one did not specify the sequence starting on a transition at $t=0$.[29] In the long time limit the two models become indistinguishable since they have the same average rate of transitions and similar fluctuations in the number of transitions, but at short times there are possible signatures. It important to note that similar systematic deviations between characteristic functions generated from data sets from the same stochastic process also appears because of the averaging over the same data sequence for 25 molecules. As a result, the strongest possible conclusion is that the data might not be consistent with a four state model, but these measurements do not give a good quantitative measure.

VI. CONCLUSION AND COMPARISONS

This paper presents a critical analysis of proposed indicators for single molecule experiments. All indicators have the potential to give qualitative insight into the dynamics of a single molecule system. The information contained in each indicator is similar, but it is convoluted differently in each measurement so that one indicator may have a relatively stronger signature than the others. An example is the long memory effects captured by the characteristic function.

The intensity correlation function is generally the weakest indicator. Measurements do not start from a transition so the deviations from equilibrium are small. Due to the initial condition, accurate measurements of the intensity correlation is sensitive to long bright or dark periods. The intensity correlation does not contain clear signatures of the dynamics of the system and only probes the total matrix $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$ which does not give insight into individual contributions.

The event correlations are generally the most useful of the three indicators. It separates the contributions from $\mathbf{K}$ from $\mathbf{\Gamma}+\mathbf{K}_d$ and the contributions from the bright and dark states. The event correlations also contain the two-event echo, which is a signature of conformational dynamics, but the indicator requires data binning, instead of averaging, which causes a loss of temporal resolution.

The characteristic function and its moments are a hybrid between the intensity correlations and event correlations. The characteristic times are determined by the same matrix as the intensity correlation $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$ but the coefficients allow exploration of $\mathbf{K}$ separately. The characteristic function allows averaging instead of binning, but it does not allow separation of bright and dark states. The average number of transitions and its variance give insight into memory effects of the system, but there are no other salient features. Application of complete information for distinct eigenvalues becomes a highly nonlinear problem for this indicator, unlike the other indicators.

In principle, Markovian systems with distinct eigenvalues only require two-time information to extract all available information about the process, but as shown in the simple example in Sec. II all available information does not give a unique kinetic scheme. The analysis can be extended to processes with a limited number of degeneracies, such as double degeneracies. These relations are theoretically interesting, but practical implementation is difficult if not impossible. The indicators can still give valuable information such as the time constant for relaxation or the time constants for the duration of an event. A major cause of the difficulties with any indicator is the large data requirements and the lack of a unique solution. A more robust numerical approach that does not depend on the inversion of averaged data is required. In the companion paper,[33] we give a demonstration of combining the single molecule indicators with a Bayesian numerical approach to extract possible kinetic schemes.

ACKNOWLEDGMENTS

This research is supported by the AT&T Research Fund Award, the NSF Career Award (Grant No. Che-0093210), and the Camille and Henry Dreyfus Teacher-Scholar award.

APPENDIX A: PROOF OF COMPLETE INFORMATION CONTENT

It is important to demonstrate the completeness of the information so that we can compare results with other indicators. The derivation follows those in the Ref. 40. The n-time correlation function is written in the Laplace domain as,

$$1^T\left\{\prod_n \delta^{(\pm)}\mathbf{S}[\mathbf{1}s_n+\mathbf{\Lambda}]^{-1}\mathbf{S}^{-1}\right\}\frac{\delta^{(\pm)}\rho_{\mathrm{eq}}}{|\delta^{(\pm)}\rho_{\mathrm{eq}}|}$$
$$=\sum_{i_1,\ldots,i_n} a^{\pm}_{\pm i_1,\ldots,\pm i_n}\prod_{m=1}^{n}\frac{1}{s_n+\lambda_{i_m}}. \tag{A1}$$

The matrix $\mathbf{S}$ is a similarity transform of the matrix $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$, and $\mathbf{\Lambda}$ is the matrix of corresponding eigenvalues, λ_i.

Since the eigenvalues are distinct this matrix is diagonal. The prefactor $a^{\pm}_{\pm i_1,\ldots,\pm i_n}$ is determined by fitting the n-time correlation function to the functional form presented above. The initial condition appears in the superscript of the prefactor $\pm$ and the $\pm i_m$ refers to the index of the eigenvalues and the state measured at the time t_m. Note that fitting the functional form can be done in the time domain. The functional form of the prefactor is

$$a^{\pm}_{\pm i_1,\ldots,\pm i_n} = 1^T\left\{\prod_n \delta^{(\pm)}\mathbf{S}\delta_{i_n}\mathbf{S}^{-1}\right\}p_{\pm}. \tag{A2}$$

In this expression, the matrices are time ordered and the matrix δ_{i_n} has all zeros except for the element $\{i_n,i_n\}$ which is unity. The vector $p_{\pm}$ is $\delta^{(\pm)}\rho_{\mathrm{eq}}/|\delta^{(\pm)}\rho_{\mathrm{eq}}|$. The matrix $\mathbf{S}\delta_i\mathbf{S}^{-1}$ can be written as an outer product of two vectors, $\mathbf{S}\delta_i\mathbf{S}^{-1}=w_iv_i^T$ with the property that $w_iv_i^Tw_iv_i^T=w_iv_i^T$ and $w_iv_i^Tw_jv_j^T=0$ for $j\neq i$. Starting from the trivial identity $w_iv_i^T=(a^{\pm}_{\pm i}/a^{\pm}_{\pm i})w_iv_i^T=(a^{\pm}_{\pm i})^{-1}w_ia^{\pm}_{\pm i}v_i^T$ with scalar $a^{\pm}_{\pm i}=1^T\delta^{(\pm)}w_iv_i^Tp_{\pm}$, which follows from the definitions above, we get $w_iv_i^T=(a^{\pm}_{\pm i})^{-1}w_i1^T\delta^{(\pm)}w_iv_i^Tp_{\pm}v_i$. Since $1^T\delta^{(\pm)}w_i$ and $v_i^Tp_{\pm}$ are scalars, they commute resulting in the expression, $w_iv_i^T=(a^{\pm}_{\pm i})^{-1}w_iv_i^Tp_{\pm}1^T\delta^{(\pm)}w_iv_i^T$. Substituting this expression for $\mathbf{S}\delta_{i_2}\mathbf{S}^{-1}$ in Eq. A2 gives a recursion relation for the values of

$$\begin{aligned}
a^{\pm}_{\pm i_1,\ldots,\pm i_n} &= 1^T\left\{\prod_n \delta^{(\pm)}\mathbf{S}\delta_{i_n}\mathbf{S}^{-1}\right\}p_{\pm}\\
&= 1^T\left\{\prod_n \delta^{(\pm)}w_{i_n}v_{i_n}^T\right\}p_{\pm}\\
&= a^{-1}_{\pm i_2}\left(1^T\left\{\prod_{n=2} \delta^{(\pm)}w_{i_n}v_{i_n}^T\right\}p_{\pm}\right)\\
&\quad\times(1^T\delta^{(\pm)}w_{i_2}v_{i_2}^T\delta^{(\pm)}w_{i_1}v_{i_1}^Tp_{\pm})\\
&= a^{\pm}_{\pm i_2,\ldots,\pm i_n}\frac{a^{\pm}_{\pm i_1,\pm i_2}}{a^{\pm}_{\pm i_2}},
\end{aligned} \tag{A3}$$

which implies that determination of $a^{\pm}_{\pm i}$ and $a^{\pm}_{\pm i,\pm j}$ from fitting $C_{\pm\pm\pm}(t_1,t_2)$ to sums of exponentials determines all of the higher order correlations. One can determine the decay rates λ_i and $a^{\pm}_{\pm i}$ from the single-time correlation function, and use the two-time correlation function to fit the coefficients of the exponents $a^{\pm}_{\pm i,\pm j}$ determined from the one-time correlation. If a limited number of degeneracies exist, such as double degeneracy, a limited number of higher order moments can capture this behavior. The result depends on the fact that eigenmodes never mix and does not generalize to arbitrary waiting time distributions even if these distributions have a single parameter.

As we discuss in previous work, the above analysis has an analogy to spectral decomposition in quantum mechanics.[26–28] By writing,

$$\begin{aligned}
&1^T\left\{\prod_n \delta^{(\pm)}\mathbf{S}[\mathbf{1}s_n+\mathbf{\Lambda}]^{-1}\mathbf{S}^{-1}\right\}p_{\pm}\\
&\quad= 1^T\mathbf{S}\left\{\prod_n \mathbf{S}^{-1}\delta^{(\pm)}\mathbf{S}[\mathbf{1}s_n+\mathbf{\Lambda}]^{-1}\right\}\mathbf{S}^{-1}p_{\pm}\\
&\quad= \mathrm{Tr}\left(\mathbf{S}^{-1}p_{\pm}1^T\mathbf{S}\left\{\prod_n \mathbf{S}^{-1}\delta^{(\pm)}\mathbf{S}[\mathbf{1}s_n+\mathbf{\Lambda}]^{-1}\right\}\right),
\end{aligned} \tag{A4}$$

the n-time intensity correlation maps into the spectral decomposition of the quantum time-correlation function with a density matrix $\rho\rightarrow\mathbf{S}^{-1}\delta^{(\pm)}p_{\pm}1^T\mathbf{S}$, eigenfrequencies $i\omega_i\rightarrow-\lambda_i$, and transition matrices, $\mathbf{B}\rightarrow\mathbf{S}^{-1}\delta^{(\pm)}\mathbf{S}$.

APPENDIX B: APPLICATION OF INFORMATION CONTENT TO THE GENERATING FUNCTION

For integer powers of m, the mth moment has the form,

$$\langle n^m(s)\rangle = \sum_{i=1}^{n}\frac{c_{m,i}}{s}1^T[\mathbf{K}[\mathbf{1}s-\mathbf{\Gamma}-\mathbf{K}-\mathbf{K}_d]^{-1}]^i\rho_{\mathrm{ic}}. \tag{B1}$$

In this expression, $c_{m,i}$ is a combinatorial factor, and we used $1^T[\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d]=0$ to get the $1/s$ term. The terms with $i<m$ can be expressed as lower order moments. From these expression it is apparent that each higher term contains new information in the form of the expression,

$$f_m(s) = \frac{1}{s}1^T[\mathbf{K}[\mathbf{1}s-\mathbf{\Gamma}-\mathbf{K}-\mathbf{K}_d]^{-1}]^m\rho_{\mathrm{ic}}. \tag{B2}$$

We can use the same projection technique with projection operators of the form $\mathbf{S}\delta_i\mathbf{S}^{-1}$ to derive a recursive relationship between terms that contain m products of the matrices $\mathbf{K}$ and those with one and two $\mathbf{K}$ matrices. One should note that $\mathbf{S}$ is a similarity transform that diagonalizes the matrix $\mathbf{\Gamma}+\mathbf{K}+\mathbf{K}_d$, which is the same transformation used in the intensity correlation calculation, but the coefficients will be different. Using the projection operator we can write

$$f_m(s) = \frac{1}{s}\sum_{i_1,\ldots,i_m} a_{i_1,\ldots,i_m}\prod_{j=i_1,\ldots,i_m}\frac{1}{s+\lambda_j}, \tag{B3}$$

where a's have a similar definition to the intensity correlation and event density, but they will be numerically different although λ's will be the same as the intensity correlation function. The same recursion relation holds,

$$a_{i_1,\ldots,i_m} = 1^T\left[\prod_{j=1}^{m}\mathbf{KS}\delta_{i_m}\mathbf{S}^{-1}\right]\rho_{\mathrm{ic}} = \frac{a_{i_1,i_2}}{a_{i_2}}a_{i_2,\ldots,i_m}. \tag{B4}$$

From the first moment we can find the a_i terms by fitting the transform to a sum of terms of the form $s^{-1}(s+\lambda_i)^{-1}$, but we cannot find the a_{i_1,i_2} terms from the simple second moments since we cannot distinguish ordering. It is important to note that the functional form does not give us any ordering of the eigenvalues, so we would fit $f_m(s)$ to

$$f_m(s) = \sum_{i_1\geqslant\ldots\geqslant i_m} b_{i_1,\ldots,i_m}\prod_{j=i_1,\ldots,i_m}\frac{1}{s+\lambda_j}, \tag{B5}$$

with $b_{i_1,\ldots,i_m}=\Sigma_{P\{i_1,\ldots,i_m\}}a_{i_1,\ldots,i_m}$, where $P\{\cdots\}$ is the permutation operator. By fitting the moments, $\langle n^m(s)\rangle$ or $f_m(s)$

as functions of the Laplace variables to the data, we can determine $b_{i_1,\ldots,i_m}$, and using the recursion relations, we write $b_{i_1,\ldots,i_m}$ in terms of products of a_{i_1,i_2} and $a_{i_1}^{-1}$. In principle, this process produces a set of equations that determine all of a_{i_1,i_2} and the properties of the HMC, but the resulting equations are nonlinear, which makes the existence of the solution difficult to ascertain. The higher order single time moments of the generating function can be used to extract the available information about these processes, but more of these moments are needed than with time correlations, either intensity or event measurements, and these moments generally cannot be reliably extracted from experiments. The problem can be avoided by determining a two-time renewal function, $\langle n(t_1)n(t_2)\rangle$, in which case the analysis becomes the same as Sec. III and only the first- and second-time moments are needed. Even using the two-time renewal function requires one to either fit the function in Laplace space, or fit functions that contain convolutions, which results in additional polynomials in terms of time that must be fit. The results is theoretically interesting, but in practice, even for a model with two eigenvalues we need to accurately measure the Laplace transform of the third moment and numerically solve a nonlinear set of equations, which is not numerically feasible. The lack of certainty of the existence of a solution, as well as numerical difficulties in finding this solution, makes the use of the characteristic function difficult even when the eigenvalues are distinct.

[1] W. E. Moerner, J. Phys. Chem. B **106**, 910 (2002).
[2] W. E. Moerner and M. Orrit, Science **283**, 1670 (1999).
[3] H. P. Lu, L. Y. Xun, and X. S. Xie, Science **282**, 1877 (1998).
[4] X. S. Xie and J. K. Trautman, Annu. Rev. Phys. Chem. **49**, 441 (1998).
[5] H. Yang, G. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie, Science **302**, 262 (2003).
[6] L. Edman, U. Mets, and R. Rigler, Proc. Natl. Acad. Sci. U.S.A. **93**, 6710 (1996).
[7] Y. Jia, A. Sytnik, L. Li, S. Vladimirov, B. S. Cooperman, and R. M. Hochstrasser, Proc. Natl. Acad. Sci. U.S.A. **94**, 7932 (1997).
[8] S. Nie and R. Zare, Annu. Rev. Biophys. Biomol. Struct. **26**, 567 (1997).
[9] T. Ha, X. W. Zhuang, H. D. K. J. W. Orr, J. R. Williamson, and S. Chu, Proc. Natl. Acad. Sci. U.S.A. **96**, 9077 (1999).
[10] X. W. Zhuang, L. E. Bartley, H. P. Babcock, R. Russell, T. J. Ha, D. Herschlag, and S. Chu, Science **288**, 2048 (2000).
[11] R. J. Davenport, G. J. L. Wuite, R. Landick, and C. Bustamante, Science **287**, 2497 (2000).
[12] S. Weiss, Nat. Struct. Biol. **7**, 724 (2000).
[13] Z. Xie, N. Srividya, T. R. Sosnick, T. Pan, and N. F. Scherer, Proc. Natl. Acad. Sci. U.S.A. **101**, 334 (2004).
[14] Z. Xie, X. Fang, T. R. Sosnick, T. Pan, and N. F. Scherer, Biochemistry **40**, 100 (2001).
[15] M. Ishikawa, J. Y. Ye, Y. Maruyama, and H. Nakatsuka, J. Phys. Chem. A **103**, 4319 (1999).
[16] J. Y. Ye, Y. Yamane, M. Yamauchi, H. Nakatsuka, and M. Ishikawa, Chem. Phys. Lett. **320**, 607 (2000).
[17] L. A. Deschenes and D. A. V. Bout, Science **292**, 255 (2001).
[18] J. Wang and P. Wolynes, Phys. Rev. Lett. **74**, 4317 (1995).
[19] J. N. Onuchic, J. Wang, and P. Wolynes, Chem. Phys. **247**, 175 (1999).
[20] C. L. Bai, C. Wang, X. S. Xie, and P. G. Wolynes, Proc. Natl. Acad. Sci. U.S.A. **96**, 1175 (1999).
[21] E. Geva and J. L. Skinner, Chem. Phys. Lett. **288**, 225 (1998).
[22] A. M. Berezhkovskii, A. Szabo, and G. H. Weiss, J. Phys. Chem. B **104**, 3776 (2000).
[23] I. V. Gopich and A. Szabo, J. Chem. Phys. **118**, 454 (2003).
[24] H. Yang and X. S. Xie, Chem. Phys. **284**, 423 (2002).
[25] X. S. Xie, J. Chem. Phys. **117**, 11024 (2002).
[26] J. S. Cao, Chem. Phys. Lett. **327**, 38 (2000).
[27] S. L. Yang and J. S. Cao, J. Phys. Chem. B **105**, 6536 (2001).
[28] S. L. Yang and J. S. Cao, J. Phys. Chem. B **117**, 10996 (2002).
[29] F. L. H. Brown, Phys. Rev. Lett. **90**, 028302 (2003).
[30] E. Barkai, Y. Jung, and R. J. Silbey, Phys. Rev. Lett. **87**, 207403 (2001).
[31] V. Chernyak, M. Schulz, and S. Mukamel, J. Chem. Phys. **111**, 7416 (1999).
[32] V. Barsegov, V. Chernyak, and S. Mukamel, J. Chem. Phys. **116**, 4240 (2002).
[33] J. B. Witkoskie and J. S. Cao, J. Chem. Phys. **121**, 6373 (2004), following paper.
[34] F. G. Ball and J. A. Rice, Math. Biosci. **112**, 189 (1992), and references within.
[35] I. Goychuk and P. Hanggi, Proc. Natl. Acad. Sci. U.S.A. **99**, 3552 (2002).
[36] A. M. Berezhkovskii, A. Szabo, and G. H. Weiss, J. Chem. Phys. **110**, 9145 (1999).
[37] M. Wagner and J. Timmer, Biophys. J. **79**, 2918 (2000).
[38] J. Wang and P. Wolynes, Phys. Rev. Lett. **74**, 4317 (1995).
[39] R. Verberk and M. Orrit, J. Chem. Phys. **119**, 2214 (2003).
[40] D. R. Fredkin and J. A. Rice, J. Appl. Probab. **23**, 208 (1986).