

A Continuous Time Representation of smFRET for the Extraction of Rapid Kinetics

by Zeliha Kilic¹, Ioannis Sgouralis², Wooseok Heo³, Kunihiro Ishii^{3,4}, Tahei Tahara^{3,4} and Steve Pressé^{1,4} (last updated August 20, 2020)

¹Center for Biological Physics, Department of Physics, Arizona State University, Tempe, AZ 85287

²Department of Mathematics, University of Tennessee, Knoxville, TN 37996

³Molecular Spectroscopy Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

⁴Ultrafast Spectroscopy Research Team, RIKEN Center for Advanced Photonics (RAP), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

⁵School of Molecular Sciences, Arizona State University, Tempe, AZ 85287

1	Introduction	3
2	Methods	5
2.1	Model Description	5
2.1.1	Dynamics of the System	5
2.1.2	Measurements	6
2.1.3	FRET Efficiency	7
2.2	Model Inference	8
2.2.1	Model Inference via HMMs	8
2.2.2	Model Inference via HMJPs	10
2.2.3	Computational Inference	10
2.3	Experimental Methods	11
2.3.1	Acquisition of Experimental Data	11
3	Results	12
3.1	Simulated Data Analysis	12
3.1.1	Acquisition of Simulated Data	12
3.1.2	Comparison of HMJPs with HMMs on Simulated Data	12
3.2	Experimental Data Analysis	19
3.2.1	Analysis Overview	19
3.2.2	Comparison of HMJPs with HMMs on Experimental Data	19
4	Discussion	20
A	Supporting Material	25
A.1	Comparison of HMJP and HMM for Simulated Data with Slow Switching Kinetics	25
A.2	Performance of HMJP for Simulated Data with Fast Switching Kinetics over a Range of Bin Sizes	26
A.3	HMJP Estimates for Additional Experimental smFRET Data	27
A.4	Modeling Summary	33
A.4.1	Hidden Markov Jump Process Model	33
A.4.2	Hidden Markov Model	34
A.4.3	Overview of the Sampling Updates	34
A.4.4	Background Photon Emission Rates	36
A.5	Notation and Analysis Options	38

Correspondence: spresse@asu.edu, Web: <https://cbp.asu.edu/content/steve-presse-lab>

Abstract

Our goal is to learn kinetic rates from single molecule FRET (smFRET) data even if these exceed the data acquisition rate. To achieve this, we develop a variant of our recently proposed *hidden Markov jump process* (HMJP) with which we learn transition kinetics from parallel measurements in donor and acceptor channels. Our HMJP generalizes the hidden Markov model (HMM) paradigm in two critical ways: (1) it deals with physical smFRET systems as they switch between conformational states in *continuous time*; (2) it estimates the transition rates between conformational states directly without having recourse to transition probabilities or assuming slow dynamics (as is necessary of the HMM). Our continuous time treatment learns transition kinetics and photon emission rates for dynamical regimes inaccessible to the HMM which treats system kinetics in discrete time. We validate the robustness of our framework on simulated data and demonstrate its performance on experimental data from FRET labeled Holliday junctions.

Keywords: smFRET | fluorescence | hidden Markov jump process | continuous time process | single molecule data analysis

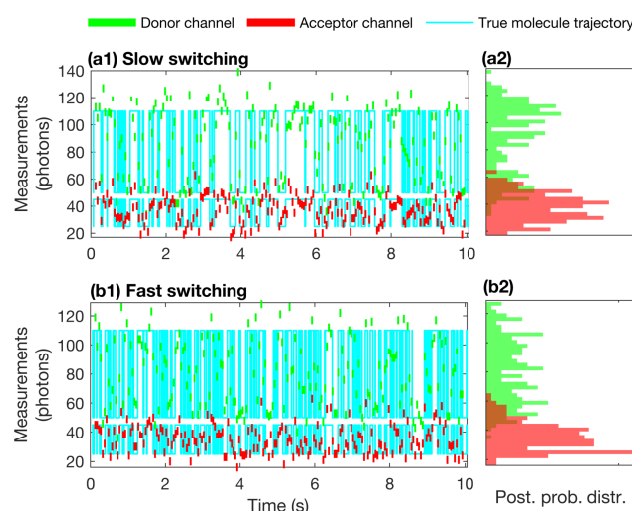


Fig. 1. An illustration of single molecule switching kinetics and corresponding measurements. In panels (a1) and (a2), we provide the simulated trajectory (cyan) of the single molecule between two states (σ_1, σ_2). The photon emission rates, i.e., number of emitted photons per unit time in the absence of noise in both donor and acceptor channels associated with the conformation states σ_1 and σ_2 are labeled with μ_{σ_1} and μ_{σ_2} , respectively. This simulated experiment provided in this figure starts at $t_0 = 0.05$ s and ends at $t_N = 10$ s with data acquisition period $\Delta t = 0.05$ s. Here in panels (a1) and (b1), for visual purposes, we assume that the measurements are acquired by a detector with fixed exposure period $\tau = 50$ ms in donor (green) and acceptor (red) channels. As the molecule switches between states σ_1, σ_2 during an integration period, the measurements represent the number of emitted photons that capture the average of the photons emitted with rates $\mu_{\sigma_1}, \mu_{\sigma_2}$ associated with the visited states. In this figure, panels (a1)-(a2) represent the simulated data where the molecule switching kinetics between the conformation states are slower than the data acquisition rate. On the other hand, in panels (b1)-(b2), we demonstrate a single molecule trajectory (cyan) when the molecule's switching kinetics are faster than the data acquisition rate. In panel (a2), slow kinetics of the molecule give rise to well separated state occupancy histograms in donor/acceptor channels around the average photon emission rates. By contrast, in panel (b2), we don't observe well separated histograms due to fast switching kinetics of the molecule.

1 Introduction

Fluorescence experiments based on single molecule Förster resonance energy transfer (smFRET) can probe the switching kinetics between conformational states defined by different inter- and intra-molecular distances (1–10). In a prototypical intra-molecular smFRET experiment, one portion of a molecule of interest is attached to a *donor* fluorophore and another to an *acceptor* fluorophore (1–18). In such an experiment, the excitation wavelength is most commonly adjusted to excite the *donor* (9, 14, 16–18). For a donor sufficiently far from the acceptor, the donor is excited and emits shorter wavelength light as compared to the longer wavelength light emitted by the acceptor in the case of energy transfer when donor and acceptor are in proximity (9, 14, 16–18). Due to the difference in the wavelength of light emitted by the donor and acceptor, the photons emitted are registered across different detectors (9, 14, 16–18). We refer to the recordings in the two detectors as the donor and acceptor channels (9, 14, 16–18). As such, the sequence of photon detections (2) encodes the kinetics according to which the distance varies between both fluorophores down to the μ s timescale (1–9). Existing approaches to analyze smFRET data include FRET efficiency histogram-based methods though such methods discard the important temporal information encoded in donor and acceptor photon arrival sequences (5–9, 11–13, 15, 16, 19). What is more, FRET efficiency histogram analysis is difficult to generalize beyond 2 conformational states as discussed in (20). In addition, on account of the arbitrary binning required to construct histograms, such methods may lead to inconclusive or erroneous estimates (5, 6, 21–27). For these reasons, modeling efforts have moved toward more direct time series analysis (7, 20, 28–31).

By relying on the hidden Markov model (HMM) paradigm (1, 3, 7, 10, 30, 32–36), time series analysis fully considers the temporal arrangement (i.e., the sequence) of photon detections and avoids histogram binning artifacts (1, 3, 7, 10, 32). This paradigm is especially fruitful in analyzing demanding experiments as the conformational states, treated as the hidden states of the HMMs, are themselves indirectly observed due to shot-noise (1, 3) and also allows for inclusion of measurement noise and specialized detector characteristics in the analysis (1, 3, 7, 10, 32).

HMMs and their variants, starting from (15) and motivating latter efforts such as the H²MMs (1, 3, 30, 35, 37) and bl-ICON (7), are the state-of-the-art in the analysis of smFRET data. However, HMMs have built into them important assumptions that we wish to lift in an effort to analyze conformational transitions occurring on time scales faster than the data acquisition rate $1/\Delta t$ supported by the detectors. Before describing these assumptions, we discuss Δt (the bin size for photon collection) otherwise known as frame rate or temporal resolution. As Δt is determined mostly by the detectors, for simplicity we assume that the same Δt applies on both channels.

In order to learn about rates exceeding the data acquisition rate, the most critical HMM assumptions that we need to lift are the following:

- HMMs make the assumption that the state switches occur *rarely* as compared to Δt . In other words, all switching rates are assumed much slower than $1/\Delta t$. For this reason, HMMs are formulated using transition probabilities rather than transition rates.
- HMMs make the assumption that transitions occur *precisely* at the end of each data acquisition period (30, 35, 38–43). In other words, intra-frame motion does not occur. For this reason, HMMs represent only instantaneous states and measurements.

The former assumption is particularly relevant to the present study as it requires, prior to the analysis, that all transitions are slower than Δt . In general, this is not only unknown but also practically difficult to quantify beforehand. After all, the objective of many experiments is to determine the switching rates in the first place. The latter assumption is also relevant to the present study as, prior to the analysis, it needs to balance two competing requirements: an upper bound on Δt set by the fastest switching rate and a lower bound on Δt supported by the detection hardware.

There have been few theoretical approaches proposed to overcome these challenges. For example, in the H²MM (3), a time grid finer than the measurement time interval is imposed and a HMM is implemented on this finer grid (3). This method is known to suffer from computational complexity (44) and it is for this reason that formulations have been sought in continuous time (44, 45). Motivated by the H²MM (3), our goal is to move to an exact continuous time treatment for smFRET systems with no approximation on the timescales of the switching kinetics.

In Fig. 1 panels (a1),(b1), we illustrate an example of smFRET measurements. In panel (a1), we show an example with slow kinetics (relative to Δt). These data are primarily contaminated with shot-noise and can be reliably analyzed within the HMM paradigm. On the other hand, in panel (b1), we demonstrate an example with fast kinetics (relative to Δt). On top of shot-noise, these data are contaminated with significant intra-frame transitions, consequently they cannot be analyzed within the HMM paradigm. As the molecule undergoes fast switching between the states during a data acquisition period, the measurement associated with that period reports on the average signal from the molecule over that period and the state of the molecule cannot be deduced from a model operating exclusively on instantaneous states such as within the HMMs paradigm. Nonetheless, the information about the fast switching is itself encoded within that measurement. It is therefore necessary to develop a method that can accommodate *continuously* evolving dynamics in order to learn the kinetics from the measurements instead of invoking an HMM formulation that, by construction, violates this critical feature.

To achieve this, a new method needs to: i) represent the dynamics of the molecule in continuous time; ii) model the acquired measurements at each data acquisition times via the average dynamics of the molecule within the associated data acquisition period; and iii) entail manageable computational cost so it allows for practical applications.

For this reason, we now turn to Markov jump processes (MJPs) describing continuous time dynamics (46–50) for which recent developments in computational statistics have suggested strategies for inferring rates using MJPs from traditional, continuous time, data (44, 45, 51). The main challenge presented by smFRET experiments, however, lie in the fact that measurements do not directly report back on the fast kinetics of molecules. Rather, frame rates report on the *average* state of the molecule over the Δt exposure window. This problem is exacerbated for molecular kinetics fast as compared to Δt . The focus of this work is really to validate a workaround, that we call the Hidden MJP (HMJP), to this critical challenge. As we will see, the HMJP is a generalization of the HMM. As such, the HMJP mathematically exactly reduces to the HMM in the limit that exposure period $\Delta t \rightarrow 0$.

To achieve this, in Section 2, we start with our HMJP smFRET model description and also, briefly for sake of comparison, summarize plain HMMs for smFRET applications. Next, in Section 3, we provide HMM and HMJP analysis comparisons for simulated measurements. In the comparison of these two methods, we present their performances in learning photon emission rates, transition probabilities (for HMMs) and kinetic rates (for HMJPs). We demonstrate how HMJPs successfully outperform HMMs especially for fast kinetic rates as compared to the data acquisition rate. Their comparison on slow switching kinetics of simulated smFRET data (where both HMM and HMJP expectedly do well) is relegated to Appendix A. Subsequently, we move onto the analysis of experimental data. We provide a comparison of HMJPs and HMMs in learning molecular trajectories, transition probabilities and kinetic rates. Lastly, in Section 4, we discuss the broader potential of HMJPs for smFRET applications.

2 Methods

Below, we provide the mathematical description of a physical system that models smFRET experiments. Subsequently we use this model in conjunction with given data to extract our estimates.

2.1 Model Description

2.1.1 Dynamics of the System

We denote the trajectory, describing the molecule’s conformational states at any given time, with $\mathcal{T}(\cdot)$. Here, $\mathcal{T}(t)$ corresponds to the molecule’s state at time t . Thus $\mathcal{T}(t)$ is a *function* over the time interval $[t_0, t_N]$. The accessible states of the molecule are labeled with σ_k and indexed with $k = 1, \dots, K$. For example, these states can be considered the $K = 2$ isomerization states of Holliday junctions (3). If the molecule is in state σ_k at time t , then we denote it by $\mathcal{T}(t) = \sigma_k$, (44, 45).

We model the transitions of the molecule as a *memoryless* process (35, 50, 52, 53) and assume exponentially distributed waiting times in each state. Together, these define our *Markov process*. In particular, as we model waiting times in continuous times, our Markov process is a *Markov jump process* (50).

In greater detail, we assume that the system chooses stochastically a state σ_k at the onset of the experiment that is denoted by $\mathcal{T}(t_0) = \sigma_k$. The probability of determining this initial state σ_k is labeled with ρ_{σ_k} . The collection of all initial probabilities is a probability vector and labeled with $\bar{\rho} = (\rho_{\sigma_1}, \rho_{\sigma_2}, \dots, \rho_{\sigma_K})$, (54–57).

Switching rates fully describe the switching kinetics of the molecule and they are labeled with $\lambda_{\sigma_k \rightarrow \sigma_{k'}}$ for all

possible states $\sigma_k, \sigma_{k'}$ where $k, k' = 1, 2, \dots, K$ and $\lambda_{\sigma_k \rightarrow \sigma_k} = 0$ by definition for $k = 1, 2, \dots, K$. For mathematical convenience, we keep track of the *escape* rates as an alternative parametrization of the switching kinetics which are given by

$$\lambda_{\sigma_k} = \sum_{k'=1}^K \lambda_{\sigma_k \rightarrow \sigma_{k'}}. \quad (1)$$

We collect all escape rates in $\bar{\lambda} = (\lambda_{\sigma_1}, \lambda_{\sigma_2}, \dots, \lambda_{\sigma_K})$. Moreover, for computational convenience, we track the normalized switching rates by the escape rates

$$\pi_{\sigma_k \rightarrow \sigma_{k'}} = \frac{\lambda_{\sigma_k \rightarrow \sigma_{k'}}}{\lambda_{\sigma_k}}. \quad (2)$$

The collection of all normalized switching rates from state σ_k is denoted by $\bar{\pi}_{\sigma_k} = (\pi_{\sigma_k \rightarrow \sigma_1}, \pi_{\sigma_k \rightarrow \sigma_2}, \dots, \pi_{\sigma_k \rightarrow \sigma_M})$. We see that each $\bar{\pi}_{\sigma_k}$ forms a probability vector (56). We can gather all transition probabilities in a matrix $\bar{\pi}$ that reads

$$\bar{\pi} = \begin{pmatrix} \bar{\pi}_{\sigma_1} \\ \bar{\pi}_{\sigma_2} \\ \vdots \\ \bar{\pi}_{\sigma_K} \end{pmatrix}. \quad (3)$$

Given $\bar{\rho}$ and $\bar{\lambda}, \bar{\pi}$, the trajectory $\mathcal{T}(\cdot)$ is obtained by a variant of the Gillespie algorithm (52) which determines a succession of states for the conformations of the system s_0, s_1, \dots, s_{M-1} and their durations $d_0, d_1, d_2, \dots, d_{M-1}$. These together define $\mathcal{T}(\cdot)$ throughout the time course $[t_0, t_N]$ by

$$\mathcal{T}(t) = \begin{cases} s_0 & \text{if } t_0 \leq t < t_0 + d_0 \\ s_1 & \text{if } t_0 + d_0 \leq t < t_0 + d_0 + d_1 \\ \vdots & \\ s_{M-1} & \text{if } t_0 + d_0 + \dots + d_{M-2} \leq t < t_0 + d_0 + d_1 + \dots + d_{M-1}. \end{cases} \quad (4)$$

For clarity, we encode $\mathcal{T}(\cdot)$ in a triplet (\vec{S}, \vec{D}, M) , where $\vec{S} = \{s_0, s_1, \dots, s_{M-1}\}$ and $\vec{D} = \{d_0, d_1, \dots, d_{M-1}\}$ and M is the size of \vec{S}, \vec{D} .

2.1.2 Measurements

The measurements in a typical smFRET experiment report on the conformational state of the molecule as it changes through time. These come in the form of two time series: $\mathbf{w}^D = (w_1^D, w_2^D, \dots, w_N^D)$ and $\mathbf{w}^A = (w_1^A, w_2^A, \dots, w_N^A)$ which are the recordings in the donor and acceptor channels, respectively. In particular, the subscripts here indicate the time level of measurements. For clarity, we assume that the measurements are time-ordered, so the $n = 1$ label coincides with the earliest acquired measurement and the $n = N$ label coincides with the latest.

We assume that measurements occur on a regular time interval denoted by Δt (although the switching of the molecule between states occurs in continuous time). Here, the values of w_n^D and w_n^A report on a number of photons collected between t_{n-1} and t_n where $t_n = t_{n-1} + \Delta t$. For completeness, we introduce an additional time level t_0 which precedes the first measurement time t_1 . This time level, t_0 defines the experiment's onset, which is not associated to any measurement; see Fig. 1.

One of the common assumptions of HMMs is that the instantaneous state of the molecule at t_n determines the measurements w_n^D and w_n^A . However, for realistic detectors, the reported values w_n^D and w_n^A are affected by the entire photon trajectory of the molecule during the n^{th} integration period represented by the time window $[t_n - \tau, t_n]$. Here, τ is the duration of each integration time for fluorescence experiments.

When we supplement our dynamical model (fully described in Section 2.1.1) with measurements, we must include a distribution describing the measurement statistics. We do so by first discussing the state specific photon emission rates in the donor and acceptor channels which we label with $\mu_{\sigma_k}^D, \mu_{\sigma_k}^A$, where the subscript highlights the state dependence of the photon emission rate. For simplicity, we gather the state specific photon emission rates in $\bar{\mu}^D = (\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \dots, \mu_{\sigma_K}^D)$, $\bar{\mu}^A = (\mu_{\sigma_1}^A, \mu_{\sigma_2}^A, \dots, \mu_{\sigma_K}^A)$.

If the molecule remains in a single state σ_k throughout an entire exposure period $[t_n - \tau, t_n]$, then the detector is triggered by $\mu_{\sigma_k}^D, \mu_{\sigma_k}^A$ and ambient contributions (background) which we label with μ_{back}^D and μ_{back}^A for the donor and acceptor channels, respectively. As such, the reported measurement, w_n^D , is similar to $\mu_{\text{back}}^D + \mu_{\sigma_k}^D$ and w_n^A is similar to $\mu_{\text{back}}^A + \mu_{\sigma_k}^A$. However, if the molecule switches between multiple states during the same exposure period, the detector is influenced by the levels of every state attained. More specifically, the n^{th} binned photon counts triggering the detector during the n^{th} exposure period, $[t_n - \tau, t_n]$, is obtained from the $\mu_{\text{back}}^D + \mu_{\mathcal{T}(\cdot)}^D$ in the donor channel and $\mu_{\text{back}}^A + \mu_{\mathcal{T}(\cdot)}^A$ in the acceptor channel over this exposure. Mathematically, this is equivalent to $\mu_{\text{back}}^D \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^D$ in the donor channel and $\mu_{\text{back}}^A \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^A$ in the acceptor channel.

With measurement noise, such as shot-noise (22, 58–61), quantification noise (62–64), or amplification noise in the case of EMCCD detectors (7, 65–67) and other degrading effects that are common in the detectors currently available, each measurement w_n^D and w_n^A depends *probabilistically* upon the triggering signal (68–70). Of course, the precise relationship depends on the detector employed in the experiment and differs between the various types of cameras, single photon detectors or other devices used. Here, we continue with a shot-limited formulation, which results in

$$w_n^D | \mathcal{T}(\cdot) \sim \text{Poisson} \left(\mu_{\text{back}}^D \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^D \right) \quad (5)$$

$$w_n^A | \mathcal{T}(\cdot) \sim \text{Poisson} \left(\mu_{\text{back}}^A \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^A \right). \quad (6)$$

2.1.3 FRET Efficiency

Later we will be making use of the notion of FRET efficiency. For this reason we define two different types of FRET efficiencies: the *characteristic* FRET efficiency (7, 15, 71–75) as well as the *apparent* FRET efficiency (7, 15, 71–75). Characteristic FRET efficiency (7, 15, 71–75) labeled with ϵ_{σ_k} for $k = 1, 2, \dots, K$ is defined as follows

$$\epsilon_{\sigma_k} = \frac{\mu_{\sigma_k}^A}{\mu_{\sigma_k}^D + \mu_{\sigma_k}^A} \quad (7)$$

and only depends on the conformational states (7, 15, 71–75). By contrast, the apparent FRET efficiency (which is often used as a proxy for ϵ_{σ_k}) is defined as follows

$$\epsilon_n^a = \frac{w_n^A}{w_n^D + w_n^A} \quad (8)$$

for $n = 1, 2, \dots, N$. Unlike the former, the latter is affected by artifacts such as measurement noise and background (15). As a result, the apparent efficiency may attain different values over time even when the molecule does not switch its conformational state.

2.2 Model Inference

Our goal is to learn the initial probabilities ρ_{σ_k} , photon emission rates $\mu_{\sigma_k}^D, \mu_{\sigma_k}^A$, switching rates $\lambda_{\sigma_k \rightarrow \sigma_{k'}}$ for all states and the trajectory of the system $\mathcal{T}(\cdot)$ during the full time course $[t_0, t_N]$ of the experiment by using the measurements $\mathbf{w}^D, \mathbf{w}^A$, and the model associated with the smFRET experiment that has just been described. First, we will explain how we learn these quantities using time series analysis within the naive HMM paradigm (7, 74, 75). Subsequently, we will present how we tackle smFRET data using our continuous time HMJPs.

2.2.1 Model Inference via HMMs

HMMs inherently assume that each measurement w_n^D and w_n^A acquired at time t_n depends only on the molecule's state at the time of data acquisition namely $\mathcal{T}(t_n)$. Therefore, we have the following approximations for the HMM formulation $\int_{t_n-\tau}^{t_n} dt \mu_{\mathcal{T}(t)}^D \approx \mu_{\mathcal{T}(t_n)}^D \tau$ and $\int_{t_n-\tau}^{t_n} dt \mu_{\mathcal{T}(t)}^A \approx \mu_{\mathcal{T}(t_n)}^A \tau$ during the exposure period $[t_n - \tau, t_n]$. Consequently, Eq. (5) and Eq. (5) become

$$w_n^D | \mathcal{T}(\cdot) \sim \text{Poisson} \left(\mu_{\text{back}}^D \tau + \mu_{\mathcal{T}(t_n)}^D \tau \right) \quad (9)$$

$$w_n^A | \mathcal{T}(\cdot) \sim \text{Poisson} \left(\mu_{\text{back}}^A \tau + \mu_{\mathcal{T}(t_n)}^A \tau \right). \quad (10)$$

Given the HMM formulation and the smFRET data $\mathbf{w}^D, \mathbf{w}^A$, we can *directly* learn the transition probabilities that govern the transitions of the molecule between its conformational states at any data acquisition time t_n for all $n = 1, 2, \dots, N$. We label a molecule's state within the HMM paradigm at time t_n with $\mathcal{T}(t_n) = c_n$. For clarity, transition probabilities determine the switching probabilities for a molecule's transitions denoted by $c_{n-1} \rightarrow c_n \rightarrow c_{n+1}$. Within the HMM formulation, $P_{c_{n-1} \rightarrow c_n}$ denotes the transition probability of the molecule from state c_{n-1} to c_n .

Given that there are K conformational states for the molecule, the number of transition probabilities $P_{\sigma_k \rightarrow \sigma_{k'}}$ becomes K^2 . We gather the transition probability from conformational state σ_k to any other conformational state including itself in $\bar{P}_{\sigma_k} = (P_{\sigma_k \rightarrow \sigma_1}, P_{\sigma_k \rightarrow \sigma_2}, \dots, P_{\sigma_k \rightarrow \sigma_K})$ that is normalized as a probability vector (namely each component sums up to 1). Eventually, we gather all transition probability vectors in a matrix called transition probability matrix denoted by $\bar{\bar{P}}$ which reads as follows

$$\bar{\bar{P}} = \begin{pmatrix} \bar{P}_{\sigma_1} \\ \bar{P}_{\sigma_2} \\ \vdots \\ \bar{P}_{\sigma_K} \end{pmatrix}. \quad (11)$$

There is a connection between the transition probability matrix $\bar{\bar{P}}$ and the molecule's switching rates $\lambda_{\sigma_k \rightarrow \sigma_{k'}}$ and escape rates λ_{σ_k} . Provided that we gather the switching rates and escape rates in the following matrix, called *generator*,

$$\bar{G} = \begin{bmatrix} -\lambda_{\sigma_1} & \lambda_{\sigma_1 \rightarrow \sigma_2} & \dots & \lambda_{\sigma_1 \rightarrow \sigma_K} \\ \lambda_{\sigma_2 \rightarrow \sigma_1} & -\lambda_{\sigma_2} & \dots & \lambda_{\sigma_2 \rightarrow \sigma_K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{\sigma_K \rightarrow \sigma_1} & \lambda_{\sigma_K \rightarrow \sigma_2} & \dots & -\lambda_{\sigma_K} \end{bmatrix} \quad (12)$$

then $\bar{\bar{P}}$ can be calculated based on $\bar{\bar{P}} = \exp(\bar{G} \Delta t)$ where $\exp(\cdot)$ stands for the matrix exponential. Knowing \bar{G} determines $\bar{\bar{P}}$ uniquely; however, knowing $\bar{\bar{P}}$ provides only a proxy for \bar{G} . Specifically, if we assume that

$\lambda_{\sigma_k} \Delta t \ll 1$ then we can use an approximation

$$\bar{\bar{P}} \approx \bar{\bar{I}} + \bar{\bar{G}} \Delta t \quad (13)$$

where $\bar{\bar{I}}$ is an identity matrix of size $K \times K$. Thus we can calculate an approximation for $\bar{\bar{G}}$ that is $\bar{\bar{G}} \approx (\bar{\bar{P}} - \bar{\bar{I}}) / \Delta t$.

Now, we provide the formulation for how to estimate the quantities including initial probabilities, $\bar{\rho}$, transition probabilities \bar{P}_{σ_k} , photon emission rates $\bar{\mu}$ and the trajectory of the system $\mathcal{T}(\cdot)$ which is encoded by $\vec{c} = (c_0, c_1, \dots, c_N)$, within an HMM paradigm. The HMM formulation is governed by the following statistical model

$$c_0 \sim \text{Categorical}(\bar{\rho}) \quad (14)$$

$$c_n | c_{n-1} \sim \text{Categorical}(\bar{P}_{c_{n-1}}) \quad (15)$$

$$w_n^D | c_n \sim \text{Poisson}(\mu_{\text{back}}^D \tau + \mu_{c_n}^D \tau) \quad (16)$$

$$w_n^A | c_n \sim \text{Poisson}(\mu_{\text{back}}^A \tau + \mu_{c_n}^A \tau). \quad (17)$$

From now on, we follow the *Bayesian paradigm* (54, 76) which requires us to prescribe prior distributions for the parameters.

We start with the prior distributions placed on the transition probabilities \bar{P}_{σ_k} for all $k = 1, 2, \dots, K$ and $\bar{\rho}$. We choose to place Dirichlet distributions with concentration parameters A and a for \bar{P}_{σ_k} and $\bar{\rho}$, respectively, that are conjugate to the Categorical distribution (30, 35, 57, 77) and formulated as follows

$$\bar{P}_{\sigma_1} \sim \text{Dirichlet}\left(\frac{A}{K}, \frac{A}{K}, \dots, \frac{A}{K}\right) \quad (18)$$

$$\bar{P}_{\sigma_2} \sim \text{Dirichlet}\left(\frac{A}{K}, \frac{A}{K}, \dots, \frac{A}{K}\right) \quad (19)$$

\vdots

$$\bar{P}_{\sigma_K} \sim \text{Dirichlet}\left(\frac{A}{K}, \frac{A}{K}, \dots, \frac{A}{K}\right) \quad (20)$$

$$\bar{\rho} \sim \text{Dirichlet}\left(\frac{\alpha}{K}, \frac{\alpha}{K}, \dots, \frac{\alpha}{K}\right). \quad (21)$$

Subsequently, we place priors on the photon emission rates $\bar{\mu}^D = (\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \dots, \mu_{\sigma_K}^D)$ and $\bar{\mu}^A = (\mu_{\sigma_1}^A, \mu_{\sigma_2}^A, \dots, \mu_{\sigma_K}^A)$. The prior that we choose to place is the Gamma distribution as it has positive support

$$\mu_{\sigma_k}^D \sim \text{Gamma}\left(\phi^D, \frac{\psi^D}{\phi^D}\right) \quad (22)$$

$$\mu_{\sigma_k}^A \sim \text{Gamma}\left(\phi^A, \frac{\psi^A}{\phi^A}\right) \quad (23)$$

with hyperparameters $\phi^D, \psi^D, \phi^A, \psi^A$.

We estimate the background photon emission rates $\mu_{\text{back}}^D, \mu_{\text{back}}^A$ by separate measurements that contain only ambient contributions as we explain in Supporting Material (A). These measurements can be obtained either after both donors and acceptors photobleach or, in a separate experiment, in which no FRET labeled molecule is present.

With all priors specified, we form the full posterior distribution (7, 30, 31, 35, 37, 77, 78)

$$\begin{aligned} & \mathbb{P}\left(\bar{\rho}, \bar{\bar{P}}, \bar{\mu}^D, \bar{\mu}^A, \mathcal{T}(\cdot) | \mathbf{w}^D, \mathbf{w}^A\right) \\ &= \mathbb{P}\left(\bar{\rho}, \bar{\bar{P}}, \bar{\mu}^D, \bar{\mu}^A, \vec{c} | \mathbf{w}^D, \mathbf{w}^A\right). \end{aligned} \quad (24)$$

Since we do not have an explicit formula for Eq. (24) we build a custom MCMC sampling scheme to generate pseudorandom numbers from Eq. (24). Details of the computational scheme are provided in Section 2.2.3.

2.2.2 Model Inference via HMJPs

The HMJP does not require any approximations on the kinetics regime allowed and instead applies directly on the formulation provided in Eq. (5) and Eq. (6). Just as with the HMM, for the HMJP formulation, we also operate within the *Bayesian paradigm* (54, 76) and thus place prior distributions on $\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A$.

First, we prescribe the prior distributions on the escape rates $\bar{\lambda} = (\lambda_{\sigma_1}, \lambda_{\sigma_2}, \dots, \lambda_{\sigma_K})$. For these, we choose Gamma distributions which are conjugate to the exponentially distributed holding times, i.e.,

$$\lambda_{\sigma_k} \sim \text{Gamma}\left(\eta, \frac{b}{\eta}\right) \quad (25)$$

for all $k = 1, 2, \dots, K$ with hyperparameters η, b . Subsequently, we prescribe independent conjugate Dirichlet distributions on the transition probabilities $\bar{\pi}_{\sigma_k}$ for all $k = 1, 2, \dots, K$

$$\bar{\pi}_{\sigma_1} \sim \text{Dirichlet}\left(0, \frac{A}{K-1}, \dots, \frac{A}{K-1}\right) \quad (26)$$

$$\bar{\pi}_{\sigma_2} \sim \text{Dirichlet}\left(\frac{A}{K-1}, 0, \dots, \frac{A}{K-1}\right) \quad (27)$$

\vdots

$$\bar{\pi}_{\sigma_K} \sim \text{Dirichlet}\left(\frac{A}{K-1}, \frac{A}{K-1}, \dots, 0\right) \quad (28)$$

with concentration hyperparameter A . Prior distributions placed on $\bar{\rho}, \bar{\mu}^D$ and $\bar{\mu}^A$ are provided in Eq. (21) and Eq. (23), respectively.

The full posterior distribution for the HMJP formulation is

$$\begin{aligned} & \mathbb{P}(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A, \mathcal{T}(\cdot) | \mathbf{w}^D, \mathbf{w}^A) \\ &= \mathbb{P}(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A, (\vec{S}, \vec{D}, M) | \mathbf{w}^D, \mathbf{w}^A). \end{aligned} \quad (29)$$

Since we do not have an analytical formula for Eq. (29), we build a custom MCMC sampling scheme. Details of the computational scheme are provided in Section 2.2.3.

2.2.3 Computational Inference

MCMC sampling from the full posterior distributions provided in Eq. (24) for the HMM and Eq. (29) for the HMJP rely on Gibbs sampling (7, 30, 31, 35, 37, 44, 77, 78). To form a large number of samples from these posteriors, we iterate the following

1. Update \vec{c} for the HMM or (\vec{S}, \vec{D}, M) for the HMJP;
2. Update transition probabilities, that is \bar{P}_{σ_k} for the HMM or $\bar{\pi}_{\sigma_k}$ and λ_{σ_k} for the HMJP;
3. Update the initial probability vector $\bar{\rho}$ for both the HMM and the HMJP;

4. Update photon emission rates $\bar{\mu}^D, \bar{\mu}^A$ for both the HMM and the HMJP.

These samples allow us to reconstruct the posterior distribution $\mathbb{P}(\bar{\rho}, \bar{P}, \bar{\mu}^D, \bar{\mu}^A, \bar{c} | \mathbf{w}^D, \mathbf{w}^A)$ for the HMM and $\mathbb{P}(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A, (\bar{S}, \bar{D}, M) | \mathbf{w}^D, \mathbf{w}^A)$ for the HMJP. The estimation of switching rates $\lambda_{\sigma_k \rightarrow \sigma_{k'}}$ with the HMJP is carried out by Eq. (2) and by Eq. (13) for the HMM.

In Supporting Material (A), we present the equation summaries for both HMJP and HMM formulations. A working code of the the implementation of our HMM and HMJP frameworks is made available through the authors' website.

2.3 Experimental Methods

Here, we introduce experimental methods. We start from sample preparation, next we move to presenting experimental procedure.

2.3.1 Acquisition of Experimental Data

Sample Preparation

The Holliday Junction strands used in this work, and whose results are shown in Figs. 3 to 7, were purchased from JBioS (Wako, Japan), of which sequences are given below

- R-strand: 5'-CGA TGA GCA CCG CTC GGC TCA ACT GGC AGT CG-3'
- H-strand: 5'-CAT CTT AGT AGC AGC GCG AGC GGT GCT CAT CG-3'
- X-strand: 5'-biotin-TCTTT CGA CTG CCA GTT GAG CGC TTG CTA GGA GGA GC-3'
- B-strand: 5'-GCT CCT CCT AGC AAG CCG CTG CTA CTA AGA TG-3'.

We note that **T** (H-strand) and **T** (B-strand) indicate thymine residues labeled with a FRET donor (ATTO-532) and an acceptor (ATTO-647N) fluorophores, respectively, at position 6 from the 5' end. The R, X, and B-strands (1 μ M, 30 μ L) and H-strand (1 μ M, 20 μ L) were mixed in TN buffer (10 mM Tris with 50 mM NaCl, pH 8.0). The mixture was annealed at 94°C for 4 min, and then gradually cooled down (2-3°C/min) to room temperature. We used a sample chamber (Grace Bio-Labs SecureSeal, GBL621502) and a coverslip that is coated by Biotin-PEG-SVA (Biotin-poly(ethylene glycol)-succinimidyl valerate) (79). Streptavidin (0.1 mg/mL in TN buffer, 100 μ L) was incubated for 20 min, which was followed by washing with TN buffer. The HJ solution (10 nM, 100 μ L) was injected for 3 s. The chamber was rinsed three times by measuring buffer (TN buffer with 10 mM MgCl₂ and 2 mM Trolox).

Experiments

Broadband light, generated by super continuum laser (Fianium SC-400-4, $f = 40$ MHz), was filtered by a bandpass filter (Semrock FF01-525/30-25) and focused on the upside coverslip surface using an objective lens (Nikon Plan Apo IR 60 \times , N.A. = 1.27). The excitation power was set to be 20 μ W at the entrance port of the microscope. The fluorescence signals were collected by the same objective lens and guided to detectors through

a multimode fiber (Thorlabs M50L02S-A). Fluorescence signals of donor and acceptor were divided by a dichroic mirror (Chroma Technology ZT633rdc) and filtered by bandpass filters (Semrock FF01-585/40-25 for donor and FF02-685/40-25 for acceptor), and then detected by hybrid detectors (Becker&Hickl HPM-100-40-C). For each photon signal detected, the routing information was appended by a router (Becker&Hickl HRT-41). The arrival time of the photon was measured by a Time-Correlated Single Photon Counting (TCSPC) module (Becker&Hickl SPC-130-EM) with time tagging mode.

3 Results

In order to show how HMJPs work and to highlight the HMJPs' advantages over HMMs, we initially benchmark our method using simulated data that mimics smFRET experiments. Simulated data are ideal for this purpose because they have a "ground truth". Generation of such data relies on the Gillespie algorithm (52). Next, we compare the strength of our HMJP method to HMMs on experimental data.

We focus on the following simulated dataset: a molecule that exhibits fast kinetics as compared to the data acquisition rate, see Fig. 1. Analysis of datasets on slow kinetics is relegated to Supporting Material (A). The results corresponding to the simulated dataset with fast kinetics for both HMJP and HMM are shown in Fig. 2. Subsequently, we show the performance of our method on experimental dataset Fig. 3. We present the results for the experimental data set in Figs. 4 to 7. Further analysis on different experimental datasets is also provided in Supporting Material (A).

Hyperparameter values used in all analyses, as well as any other choices made are presented in Supporting Material (A). For clarity, we only have access to the data demonstrated with the green and red dashes of panels (a1) and (b1) of Fig. 1 not the cyan (ground truth) trajectories. These trajectories are unknown and to be determined along with other model parameters.

3.1 Simulated Data Analysis

3.1.1 Acquisition of Simulated Data

In the generation of our simulated data, we assumed $K = 2$ attainable states, such as on/off or folded/unfolded states for illustrative purposes only (our method trivially generalizes to more states). We assumed photon emission rates which we set at $\mu_{\sigma_1}^D = 800$ photons/s, $\mu_{\sigma_2}^D = 2000$ photons/s and $\mu_{\sigma_1}^A = 300$ photons/s, $\mu_{\sigma_2}^A = 700$ photons/s where $\mu_{\text{back}}^D = \mu_{\text{back}}^A = 200$ photons/s. Additionally, we defined a data acquisition period of $\Delta t = 0.05$ s and consider the exposure period by setting τ equal to 100% of Δt . The onset and concluding time of the simulated data are at $t_0 = 0.095$ s and $t_N = 10.05$ s, respectively.

We use the following structure for the switching rates $\lambda_{\sigma_1 \rightarrow \sigma_2}, \lambda_{\sigma_2 \rightarrow \sigma_1}$ in order to specify system kinetics, with a parameter τ_f that sets the system kinetics time scale,

$$\lambda_{\sigma_1 \rightarrow \sigma_2} = \frac{2.2}{\tau_f}, \quad \lambda_{\sigma_2 \rightarrow \sigma_1} = \frac{2.1}{\tau_f} \quad (30)$$

in the analysis of fast switching rates. We simulate a case with $\tau_f = 0.1$ s in Fig. 1 panel (b1), which involves system kinetics that are faster than the data acquisition rate.

3.1.2 Comparison of HMJPs with HMMs on Simulated Data

Here, we compare HMJPs and HMMs on the analysis of simulated measurements given in Fig. 1 panel (a1).

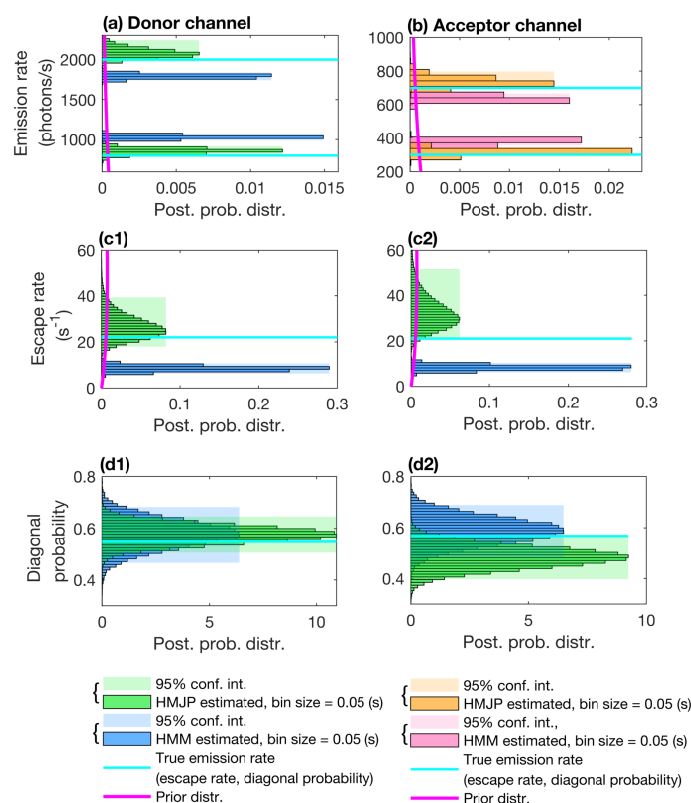


Fig. 2. HMJP and HMM photon emission rate, escape rate and transition probability estimates for fast switching kinetics in simulated measurements. Here, we provide posterior photon emission rate, escape rate and transition probability estimates obtained with HMJP and HMM when the switching rate is faster than the data acquisition rate $1/\Delta t = 20$ (1/s). We expect HMMs to perform poorly in estimating the true photon emission rates, escape rates and transition probabilities when the system switching is fast. In this figure's panels (a)-(b), we superposed the posterior distributions over photon emission rates for HMJP (green for donor, orange for acceptor) and HMM (blue for donor, pink for acceptor) along with their 95% confidence intervals and the true photon emission rates (dashed cyan line). Next, in panels (c1)-(c2) and (d1)-(d2), we superposed the posterior distributions over escape rates and transition probabilities (green for HMJP and blue for HMM) along with their 95% confidence intervals, true escape rates and true transition probabilities (dashed cyan lines), respectively. In this figure's panels (c1)-(c2) correspond to the posterior distribution of escape rates that are labeled with $\lambda_{\sigma_k \rightarrow \sigma_{k'}}$ for all $k, k' = 1, 2$ with $k \neq k'$. Panels (d1)-(d2) correspond to the transition probabilities labeled as $P_{\sigma_k \rightarrow \sigma_k}$ for all $k = 1, 2$. Here, simulated measurements are generated with the same parameters as those provided in Fig. 1 panels (b1)-(b2).

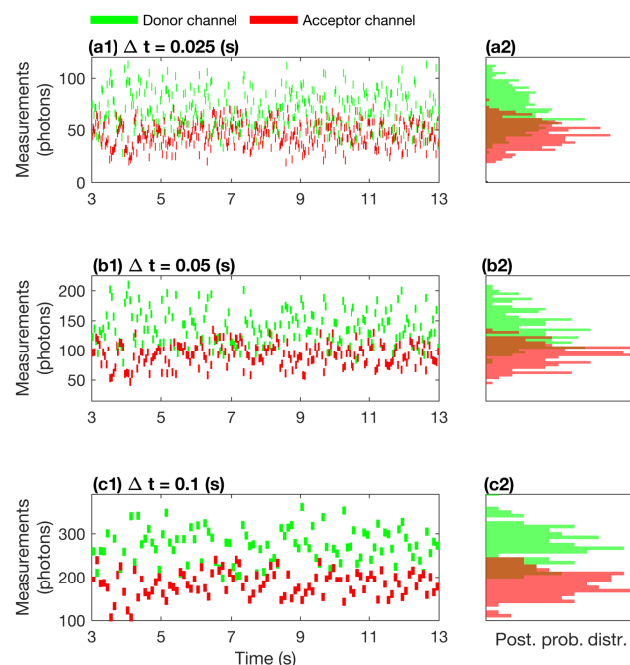


Fig. 3. An illustration of experimental smFRET measurements for data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. In panels (a1), (b1) and (c1), we provide the measurements for the same smFRET experiment coinciding with the data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we assume that the measurements are acquired by detectors with fixed exposure periods coinciding with the data acquisition periods $\tau = 0.025, 0.05$ s and $\tau = 0.1$ s in donor (green) and acceptor (red) channels. Panels (a2), (b2) and (c2) are especially interesting to our analysis. Here, we don't observe well separated histograms due to fast switching kinetics of the molecule as shown in Fig. 1 panel (b2) for simulated experiments.

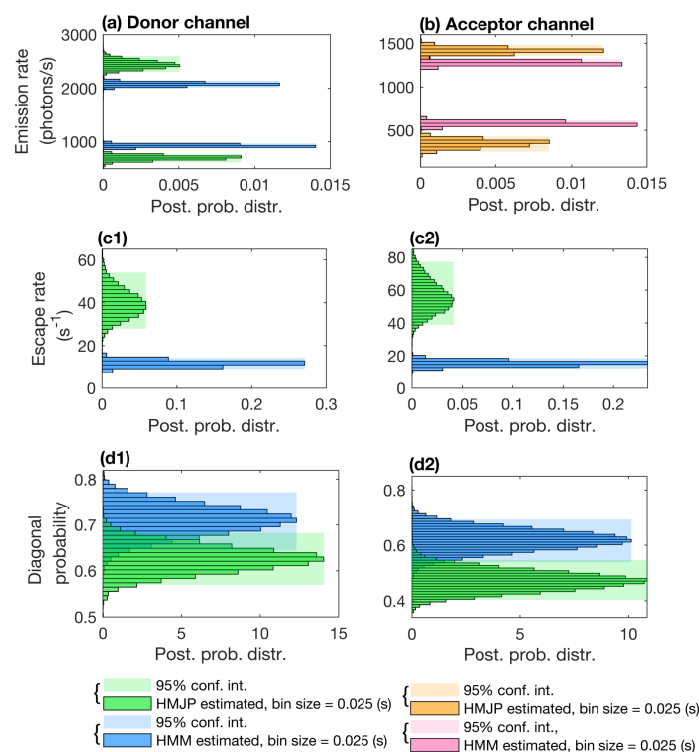


Fig. 4. HMJP and HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025$ s. Here, we provide posterior photon emission rate (panels (a)-(b)), escape rate (panels (c1)-(c2)) and transition probability (panels (d1)-(d2)) estimates obtained with HMJP and HMM for the experimental data shown in Fig. 3 panel (a1). Here, we follow a similar color convention to that of Fig. 2. However, unlike Fig. 3, ground truth information is not available for the photon emission rates, escape rates and transition probabilities as the analysis is carried out for experimental data.

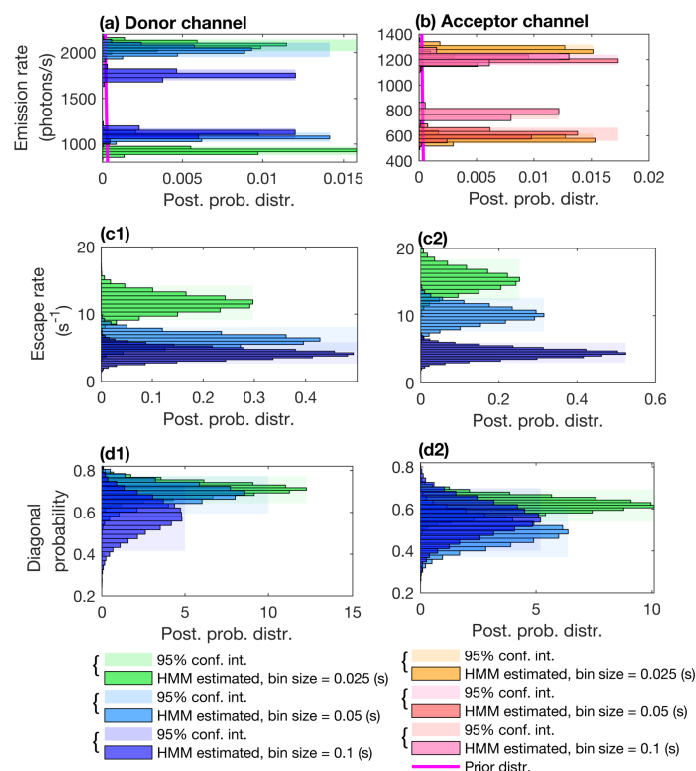


Fig. 5. HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate, escape rate and diagonal transition probability estimates obtained with HMM for the measurements shown in Fig. 3 panels (a1), (b1) and (c1). We expect HMM posterior estimates associated with 3 different exposure periods not to be consistent as the switching kinetics approach the data acquisition rate. In panels (a)-(b), we superposed the posterior distributions for the measurements associated with exposure periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s over photon emission rates for HMM along with their 95% confidence intervals. Next, in panels (c1)-(c2), we superposed the posterior distributions over escape rates along with their 95% confidence intervals. Finally, in panels (d1)-(d2), we superposed the posterior distributions over transition probabilities and their 95% confidence intervals. Posterior distributions over photon emission rate, escape rate and diagonal transition probability estimates for the integration periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s are demonstrated in green, blue and purple, respectively.

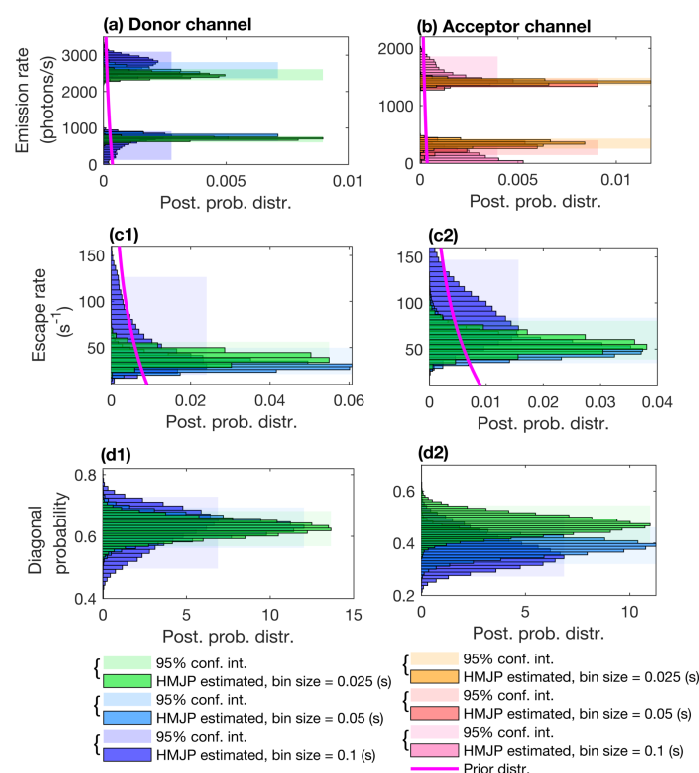


Fig. 6. HMJP photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate, escape rate and diagonal transition probability estimates obtained with HMJP for the measurements shown in Fig. 3 panels (a1), (b1) and (c1). We expect HMJP posterior estimates associated with 3 different exposure periods to be consistent across examples even if the switching kinetics exceed data acquisition rate. In panels (a)-(b), we superposed the posterior distributions for the measurements associated with exposure periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s over photon emission rates for the HMJP along with their 95% confidence intervals. Next, in panels (c1)-(c2), we superposed the posterior distributions over escape rates along with their 95% confidence intervals. Finally, in panels (d1)-(d2), we superposed the posterior distributions over transition probabilities and their 95% confidence intervals. Here, we follow a similar color convention to that of Fig. 5.

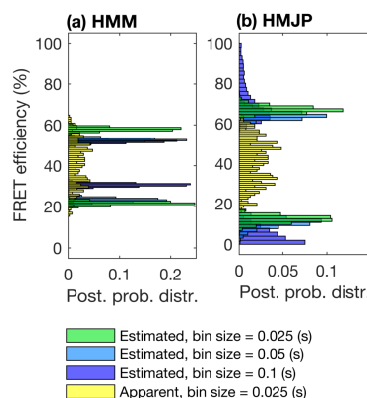


Fig. 7. HMM and HMJP FRET efficiency estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior FRET efficiency estimates obtained with HMM and HMJP for the measurements shown in Fig. 3 panels (a1), (b1) and (c1). We expect inconsistent posterior FRET efficiency estimates for HMMs associated with 3 different exposure periods; see panel (a). On the other hand, we expect HMJPs to provide consistent posterior estimates over FRET efficiencies for 3 different exposure periods; see panel (b). Here, in panels (a)-(b), we superposed the posterior distributions over FRET efficiencies along with the apparent FRET efficiency (yellow). In this figure, we follow a similar color convention to that of Fig. 5.

We expect HMJPs to perform better than HMMs as the simulated data are generated with switching rates 5% faster than the data acquisition rate (and thus many transitions occur during the data acquisition time). In this regime of switching rates, the approximation in Eq. (13) required by HMMs fails.

Given the measurements, we estimate the posterior distribution over the quantities of interest including the trajectory, $\mathcal{T}(\cdot)$, initial transition probability matrix, $\bar{\rho}$, transition probabilities $\bar{\pi}$ (and \bar{P} in HMM), photon emission rates, $\bar{\mu}^D, \bar{\mu}^A$ and escape rates, $\bar{\lambda}$. To achieve these estimates, we use HMJP and HMM samplers as introduced in Section 2.2.3 to generate pseudorandom numbers from the posterior distributions $\mathbb{P}(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A, (\bar{S}, \bar{D}, M) | \mathbf{w}^D, \mathbf{w}^A)$ and $\mathbb{P}(\bar{\rho}, \bar{P}, \bar{\mu}, \bar{c} | \mathbf{w}^D, \mathbf{w}^A)$, respectively.

We first show emission and escape rates estimates in Fig. 2. In Fig. 2 panels (a)-(b), (c1)-(c2), (d1)-(d2), we observe that the ground truths for photon emission rates, escape rates and transition probabilities lie within the 95% credible intervals for the corresponding HMJP estimates. By contrast, the HMM performs poorly due to the failure of the approximation in Eq. (13) resulting in unsatisfactory photon emission rate and escape rate estimates.

In particular, from Fig. 2 panels (a)-(b), we see that the HMM overestimates (by about 12% namely HMM provides estimates that are approximately 1.12 times the ground truth values for $\mu_{\sigma_1}^D$ and $\mu_{\sigma_1}^A$) $\mu_{\sigma_1}^D$ and $\mu_{\sigma_1}^A$ and underestimates (about 5%) $\mu_{\sigma_2}^D$ and $\mu_{\sigma_2}^A$. The failure of the HMM is more pronounced when looking at λ_{σ_1} and λ_{σ_2} . For example, in Fig. 2 panels (c1)-(c2), the HMM provides narrow distributions over escape rates by grossly underestimating both λ_{σ_1} and λ_{σ_2} (by about 50%). On the other hand, in Fig. 2 panels (c1)-(c2), we see that HMJP posterior distribution mode coincides with the ground truth.

The HMM's inability to capture fast kinetics is also reflected by its wide posterior distributions over transition probabilities $P_{\sigma_1 \rightarrow \sigma_1}, P_{\sigma_2 \rightarrow \sigma_2}$ as shown in Fig. 2 panels (d1)-(d2).

In Supporting Material (A) Fig. S1, we provide the analysis on slow switching kinetics for the HMJP and HMM estimate comparisons as in Fig. 2 where both HMJP and HMM perform well.

Subsequently, in Supporting Material (A) Fig. S2, we provide the robustness of HMJP posterior estimates over photon emission rates, escape rates and transition probabilities for the data set provided in Fig. 1 panel (b1). We observe that when the kinetic rates are more than 4 times the data acquisition rate then the HMJP starts failing in providing consistent posterior estimates over photon emission rates, escape rates and transition probabilities.

3.2 Experimental Data Analysis

We now move onto the analysis of experimental data sets.

3.2.1 Analysis Overview

Having benchmarked our HMJP on the simulated smFRET data set, we now move on to assessing the HMJP's performance on experimental smFRET data provided in Fig. 3, for Holliday junctions (HJ) as described in sample preparation. We deliberately start from single photon arrival data sets. Such data sets give us the flexibility to bin data as we please in order to mimic experimental data collected over broad range binning conditions as is the case in (5, 6, 21–27). In particular, we bin using 3 data acquisition periods: $\Delta t = 0.025$ s (with results shown in Figs. 4 to 7), $\Delta t = 0.05$ s (with results shown in Figs. 4 to 7) and $\Delta t = 0.1$ s (with results shown in Figs. 4 to 7). We pick these bin sizes for a particular reason. We estimate for the smallest bins the rates $\sim 35 - 45$ (1/s). Then, we consider data acquisition rates that are twice and four times bigger. We also set the exposure period, τ to 100% of the Δt namely $\tau = 0.025, 0.05$ s and $\tau = 0.1$ s. As we will see, the HMJP returns values for the rates that are consistent across bin sizes as would be expected for a method that can learn transition rates slower or faster than the data acquisition rate. By contrast, the failure of HMM becomes apparent due to the inconsistency of the estimates for the switching kinetics as a function of data bin size. The HJs we use have two well characterized conformational states (3, 80). Here, our goal is to estimate the HJ emission and switching rates in addition to the trajectory of HJs simultaneously.

In Fig. 3 panels (a1), (b1), (c1), we show the data under all 3 different bin sizes. From the data, we estimate the posterior distribution over the trajectory, $\mathcal{T}(\cdot)$, initial transition probability matrix, $\bar{\rho}$, transition probabilities $\bar{\pi}$ (and \bar{P} in HMM), photon emission rates, $\bar{\mu}^D, \bar{\mu}^A$ and escape rates, $\bar{\lambda}$. In Fig. 4, we show a direct comparison of the HMJP and HMM under the smallest bin size, where we anticipate both HMM and HMJP to be consistent with one another. As we move to larger bin sizes, the HMJP remains consistent with the result that we obtain for the smallest bin but the HMM eventually shows a lack of consistency as the bin size increases. This is apparent from Fig. 5, where we look at increasing bin size for HMM and in Fig. 6 where we look at increasing bin size for the HMJP. Finally in Fig. 7, we show the estimates over FRET efficiencies where we increase bin sizes for the case of both HMM and HMJP.

3.2.2 Comparison of HMJPs with HMMs on Experimental Data

We begin with an HMJP and HMM comparison for the smallest bin size, $\Delta t = 0.025$ s; see Fig. 3 panel (a1), on the photon emission rates, escape rates and transition probability estimates provided in Fig. 4. We point out that the simulated dataset provided in Fig. 1 mimic the dataset provided in Fig. 3. Even though, for experimental data Fig. 3 panels (a1), “ground truth” photon emission rates, escape rates and transition probabilities are not known, we suspect that similar observations provided in Fig. 2 hold for Fig. 4.

In Fig. 4 panels (a)-(b), we observe that HMM overestimates μ_{σ_1} and underestimates μ_{σ_2} compared to HMJP estimates for μ_{σ_1} and μ_{σ_2} . In addition, as can be seen in Fig. 4 panels (c1)-(c2), HMM provides narrow distributions over escape rates for both λ_{σ_1} and λ_{σ_2} . This contradicts with the escape rate posterior of the HMJP; see panels (c1)-(c2) of Fig. 4. We speculate that as HMM is unable to resolve fast switching kinetics,

missed kinetic transitions are reflected in terms of underestimation/overestimation of photon emission rates or escape rates; see Fig. 4 panels (a)-(b) and (c1)-(c2). In Fig. 4 panels (d1)-(d2), we observe that HMM and HMJP estimates for $P_{\sigma_1 \rightarrow \sigma_1}$ are overlapping although they have distinct posterior distributions for $P_{\sigma_2 \rightarrow \sigma_2}$.

Next, we show in Fig. 5 that the HMMs' photon emission rates, escape rates and transition probability estimates grow inconsistent for larger bin sizes, $\Delta t = 0.05, 0.1$ s (see Fig. 3 panels (b)-(c)). As a result of binning, we have multiple different apparent photon emission rates that appear. This could be overinterpreted as noise by the HMM; see panels (a)-(b). Indeed, HMMs interpret the various photon emission rates as an increased noise variance. As a related artifact, the HMM misses multiple transitions and therefore grossly underestimates escape rates.

As we see in panels (c1)-(c2), HMMs start underestimating escape rates as bin sizes become larger from $\Delta t = 0.025$ s to $\Delta t = 0.05, 0.1$ s. We emphasize that these escape estimates are approximate escape rate estimates obtained from the transition probability estimates for the HMM (see Eq. (13)) because a HMM does not directly report on kinetic rates. We observe in Fig. 5 panels (d1)-(d2) that the uncertainty around the HMM transition probability estimates increase and are furthermore not consistent as bin sizes increase. Subsequently, we repeat the exact same analysis as we did for the HMM for the HMJP. In particular there is an exact correspondence between Fig. 5 panels (a)-(b), (c1)-(c2), (d1)-(d2) and Fig. 6 panels (a)-(b), (c1)-(c2), (d1)-(d2). Details are provided in the captions however the important message is that HMJP escape rate estimates (see Fig. 6 panels (c1)-(c2)) has remained consistent across increased bin sizes.

Finally, in Fig. 7 panels (a), (b), we provide the posterior distributions over FRET efficiencies for the HMM and the HMJP. Here, we observe that HMMs provide inconsistent posterior FRET efficiency estimates for larger bin sizes; see panel (a). On the other hand HMJPs provide consistent FRET efficiency estimates; see panel (b).

In Supporting Material (A), we provide similar analysis for two other *experimental* smFRET data sets; see Fig. S3, Fig. S4. These analyzed data sets provided in Fig. S3, Fig. S4 were acquired under same experimental conditions as presented here. By analyzing these additional data sets, we justify the robustness of our method; see Figs. S4 to S10.

4 Discussion

Single molecule FRET data has the potential of revealing switching kinetics occurring on time scales at or even exceeding the measurement time scale (1, 3, 10, 32). This is especially helpful for smFRET data collected in a binned fashion (5, 6, 21–27). To achieve this, we have extended the HMM paradigm, previously used in the analysis of smFRET, by treating physical processes as they occur in nature (1, 3, 5–7, 10, 21–27, 32); that is, as continuous time processes using a Markov jump process framework (15, 47, 71–75, 81, 82).

The framework that we present here can treat multiple different data collection regimes. For example, it can treat different exposure periods, it can also treat different measurement models and variable background levels. Our framework reduces to existing frameworks, such as the HMM, in specific limits. For example, when the exposure period (τ) is as long as the data acquisition period (Δt), and the only measurements available are those precisely at the data acquisition time, then our model reduces to the HMM. Similarly, our method is more general than the H²MM (3) a previously developed method that treats the continuously evolving molecule trajectory by discretizing this process on a finer time scale than the data acquisition period. In the infinitely small discretization of this grid, the H²MM recovers our HMJP (3).

Finally, our method for smFRET data analysis goes beyond existing methods such as the H²MMs (3) or others (1) by not only making point estimates for the molecular trajectory, switching kinetics and photon emission rates. Rather, the Bayesian paradigm (54, 76) within which we operate allows us to compute full posterior distributions

over these estimated quantities.

As a final note, our framework might also be extended in multiple different ways. Trivially, we can change the detector model of Eq. (5) and Eq. (6) to incorporate different kinds of detector models such as for sCMOS cameras used in (7, 62, 65) this is equivalent to a change of Eq. (5) and Eq. (6). In addition, we can extend our framework to account for photophysical states of the fluorophores as well as accounting for spectral cross bleeding between donor and acceptor channels (7) by changing the measurement model such that the photon emission rates are affected by the photophysical states of the fluorophores. A non trivial extension of our work is to consider a single photon arrival framework for smFRET as opposed to dealing with binned data (1, 3). Additionally, it would be possible to leverage the strengths of Bayesian nonparametrics to learn the number of conformational states within the HMJP (7, 30, 31, 35, 37, 49, 77, 78, 83). Overcoming these difficulties is the focus of future work.

Acknowledgements

SP acknowledges support from NSF CAREER grant MCB-1719537 and NIH NIGMS (R01GM134426). TT acknowledges support from JSPS KAKENHI Grant Number JP19F19340. ASU cluster AGAVE and Saguaro are the main computational resources utilized in this study.

Author contributions

ZK analyzed data and developed analysis tools; ZK and IS developed computational tools; WH, KI, TT contributed experimental data; ZK, IS, SP conceived research; SP oversaw all aspects of the projects.

Author declaration

There is no competing interests.

Supporting Material Datasets

Data sets are supplied in .txt format.

1. I. V. Gopich and A. Szabo, "Decoding the pattern of photon colors in single-molecule fret," *The Journal of Physical Chemistry B*, vol. 113, pp. 10965–10973, 2009.
2. I. V. Gopich and A. Szabo, "Fret efficiency distributions of multistate single molecules," *The Journal of Physical Chemistry B*, vol. 114, pp. 15221–15226, 2010.
3. M. Pirchi, R. Tsukanov, R. Khamis, T. E. Tomov, Y. Berger, D. C. Khara, H. Volkov, G. Haran, and E. Nir, "Photon-by-photon hidden markov model analysis for microsecond single-molecule fret kinetics," *The Journal of Physical Chemistry B*, vol. 120, pp. 13065–13075, 2016.
4. L. P. Watkins and H. Yang, "Detection of intensity change points in time-resolved single-molecule measurements," *The Journal of Physical Chemistry B*, vol. 109, pp. 617–628, 2005.
5. M. Hajdziona and A. Molski, "Estimation of single-molecule blinking parameters using photon counting histogram," *Chemical Physics Letters*, vol. 470, pp. 363–366, 2009.
6. B. Schuler, "Single-molecule fluorescence spectroscopy of protein folding," *ChemPhysChem*, vol. 6, pp. 1206–1220, 2005.
7. I. Sgouralis, S. Madaan, F. Djutanta, R. Kha, R. F. Hariadi, and S. Pressé, "A bayesian nonparametric approach to single molecule förster resonance energy transfer," *The Journal of Physical Chemistry B*, vol. 123, pp. 675–688, 2018.
8. J. Guan, C. Jia, Y. Li, Z. Liu, J. Wang, Z. Yang, C. Gu, D. Su, K. N. Houk, D. Zhang, *et al.*, "Direct single-molecule dynamic detection of chemical reactions," *Science advances*, vol. 4, p. eaar2177, 2018.
9. R. B. Sekar and A. Periasamy, "Fluorescence resonance energy transfer (fret) microscopy imaging of live cell protein localizations,"

- The Journal of cell biology*, vol. 160, pp. 629–633, 2003.
10. K. R. Haas, H. Yang, and J.-W. Chu, "Expectation-maximization of the potential of mean force and diffusion coefficient in langevin dynamics from single molecule fret data photon by photon," *The Journal of Physical Chemistry B*, vol. 117, pp. 15591–15605, 2013.
11. T. Ha and P. Tinnefeld, "Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging," *Annual review of physical chemistry*, vol. 63, 2012.
12. M. Andrec, R. M. Levy, and D. S. Talaga, "Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden markov models," *The Journal of Physical Chemistry A*, vol. 107, pp. 7454–7464, 2003.
13. B. G. Keller, A. Kobitski, A. Jäschke, G. U. Nienhaus, and F. Noé, "Complex rna folding kinetics revealed by single-molecule fret and hidden markov models," *Journal of the American Chemical Society*, vol. 136, pp. 4534–4543, 2014.
14. T. Forster, "Delocalization excitation and excitation transfer," *Modern quantum chemistry*, 1965.
15. R. Roy, S. Hohng, and T. Ha, "A practical guide to single-molecule fret," *Nature methods*, vol. 5, pp. 507–516, 2008.
16. T. W. Gadella, *FRET and FLIM techniques*. 2011.
17. A. Periasamy and R. Day, *Molecular imaging: FRET microscopy and spectroscopy*. 2011.
18. D. C. Harris, *Quantitative chemical analysis*. 2010.
19. M. Tavakoli, S. Jazani, I. Sgouralis, O. M. Shafraz, S. Sivasankar, B. Donaphon, M. Levitus, and S. Pressé, "Pitching single-focus confocal data analysis one photon at a time with bayesian nonparametrics," *Phys. Rev. X*, vol. 10, p. 011021, 2020.
20. I. V. Gopich and A. Szabo, "Single-macromolecule fluorescence resonance energy transfer and free-energy profiles," *The Journal of Physical Chemistry B*, vol. 107, pp. 5058–5063, 2003.
21. M. Antonik, S. Felekyan, A. Gaiduk, and C. A. Seidel, "Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis," *The Journal of Physical Chemistry B*, vol. 110, pp. 6970–6978, 2006.
22. E. Nir, X. Michalet, K. M. Hamadani, T. A. Laurence, D. Neuhauser, Y. Kovchegov, and S. Weiss, "Shot-noise limited single-molecule fret histograms: comparison between theory and experiments," *The Journal of Physical Chemistry B*, vol. 110, pp. 22103–22124, 2006.
23. L. P. Watkins, H. Chang, and H. Yang, "Quantitative single-molecule conformational distributions: a case study with poly-(l-proline)," *The Journal of Physical Chemistry A*, vol. 110, pp. 5191–5203, 2006.
24. K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, and W. A. Eaton, "Characterizing the unfolded states of proteins using single-molecule fret spectroscopy and molecular simulations," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 1528–1533, 2007.
25. S. Kalinin, A. Valeri, M. Antonik, S. Felekyan, and C. A. Seidel, "Detection of structural dynamics by fret: a photon distribution and fluorescence lifetime analysis of systems with multiple states," *The Journal of Physical Chemistry B*, vol. 114, pp. 7983–7995, 2010.
26. H. S. Chung, I. V. Gopich, K. McHale, T. Cellmer, J. M. Louis, and W. A. Eaton, "Extracting rate coefficients from single-molecule photon trajectories and fret efficiency histograms for a fast-folding protein," *The journal of physical chemistry A*, vol. 115, pp. 3642–3656, 2011.
27. B. Schuler, "Single-molecule fret of protein structure and dynamics-a primer," *Journal of nanobiotechnology*, vol. 11, p. S2, 2013.
28. I. Gopich and A. Szabo, "Theory of photon statistics in single-molecule förster resonance energy transfer," *The Journal of chemical physics*, vol. 122, p. 014707, 2005.
29. I. V. Gopich and A. Szabo, "Single-molecule fret with diffusion and conformational dynamics," *The Journal of Physical Chemistry B*, vol. 111, pp. 12925–12932, 2007.
30. I. Sgouralis and S. Pressé, "Icon: an adaptation of infinite hmms for time traces with drift," *Biophysical journal*, vol. 112, pp. 2117–2126, 2017.
31. M. Tavakoli, J. N. Taylor, C.-B. Li, T. Komatsuzaki, and S. Pressé, "Single molecule data analysis: An introduction," *arXiv preprint arXiv:1606.00403*, 2016.
32. M. Hajdziona and A. Molski, "Maximum likelihood-based analysis of single-molecule photon arrival trajectories," *The Journal of Chemical Physics*, vol. 134, p. 054112, 2011.
33. S. Pressé, J. Lee, and K. A. Dill, "Extracting conformational memory from single-molecule kinetic data," *The Journal of Physical Chemistry B*, vol. 117, pp. 495–502, 2013.
34. S. Pressé, J. Peterson, J. Lee, P. Elms, J. L. MacCallum, S. Marqusee, C. Bustamante, and K. Dill, "Single molecule conformational memory extraction: p5ab rna hairpin," *The Journal of Physical Chemistry B*, vol. 118, pp. 6597–6603, 2014.
35. I. Sgouralis and S. Pressé, "An introduction to infinite hmms for single-molecule data analysis," *Biophysical journal*, vol. 112, pp. 2021–2029, 2017.
36. S. Kilic, S. Felekyan, O. Doroshenko, I. Boichenko, M. Dimura, H. Vardanyan, L. C. Bryan, G. Arya, C. A. Seidel, and B. Fierz, "Single-molecule fret reveals multiscale chromatin dynamics modulated by hp1 α ," *Nature communications*, vol. 9, pp. 1–14,

- 2018.
37. I. Sgouralis, M. Whitmore, L. Lapidus, M. J. Comstock, and S. Pressé, "Single molecule force spectroscopy at high data acquisition: A bayesian nonparametric analysis," *The Journal of chemical physics*, vol. 148, p. 123320, 2018.
38. C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
39. L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, pp. 257–286, 1989.
40. L. Rabiner and B. Juang, "An introduction to hidden markov models," *ieee assp magazine*, vol. 3, pp. 4–16, 1986.
41. L. R. Rabiner, *Multirate digital signal processing*. 1996.
42. Z. Kilic and S. Pressé, "Bayesian nonparametric analysis of transcriptional processes," *Biophysical Journal*, vol. 116, p. 299a, 2019.
43. Z. Kilic, I. Sgouralis, and S. Pressé, "Rna polymerase dynamics and other single-molecule continuous time problems," *Biophysical Journal*, vol. 118, p. 544a, 2020.
44. V. Rao and Y. W. Teh, "Fast mcmc sampling for markov jump processes and extensions," *The Journal of Machine Learning Research*, vol. 14, pp. 3295–3320, 2013.
45. B. Zhang and V. Rao, "Efficient parameter sampling for markov jump processes," *Journal of Computational and Graphical Statistics*, vol. 0, pp. 1–18, 2020.
46. P. Metzner, E. Dittmer, T. Jahnke, and C. Schütte, "Generator estimation of markov jump processes," *Journal of Computational Physics*, vol. 227, pp. 353–375, 2007.
47. T. Burzykowski, J. P. Szubiakowski, and T. Ryden, "Statistical analysis of data from single molecule experiment," in *IV Workshop on Atomic and Molecular Physics*, vol. 5258, pp. 171–177, 2003.
48. A. Hobolth and E. A. Stone, "Simulation from endpoint-conditioned, continuous-time markov chains on a finite state space, with applications to molecular evolution," *The annals of applied statistics*, vol. 3, p. 1204, 2009.
49. J. H. Huggins, K. Narasimhan, A. Saeedi, and V. K. Mansinghka, "Jump-means: Small-variance asymptotics for markov jump processes," *arXiv preprint arXiv:1503.00332*, 2015.
50. N. G. Van Kampen, *Stochastic processes in physics and chemistry*, vol. 1. 1992.
51. C. H. L. Beentjes and R. E. Baker, "Uniformization techniques for stochastic simulation of chemical reaction networks," *The Journal of Chemical Physics*, vol. 150, p. 154107, 2019.
52. D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of computational physics*, vol. 22, pp. 403–434, 1976.
53. D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, vol. 81, pp. 2340–2361, 1977.
54. D. Sivia and J. Skilling, *Data analysis: a Bayesian tutorial*. 2006.
55. B. Buelens, P. Daas, J. Burger, M. Puts, and J. van den Brakel, *Selectivity of Big data*. 2014.
56. A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. 2002.
57. C. Robert and G. Casella, *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.
58. S. Kalinin, E. Sisamakias, S. W. Magennis, S. Felekyan, and C. A. Seidel, "On the origin of broadening of single-molecule fret efficiency distributions beyond shot noise limits," *The journal of physical chemistry B*, vol. 114, pp. 6197–6206, 2010.
59. K. I. Mortensen, L. S. Churchman, J. A. Spudich, and H. Flyvbjerg, "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nature methods*, vol. 7, p. 377, 2010.
60. F. Charrière, T. Colomb, F. Montfort, E. Cuche, P. Marquet, and C. Depeursinge, "Shot-noise influence on the reconstructed phase image signal-to-noise ratio in digital holographic microscopy," *Applied optics*, vol. 45, pp. 7667–7673, 2006.
61. M. Gross, P. Goy, and M. Al-Koussa, "Shot-noise detection of ultrasound-tagged photons in ultrasound-modulated optical imaging," *Optics letters*, vol. 28, pp. 2482–2484, 2003.
62. F. Huang, T. M. Hartwich, F. E. Rivera-Molina, Y. Lin, W. C. Duim, J. J. Long, P. D. Uchil, J. R. Myers, M. A. Baird, W. Mothes, *et al.*, "Video-rate nanoscopy using sccmos camera-specific single-molecule localization algorithms," *Nature methods*, vol. 10, pp. 653–658, 2013.
63. V. Krishnaswami, C. J. Van Noorden, E. M. Manders, and R. A. Hoebe, "Towards digital photon counting cameras for single-molecule optical nanoscopy," *Optical Nanoscopy*, vol. 3, p. 1, 2014.
64. Y. Lin, J. J. Long, F. Huang, W. C. Duim, S. Kirschbaum, Y. Zhang, L. K. Schroeder, A. A. Rebane, M. G. M. Velasco, A. Virrueta, *et al.*, "Quantifying and optimizing single-molecule switching nanoscopy at high speeds," *PloS one*, vol. 10, 2015.
65. M. F. Juetter, D. S. Terry, M. R. Wasserman, R. B. Altman, Z. Zhou, H. Zhao, and S. C. Blanchard, "Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale," *Nature methods*, vol. 13, pp. 341–344, 2016.
66. T. Yoshida, S. Alfaqaan, N. Sasaoka, and H. Imamura, "Application of fret-based biosensor "ateam" for visualization of atp levels in the mitochondrial matrix of living mammalian cells," in *Mitochondria*, pp. 231–243, 2017.
67. J. Lee, S. Park, and S. Hohng, "Accelerated fret-paint microscopy," *Molecular brain*, vol. 11, p. 70, 2018.
68. M. A. Little, *Machine Learning for Signal Processing: Data Science, Algorithms, and Computational Statistics*. 2019.

69. A. Lee, K. Tsekouras, C. Calderon, C. Bustamante, and S. Pressé, "Unraveling the thousand word picture: an introduction to super-resolution data analysis," *Chemical reviews*, vol. 117, pp. 7276–7330, 2017.
70. B. A. Hamilton, "The field guide to data science," 2015.
71. M. R. Wasserman, J. L. Alejo, R. B. Altman, and S. C. Blanchard, "Multiperspective smfret reveals rate-determining late intermediates of ribosomal translocation," *Nature structural & molecular biology*, vol. 23, p. 333, 2016.
72. M. Lu, X. Ma, L. R. Castillo-Menendez, J. Gorman, N. Alsahafi, U. Ermel, D. S. Terry, M. Chambers, D. Peng, B. Zhang, *et al.*, "Associating hiv-1 envelope glycoprotein structures with states on the virus observed by smfret," *Nature*, vol. 568, pp. 415–419, 2019.
73. A. Mazouchi, Z. Zhang, A. Bahram, G.-N. Gomes, H. Lin, J. Song, H. S. Chan, J. D. Forman-Kay, and C. C. Gradinaru, "Conformations of a metastable sh3 domain characterized by smfret and an excluded-volume polymer model," *Biophysical journal*, vol. 110, pp. 1510–1522, 2016.
74. Y. Zhu, L. He, Y. Liu, Y. Zhao, and X. C. Zhang, "smfret probing reveals substrate-dependent conformational dynamics of e. coli multidrug mdfa," *Biophysical journal*, vol. 116, pp. 2296–2303, 2019.
75. E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Chung, X. Michalet, and S. Weiss, "Toward dynamic structural biology: Two decades of single-molecule förster resonance energy transfer," *Science*, vol. 359, p. eaan1133, 2018.
76. E. T. Lee and J. Wang, *Statistical methods for survival data analysis*, vol. 476. 2003.
77. S. Jazani, I. Sgouralis, O. M. Shafraz, M. Levitus, S. Sivasankar, and S. Pressé, "An alternative framework for fluorescence correlation spectroscopy," *Nature communications*, vol. 10, pp. 1–10, 2019.
78. S. Jazani, I. Sgouralis, and S. Pressé, "A method for single molecule tracking using a conventional single-focus confocal setup," *The Journal of Chemical Physics*, vol. 150, p. 114108, 2019.
79. S. D. Chandross, A. C. Haagsma, Y. K. Lee, J.-H. Hwang, J.-M. Nam, and C. Joo, "Surface passivation for single-molecule protein studies," *JoVE (Journal of Visualized Experiments)*, p. e50549, 2014.
80. J.-Y. Kim, C. Kim, and N. K. Lee, "Real-time submillisecond single-molecule fret dynamics of freely diffusing molecules with liposome tethering," *Nature communications*, vol. 6, pp. 1–9, 2015.
81. A. Krishnamoorti, R. C. Cheng, V. Berka, and M. Maduke, "Clc conformational landscape as studied by smfret," *Biophysical Journal*, vol. 116, p. 555a, 2019.
82. S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. Lilley, and T. Ha, "Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction," *Science*, vol. 318, pp. 279–283, 2007.
83. A. Georgoulas, J. Hillston, and G. Sanguinetti, "Unbiased bayesian inference for population markov jump processes via random truncations," *Statistics and computing*, vol. 27, pp. 991–1002, 2017.
84. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. 2013.

A Supporting Material

A.1 Comparison of HMJP and HMM for Simulated Data with Slow Switching Kinetics

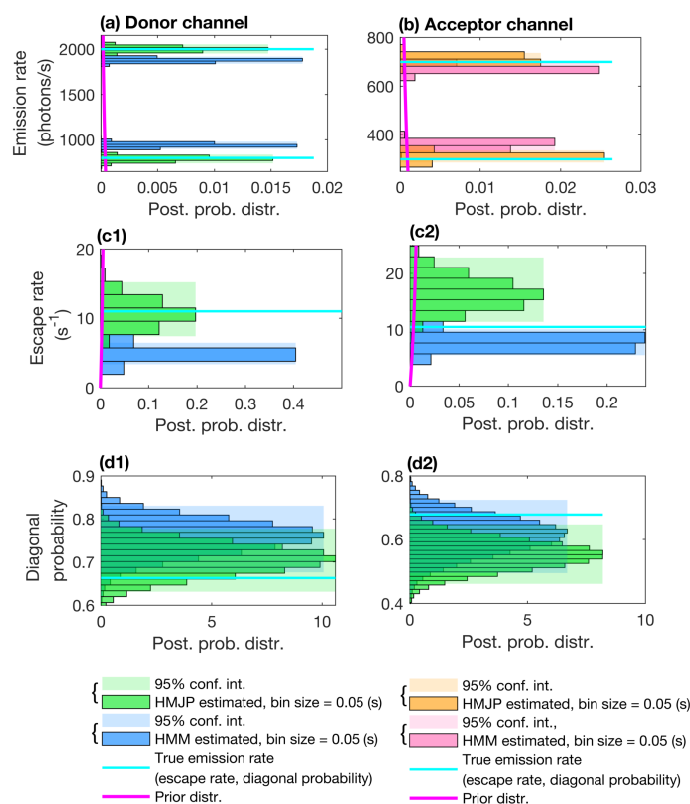


Fig. A.1. HMJP and HMM photon emission rate, escape rate and transition probability estimates for slow switching kinetics in simulated measurements. Here, we provide posteriors over the photon emission rate, escape rate and transition probability estimates obtained with HMJPs and HMMs when the switching rate is slower than the data acquisition rate $1/\Delta t = 20$ (1/s). We expect HMMs to perform poorly in estimating the true photon emission rates, escape rates and instead perform better in estimating transition probabilities when the system switching is slow. Here, we provide the posterior distributions with the same color convention as in Fig. 2. Simulated measurements used here are generated with the same parameters as those provided in Fig. 1 panels (a1)-(a2).

A.2 Performance of HMJP for Simulated Data with Fast Switching Kinetics over a Range of Bin Sizes

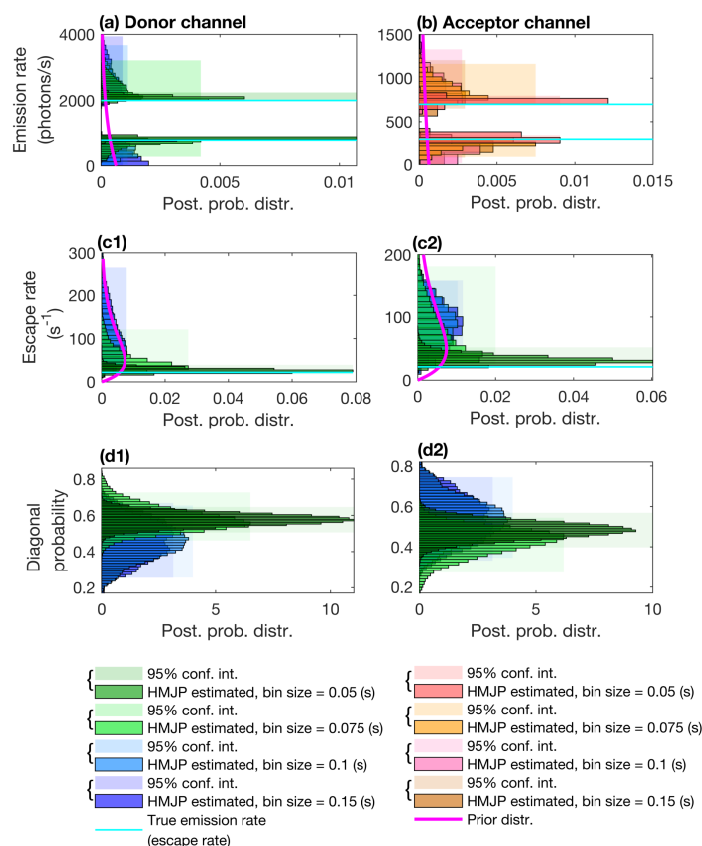


Fig. A.2. HMJP photon emission rate, escape rate and transition probability estimates for fast switching kinetics in simulated measurements. Here, we provide posterior photon emission rate, escape rate and diagonal transition probability estimates obtained with HMJP for the measurements shown in Fig. 1 panels (b1). Here, we expect HMJP posterior estimates to be consistent for the first 3 data acquisition periods ($\Delta t = 0.05, 0.075, 0.1$ s) and not consistent for $\Delta t = 0.15$ s. In this figure's each panel, we have the same content information as in Fig. A.1. Posterior distributions over photon emission rate, escape rate and diagonal transition probability estimates for the integration periods $\Delta t = 0.05, 0.075, 0.1$ s and $\Delta t = 0.15$ s are demonstrated in dark green, green, blue and purple, respectively.

A.3 HMJP Estimates for Additional Experimental smFRET Data

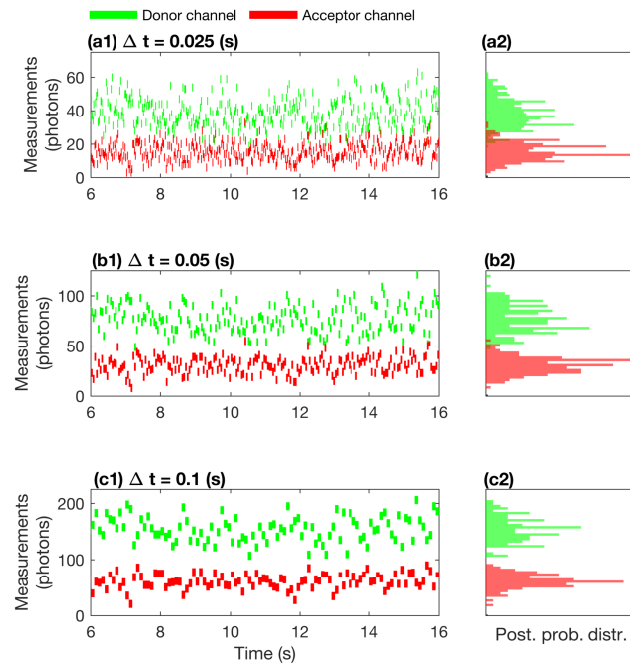


Fig. A.3. An illustration of experimental smFRET measurements for data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. In panels (a1), (b1) and (c1), we provide the measurements for the same smFRET experiment coinciding with the data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we assume that the measurements are acquired by detectors with fixed exposure periods coinciding with the data acquisition periods $\tau = 0.025, 0.05$ s and $\tau = 0.1$ s in donor (green) and acceptor (red) channels.

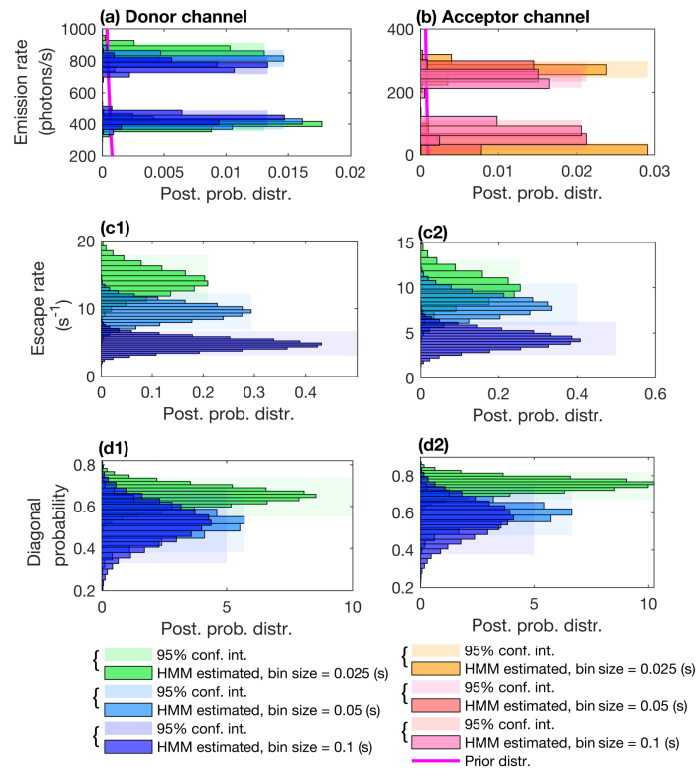


Fig. A.4. HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate (panels (a)-(b)), escape rate (panels (c1)-(c2)) and transition probability (panels (d1)-(d2)) estimates obtained with HMM for the experimental data shown in Fig. A.3 panels (a1),(b1) and (c1). Here, we follow a similar color convention as in Fig. 5.

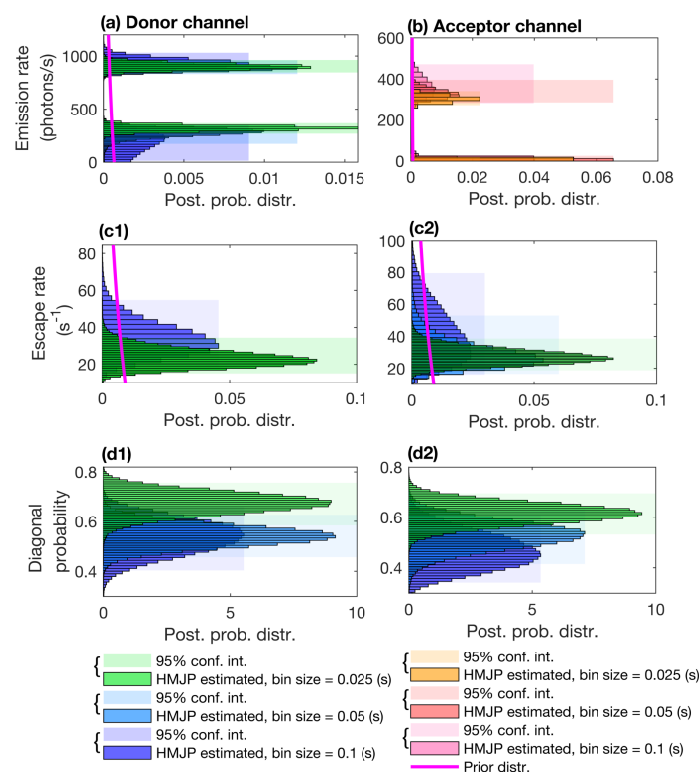


Fig. A.5. HMJP photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate (panels (a)-(b)), escape rate (panels (c1)-(c2)) and transition probability (panels (d1)-(d2)) estimates obtained with HMJP for the experimental data shown in Fig. A.3 panels (a1), (b1) and (c1). Here, we follow a similar color convention as in Fig. A.4.

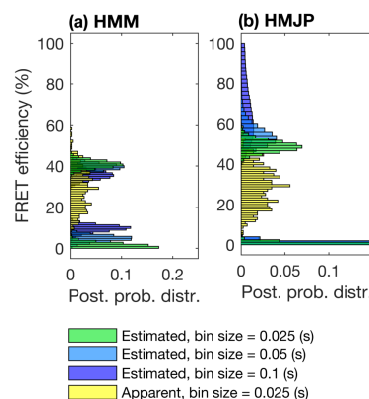


Fig. A.6. HMJP and HMM FRET efficiency estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate and FRET efficiency estimates obtained with HMJP for the measurements shown in Fig. A.3 panels (a1), (b1) and (c1). We expect inconsistent posterior FRET efficiency estimates for HMM associated with 3 different exposure periods; see panel (a). On the other hand, we expect HMJP to provide consistent posterior estimates over FRET efficiencies for 3 different exposure periods; see panel (b). Here, in panels (a)-(b), we superposed the posterior distributions over FRET efficiencies along with the apparent FRET efficiency (yellow). In this figure, we follow a similar color convention to that of Fig. 5.

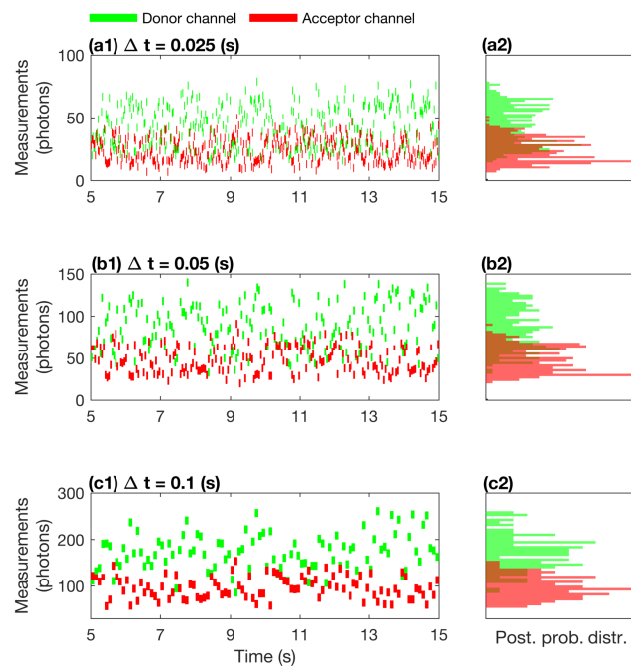


Fig. A.7. An illustration of experimental smFRET measurements for data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. In panels (a1), (b1) and (c1), we provide the measurements for the same smFRET experiment coinciding with the data acquisition periods $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we assume that the measurements are acquired by detectors with fixed exposure periods coinciding with the data acquisition periods $\tau = 0.025, 0.05$ s and $\tau = 0.1$ s in donor (green) and acceptor (red) channels.

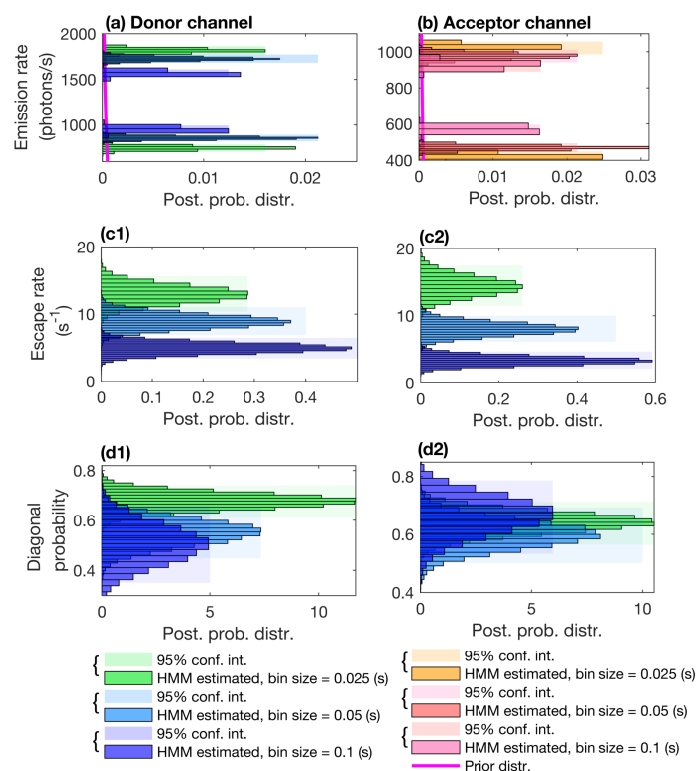


Fig. A.8. HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate (panels (a)-(b)), escape rate (panels (c1)-(c2)) and transition probability (panels (d1)-(d2)) estimates obtained with HMM for the experimental data shown in Fig. A.7 panels (a1),(b1) and (c1). Here, we follow a similar color convention as in Fig. 5.

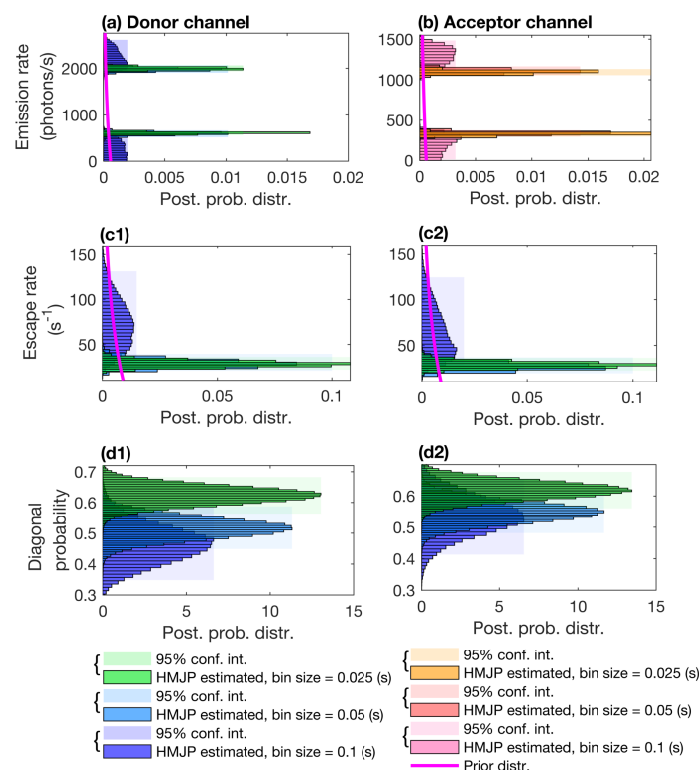


Fig. A.9. HMJP photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior emission rate (panels (a)-(b)), escape rate (panels (c1)-(c2)) and transition probability (panels (d1)-(d2)) estimates obtained with HMJP for the experimental data shown in Fig. A.7 panels (a1),(b1) and (c1). Here, we follow a similar color convention as in Fig. A.4.

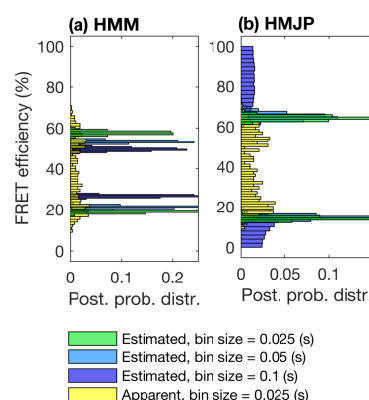


Fig. A.10. HMJP and HMM FRET efficiency estimates for experimental smFRET measurements with $\Delta t = 0.025, 0.05$ s and $\Delta t = 0.1$ s. Here, we provide posterior photon emission rate and FRET efficiency estimates obtained with HMJP for the measurements shown in Fig. A.7 panels (a1), (b1) and (c1). We expect inconsistent posterior FRET efficiency estimates for HMM associated with 3 different exposure periods; see panel (a). On the other hand, we expect HMJP to provide consistent posterior estimates over FRET efficiencies for 3 different exposure periods; see panel (b). Here, in panels (a)-(b), we superposed the posterior distributions over FRET efficiencies along with the apparent FRET efficiency (yellow). In this figure, we follow a similar color convention to that of Fig. 5.

A.4 Modeling Summary

A.4.1 Hidden Markov Jump Process Model

For $K = 2$, the full set of HMJP equations is

$$\bar{\rho} \sim \text{Dirichlet}\left(\frac{\alpha}{2}, \frac{\alpha}{2}\right) \quad (\text{A.1})$$

$$\lambda_{\sigma_1} \sim \text{Gamma}\left(\eta, \frac{b}{\eta}\right) \quad (\text{A.2})$$

$$\lambda_{\sigma_2} \sim \text{Gamma}\left(\eta, \frac{b}{\eta}\right) \quad (\text{A.3})$$

$$\mu_{\sigma_1}^D \sim \text{Gamma}\left(\phi^D, \frac{\psi^D}{\phi^D}\right) \quad (\text{A.4})$$

$$\mu_{\sigma_2}^D \sim \text{Gamma}\left(\phi^D, \frac{\psi^D}{\phi^D}\right) \quad (\text{A.5})$$

$$\mu_{\sigma_1}^A \sim \text{Gamma}\left(\phi^A, \frac{\psi^A}{\phi^A}\right) \quad (\text{A.6})$$

$$\mu_{\sigma_2}^A \sim \text{Gamma}\left(\phi^A, \frac{\psi^A}{\phi^A}\right) \quad (\text{A.7})$$

$$\bar{\pi}_{\sigma_1} \sim \text{Dirichlet}(0, A) \quad (\text{A.8})$$

$$\bar{\pi}_{\sigma_2} \sim \text{Dirichlet}(A, 0) \quad (\text{A.9})$$

$$s_0 | \bar{\rho} \sim \text{Categorical}(\bar{\rho}) \quad (\text{A.10})$$

$$d_m | s_m, \lambda_{\sigma_1}, \lambda_{\sigma_2} \sim \text{Exponential}(\lambda_{s_m}) \quad \text{for } m = 0, 1, 2, \dots, M-1 \quad (\text{A.11})$$

$$s_{m+1} | s_m, \bar{\pi}_{\sigma_1}, \bar{\pi}_{\sigma_2} \sim \text{Categorical}(\bar{\pi}_{s_m}) \quad \text{for } m = 0, 1, 2, \dots, M-1 \quad (\text{A.12})$$

$$w_n^D | \mathcal{T}(\cdot), \mu_{\text{back}}^D, \mu_{\sigma_1}^D, \mu_{\sigma_2}^D \sim \text{Poisson}\left(\mu_{\text{back}}^D \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^D\right) \quad \text{for } m = 1, 2, \dots, M \quad (\text{A.13})$$

$$w_n^A | \mathcal{T}(\cdot), \mu_{\text{back}}^A, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A \sim \text{Poisson}\left(\mu_{\text{back}}^A \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{T}(t)}^A\right) \quad \text{for } m = 1, 2, \dots, M \quad (\text{A.14})$$

where $\mathcal{T}(\cdot)$ is formulated as follows

$$\mathcal{T}(t) = \begin{cases} s_0 & \text{if } t_0 \leq t < t_0 + d_0 \\ s_1 & \text{if } t_0 + d_1 \leq t < t_0 + d_0 + d_1 \\ \vdots & \\ s_{M-1} & \text{if } t_0 + d_0 + \dots + d_{M-2} \leq t < t_0 + d_0 + d_1 + \dots + d_{M-1} \end{cases} \quad (\text{A.15})$$

with M determined based on the first time

$$t_0 + \sum_{m=0}^{M-1} d_m \geq t_N. \quad (\text{A.16})$$

A.4.2 Hidden Markov Model

For $K = 2$, the full set of HMM equations is

$$\bar{\rho} \sim \text{Dirichlet} \left(\frac{\alpha}{2}, \frac{\alpha}{2} \right) \quad (\text{A.17})$$

$$\bar{P}_{\sigma_1} \sim \text{Dirichlet} \left(\frac{A}{2}, \frac{A}{2} \right) \quad (\text{A.18})$$

$$\bar{P}_{\sigma_2} \sim \text{Dirichlet} \left(\frac{A}{2}, \frac{A}{2} \right) \quad (\text{A.19})$$

$$\mu_{\sigma_1}^D \sim \text{Gamma} \left(\phi^D, \frac{\psi^D}{\phi^D} \right) \quad (\text{A.20})$$

$$\mu_{\sigma_2}^D \sim \text{Gamma} \left(\phi^D, \frac{\psi^D}{\phi^D} \right) \quad (\text{A.21})$$

$$\mu_{\sigma_1}^A \sim \text{Gamma} \left(\phi^A, \frac{\psi^A}{\phi^A} \right) \quad (\text{A.22})$$

$$\mu_{\sigma_2}^A \sim \text{Gamma} \left(\phi^A, \frac{\psi^A}{\phi^A} \right) \quad (\text{A.23})$$

$$c_0 | \bar{\rho} \sim \text{Categorical}(\bar{\rho}) \quad (\text{A.24})$$

$$c_n | c_{n-1}, \bar{P}_{\sigma_1}, \bar{P}_{\sigma_2} \sim \text{Categorical}(\bar{P}_{c_{n-1}}) \quad \text{for } n = 2, \dots, N \quad (\text{A.25})$$

$$w_n^D | c_n, \mu_{\sigma_1}^D, \mu_{\sigma_2}^D \sim \text{Poisson}(\mu_{\text{back}}^D \tau + \mu_{c_n}^D \tau) \quad \text{for } m = 1, 2, \dots, M \quad (\text{A.26})$$

$$w_n^A | c_n, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A \sim \text{Poisson}(\mu_{\text{back}}^A \tau + \mu_{c_n}^A \tau) \quad \text{for } m = 1, 2, \dots, M \quad (\text{A.27})$$

A.4.3 Overview of the Sampling Updates

In order to produce samples from the full posterior distributions $\mathbb{P}(\bar{\rho}, \bar{P}, \bar{\mu}^D, \bar{\mu}^A, \vec{c} | \mathbf{w}^D, \mathbf{w}^A)$ for the HMM and $\mathbb{P}(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \bar{\mu}^D, \bar{\mu}^A, (\vec{S}, \vec{D}, M) | \mathbf{w}^D, \mathbf{w}^A)$ for the HMJP, we use Gibbs sampling (7, 30, 31, 35, 37, 44, 77, 78). Specifically, we repeat the following steps:

1. Update \vec{c} for the HMM or (\vec{S}, \vec{D}, M) for the HMJP;
2. Update transition probabilities, that is \bar{P}_{σ_k} for the HMM or $\bar{\pi}_{\sigma_k}$ and λ_{σ_k} for the HMJP;
3. Update the initial probability vector $\bar{\rho}$ for both the HMM and the HMJP;
4. Update photon emission rates $\bar{\mu}^D, \bar{\mu}^A$ for both the HMM and the HMJP.

Here, we present the equation summaries for sampling photon emission rates.

Sampling photon emission rates for HMJP In order to produce proposals for the Metropolis Hastings algorithm, we use Hamiltonian Monte Carlo (HMC). This method is very fruitful as Hamiltonian dynamics preserve volume in (q, p) space therefore we can use the trajectories to define complex mappings. Here the target distribution we have is the following

$$\mathbb{P}(\bar{\mu}^A, \bar{\mu}^D | \mathbf{w}^D, \mathbf{w}^A, (\vec{S}, \vec{D}, M)). \quad (\text{A.28})$$

There are two main variables, these are the position that is labeled with q and the momentum, p . Here we collect the parameter of interest to be sampled in q such that

$$q = (\bar{\mu}^D, \bar{\mu}^A) = (\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A). \quad (\text{A.29})$$

We have the following potential

$$U(q) = -\log \left(\mathbb{P} \left(\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A \mid \mathbf{w}^D, \mathbf{w}^A, (\vec{S}, \vec{D}, M) \right) \right) \quad (\text{A.30})$$

$$= -\log \left(\mathbb{P} \left(\mathbf{w}^D \mid \mu_{\sigma_1}^D, \mu_{\sigma_2}^D, (\vec{S}, \vec{D}, M) \right) \right) - \log \left(\mathbb{P} \left(\mu_{\sigma_1}^D, \mu_{\sigma_2}^D \right) \right) \quad (\text{A.31})$$

$$- \log \left(\mathbb{P} \left(\mathbf{w}^A \mid \mu_{\sigma_1}^A, \mu_{\sigma_2}^A, (\vec{S}, \vec{D}, M) \right) \right) - \log \left(\mathbb{P} \left(\mu_{\sigma_1}^A, \mu_{\sigma_2}^A \right) \right) \quad (\text{A.32})$$

$$= -\sum_{n=1}^N \left(w_n^D \log \left(\mu_{\text{back}}^D \tau + \tau \mu_{\sigma_1}^D \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \tau \mu_{\sigma_2}^D \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F \right) \right) \quad (\text{A.33})$$

$$+ \sum_{n=1}^N \left(\mu_{\text{back}}^D \tau + \tau \mu_{\sigma_1}^D \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \tau \mu_{\sigma_2}^D \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F \right) \quad (\text{A.34})$$

$$+ \sum_{j=1}^2 \left((1 - \phi^D) \log \left(\mu_{\sigma_j}^D \right) + \frac{\phi^D}{\psi^D} \mu_{\sigma_j}^D \right) \quad (\text{A.35})$$

$$- \sum_{n=1}^N \left(w_n^A \log \left(\mu_{\text{back}}^A \tau + \tau \mu_{\sigma_1}^A \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \tau \mu_{\sigma_2}^A \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F \right) \right) \quad (\text{A.36})$$

$$+ \sum_{n=1}^N \left(\mu_{\text{back}}^A \tau + \tau \mu_{\sigma_1}^A \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \tau \mu_{\sigma_2}^A \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F \right) \quad (\text{A.37})$$

$$+ \sum_{j=1}^2 \left((1 - \phi^A) \log \left(\mu_{\sigma_j}^A \right) + \frac{\phi^A}{\psi^A} \mu_{\sigma_j}^A \right). \quad (\text{A.38})$$

We also need the gradient of $U(q)$ namely, $\frac{\partial U}{\partial \mu_{\sigma_1}^D}, \frac{\partial U}{\partial \mu_{\sigma_2}^D}, \frac{\partial U}{\partial \mu_{\sigma_1}^A}, \frac{\partial U}{\partial \mu_{\sigma_2}^A}$.

$$\begin{aligned} \bullet \frac{\partial U}{\partial \mu_{\sigma_1}^D} &= -\sum_{n=1}^N \left(\frac{w_n \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F}{\mu_{\text{back}}^D + \mu_{\sigma_1}^D \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \mu_{\sigma_2}^D \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F} - \tau \sum_{n=1}^{N_n} ((\chi_{\sigma_1}(s_k)) t_k^F) \right) + \frac{(1-\phi^D)}{\mu_{\sigma_1}^D} + \frac{\phi^D}{\psi^D} \\ \bullet \frac{\partial U}{\partial \mu_{\sigma_2}^D} &= -\sum_{n=1}^N \left(\frac{w_n \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F}{\mu_{\text{back}}^D + \mu_{\sigma_1}^D \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \mu_{\sigma_2}^D \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F} - \tau \sum_{n=1}^{N_n} ((\chi_{\sigma_2}(s_k)) t_k^F) \right) + \frac{(1-\phi^D)}{\mu_{\sigma_2}^D} + \frac{\phi^D}{\psi^D} \\ \bullet \frac{\partial U}{\partial \mu_{\sigma_1}^A} &= -\sum_{n=1}^N \left(\frac{w_n \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F}{\mu_{\text{back}}^A + \mu_{\sigma_1}^A \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \mu_{\sigma_2}^A \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F} - \tau \sum_{n=1}^{N_n} ((\chi_{\sigma_1}(s_k)) t_k^F) \right) + \frac{(1-\phi^A)}{\mu_{\sigma_1}^A} + \frac{\phi^A}{\psi^A} \\ \bullet \frac{\partial U}{\partial \mu_{\sigma_2}^A} &= -\sum_{n=1}^N \left(\frac{w_n \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F}{\mu_{\text{back}}^A + \mu_{\sigma_1}^A \sum_{k=1}^{N_n} (\chi_{\sigma_1}(s_k)) t_k^F + \mu_{\sigma_2}^A \sum_{k=1}^{N_n} (\chi_{\sigma_2}(s_k)) t_k^F} - \tau \sum_{n=1}^{N_n} ((\chi_{\sigma_2}(s_k)) t_k^F) \right) + \frac{(1-\phi^A)}{\mu_{\sigma_2}^A} + \frac{\phi^A}{\psi^A}. \end{aligned}$$

Sampling photon emission rates for HMM Similarly with above, we use HMC on the target distribution

$$\mathbb{P}(\bar{\mu}^A, \bar{\mu}^D | \mathbf{w}^D, \mathbf{w}^A, \vec{c}). \quad (\text{A.39})$$

There are two main variables, these are the position that is labeled with q and the momentum, p . Here we collect the parameter of interest to be sampled in q such that

$$q = (\bar{\mu}^D, \bar{\mu}^A) = (\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A). \quad (\text{A.40})$$

We have the following potential

$$U(q) = -\log(\mathbb{P}(\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \mu_{\sigma_1}^A, \mu_{\sigma_2}^A | \mathbf{w}^D, \mathbf{w}^A, \vec{c})) \quad (\text{A.41})$$

$$= -\log(\mathbb{P}(\mathbf{w}^D | \mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \vec{c})) - \log(\mathbb{P}(\mu_{\sigma_1}^D, \mu_{\sigma_2}^D)) \quad (\text{A.42})$$

$$- \log(\mathbb{P}(\mathbf{w}^A | \mu_{\sigma_1}^A, \mu_{\sigma_2}^A, \vec{c})) - \log(\mathbb{P}(\mu_{\sigma_1}^A, \mu_{\sigma_2}^A)) \quad (\text{A.43})$$

$$= -\sum_{n=1}^N (w_n^D \log(\mu_{\text{back}}^D \tau + \tau \mu_{\sigma_1}^D (\chi_{\sigma_1}(c_n)) + \tau \mu_{\sigma_2}^D (\chi_{\sigma_2}(c_n)))) \quad (\text{A.44})$$

$$+ \sum_{n=1}^N (\mu_{\text{back}}^D \tau + \tau \mu_{\sigma_1}^D (\chi_{\sigma_1}(c_n)) t_k^F + \tau \mu_{\sigma_2}^D (\chi_{\sigma_2}(c_n))) \quad (\text{A.45})$$

$$+ \sum_{j=1}^2 \left((1 - \phi^D) \log(\mu_{\sigma_j}^D) + \frac{\phi^D}{\psi^D} \mu_{\sigma_j}^D \right) \quad (\text{A.46})$$

$$- \sum_{n=1}^N (w_n^A \log(\mu_{\text{back}}^A \tau + \tau \mu_{\sigma_1}^A (\chi_{\sigma_1}(c_n)) + \tau \mu_{\sigma_2}^A (\chi_{\sigma_2}(c_n)))) \quad (\text{A.47})$$

$$+ \sum_{n=1}^N (\mu_{\text{back}}^A \tau + \tau \mu_{\sigma_1}^A (\chi_{\sigma_1}(c_n)) t_k^F + \tau \mu_{\sigma_2}^A (\chi_{\sigma_2}(c_n))) \quad (\text{A.48})$$

$$+ \sum_{j=1}^2 \left((1 - \phi^A) \log(\mu_{\sigma_j}^A) + \frac{\phi^A}{\psi^A} \mu_{\sigma_j}^A \right). \quad (\text{A.49})$$

We also need the gradient of $U(q)$ namely, $\frac{\partial U}{\partial \mu_{\sigma_1}^D}, \frac{\partial U}{\partial \mu_{\sigma_2}^D}, \frac{\partial U}{\partial \mu_{\sigma_1}^A}, \frac{\partial U}{\partial \mu_{\sigma_2}^A}$.

- $\frac{\partial U}{\partial \mu_{\sigma_1}^D} = -\sum_{n=1}^N \left(\frac{w_n^D (\chi_{\sigma_1}(c_n))}{\mu_{\text{back}}^D + \mu_{\sigma_1}^D (\chi_{\sigma_1}(c_n)) + \mu_{\sigma_2}^D (\chi_{\sigma_2}(c_n))} - \tau (\chi_{\sigma_1}(c_n)) \right) + \frac{(1-\phi^D)}{\mu_{\sigma_1}^D} + \frac{\phi^D}{\psi^D}$
- $\frac{\partial U}{\partial \mu_{\sigma_2}^D} = -\sum_{n=1}^N \left(\frac{w_n^D (\chi_{\sigma_2}(c_n))}{\mu_{\text{back}}^D + \mu_{\sigma_1}^D (\chi_{\sigma_1}(c_n)) + \mu_{\sigma_2}^D (\chi_{\sigma_2}(c_n))} - \tau (\chi_{\sigma_2}(c_n)) \right) + \frac{(1-\phi^D)}{\mu_{\sigma_2}^D} + \frac{\phi^D}{\psi^D}$
- $\frac{\partial U}{\partial \mu_{\sigma_1}^A} = -\sum_{n=1}^N \left(\frac{w_n^A (\chi_{\sigma_1}(c_n))}{\mu_{\text{back}}^A + \mu_{\sigma_1}^A (\chi_{\sigma_1}(c_n)) + \mu_{\sigma_2}^A (\chi_{\sigma_2}(c_n))} - \tau ((\chi_{\sigma_1}(c_n))) \right) + \frac{(1-\phi^A)}{\mu_{\sigma_1}^A} + \frac{\phi^A}{\psi^A}$
- $\frac{\partial U}{\partial \mu_{\sigma_2}^A} = -\sum_{n=1}^N \left(\frac{w_n^A (\chi_{\sigma_2}(c_n))}{\mu_{\text{back}}^A + \mu_{\sigma_1}^A (\chi_{\sigma_1}(c_n)) + \mu_{\sigma_2}^A (\chi_{\sigma_2}(c_n))} - \tau ((\chi_{\sigma_2}(c_n))) \right) + \frac{(1-\phi^A)}{\mu_{\sigma_2}^A} + \frac{\phi^A}{\psi^A}.$

A.4.4 Background Photon Emission Rates

We estimate the values of $\mu_{\text{back}}^D, \mu_{\text{back}}^A$ from portions of the time traces without signal, where $\mu_{\sigma_k}^D = \mu_{\sigma_k}^A = 0$ photons/s. We analyze these portions, labeled with $\mathbf{w}_{\text{back}}^D = (w_{1,\text{back}}^D, w_{2,\text{back}}^D, \dots, w_{N_{\text{back}},\text{back}}^D)$ and $\mathbf{w}_{\text{back}}^A =$

$(w_{1,\text{back}}^A, w_{2,\text{back}}^A, \dots, w_{N_{\text{back}},\text{back}}^A)$, to estimate the background photon emission rates based on the formulation provided below

$$\mu_{\text{back}}^D \sim \text{Gamma}\left(\xi^D, \frac{\nu^D}{\xi^D}\right) \quad (\text{A.50})$$

$$\mu_{\text{back}}^A \sim \text{Gamma}\left(\xi^A, \frac{\nu^A}{\xi^A}\right) \quad (\text{A.51})$$

$$w_{n,\text{back}}^D | \mu_{\text{back}}^D \sim \text{Poisson}(\mu_{\text{back}}^D \tau) \quad (\text{A.52})$$

$$w_{n,\text{back}}^A | \mu_{\text{back}}^A \sim \text{Poisson}(\mu_{\text{back}}^A \tau) \quad (\text{A.53})$$

for all $n = 1, 2, \dots, N_{\text{back}}$ with hyperparameters $\xi^D, \nu^D, \xi^A, \nu^A$.

Therefore, we have the following posterior distribution for the background photon emission rates for the donor channel

$$\mathbb{P}(\mu_{\text{back}}^D | \mathbf{w}_{\text{back}}^D) \propto \mathbb{P}(\mathbf{w}_{\text{back}}^D | \mu_{\text{back}}^D) \mathbb{P}(\mu_{\text{back}}^D) = \left(\prod_{n=1}^N \mathbb{P}(w_{n,\text{back}}^D | \mu_{\text{back}}^D) \right) \mathbb{P}(\mu_{\text{back}}^D) \quad (\text{A.54})$$

$$\propto \exp\left(-\mu_{\text{back}}^D \left(\frac{\xi^D}{\nu^D} + \sum_{n=1}^N \tau\right)\right) \mu_{\text{back}}^D \left(\xi^D + \sum_{n=1}^N w_{n,\text{back}}^D - 1\right) \quad (\text{A.55})$$

$$= \text{Gamma}\left(\mu_{\text{back}}^D; \xi^D + \sum_{n=1}^N w_{n,\text{back}}^D, \frac{1}{\frac{\xi^D}{\nu^D} + \sum_{n=1}^N \tau}\right). \quad (\text{A.56})$$

We can then estimate μ_{back}^D via the mean of the above posterior distribution, which is equal to $\frac{\xi^D + \sum_{n=1}^N w_{n,\text{back}}^D}{\frac{\xi^D}{\nu^D} + \sum_{n=1}^N \tau}$. As

the prior becomes non-informative (54, 76, 84); that is, at the limit $\xi^D \rightarrow 0$, our estimate reduces to

$$\hat{\mu}_{\text{back}}^D = \frac{\sum_{n=1}^N w_{n,\text{back}}^D}{N\tau}. \quad (\text{A.57})$$

We obtain the estimator $\hat{\mu}_{\text{back}}^A$ in the acceptor channel similarly.

A.5 Notation and Analysis Options

Table A.1. *Notation Conventions*

Description	Variable	Units
Total number of measurements	N	—
Total number of conformational states	K	—
Total number of jumps	M	—
Sequence of system states $\{c_0, c_1, \dots, c_N\}$	\vec{c}	—
HMM trajectory	$\mathcal{T}(\cdot) = \vec{c}$	—
State of the system in HMM at time t_n	c_n	—
State of the system in HMJP at n^{th} jump	s_n	—
Holding time of the system in HMJP at n^{th} jump	d_{n+1}	s
HMJP trajectory	$\mathcal{T}(\cdot) = (\vec{S}, \vec{D}, M)$	—
Initial transition probability matrix	$\bar{\rho}$	—
Transition probability matrix in HMM	\bar{P}	—
Probability of transitioning from state σ_k to $\sigma_{k'}$ in HMM	$P_{\sigma_k \rightarrow \sigma_{k'}}$	—
Transition probability for the embedded discrete Markov chain in HMJP	$\bar{\pi}$	—
Escape rate for the k^{th} state of the system	$\lambda_{\sigma_K} = -q_{\sigma_k \rightarrow \sigma_k}$	1/s
Absolute escape rate vector	$\bar{\lambda}$	—
photon emission rates	$\bar{\mu}^D, \bar{\mu}^A$	photons/s
Background photon emission rate estimates	$\hat{\mu}_{back}^D, \hat{\mu}_{back}^A$	photons/s
Generator matrix in HMJP	\bar{G}	1/s
Rate of transitioning from state σ_k to $\sigma_{k'}$ in HMJP	$\lambda_{\sigma_k \rightarrow \sigma_{k'}}$	1/s
Initial time of the experiment	t_0	s
End time of the experiment	t_N	s
Fixed exposure period	τ	s
Fixed inactive period of detector	τ_D	s
Fixed data acquisition period	Δt	s
Data acquisition rate	$1/\Delta t$	1/s
Representative of n^{th} data acquisition period	$[t_{n-1}, t_n]$	—
Representative of n^{th} exposure period	$[t_n - \tau, t_n]$	—
Concentration parameter for $\bar{\rho}$	α	—
Concentration parameter for $\bar{\pi}$ and \bar{P}	A	—
Shape parameter	η	—
Scale parameter	b	—
Emission rate shape	ϕ^D, ϕ^A	—
Emission rate scale	$\psi^D/\phi^D, \psi^A/\phi^A$	—
Background emission rate shape	ξ^D, ξ^A	—
Background emission rate scale	$\nu^D/\xi^D, \nu^A/\xi^A$	—
Indicator	$\chi_u(v) = \begin{cases} 1 & v \neq u \\ 0 & v = u \end{cases}$	—

Table A.2. *Parameter Choices and Units*

-	ψ^D, ψ^A	ϕ^D, ϕ^A	η	b	A	α	Δt
Units	photons/s	-	-	1/s	-	-	s
Fig. 1	total photon counts/total duration	1	2	100	1	1	0.05
Fig. 2	total photon counts/total duration	1	2	100	1	1	0.05
Fig. 3	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. 4	total photon counts/total duration	1	2	100	1	1	0.025
Fig. 5	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. 6	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. 7	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.1	total photon counts/total duration	1	2	300	1	1	0.05
Fig. A.1	total photon counts/total duration	1	2	300	1	1	0.05
Fig. A.2	total photon counts/total duration	1	2	300	1	1	0.05, 0.075, 0.1, 0.15
Fig. A.3	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.4	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.5	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.6	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.7	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.8	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.9	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1
Fig. A.10	total photon counts/total duration	1	2	100	1	1	0.025, 0.05, 0.1