Extraction of rapid kinetics from smFRET measurements using integrative detectors

Zeliha Kilic1, Ioannis Sgouralis2, Wooseok Heo3, Kunihiko Ishii3,4, Tahei Tahara3,4, Steve Pressé1,5,6,*

1Center for Biological Physics, Department of Physics, Arizona State University, Tempe, AZ 85287, USA
2Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA
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, USA
6Lead contact

SUMMARY

Hidden Markov models (HMMs) are used to learn single-molecule kinetics across a range of experimental techniques. By their construction, HMMs assume that single-molecule events occur on slower timescales than those of data acquisition. To move beyond that HMM limitation and allow for single-molecule events to occur on any timescale, we must treat single-molecule events in continuous time as they occur in nature. We propose a method to learn kinetic rates from single-molecule Förster resonance energy transfer (smFRET) data collected by integrative detectors, even if those rates exceed data acquisition rates. To achieve that, we exploit our recently proposed “hidden Markov jump process” (HMJP), with which we learn transition kinetics from parallel measurements in donor and acceptor channels. HMJPs generalize the HMM paradigm in two critical ways: (1) they deal with physical smFRET systems as they switch between conformational states in continuous time, and (2) they estimate transition rates between conformational states directly without having recourse to transition probabilities or assuming slow dynamics. Our continuous-time treatment learns the transition kinetics and photon emission rates for dynamic regimes that are inaccessible to HMMs, which treat system kinetics in discrete time. We validate
our framework’s robustness on simulated data and demonstrate its performance on experimental data from FRET-labeled Holliday junctions.

**Graphical Abstract**

Kilic et al. apply their recently developed HMM generalization toward unraveling single-molecule FRET kinetics occurring on timescales exceeding data acquisition rates of integrative detectors. This HMM generalization treats single-molecule kinetics in continuous time, as occurs in nature, and treats integrative-detector measurements as occurring in discrete time.

**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.\(^1\)\(^–\)\(^7\) In a prototypical intra-molecular smFRET experiment, one portion of a molecule of interest is attached to a *donor* fluorophore and another to a *acceptor* fluorophore.\(^1\)\(^–\)\(^15\) In such experiments, the excitation wavelength is most commonly adjusted to excite the donor,\(^6\)\(^,\)\(^11\)\(^,\)\(^13\)\(^–\)\(^15\) and for a donor sufficiently far from the acceptor, the donor is excited and emits shorter wavelength light as compared with the longer wavelength light emitted by the acceptor in the case of energy transfer when a donor and acceptor are in proximity.\(^6\)\(^,\)\(^11\)\(^,\)\(^13\)\(^–\)\(^15\) Because of the difference in the wavelength of light emitted by donors and acceptors, photons emitted are registered across different detectors.\(^6\)\(^,\)\(^11\)\(^,\)\(^13\)\(^–\)\(^15\) We refer to the recordings in the two detectors as the donor and acceptor channels.\(^6\)\(^,\)\(^11\)\(^,\)\(^13\)\(^–\)\(^15\) As such, the sequence of photon detection\(^2\) encodes the kinetics; according to which, the distance varies between both fluorophores down to microsecond timescales.\(^1\)\(^–\)\(^6\) Existing approaches for analyzing smFRET data include FRET efficiency histogram-based
methods, although such methods discard the important temporal information encoded in donor and acceptor photon arrival sequences.\textsuperscript{4–6,8–10,12,13,16} What is more, FRET efficiency histogram analysis is difficult to generalize beyond two conformational states.\textsuperscript{17} In addition, because arbitrary binning is required to construct histograms, such methods often lead to inconclusive or erroneous estimates.\textsuperscript{4,18–20} For these reasons, modeling efforts have moved toward more direct time-series analyses.\textsuperscript{5,17,21–24}

By relying on the hidden Markov model (HMM) paradigm,\textsuperscript{1,3,5,23–32} time series analysis fully considers the temporal arrangement (i.e., the sequence) of photon detections and avoids histogram binning artifacts.\textsuperscript{1,3,5,25} This paradigm is especially fruitful in analyzing demanding experiments, such as conformational states, which are treated as the hidden states of the HMMs, are themselves indirectly observed because of shot noise,\textsuperscript{1,3} and also allows for inclusion of measurement noise and specialized detector characteristics in the analysis.\textsuperscript{1,3,5,25}

Starting with Ref.\textsuperscript{12} and later with others,\textsuperscript{1,3,5,23,28,33} HMMs and their variants have had important assumptions built into them. Those are the assumptions we wish to remove in an effort to analyze conformational transitions occurring on timescales faster than the data acquisition rate $1/\Delta t$ supported by integrative detectors, such as cameras. Before describing these assumptions, we discuss $\Delta t$ (the bin size for photon collection), otherwise known as the frame rate or temporal resolution. Because $\Delta t$ is determined mostly by the integrative detectors, for simplicity, we assume that the same $\Delta t$ applies to both channels.

To learn rates that exceed data-acquisition rates, the most critical HMM assumptions we need to remove are the following:

- HMMs make the assumption that the state switches occur \textit{rarely} as compared to $\Delta t$. In other words, all switching rates are assumed to be much slower than $1/\Delta t$. For that reason, HMMs are formulated using transition probabilities rather than transition rates.

- HMMs approximate the transitions as occurring \textit{precisely} at the end of each data acquisition period.\textsuperscript{23,28,34–39} In other words, intra-frame motion does not occur. For that reason, HMMs represent only instantaneous states and measurements.

The first assumption is particularly relevant to the present study because it requires, before an analysis, that all transitions be slower than $\Delta t$. In general, this is not only unknown but also difficult to assess beforehand. After all, the objective of many experiments is to determine the switching rates, in the first place. The second assumption is also relevant to the present study because, before an analysis, it needs to balance two competing requirements: an upper bound on $\Delta t$ set by the fastest switching rate, and a lower bound on $\Delta t$ supported by the detection hardware.

One method, termed H$^2$MM (a variant of HMM), was specifically developed to overcome these challenges.\textsuperscript{3} In H$^2$-MMs, HMMs are applied on time grids finer than the measurement time interval. However, because of the finer grid, H$^2$ MM and their variants suffer from high computational costs.\textsuperscript{40} For that reason, continuous time formulations,\textsuperscript{40,41} which allow transitions to occur at any time, have been sought. Formulations such as those are important
because they would provide a framework to develop robust and general methods for extracting kinetics that are faster than the data acquisition time and may provide ways to improve on previous efforts to extract fast kinetics, which have been limited in various ways, e.g., limited to two conformational states.\cite{31} In particular, an important goal of ours herein is to adapt continuous-time formulations\cite{40,41} to the averaging inherent to integrative detector modalities.

We mention in passing that alternate approaches have been developed to analyze fast kinetics by smFRET,\cite{1,9,42,43} although they have been tied to the analysis of single-photon arrivals. Such methods do not rely on the HMM paradigm because they do not analyze binned data collected by integrative detectors, even though those methods are also limited by the photon detection rate. As such, they do not suffer from the limitations of HMMs.

However, all the aforementioned methods, whether in the analysis of single-photon data or binned data,\cite{1,3,9,42,43} rely on maximum likelihood,\cite{44} often with assumptions required to render computations analytically tractable (such as assuming there are only transitions between two linearly arranged conformational states).\cite{45} Because of working within a maximum-likelihood approach, only point estimates of the quantities of interest can be directly drawn from the data with additional statistical strategies, such as Fisher information or bootstrapping, used to extract error bars.\cite{44}

Maximum-likelihood methods lie in sharp contrast to the Bayesian methods we will employ, which will allow us to directly propagate measurement uncertainty into an uncertainty over parameter estimates.

In Figures 1A and 1C, we illustrate an example of smFRET measurements acquired with integrative detectors, such as modern scientific-grade cameras. In Figure 1A, we show an example with slow kinetics (relative to \( \Delta t \)). These data are primarily contaminated with shot noise and can be reliably analyzed within the HMM paradigm. In contrast, in Figure 1C, we demonstrate an example with fast kinetics (relative to \( \Delta t \)). In addition to 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 the 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 HMM paradigm. Nonetheless, the information on fast switching is itself encoded within that measurement, even though some of the information is indeed lost. It is, therefore, necessary to develop a method that can accommodate continuously evolving dynamics to learn the kinetics from the measurements, instead of invoking an HMM formulation that, by construction, violates that critical feature.

To achieve that, a new method needs to (1) represent the dynamics of the molecule in continuous time, (2) model the acquired measurements at each data acquisition times via the average dynamics of the molecule within the associated data-acquisition period, and (3) entail manageable computational cost so it allows for practical applications and propagation of errors from measurement uncertainty.
For this reason, we now turn to Markov jump processes (MJPs), which describe continuous time dynamics, and propose inference strategies for the MJPs.

The focus of this work is to adapt inverse strategies on MJPs to deal with the integrative nature of the detectors (that average smFRET signal over the exposure windows) to extract fast kinetics. Because we extract MJPs hidden by the integrative nature of the measurement, we call our method the hidden MJP (HMJP). As we will see, HMJPs are direct generalizations of HMMs, and as such, HMJPs will rigorously reduce to the HMM as the exposure period approaches zero, $\Delta t \to 0$.

Briefly, in the results section, we provide HMM and HMJP analysis comparisons on simulated measurements of smFRET collected by integrative detectors. In comparing the HMM and HMJP, we present their respective performance in acquiring 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 with the data-acquisition rate. Their comparison on slow-switching kinetics of simulated smFRET data (in which both the HMM and HMJP, as expected, do well) can be found in supplemental information. Subsequently, we move onto the analysis of experimental data. We provide a comparison of HMJPs and HMMs in their learning of molecular trajectories, transition probabilities, and kinetic rates. Next, in our discussion, we describe the broader potential for HMJPs for smFRET applications. Lastly, in our experimental procedures section, we provide our HMJP smFRET model description and briefly, for sake of comparison, summarize plain HMMs for smFRET applications.

**RESULTS**

**Simulated data analysis overview**

To show how HMJPs work and to highlight the advantages of HMJPs over that of 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. Next, we compare the strength of our HMJP method to that of HMMs on experimental data.

We focus on the following simulated datasets: a molecule exhibiting fast kinetics, as compared with the data-acquisition rate, and with two and three conformational states (see Figures 1 and 2, respectively).

The simpler analysis of datasets with slow kinetics can be found in supplemental information. The results corresponding to the simulated dataset with fast kinetics for both HMJP and HMM are shown in Figures 3 and 4. Subsequently, we show the performance of our method on an experimental dataset (Figure 5). We present the results for the experimental dataset in Figures 6, 7, 8, and 9. Further analysis on different experimental datasets, additional analysis of datasets with more than two conformational states, and an analysis of simulated datasets generated based on electron multiplying charge-coupled device (EMCCD) models are also provided in supplemental information.
Hyperparameter values used in all analyses, as well as any other choices made, are presented in supplemental information. For clarity, we only have access to the data demonstrated with the green and red dashes of Figures 1A and 1C (not the cyan [ground truth] trajectories). Those ground truth trajectories are unknown and are determined along with other model parameters.

**Acquisition of simulated data**—In the generation of our simulated data, provided in Figure 1C, we assumed \( K = 2 \) attainable states, indexed as \( \sigma_1 \) and \( \sigma_2 \). These could be folded or unfolded states for illustrative purposes (our method trivially generalizes to more states).

We assumed photon emission rates indexed with \( \mu_{\sigma_k}^D \) and \( \mu_{\sigma_k}^A \) for the donor and acceptor in each attainable state \( \sigma_k \), which we set to \( \mu_{\sigma_1}^D = 800 \text{photons/s} \), \( \mu_{\sigma_2}^D = 2000 \text{photons/s} \), and \( \mu_{\sigma_1}^A = 300 \text{photons/s} \), \( \mu_{\sigma_2}^A = 700 \text{photons/s} \) where the background photon emission rates are set to \( \mu_{\text{back}}^D = \mu_{\text{back}}^A = 200 \text{photons/s} \). Additionally, we defined a data-acquisition period of \( \Delta t = 0.05 \) and consider the exposure period by setting \( \tau \) equal to 100% of \( \Delta t \). The onset and concluding time of the simulated data are at \( t_0 = 0.095 \text{s} \) and \( t_N = 10.05 \text{s} \), respectively.

We use the parameter \( \tau_f \) as follows, to set a timescale in expressing the switching rates between both states \( (\lambda_{\sigma_1 \rightarrow \sigma_2}, \lambda_{\sigma_2 \rightarrow \sigma_1}) \),

\[
\lambda_{\sigma_1 \rightarrow \sigma_2} = \frac{2.2}{\tau_f}, \lambda_{\sigma_2 \rightarrow \sigma_1} = \frac{2.1}{\tau_f} \tag{Equation 1}
\]

in the analysis of fast and slow switching rates.

We simulate the case with \( \tau_f = 0.1 \text{s} \) in Figure 1C, which involves system kinetics exceeding the data acquisition rate. These kinetic rates are chosen to simulate data closer to the experimental data (shown later in Figure 5) to benchmark our framework.

In this simulated data, we set the data acquisition period to \( \Delta t = 0.05 \) such that our \( \lambda_{\sigma_1 \rightarrow \sigma_2} \) and \( \lambda_{\sigma_2 \rightarrow \sigma_1} \) exceed the data acquisition rate \( (1/\Delta t) \).

Next, we simulated a dataset with fast dynamics (Figure 2), in which \( K = 3 \) attainable states (indexed \( \sigma_1 \) to \( \sigma_3 \) ) were assumed for the underlying dynamics of the single molecule. Further, we set photon emission rates for the donor and acceptor in each state as

\[
\mu_{\sigma_1}^D = 800 \text{photons/s}, \mu_{\sigma_2}^D = 300 \text{photons/s}, \mu_{\sigma_3}^D = 6000 \text{photons/s}, \mu_{\sigma_1}^A = 300 \text{photons/s}, \\
\mu_{\sigma_2}^A = 2000 \text{photons/s}, \mu_{\sigma_3}^A = 5000 \text{photons/s}, \mu_{\text{back}}^D = \mu_{\text{back}}^A = 200 \text{photons/s}.
\]

The data acquisition period here was also set to \( \Delta t = 0.05 \) with the exposure period \( \tau \) set to 100% of the data acquisition period. In the dataset demonstrated in Figure 2, the onset of the simulated experiment is \( t_0 = 0 \text{s} \), where the final time is \( t_N = 7 \text{s} \). The switching rates used in the generation of the dataset in Figure 2 are
\[ \lambda_{o_1 \rightarrow o_2} = \frac{2.2}{\tau_f}, \quad \lambda_{o_1 \rightarrow o_3} = 0, \quad \lambda_{o_2 \rightarrow o_1} = \frac{2.1}{\tau_f}, \quad \lambda_{o_2 \rightarrow o_3} = 0, \quad \lambda_{o_3 \rightarrow o_1} = \frac{2}{\tau_f}, \]

(Equation 2)

with \( \tau_f = 0.1 \) s.

**Comparison of HMJPs with HMMs on simulated data**—Here, we compare HMJPs and HMMs on the analysis of simulated measurements given in Figures 1C and 2A.

We start with the analysis of the data provided in Figure 1C. We expect HMJPs to perform better than HMMs do because 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 Equation 16 required by HMMs fails.

Given the measurements, we estimate the posterior distribution over the quantities of interest, including the trajectory, \( \bar{T}(\cdot) \); 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 Computational inference section to generate pseudorandom numbers from the posterior distributions

\[ p(\bar{\rho}, \bar{\pi}, \bar{\lambda}, \mu^D, \mu^A, \bar{S}, \bar{D}, M | w^D, w^A) \]

and

\[ p(\bar{\rho}, \bar{P}, \bar{\mu}, \bar{\gamma} | w^D, w^A), \]

respectively.

We first show emission and escape rate estimates in Figure 3. In Figures 3A–3F, we observe that the ground truths for photon emission rates, escape rates, and transition probabilities lie within the 95% confidence intervals for the corresponding HMJP estimates. For clarity, we compute the 95% confidence interval using only samples produced after burn-in. By contrast, the HMM performs poorly because of the failure of the approximation in Equation 16, resulting in an unsatisfactory photon emission rate and unsatisfactory escape rate estimates.

In particular, from Figures 3A and 3B, we see that the HMM overestimates \( \mu^D_{o_1} \) and \( \mu^A_{o_1} \) (by about 12%, namely, HMM provides estimates that are approximately 1.12 times the ground-truth values for \( \mu^D_{o_1} \) and \( \mu^A_{o_1} \)) and underestimates (by about 5%) \( \mu^D_{o_2} \) and \( \mu^A_{o_2} \).

The failure of the HMM is more pronounced when looking at \( \lambda_{o_1} \) and \( \lambda_{o_2} \). For example, in Figures 3C and 3D, the HMM provides narrow distributions over escape rates by grossly underestimating both \( \lambda_{o_1} \) and \( \lambda_{o_2} \) (by about 50%). On the other hand, in Figures 3C and 3D, we see that HMJP posterior distribution mode coincides with the ground truth.

The HMM inability to capture fast kinetics is also reflected by its wide posterior distributions over transition probabilities \( P_{o_1 \rightarrow o_1}, P_{o_2 \rightarrow o_2} \) as shown in Figures 3E and 3F.

We next move to the analysis of the data provided in Figure 2A. Just as with the analysis of Figure 1C, we observe similar results in Figure 4. In particular, from Figures 4A and 4B, we...
observe that HMMs fail to estimate $\mu_{D_1}^D$, $\mu_{A_1}^A$, $\mu_{D_3}^D$, and $\mu_{A_3}^A$, such that the HMMs overestimate (by about 30%) $\mu_{D_1}^D$ and $\mu_{A_1}^A$, whereas they underestimate (by about 20%) $\mu_{D_3}^D$ and $\mu_{A_3}^A$. By contrast, the mode of the HMJP posterior over photon emission coincides with the ground truth for both donor and acceptor channels. Next, HMMs provide broad, uncertain transition-probability estimates as shown in Figures 4F–4H. These poor HMM transition probability estimates give rise to the tight posterior escape rate estimates shown in Figures 4C–4E, based on Equation 16. By contrast, the HMJP posteriors over transition probabilities, as well as escape rates, have modes coinciding with the ground truths.

In supplemental information and Figure S1, we provide an analysis of slow-switching kinetics for the HMJP and HMM estimate comparisons shown in Figure 3, where both HMJP and HMM perform well.

Subsequently, in supplemental information and Figure S2, we provide the robustness of HMJP posterior estimates over photon emission rates, escape rates, and transition probabilities for the dataset provided in Figure 1C. We observe that, when the kinetic rates are more than four times the data acquisition rate, the HMJP starts failing in providing consistent posterior estimates over photon emission rates, escape rates, and transition probabilities. We have run further analyses to test the HMJP performance in estimating posterior distributions for photon emission rates and switching rates on the data generated for five cases.

First, we considered fast dynamics and different exposure periods (see Figures S3 and S4). In this first case, we tested the robustness of the HMJP framework on a two-state system in estimating photon emission rates and switching rates in which the ground-truth switching rates differ by more than the ones used to generate the data shown in Figure 1. In addition, we performed this test for the estimates of the HMJP for two different exposure periods. We demonstrate in Figures S3 and S4 that the HMJP posteriors over photon emission rates and switching rates have ground-truth values that lie in the posterior’s 95% confidence intervals.

Second, we generated data with slow-fast (in which we have mixed regime of few or many transitions occurring within a data acquisition period) dynamics and different exposure periods (see Figures S5 and S6). Here, we run the same test as presented in the first case but for switching kinetics, and we observe both slower and faster kinetics for the simulated single molecule than the data acquisition rate. The HMJP still performed well in estimating both photon emission rates and switching kinetics (see Figures S5 and S6).

Third, we generated data with fast dynamics by using three different sets of photon emission rates (see Figures S7–S9). Here, we tested the robustness of HMJP in estimating posterior photon emission rates and switching kinetics for various differences in photon emission rates across each conformational states of the simulated single-molecule trajectory. In this test, the HMJP still performed quite well in estimating photon emission rates and switching kinetics when the switching kinetics exceeded the data acquisition rate on datasets generated with a range of photon emission rates.
Fourth, we tested the HMJP on its ability to deduce the correct parameters when an incorrect model is specified. For example, we specified three states for data generated using two states and found that the additional state learned was redundant (see Figures S10–S12).

Fifth, we tested the performance of our model on data simulated using a specific time for the integrative detector (EMCCDs). We have run our analysis on simulated data generated with three conformational state dynamics. Further, we set photon emission rates for donor and acceptor at $\mu_{\sigma_1}^D = 800 \text{photons/s}$, $\mu_{\sigma_2}^D = 3000 \text{photons/s}$, $\mu_{\sigma_3}^D = 6000 \text{photons/s}$, and $\mu_{\sigma_1}^A = 300 \text{photons/s}$, $\mu_{\sigma_2}^A = 2000 \text{photons/s}$, and $\mu_{\sigma_3}^A = 5000 \text{photons/s}$, where

$$\mu_{\text{back}}^D = \mu_{\text{back}}^A = 200 \text{photons/s}.$$  

The data acquisition period here was also set to $\Delta t = 0.05 \text{s}$ with the exposure period $\tau$ set to 100% of the data acquisition period. In the dataset demonstrated in Figure S13, the onset of the simulated experiment is $t_0 = 0 \text{s}$ where the final time is $t_N = 7 \text{s}$. The switching rates used in the generation of the dataset in Figure S13, used $\tau_f = 0.1 \text{s}$ and

$$\lambda_{\sigma_1 \rightarrow \sigma_2} = \frac{1}{\tau_f}, \lambda_{\sigma_1 \rightarrow \sigma_3} = \frac{1.2}{\tau_f}, \lambda_{\sigma_2 \rightarrow \sigma_1} = \frac{1.5}{\tau_f}, \lambda_{\sigma_2 \rightarrow \sigma_3} = \frac{0.6}{\tau_f}, \lambda_{\sigma_3 \rightarrow \sigma_1} = \frac{0.5}{\tau_f}.$$  

(Equation 3)

We presented the results for the posterior switching-rate estimates and the photon emission rates in donor and acceptor channels in Figures S14–S16, respectively. We observe that our framework performs very well in estimating the model parameters in the case of a different camera model along with fast-switching dynamics based on the provided posterior estimates.

We now move onto the analysis of experimental datasets.

**Experimental data analysis overview**

Having benchmarked our HMJP on the simulated smFRET dataset, we now move on to assessing the HMJP performance on experimental smFRET data provided in Figure 5, for Holliday junctions (HJs), as described in the sample preparation. We deliberately start from single photon arrival datasets. Such datasets give us the flexibility to bin data as we please to mimic experimental data collected over a broad range of binning conditions as is the case in several studies. In particular, we bin using three data acquisition periods: $\Delta t = 0.025 \text{s}$ (with results shown in Figures 6, 7, 8, and 9), $\Delta t = 0.05 \text{s}$ (with results shown in Figures 6, 7, 8, and 9), and $\Delta t = 0.1 \text{s}$ (with results shown in Figures 6, 7, 8, and 9). We picked these bin sizes for a particular reason: we estimate rates of ~35–45 (1/s) for the smallest bins. Then, we consider data acquisition rates that are twice and four times larger. We also set the exposure period, $\tau$ to 100% of the $\Delta t$, namely, $\tau = 0.025$, 0.05 s and $\tau = 0.1$ s.

We emphasize that the fast transitions are averaged away when the data acquisition period is larger, especially for $\Delta t = 0.1 \text{s}$ (in other words, when we have the largest time bin provided in Figure 5E). Thus, the largest time bin may, counter-intuitively, appear as though it had the best separated peaks (see the histograms in Figure 5). However, despite being the sharpest, it has the least amount of information.
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 because of the inconsistency of the estimates for the switching kinetics as a function of the data bin size. The HJs we used have two well-characterized conformational states. Here, our goal was to estimate the HJ emission and switching rates, in addition to the trajectory of HJs, simultaneously.

In Figures 5A, 5C, and 5E, we show the data under all three different bin sizes. From the data, we estimate the posterior distribution over the trajectory, $\mathcal{T}(\cdot)$; initial transition probability matrix, $\rho$; transition probabilities, $\pi$ (and $\mathcal{P}$ in HMM); photon emission rates, $\mu^D$, $\mu^A$; and escape rates, $\lambda$. In Figure 6, we show a direct comparison of the HMJP and HMM under the smallest bin size, in which 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 results that we obtained for the smallest bin, but the HMM eventually shows a lack of consistency as the bin sizes increase. This is apparent from Figure 7, in which we look at the increasing bin size for HMM, and in Figure 8, in which we look at the increasing bin size for HMJP. Finally, in Figure 9, we show the estimates over FRET efficiencies, in which we increase bin sizes for both HMM and HMJP.

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 Figure 5A), on the photon emission rates, escape rates, and transition probability estimates provided in Figure 6. We show that the simulated dataset provided in Figure 1 mimics the dataset provided in Figure 5.

Even though, for experimental data shown in Figure 5A, the “ground truth” photon emission rates, escape rates, and transition probabilities are not known, we suspect that similar observations provided in Figure 3 hold for those in Figure 6.

In Figures 6A and 6B, we observe that HMM overestimates $\mu_{\sigma_1}$ and underestimates $\mu_{\sigma_2}$ compared with HMJP estimates for $\mu_{\sigma_1}$ and $\mu_{\sigma_2}$. In addition, as shown in Figures 6C and 6D, HMM provides narrow distributions over the escape rates for both $\lambda_{\sigma_1}$ and $\lambda_{\sigma_2}$.

This contradicts with the escape rate posterior of the HMJP (see Figures 6C and 6D). We speculate that, because HMM is unable to resolve fast-switching kinetics, missed kinetic transitions are reflected in terms of underestimation or overestimation of photon emission rates or escape rates (see Figures 6A–6D). In Figures 6E and 6F, we observe that HMM and HMJP estimates for $P_{\sigma_1 \rightarrow \sigma_1}$ overlap, although they have distinct posterior distributions for $P_{\sigma_2 \rightarrow \sigma_2}$.

Next, we show that, in Figure 7, the HMM photon emission rates, escape rates, and transition probability estimates grow inconsistent for larger bin sizes, $\Delta t = 0.05,0.1$s (see Figures 5B and 5C). As a result of binning, multiple different apparent photon emission rates appear. This could be overinterpreted as noise by the HMM (see Figures 7A and 7B).
Indeed, HMMs interpret the various photon emission rates as increased noise variance. As a related artifact, the HMM misses multiple transitions and, therefore, grossly underestimates the escape rates.

As shown in Figures 7C and 7D, HMMs start underestimating the 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 Equation 16) because an HMM does not directly report kinetic rates.

In Figures 7E and 7F, the uncertainty around the HMM transition probability estimates increases and are, furthermore, not consistent as bin sizes increase. Subsequently, we repeat the exact same analysis we performed with the HMM, with the HMJP. In particular, there is an exact correspondence between Figures 7A–7F and 8A–8F. Details are provided in the captions; however, the important message is that the HMJP escape rate estimates (see Figures 8C and 8D) remain consistent across the increased bin sizes.

Finally, in Figures 9A and 9B, we provide the posterior distributions over FRET efficiencies for the HMM and HMJP. Here, the HMMs provide inconsistent posterior FRET efficiency estimates for larger bin sizes (see Figure 9A). In contrast, HMJPs provide consistent FRET efficiency estimates (see Figure 9B).

In supplemental information, we provide similar analyses for two other experimental smFRET datasets (see Figures S17 and S21). The datasets provided in Figures S17 and S21 were acquired under the same experimental conditions as those presented here. By analyzing those additional datasets, we justify the robustness of our method (see Figures S18–S20 and S22–S24).

**DISCUSSION**

Here, we provide the general framework from which switching kinetics, occurring on timescales at or even exceeding measurement timescales,\textsuperscript{1,3,25} can be obtained from smFRET data collected on integrative detectors.\textsuperscript{4,18–20} To achieve that, we have extended the HMM paradigm previously used in the analysis of smFRET, by treating physical processes as they occur in nature;\textsuperscript{1,3–5,18–20,25} that is, as continuous-time processes using a MJP framework.\textsuperscript{12,47,55–59}

By taking advantage of algorithmic advancements in inference on continuous-time processes that are scarcely half a decade old,\textsuperscript{40} the HMJP scales as $O(MK)$, in which $M$ is the number of jumps, and $K$ is the number of states (see Robert and Casella\textsuperscript{60} and Kilic et al.\textsuperscript{61}). This scaling is, in itself, a testament to the powerful method of uniformization that only introduces jumps if they are warranted based on the data. That is in contrast to naive grid-based methods, which start with a very fine grid and must commit to $O(M'K)$ scaling, in which $M'$ is the number of grid points.

Our framework can treat multiple different data-collection regimes. For example, it can treat different molecular states, exposure periods, measurement models, and background levels.
Of note is the fact that our framework’s measurement model reduces to the HMM measurement model in the expected limits. For example, when the exposure period ($\tau$) is as long as the data acquisition period ($\Delta t$), and the only measurements available are those precisely at the data acquisition time, then the HMJP is equivalent to the HMM. When not in that limit, HMJPs outperform HMMs in estimating photon emission and switching rates, although HMJPs do begin failing when single-molecule switching rates exceed approximately four times the data-acquisition rate. What is more, existing methods become equivalent to ours in the appropriate limit. For example, in the limit that the time is discretized on an infinitely small grid, the $H^2$ MM$^3$ treatment of processes occurring on timescales below the data acquisition rate reduces to the continuous-time HMJP descriptions, that is, the HMJP operates within a Bayesian paradigm in which the full posterior distributions over unknown quantities are determined.$^{62,63}$

Interest in unraveling kinetics that exceed the data acquisition time is not new.$^{31,32,64–66}$ However, prior attempts at determining such kinetics have suffered from a lack of generality. For example, the method in Stigler and Rief$^{32}$ accounted for the effect of missed events on the switching rates by allowing for virtual conformational states (known as augmented network models) at which the molecule spends very short dwell times, such that the virtual confrontational states themselves cannot communicate. Therefore, their formulation is limited to a very specific conformational-state network model. Furthermore, the method that study$^{32}$ only provides point estimates for the model parameters, whereas the HMJP framework relies on the Bayesian paradigm$^{62,63}$ and, as such, it provides full posterior distributions for all unknown parameters, rather than their point estimates.

HMJPs, just like HMMs, do not, however, pertain to the analysis of single-photon arrival in the smFRET,$^{1,9,42,43}$ and, for that reason, limitations of the HMM are not intrinsic to single-photon-arrival analysis methods. However, because existing smFRET analysis, starting from single photon arrival, operates within a maximum-likelihood paradigm, it can, in principle, be reframed within a Bayesian framework,$^{67}$ such as the HMJP. Doing so, provides such methods with the ability to determine full posterior distributions, not just point estimates, around unknown quantities, thereby propagating uncertainty from the measurement into an uncertainty over the parameters of interest.

Although we have focused on determining parameters such as kinetic rate, it is also possible to deduce molecular state trajectories in time, whether from binned or single-photons data. Directly building off the HMM, the HMJP determines both the realization of the latent variables (molecular state trajectories) and kinetic parameters simultaneously and, thus, self-consistently. By contrast, the existing literature on single-photon arrival, for example as presented in Zosel et al.,$^{68}$ does so as two successive (disjointed) steps. For example, kinetic rates are first extracted using the method of Gopich and Szabo,$^{1}$ and the molecular state trajectories are then deduced using the concept of Viterbi paths$^{62,63}$ on data binned a posteriori. Such approximations are not intrinsic to smFRET single-photon arrival analysis and, in principle, simultaneous determination of molecular state trajectories and kinetic parameters should be possible.
As a final note, our framework might also be extended in multiple different ways. Trivially, we can change the detector model of Equations 8 and 9 to incorporate different kinds of detector models, such as that for scientific complementary metal oxide semiconductor (sCMOS) cameras used in many studies,5,69,70 which is equivalent to a change in Equations 8 and 9, as in the case of incorporating the EMCCD camera model;53 the details of which are presented in supplemental information. In addition, we can extend our framework to account for photophysical states of the fluorophores as well as to account for spectral cross-bleeding between donor and acceptor channels5 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 That is, to re-pitch the single-photon arrival frameworks within a Bayesian paradigm in which a full posterior (over latent variables and parameters) is sampled. In addition, it should be possible to leverage the strengths of Bayesian nonparametrics to acquire the number of conformational states within the HMJP.5,22,24,28,33,49,71–73 Overcoming these difficulties is the focus of future work.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact—The lead contact is Steve Pressé (spresse@asu.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability—All code is available on the lead contact’s website (https://cbp.asu.edu/content/steve-presse-lab) and upon request (Massachusetts Institute of Technology [MIT] license). Data are available upon request by contacting the lead contact.

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

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 \( \sigma_k \) and indexed with \( k = 1, \ldots, K \). For example, those states can be considered the \( K = 2 \) isomerization states of HJs.3 If the molecule is in state \( \sigma_k \) at time \( t \), then, we denote it by \( \mathcal{T}(t_0) = \sigma_k \).40,41

We model the transitions of the molecule as a memoryless process28,50,52,74 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 an MJP.50

In greater detail, we assume that the system chooses stochastically a state \( \sigma_k \) at the onset of the experiment that is denoted by \( \mathcal{T}(t_0) = \sigma_k \). The probability of determining this initial state
\( \sigma_k \) is labeled with \( \rho_{\sigma K} \). The collection of all initial probabilities is a probability vector and labeled with \( \bar{\rho} = (\rho_{\sigma_1}, \rho_{\sigma_2}, \ldots, \rho_{\sigma_K}) \).\^{60,62,75,76}

Switching the rates fully describes the switching kinetics of the molecule, and they are labeled with \( \lambda_{\sigma k} \to \sigma_{k'} \) for all possible states \( \sigma_k, \sigma_{k'} \) where \( k, k' = 1, 2, \ldots, K \) and by convention \( \lambda_{\sigma k} \to \sigma_{k'} = 0 \) by the rate-matrix definition for \( k = 1, 2, \ldots, K \). For mathematical convenience, we keep track of the escape rates as an alternative parameterization of the switching kinetics, which are given by

\[
\lambda_{\sigma k} = \sum_{k' = 1}^{K} \lambda_{\sigma k} \to \sigma_{k'}.
\]  

(Equation 4)

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

\[
\pi_{\sigma k} \to \sigma_{k'} = \frac{\lambda_{\sigma k} \to \sigma_{k'}}{\lambda_{\sigma k}}.
\]  

(Equation 5)

The collection of all normalized switching rates from state \( \sigma_k \) is denoted by \( \bar{\pi}_{\sigma k} = (\pi_{\sigma k} \to \sigma_1, \pi_{\sigma k} \to \sigma_2, \ldots, \pi_{\sigma k} \to \sigma_M) \). We see that each \( \bar{\pi}_{\sigma k} \) forms a probability vector.\^{76}

We can gather all transition probabilities in a matrix \( \bar{\pi} \) that reads

\[
\bar{\pi} = \begin{pmatrix}
\pi_{\sigma 1} \\
\pi_{\sigma 2} \\
\vdots \\
\pi_{\sigma K}
\end{pmatrix}.
\]  

(Equation 6)

Given \( \bar{\rho} \) and \( \lambda, \bar{\pi} \), the trajectory \( \mathcal{F}(\cdot) \) is obtained by a variant of the Gillespie algorithm,\^{52}

which determines a succession of states for the conformations of the system \( s_0, s_1, \ldots, s_{M-1} \) and their durations \( d_0, d_1, d_2, \ldots, d_{M-1} \). These together define \( \mathcal{F}(\cdot) \) throughout the time course \([t_0, t_M]\) by

\[
\mathcal{F}(t) = \begin{cases}
  s_0 & \text{if } t_0 \leq t < t_0 + d_0 \\
  s_1 & \text{if } t_0 + d_1 \leq t < t_0 + d_0 + d_1 \\
  \vdots & \\
  s_{M-1} & \text{if } t_0 + d_0 + \cdots + d_{M-2} \leq t < t_0 + d_0 + d_1 + \cdots + d_{M-1}.
\end{cases}
\]  

(Equation 7)

For clarity, we encode \( \mathcal{F}(\cdot) \) in a triplet \((\vec{S}, \vec{D}, M)\) in which \( \vec{S} = \{s_0, s_1, \ldots, s_{M-1}\} \), \( \vec{D} = \{d_0, d_1, \ldots, d_{M-1}\} \), and \( M \) are the size of \( \vec{S} \) and \( \vec{D} \).

Cell Rep Phys Sci. Author manuscript; available in PMC 2021 June 16.
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: \( w_D = (w_1^D, w_2^D, \ldots, w_N^D) \) and \( w_A = (w_1^A, w_2^A, \ldots, w_N^A) \), 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 \( \Delta t \) (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 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 with any measurement (see Figure 1).

One common assumption 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, \( \tau \) is the duration of each integration time for fluorescence experiments.

When we supplement our dynamic model (fully described in the Dynamics of the system section) 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 \( \mu_{\sigma_k}^D \) and \( \mu_{\sigma_k}^A \), in which the subscript highlights the state dependence of the photon emission rate. For simplicity, we gather the state-specific photon emission rates in \( \bar{\mu}_D = (\mu_{\sigma_1}^D, \mu_{\sigma_2}^D, \ldots, \mu_{\sigma_K}^D) \) and \( \bar{\mu}_A = (\mu_{\sigma_1}^A, \mu_{\sigma_2}^A, \ldots, \mu_{\sigma_K}^A) \).

If the molecule remains in a single-state \( \sigma_k \) throughout an entire exposure period \([t_n - \tau, t_n]\), then the detector is triggered by \( \mu_{\sigma_k}^D + \mu_{\text{back}}^D \) and \( \mu_{\sigma_k}^A + \mu_{\text{back}}^A \) 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 \( \mu_{\text{back}}^D + \mu_{\sigma_k}^D(\cdot) \) in the donor channel and \( \mu_{\text{back}}^A + \mu_{\sigma_k}^A(\cdot) \) in the acceptor channel over this exposure. Mathematically, this is equivalent to \( \mu_{\text{back}}^D \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\sigma_k}^D(\cdot) \) in the donor channel and \( \mu_{\text{back}}^A \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\sigma_k}^A(\cdot) \) in the acceptor channel.

With measurement noise, such as shot noise, quantification noise, or amplification noise in the case of EMCCD detectors 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.\(^{85–87}\) 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

\[
\begin{align*}
\mathbb{P}(w_n^D | \mathcal{F}(\cdot)) &= \text{Poisson} \left( \mu_{\text{back}}^D \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{F}(\cdot)}^D(t) \right) \\
\mathbb{P}(w_n^A | \mathcal{F}(\cdot)) &= \text{Poisson} \left( \mu_{\text{back}}^A \tau + \int_{t_n - \tau}^{t_n} dt \mu_{\mathcal{F}(\cdot)}^A(t) \right)
\end{align*}
\]  

(Equation 8)

(Full expression for \( \mathcal{F}(\cdot) \) is omitted for brevity.)

**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\(^{5,12,55–57}\) as well as the apparent FRET efficiency.\(^{5,12,55–57}\) Characteristic FRET efficiency\(^{5,12,55–57}\) labeled with \( \varepsilon_{\sigma_k} \) for \( k = 1, 2, \ldots, K \) is defined as follows:

\[
\varepsilon_{\sigma_k} = \frac{\mu_{\sigma_k}^A}{\mu_{\sigma_k}^D + \mu_{\sigma_k}^A}
\]  

(Equation 10)

and only depends on the conformational states.\(^{5,12,55–57}\) By contrast, the apparent FRET efficiency (which is often used as a proxy for \( \varepsilon_{\sigma_k} \)) is defined as follows:

\[
\varepsilon_n^D = \frac{w_n^A}{w_n^D + w_n^A}
\]  

(Equation 11)

for \( n = 1, 2, \ldots, N \). Unlike the former, the latter is affected by artifacts, such as measurement noise and background.\(^\text{12}\) As a result, the apparent efficiency may attain different values over time, even when the molecule does not switch its conformational state.

Our goal is to learn the initial probabilities, \( \rho_{\sigma_k} \); photon emission rates, \( \mu_{\sigma_k}^D, \mu_{\sigma_k}^A \); and switching rates, \( \lambda_{\sigma_k \sigma_k'} \) for all states and the trajectory of the system \( \mathcal{F}(\cdot) \) 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 a time-series analysis within the naive HMM paradigm.\(^\text{5}\) Subsequently, we will present how we tackle smFRET data using our continuous-time HMMs.

**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 \( \mathcal{F}(t_n) \). Therefore, we have the following approximations for the HMM formulation.
\[
\int_{t_n-\tau}^{t_n} d\mu_{\mathcal{T}}(\cdot) \approx \mu_{\mathcal{T}}(t_n) \tau
\]
during the exposure period \([t_n - \tau, t_n]\).

Consequently, Equations 8 and 9 become
\[
w_{nD}^D | \mathcal{T}(\cdot) \sim \text{Poisson} \left( \mu_{\text{back}} + \mu_{\mathcal{T}}(t_n) \tau \right)
\]
(Equation 12)
\[
w_{nA}^A | \mathcal{T}(\cdot) \sim \text{Poisson} \left( \mu_{\text{back}} + \mu_{\mathcal{T}}(t_n) \tau \right)
\]
(Equation 13)

Given the HMM formulation and the smFRET data \(w^D, w^A\), 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 \(n = 1, 2, \ldots, N\). We label a molecule’s state within the HMM paradigm at time \(t_n\) with \(T(t_n) = c_n\). 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, \(P_{c_{n-1} \rightarrow c_n} = c_n \) 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 \(\sigma_k\) to any other conformational state, including itself in
\[
P_{\sigma_k} = (P_{\sigma_k \rightarrow \sigma_1}, P_{\sigma_k \rightarrow \sigma_2}, \ldots, P_{\sigma_k \rightarrow \sigma_K})
\]
which is normalized as a probability vector (namely each component has a sum of 1). Eventually, we gather all transition probability vectors in a matrix called the transition-probability matrix, denoted by \(\overline{P}\), which reads as follows:

\[
\overline{P} = \begin{bmatrix}
\overline{P}_{\sigma_1} \\
\overline{P}_{\sigma_2} \\
\vdots \\
\overline{P}_{\sigma_K}
\end{bmatrix}
\]
(Equation 14)

There is a connection between the transition probability matrix \(\overline{P}\) and the molecule’s switching rates \(\lambda_{\sigma_k \rightarrow \sigma_k'}\) and escape rate \(\lambda_{\sigma_k}\), provided that we gather the switching rates and escape rates in the following matrix, called the generator,

\[
\overline{G} = \begin{bmatrix}
-\lambda_{\sigma_1} & \lambda_{\sigma_1 \rightarrow \sigma_2} & \cdots & \lambda_{\sigma_1 \rightarrow \sigma_K} \\
\lambda_{\sigma_2 \rightarrow \sigma_1} & -\lambda_{\sigma_2} & \cdots & \lambda_{\sigma_2 \rightarrow \sigma_K} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{\sigma_K \rightarrow \sigma_1} & \lambda_{\sigma_K \rightarrow \sigma_2} & \cdots & -\lambda_{\sigma_K}
\end{bmatrix}
\]
(Equation 15)

then, \(\overline{P}\) can be calculated based on \(\overline{P} = \exp(\overline{G} \Delta t)\) where \(\exp(\cdot)\) stands for the matrix exponential. Knowing \(\overline{G}\) determines \(\overline{P}\) uniquely; however, knowing \(\overline{P}\) provides only a proxy for \(\overline{G}\). Specifically, if we assume that \(\lambda_{\sigma_k} \Delta t \ll 1\) then we can use an approximation.
\[
\bar{P} \approx \bar{I} + \bar{G} \Delta t 
\]  
(Equation 16)

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

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

\[
c_0 \sim \text{Categorical} (\bar{\rho}) \quad \text{(Equation 17)}
\]

\[
c_n \mid c_{n-1} \sim \text{Categorical} (\bar{P}_{c_{n-1}}) \quad \text{(Equation 18)}
\]

\[
u_n^D \mid c_n \sim \text{Poisson} (\mu_{\text{back}}^D + \mu_n^D) \quad \text{(Equation 19)}
\]

\[
u_n^A \mid c_n \sim \text{Poisson} (\mu_{\text{back}}^A + \mu_n^A) \quad \text{(Equation 20)}
\]

From now on, we follow the Bayesian paradigm,\(^{62,63}\) which requires us to prescribe prior distributions for the parameters.

We start with the prior distributions placed on the transition probabilities \(\bar{P}_{\sigma_k}\) for all \(k = 1, 2, \ldots, K\) and \(\bar{\rho}\). We choose to place Dirichlet distributions with concentration parameters \(A\) and \(a\) for \(\bar{P}_{\sigma_k}\) and \(\bar{\rho}\), respectively, that are conjugate to the Categorical distribution\(^{23,28,60,71}\) and formulated as follows

\[
\bar{P}_{\sigma_1} \sim \text{Dirichlet} \left( A, A, \ldots, A \right) \quad \text{(Equation 21)}
\]

\[
\bar{P}_{\sigma_2} \sim \text{Dirichlet} \left( A, A, \ldots, A \right) \quad \text{(Equation 22)}
\]

\[
\vdots
\]

\[
\bar{P}_{\sigma_K} \sim \text{Dirichlet} \left( A, A, \ldots, A \right) \quad \text{(Equation 23)}
\]

\[
\bar{\rho} \sim \text{Dirichlet} \left( a, a, \ldots, a \right) \quad \text{(Equation 24)}
\]

Subsequently, we place priors on the photon emission rates \(\bar{\mu}^D = (\mu_n^D, \mu_{\sigma_2}^D, \ldots, \mu_{\sigma_K}^D)\) and \(\bar{\mu}^A = (\mu_n^A, \mu_{\sigma_2}^A, \ldots, \mu_{\sigma_K}^A)\). The prior that we choose to place is the gamma distribution because it has positive support, meaning that it has a range on positive real numbers.
with hyperparameters $\varphi^D, \psi^D, \varphi^A, \psi^A$.

Here, we note that prior distributions are used to place non-zero probability where we anticipate posterior probability density. Typically, we use at least $\approx 150$ data points to overcome the effects of bias from the prescribed prior distributions. This, in turn, allows us to converge to a posterior distribution centered on the ground-truth values.

We estimate the background photon emission rates $\mu^D_{\text{back}}, \mu^A_{\text{back}}$ by separate measurements that contain only ambient contributions, as we explain in supplemental information (5). 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:

$$P\left(\vec{\rho}, \vec{P}, \mu^D, \mu^A, \vec{\tau} \mid \vec{w}^D, \vec{w}^A\right) = P\left(\vec{\rho}, \vec{P}, \mu^D, \mu^A, \vec{\tau} \mid \vec{w}^D, \vec{w}^A\right).$$

(Equation 27)

Because we do not have an explicit formula for Equation 27, we build a custom Markov chain Monte Carlo (MCMC) sampling scheme to generate pseudorandom numbers from Equation 27. Details of the computational scheme are provided in Computational inference.

**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 Equations 8 and 9. Just as with the HMM, for the HMJP formulation, we also operate within the Bayesian paradigm and, thus, place prior distributions on $\vec{\rho}, \vec{P}, \vec{\tau}, \mu^D, \mu^A$.

First, we prescribe the prior distributions on the escape rates $\vec{\lambda} = \{\lambda_{\sigma_1}, \lambda_{\sigma_2}, \ldots, \lambda_{\sigma_K}\}$. For those, we choose the gamma distributions that are conjugate to the exponentially distributed holding times, i.e.,

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

(Equation 28)

for all $k = 1, 2\ldots, K$ with hyperparameters $\eta$ and $b$. Here, hyperparameters $\eta$ and $\frac{b}{\eta}$ are known as the shape and rate parameters of the gamma distribution, respectively. The product of these hyperparameters, which is $b$, sets the breadth of the gamma distribution.

Subsequently, we prescribe independent conjugate Dirichlet distributions on the transition probabilities $\pi_{\sigma_k}$ for all $k = 1, 2\ldots, K$:
Prior distributions placed on $\overline{\rho}$ and $\overline{\mu}^D$ and $\overline{\mu}^A$ are provided in Equations 24 and 26, respectively. The full posterior distribution for the HMJP formulation is

$$p(\overline{\rho}, \overline{\pi}, \lambda, \overline{\mu}^D, \overline{\mu}^A, \mathcal{S}(\cdot) | w^D, w^A) = p(\overline{\rho}, \overline{\pi}, \lambda, \overline{\mu}^D, \overline{\mu}^A, (\mathcal{S}, \mathcal{D}, M) | w^D, w^A).$$  \hspace{1cm} (Equation 32)

Because we do not have an analytical formula for Equation 32, we built a custom MCMC sampling scheme. Details of the computational scheme are provided in Computational inference.

**Computational inference**—MCMC sampling from the full posterior distributions provided in Equation 27 for the HMM and Equation 32 for the HMJP rely on Gibbs sampling.\(^{5,23,24,28,33,40,71,72}\) To form a large number of samples from these posteriors, we iterate the following:

1. Update $\overline{\mu}$ for the HMM or $(\mathcal{S}, \mathcal{D}, M)$ for the HMJP;
2. Update the transition probabilities, that is, $\overline{\sigma}_k$ for the HMM or $\overline{\pi}_K$ and $\lambda_{\sigma_K}$ for the HMJP;
3. Update the initial probability vector $\overline{\rho}$ for both the HMM and the HMJP; and
4. Update the photon emission rates $\overline{\mu}^D$ and $\overline{\mu}^A$ for both the HMM and the HMJP.

These samples allow us to reconstruct the posterior distribution $p(\overline{\rho}, \overline{\pi}, \lambda, \overline{\mu}^D, \overline{\mu}^A, \overline{\sigma} | w^D, w^A)$ for the HMM and $p(\overline{\rho}, \overline{\pi}, \overline{\lambda}, \overline{\mu}^D, \overline{\mu}^A, (\mathcal{S}, \mathcal{D}, M) | w^D, w^A)$ for the HMJP. The estimation of switching rates $\lambda_{\sigma_k} \rightarrow \sigma_k'$ with the HMJP is performed by Equations 5 and 16 for the HMM.

In supplemental information, we present the equation summaries for both HMJP and HMM formulations in Note S1. Furthermore, notations and parameter choices are provided in Tables S1 and S2. A working code of the implementation of our HMM and HMJP frameworks is made available through the authors’ website.
Acquisition of experimental data

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

The HJ strands used in this work, and whose results are shown in Figures 5, 6, 7, 8, and 9, were purchased from JBioS (Wako, Japan); of which, the sequences are given as follows:

R-strand: 5′-CGA TGA GCA CCG CTC GGC TCA ACT GGC AGT CG-3′
H-strand: 5′-CAT CT TAGT 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 AGG CCG CTG CTA CTA AGA TG-3′

We note that T (the H-strand) and T (the 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 μM, 30 μL) and the H-strand (1 μM, 20 μL) 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 (2–3°C/min) to room temperature. We used a sample chamber (Grace Bio-Labs SecureSeal, GBL621502) and a coverslip that was coated with Biotin-PEG-SV A (Biotin-poly(ethylene glycol)-succinimidyl valerate). Streptavidin (0.1 mg/mL in TN buffer, 100 μL) was incubated for 20 min, which was followed by washing with the TN buffer. The HJ solution (10 nM, 100 μL) was injected for 3 s. The chamber was rinsed three times by measuring the buffer (TN buffer with 10 mM MgCl2 and 2 mM Trolox).

Broadband light, generated by super continuum laser (Fianium SC-400-4, f = 40MHz), 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 IR60 ×, numerical aperture [NA] = 1.27). The excitation power was set to 20 μW 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 the donor and FF02-685/40-25 for the 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 with a time-correlated single photon counting (TCSPC) module (Becker&Hickl SPC-130-EM) in time-tagging mode.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.
ACKNOWLEDGMENTS

S.P. acknowledges support from an NSF CAREER grant MCB-1719537 and an NIH NIGMS grants R01GM134426 and R01GM130745. T.T. acknowledges support from JSPS KAKENHI grant JP19F19340. Arizona State University cluster AGA VE and Saguaro are the main computational resources used in this study.

REFERENCES

1. Gopich IV, and Szabo A (2009). Decoding the pattern of photon colors in single-molecule FRET. J. Phys. Chem. B 113, 10965–10973. [PubMed: 19588948]
2. Gopich IV, and Szabo A (2010). FRET efficiency distributions of multistate single molecules. J. Phys. Chem. B 114, 15221–15226. [PubMed: 21028764]
3. Pirchi M, Tsukanov R, Khamis R, Tomov TE, Berger Y, Khara DC, Volkov H, Haran G, and Nir E (2016). Photon-by-photon hidden Markov model analysis for microsecond single-molecule FRET kinetics. J. Phys. Chem. B 120, 13065–13075. [PubMed: 27977207]
4. Schuler B (2005). Single-molecule fluorescence spectroscopy of protein folding. ChemPhysChem 6, 1206–1220. [PubMed: 15991265]
5. Sgouralis I, Madaan S, Djutanta F, Kha R, Hariadi RF, and Pressé S (2019). A bayesian nonparametric approach to single molecule Förster resonance energy transfer. J. Phys. Chem. B 123, 675–688. [PubMed: 30571128]
6. Sekar RB, and Periasamy A (2003). Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. J. Cell Biol 160, 629–633. [PubMed: 12615908]
7. Hellenkamp B, Schmid S, Doroshenko O, Opanasyuk O, Kühnemuth R, Rezaei Adariani S, Ambrose B, Azmauriyan M, Barth A, Birkedal V, et al. (2018). Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. Nat. Methods 15, 669–676. [PubMed: 30171252]
8. Ha T, and Tinnefeld P (2012). Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging. Annu. Rev. Phys. Chem. 63, 595–617. [PubMed: 22404588]
9. Andrec M, Levy RM, and Talaga DS (2003). Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden Markov models. J. Phys. Chem. A 107, 7454–7464. [PubMed: 19626138]
10. Keller BG, Kobitski A, Jäschke A, Nienhaus GU, and Noé F (2014). Complex RNA folding kinetics revealed by single-molecule FRET and hidden Markov models. J. Am. Chem. Soc 136, 4534–4543. [PubMed: 24568646]
11. Förster T (1965). Delocalization excitation and excitation transfer. In Modern Quantum Chemistry, Sunanoglu O, ed. (Academic Press), pp. 93–137.
12. Roy R, Hohng S, and Ha T (2008). A practical guide to single-molecule FRET. Nat. Methods 5, 507–516. [PubMed: 18511918]
13. Gadella TW (2011). FRET and FLIM Techniques, Volume 33, First Edition (Elsevier Science).
14. Periasamy A, and Day R (2005). Molecular Imaging: FRET Microscopy and Spectroscopy (Academic Press).
15. Harris DC (2010). Quantitative Chemical Analysis, Eighth Edition (W.H. Freeman).
16. Tavakoli M, Jazani S, Sgouralis I, Shafraz OM, Sivasankar S, Donaphon B, Levitus M, and Pressé S (2020). Pitching single-focus confocal data analysis one photon at a time with Bayesian nonparametrics. Phys. Rev. X 10, 011021.
17. Gopich IV, and Szabo A (2003). Single-macromolecule fluorescence resonance energy transfer and free-energy profiles. J. Phys. Chem. B 107, 5058–5063.
18. Antonik M, Felekyan S, Gaiduk A, and Seidel CA (2006). Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. J. Phys. Chem. B 110, 6970–6978. [PubMed: 16571010]
19. Nir E, Micheletti X, Hamadani KM, Laurence TA, Neuhauser D, Kovchegov Y, and Weiss S (2006). Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. J. Phys. Chem. B 110, 22103–22124. [PubMed: 17078646]
20. Schuler B (2013). Single-molecule FRET of protein structure and dynamics—a primer. J. Nanobiotechnology 11 (Suppl 1), S2. [PubMed: 24565277]
21. Gopich I, and Szabo A (2005). Theory of photon statistics in single-molecule Förster resonance energy transfer. J. Chem. Phys 122, 14707. [PubMed: 15638691]
22. Gopich IV, and Szabo A (2007). Single-molecule FRET with diffusion and conformational dynamics. J. Phys. Chem. B 111, 12925–12932. [PubMed: 17929964]
23. Spouralis I, and Pressé S (2017). Icon: an adaptation of infinite hmsms for time traces with drift. Biophys. J 112, 2117–2126. [PubMed: 28538149]
24. Tavakoli M, Taylor JN, Li C-B, Komatsuzaki T, and Pressé S (2016). Single molecule data analysis: an introduction. arXiv, arXiv:1606.00403v2. https://arxiv.org/abs/1606.00403
25. Hajdziona M, and Molski A (2011). Maximum likelihood-based analysis of single-molecule photon arrival trajectories. J. Chem. Phys 134, 054112. [PubMed: 21303097]
26. Pressé S, Lee J, and Dill KA (2013). Extracting conformational memory from single-molecule kinetic data. J. Phys. Chem. B 117, 495–502. [PubMed: 23259771]
27. Pressé S, Peterson J, Lee J, Elms P, MacCallum JL, Marqusee S, Bustamante C, and Dill K (2014). Single molecule conformational memory extraction: p5ab RNA hairpin. J. Phys. Chem. B 118, 6597–6603. [PubMed: 24898871]
28. Spouralis I, and Pressé S (2017). An introduction to infinite hmsms for single-molecule data analysis. Biophys. J 112, 2021–2029. [PubMed: 28538142]
29. Kilic S, Felekyan S, Doroshenko O, Boichenko I, Dimura M, Vardanyan H, Bryan LC, Arya G, Seidel CAM, and Fierz B (2018). Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1a. Nat. Commun 9, 235. [PubMed: 29339721]
30. Kilic Z, Spouralis I, and Pressé S (2021). Residence time analysis of RNA polymerase transcription dynamics: a Bayesian sticky HMM approach. Biophys. J S0006–3495, 00207–1.
31. Kinz-Thompson CD, and Gonzalez RL Jr. (2018). Increasing the time resolution of single-molecule experiments with bayesian inference. Biophys. J 114, 289–300. [PubMed: 29401427]
32. Stigler J, and Rief M (2012). Hidden Markov analysis of trajectories in single-molecule experiments and the effects of missed events. ChemPhysChem 13, 1079–1086. [PubMed: 22392881]
33. Spouralis I, Whitmore M, Lapidus L, Comstock MJ, and Pressé S (2018). Single molecule force spectroscopy at high data acquisition: a Bayesian nonparametric analysis. J. Chem. Phys 148, 123320. [PubMed: 29604816]
34. Bishop CM (2006). Pattern Recognition and Machine Learning (Springer).
35. Rabiner LR (1989). A tutorial on hidden markov models and selected applications in speech recognition. Proc. IEEE 77, 257–286.
36. Rabiner L, and Juang B (1986). An introduction to hidden Markov models. IEEE ASSP Mag. 3, 4–16.
37. Crochiere RE, and Rabiner LR (1996). Multirate digital signal processing, First Edition (Prentice-Hall).
38. Kilic Z, and Pressé S (2019). Bayesian nonparametric analysis of transcriptional processes. Biophys. J 116, 299a.
39. Kilic Z, Spouralis I, and Pressé S (2020). RNA polymerase dynamics and other single-molecule continuous time problems. Biophys. J 118, 544a.
40. Rao V, and Teh YW (2013). Fast mcmc sampling for markov jump processes and extensions. J. Mach. Learn. Res 14, 3295–3320.
41. Zhang B, and Rao V (2020). Efficient parameter sampling for markov jump processes. J. Comput. Graph. Stat 0, 1–18.
42. Chung HS, and Eaton WA (2013). Single-molecule fluorescence probes dynamics of barrier crossing. Nature 502, 685–688. [PubMed: 24153185]
43. Yoo J, Kim J-Y, Louis JM, Gopich IV, and Chung HS (2020). Fast three-color single-molecule FRET using statistical inference. Nat. Commun 11, 3336. [PubMed: 32620782]
44. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, and Rubin DB (2013). Bayesian Data Analysis, Third Edition (CRC Press).
45. Gopich IV (2015). Accuracy of maximum likelihood estimates of a two-state model in single-molecule FRET. J. Chem. Phys 142, 034110. [PubMed: 25612692]
46. Metzner P, Horenko I, Schütte C, and Schütte C (2007). Generator estimation of Markov jump processes based on incomplete observations nonequidistant in time. Phys. Rev. E Stat. Nonlin. Soft Matter Phys 76, 066702. [PubMed: 18233938]
47. Burzyskowski T, Szubiakowski JP, and Ryden T (2003). Statistical analysis of data from single molecule experiment. In Proceedings of the IV Workshop on Atomic and Molecular Physics, 5258, Heldt J, ed, pp. 171–177.
48. Hobolth A, and Stone EA (2009). Simulation from endpoint-conditioned, continuous-time markov chains on a finite state space, with applications to molecular evolution. Ann. Appl. Stat 3, 1204. [PubMed: 20148133]
49. Huggins JH, Narasimhan K, Saeedi A, and Mansinghka VK (2015). Jump-means: small-variance asymptotics for markov jump processes. arXiv, arXiv:1503.00332. https://arxiv.org/abs/1503.00332.
50. Cinlar E (2013). Introduction to stochastic processes (Courier Corporation).
51. Beentjes CHL, and Baker RE (2019). Uniformization techniques for stochastic simulation of chemical reaction networks. J. Chem. Phys 150, 154107. [PubMed: 31005097]
52. Gillespie DT (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys 22, 403–434.
53. Hirsch M, Wareham RJ, Martin-Fernandez ML, Hobson MP, and Rolfe DJ (2013). A stochastic model for electron multiplication charge-coupled devices—from theory to practice. PLoS ONE 8, e53671. [PubMed: 23382848]
54. Kim JY, Kim C, and Lee NK (2015). Real-time submillisecond single-molecule FRET dynamics of freely diffusing molecules with liposome tethering. Nat. Commun 6, 6992. [PubMed: 25908552]
55. Wasserman MR, Alejo JL, Altman RB, and Blanchard SC (2016). Multiperspective smFRET reveals rate-determining late intermediates of ribosomal translocation. Nat. Struct. Mol. Biol 23, 333–341. [PubMed: 26926435]
56. Lu M, Ma X, Castillo-Menendez LR, Gorman J, Alsahafi N, Ermel U, Terry DS, Chambers M, Peng D, Zhang B, et al. (2019). Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. Nature 568, 415–419. [PubMed: 30971821]
57. Mazouchi A, Zhang Z, Bahram A, Gomes G-N, Lin H, Song J, Chan HS, Forman-Kay JD, and Gradinaru CC (2016). Conformations of a metastable sh3 domain characterized by smfret and an excluded-volume polymer model. Biophys. J 110, 1510–1522. [PubMed: 27074677]
58. Krishnamoorti A, Cheng RC, Berka V, and Maduke M (2019). C1c conformational landscape as studied by smfret. Biophys. J 116, 555a.
59. Hohng S, Zhou R, Nahas MK, Yu J, Schulten K, Lilley DM, and Ha T (2007). Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. Science 318, 279–283. [PubMed: 17932299]
60. Robert C, and Casella G (2013). Monte Carlo statistical methods (Springer Science & Business Media).
61. Kilic Z, Sgouralis I, and Pressé S (2021). Generalizing hmms to continuous time for fast kinetics: Hidden markov jump processes. Biophys. J 120, 409–423. [PubMed: 33421415]
62. Sivia D, and Skilling J (2006). Data Analysis: A Bayesian Tutorial, Second Edition (Oxford University Press).
63. Lee ET, and Wang J (2003). Statistical Methods for Survival Data Analysis, Third Edition Volume 476 (Wiley).
64. Qin F, Auerbach A, and Sachs F (1996). Estimating single-channel kinetic parameters from idealized patch-clamp data containing missed events. Biophys. J 70, 264–280. [PubMed: 8770203]
65. Roux B, and Sauvé R (1985). A general solution to the time interval omission problem applied to single channel analysis. Biophys. J 48, 149–158. [PubMed: 2410047]
66. Rollins GC, Shin JY, Bustamante C, and Pressé S (2015). Stochastic approach to the molecular counting problem in superresolution microscopy. Proc. Natl. Acad. Sci. USA 112, E110–E118. [PubMed: 25535361]
67. Bryan JS 4th, Sgouralis I, and Pressé S (2020). Inferring effective forces for Langevin dynamics using Gaussian processes. J. Chem. Phys 152, 124106. [PubMed: 32241120]

68. Zosel F, Mercadante D, Nettels D, and Schuler B (2018). A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction. Nat. Commun 9, 3332. [PubMed: 30127362]

69. Juette MF, Terry DS, Wasserman MR, Altman RB, Zhou Z, Zhao H, and Blanchard SC (2016). Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. Nat. Methods 13, 341–344. [PubMed: 26878382]

70. Huang F, Hartwich TM, Rivera-Molina FE, Lin Y, Duim WC, Long JJ, Uchil PD, Myers JR, Baird MA, Mothes W, et al. (2013). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. Nat. Methods 10, 653–658. [PubMed: 23708387]

71. Jazani S, Sgouralis I, Shafraz OM, Levitus M, Sivasankar S, and Pressé S (2019). An alternative framework for fluorescence correlation spectroscopy. Nat. Commun 10, 3662. [PubMed: 31413259]

72. Jazani S, Sgouralis I, and Pressé S (2019). A method for single molecule tracking using a conventional single-focus confocal setup. J. Chem. Phys 150, 114108. [PubMed: 30902018]

73. Georgoulas A, Hillston J, and Sanguinetti G (2017). Unbiased Bayesian inference for population Markov jump processes via random truncations. Stat. Comput 27, 991–1002. [PubMed: 28690370]

74. Gillespie DT (1977). Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem 81, 2340–2361.

75. Buelens B, Daas P, Burger J, Puts M, and van den Brakel J (2014). Selectivity of Big Data. In Proceedings of the Statistics Netherlands Discussion Paper 201411 (Statistics Netherlands).

76. Papoulis A, and Pillai SU (2002). Probability, Random Variables, and Stochastic Processes, Fourth Edition (McGraw Hill).

77. Kalinin S, Sisamakis E, Magennis SW, Felekyan S, and Seidel CA (2010). On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits. J. Phys. Chem. B 114, 6197–6206. [PubMed: 20397670]

78. Mortensen KI, Churchman LS, Spudich JA, and Flyvbjerg H (2010). Optimized localization analysis for single-molecule tracking and super-resolution microscopy. Nat. Methods 7, 377–381. [PubMed: 20364147]

79. Charrière F, Colomb T, Montfort F, Cuche E, Marquet P, and Depeursinge C (2006). Shot-noise influence on the reconstructed phase image signal-to-noise ratio in digital holographic microscopy. Appl. Opt 45, 7667–7673. [PubMed: 17068602]

80. Gross M, Goy P, and Al-Koussa M (2003). Shot-noise detection of ultrasound-tagged photons in ultrasound-modulated optical imaging. Opt. Lett 28, 2482–2484. [PubMed: 14690121]

81. Krishnaswami V, Van Noorden CJ, Manders EM, and Hoebe RA (2014). Towards digital photon counting cameras for single-molecule optical nanoscopy. Opt. Nanoscopy 3, 1.

82. Lin Y, Long JJ, Huang F, Duim WC, Kirschbaum S, Zhang Y, Schroeder LK, Rebane AA, Velasco MGM, Virrueta A, et al. (2015). Quantifying and optimizing single-molecule switching nanoscopy at high speeds. PLoS ONE 10, e0128135. [PubMed: 26011109]

83. Yoshida T, Alfaqan S, Sasaoka N, and Imamura H (2017). Application of FRET-based biosensor “ateam” for visualization of atp levels in the mitochondrial matrix of living mammalian cells. Methods Mol. Biol 1567, 231–243. [PubMed: 28276022]

84. Lee J, Park S, and Hohng S (2018). Accelerated FRET-PAINT microscopy. Mol. Brain 11, 70. [PubMed: 30466455]

85. Little MA (2019). Machine Learning for Signal Processing: Data Science, Algorithms, and Computational Statistics (Oxford University Press).

86. Lee A, Tsekouras K, Calderon C, Bustamante C, and Pressé S (2017). Unraveling the thousand word picture: an introduction to super-resolution data analysis. Chem. Rev 117, 7276–7330. [PubMed: 28414216]

87. Booz Allen Hamilton (2015). The Field Guide to Data Science, Second Edition (Booz Allen Hamilton).

88. Chandraadoss SD, Haagsma AC, Lee YK, Hwang J-H, Nam J-M, and Joo C (2014). Surface passivation for single-molecule protein studies. J. Vis. Exp 86, 50549.
Highlights

Generalization of HMMs is now available for processes evolving continuously in time.

With newly developed Markov jump processes, rapid kinetics can be extracted from smFRET.

Kinetics faster than the data acquisition rate for integrative detectors are acquired.
Figure 1. An illustration of single molecule switching kinetics and corresponding measurements (A–D) We provide the simulated trajectory (cyan) of the single molecule between two states (σ1 and σ2) (A and B). The photon emission rates, i.e., the number of emitted photons per unit of time in the absence of noise in both donor and acceptor channels, associated with the conformation states σ1 and σ2, are labeled with μσ1 and μσ2, respectively. The simulated experiment provided in this figure starts at Δt = 0.05 s and ends at tN = 10 s with the data acquisition period being Δt = 0.05 s. Here, in (A) and (C), for visual purposes, we assume that the measurements are acquired by a detector with a fixed exposure period of τ = 50 ms in donor (green) and acceptor (red) channels. As the molecule switches between states (σ1 and σ2) during an integration period, the measurements represent the number of total emitted photons and reflect the rates μσ1, μσ2, which correspond to the visited conformational states. (A) and (B) represent the simulated data in which the molecule-switching kinetics between the conformation states are slower than the data acquisition rate. On the other hand, in (C) and (D), we demonstrate a single-molecule trajectory (cyan) when the molecule’s switching...
kinetics are faster than the data acquisition rate. In (B), the slow kinetics of the molecule give rise to well-separated state-occupancy histograms in donor and acceptor channels around the average photon emission rates. By contrast, in (D), we don’t observe well-separated histograms because of the fast-switching kinetics of the molecule.
Figure 2. An illustration of experimental smFRET measurements for data acquisition periods $\Delta t = 0.05$ s

(A and B) We provide the measurements for the simulated trajectory of the single molecule between three states ($\sigma_1$, $\sigma_2$, and $\sigma_3$) with the same color scheme as shown in Figure 1 (A). Here, we assume that the measurements are acquired by integrative detectors with fixed exposure periods that coincide with a data acquisition period of $\tau = 0.05$ s in the donor (green) and acceptor (red) channels.
Figure 3. The HMJP and HMM photon emission rate, escape rate, and transition-probability estimates for fast-switching kinetics in simulated measurements.

(A and B) Here, we provide the 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 (A) and (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 (also known as credible intervals) and the true photon emission rates (cyan line).

(C–F) We superposed the posterior distributions over the 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 (cyan lines), respectively. (C) and (D) correspond to the posterior distribution of the escape rates that are labeled with Kilic et al. Page 30

Cell Rep Phys Sci. Author manuscript; available in PMC 2021 June 16.
\( \lambda_{\sigma_k} \rightarrow \sigma_{k'} \) for all \( k, k' = 1, 2 \) with \( k \neq k' \). (E) and (F) 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 Figures 1C and 1D.
Figure 4. HMJP and HMM photon emission rate, escape rate and transition probability estimates for fast switching kinetics in simulated measurements for 3 conformational states (A–H) 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 faster than the data acquisition rate 1/Δt for three conformational states. We expect HMMs to perform poorly in estimating the true photon emission rates, escape rates, and transition probabilities when the system switching is fast. Here, we provide the posterior distributions with the same color convention as in Figure 3. Simulated measurements used here are generated with the same parameters as those provided in Figures 2A and 2B.
Figure 5. An illustration of experimental smFRET measurements for data acquisition periods $\Delta t = 0.025$ s, $\Delta t = 0.05$ s, and $\Delta t = 0.1$ s.

(A–F) In (A), (C), and (E), we provide the measurements for the same smFRET experiment coinciding with the data acquisition periods $\Delta t = 0.025$ s, $\Delta t = 0.05$ 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 $\tau = 0.025$ s, $\tau = 0.05$ s, and $\tau = 0.1$ s in donor (green) and acceptor (red) channels. (B), (D), and (F) are especially interesting to our analysis. Here, we don’t observe well-separated histograms because of the fast-switching kinetics of the molecule, as shown in Figure 1D for simulated experiments.
Figure 6. HMJP and HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with $\Delta t = 0.025$ s (A–F) Here, we provide posterior photon emission rate (A and B), escape rate (C and D), and transition probability (E and F) estimates obtained with HMJP and HMM for the experimental data shown in Figure 5A. Here, we follow a similar color convention to that of Figure 3. However, unlike Figure 4, ground-truth information is not available for the photon emission rates, escape rates, and transition probabilities because the analysis is performed with experimental data.
Figure 7. HMM photon emission rate, escape rate and transition probability estimates for experimental smFRET measurements with Δt = 0.025 s, Δt = 0.05 s, and Δt = 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 Figures 5A, 5C, and 5E. We do not expect HMM posterior estimates associated with three different exposure periods to be consistent because of the switching-kinetics approach for the data acquisition rate.

(A and B) We superposed the posterior distributions for the measurements associated with exposure periods Δt = 0.025 s, Δt = 0.05 s, and Δt = 0.1 s over the photon emission rates for HMM, along with their 95% confidence intervals.

(C and D) We superposed the posterior distributions over the escape rates along with their 95% confidence intervals.

(E and F) 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 $\Delta t = 0.025$ s, $\Delta t = 0.05$ s, and $\Delta t = 0.1$ s are demonstrated in green, blue, and purple, respectively.
Here, we provide posterior photon emission rate, escape rate, and diagonal transition probability estimates obtained with HMJP for the measurements shown in Figures 5A, 5C, and 5E. We expect HMJP posterior estimates associated with three different exposure periods to be consistent across examples, even if the switching kinetics exceed data acquisition rate.

(A and B) We superposed the posterior distributions for the measurements associated with exposure periods $\Delta t = 0.025$ s, $\Delta t = 0.05$ s, and $\Delta t = 0.1$ s over the photon emission rates for the HMJP along with their 95% confidence intervals.

(C and D) We superposed the posterior distributions over the escape rates along with their 95% confidence intervals.

(E and F) We superposed the posterior distributions over transition probabilities and their 95% confidence intervals. Here, we follow a similar color convention as that in Figure 7.
Figure 9. HMM and HMJP FRET efficiency estimates for experimental smFRET measurements with Δt = 0.025 s, Δt = 0.05 s, and Δt = 0.1 s

(A and B) Here, we provide posterior FRET efficiency estimates obtained with HMM and HMJP for the measurements shown in Figures 5A, 5C, and 5E. We expect inconsistent posterior FRET efficiency estimates for HMMs associated with three different exposure periods (A). On the other hand, we expect HMJPs to provide consistent posterior estimates over FRET efficiencies for three different exposure periods (B). Here, in (A) and (B), we superposed the posterior distributions over the FRET efficiencies along with the apparent FRET efficiency (yellow). In this figure, we follow a similar color convention to that used in Figure 7.