Single Photon smFRET. III. Application to Pulsed Illumination

Matthew Safar[1,2], Ayush Saurabh[1,3], Bidyut Sarkar[4,5], Mohamadreza Fazel[1,3], Kunihiko Ishii[4,5], Tahei Tahara[4,5], Ioannis Sgouralis[6], and Steve Pressé[1,3,7]

[1]Center for Biological Physics,
Arizona State University, Tempe, AZ, USA
[2]Department of Mathematics and Statistical Science,
Arizona State University, Tempe, AZ, USA
[3]Department of Physics,
Arizona State University, Tempe, AZ, USA
[4]Molecular Spectroscopy Laboratory, RIKEN, 2-1 Hirosawa,
Wako, Saitama 351-0198, Japan
[5]Ultrafast Spectroscopy Research Team, RIKEN Center for Advanced Photonics (RAP), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
[6]Department of Mathematics, University of Tennessee Knoxville,
Knoxville, TN, USA
[7]School of Molecular Sciences, Arizona State University,
Phoenix, AZ, USA

November 8, 2022

Contents

1 Terminology Convention **2**

2 Introduction **3**

3 Forward Model and Inverse Strategy **4**
3.1 Inference Procedure: Parametric Sampler 7
3.2 Inference Procedure: Nonparametrics Sampler 8

4 Results **9**
4.1 Simulated Data Analysis 9
4.2 Experimental Data Analysis: Holliday Junction (HJ) 12

4.3 Experimental Data Acquisition 14

5 Discussion **15**

6 Code Availability **16**

7 Acknowledgments **16**

Bibliography **17**

Abstract

Förster resonance energy transfer (FRET) using pulsed illumination has been pivotal in leveraging lifetime information in FRET analysis. However, there remain major challenges in quantitative single photon, single molecule FRET (smFRET) data analysis under pulsed illumination including: 1) simultaneously deducing kinetics and number of system states; 2) providing uncertainties over estimates, particularly uncertainty over the number of system states; 3) taking into account detector noise sources such as crosstalk, and the instrument response function contributing to uncertainty; in addition to 4) other experimental noise sources such as background. Here, we implement the Bayesian nonparametric framework described in the first companion manuscript that addresses all aforementioned issues in smFRET data analysis specialized for the case of pulsed illumination. Furthermore, we apply our method to both synthetic as well as experimental data acquired using Holliday junctions.

Why It Matters

In the first companion manuscript of this series, we developed new methods to analyze noisy smFRET data. These methods eliminate the requirement of *a priori* specifying the dimensionality of the physical model describing a molecular complex's kinetics. Here, we apply these methods to experimentally obtained datasets with samples illuminated by laser pulses at regular time intervals. In particular, we study conformational dynamics of Holliday junctions.

1 Terminology Convention

To be consistent throughout our three part manuscript, we precisely define some terms as follows

1. a macromolecular complex under study is always referred to as a *system*,
2. the configurations through which a system transitions are termed *system states*, typically labeled using σ,
3. FRET dyes undergo quantum mechanical transitions between *photophysical states*, typically labeled using ψ,

4. a system-FRET combination is always referred to as a *composite*,
5. a composite undergoes transitions among its *superstates*, typically labeled using ϕ,
6. all transition rates are typically labeled using λ,
7. the symbol N is generally used to represent the total number of discretized time windows, typically labeled with n, and
8. the symbol w_n is generally used to represent the observations in the n-th time window.

2 Introduction

Amongst the many fluorescence methods available [1–7], single molecule Förster resonance energy transfer (smFRET) has been useful in probing interactions and conformational changes on nanometer scales [8–12]. This is typically achieved by estimating FRET efficiencies (and system states) at all instants of an smFRET trace and subsequently estimating transition rates. Furthermore, among different FRET modalities, FRET efficiencies are most accurately determined under pulsed illumination [13–15], where the FRET dyes are illuminated by short laser bursts at known times.

Under this illumination procedure, photon arrival times are recorded with respect to the immediately preceding pulse, thereby facilitating accurate estimation of fluorescence lifetimes as well as FRET rates. As such, in this manuscript, we will focus on single photon smFRET analysis under pulsed illumination.

Under pulsed illumination, information on kinetic parameters present in smFRET data is traditionally learned by: binned photon methods thereby eliminating lifetime information altogether [16–18]; bulk correlative methods [19–21]; and single photon methods [14, 22, 23]. However, these methods are parametric, *i.e.*, require fixing the number of system states *a priori*, and necessarily only learn system kinetics even though information on the number of system states is encoded in the data.

In this paper, we implement a general smFRET analysis framework presented in the Sec. 2.5.1 of the first companion manuscript [24] for the case of pulsed illumination to learn full distributions. In other words, probability distributions over parameters taking into account uncertainties from all existing sources such as crosstalk and background. These parameters include the system transition probabilities, and photophysical rates, that is, donor and acceptor relaxation and FRET rates, with special attention paid to uncertainty arising from sources such as inherent stochasticity in photon arrival times and detectors. As our main concern is deducing the number of system states using single photon arrivals while incorporating detector effects, we leverage the formalism of infinite hidden Markov models (iHMM) [25–30] within the Bayesian nonparametric (BNP) paradigm [25, 26, 31–38]. The iHMM framework assumes an *a priori* infinite number of system states with associated transition probabilities, where the number of system states warranted by input data is enumerated by those states most visited over the course of the system state trajectory.

Next, to benchmark our BNP-FRET sampler, we analyzed synthetic and experimental smFRET data acquired using a single confocal microscope with pulsed illumination optimized to excite donor dyes.

In particular, we employ a broad range of experimental data acquired from Holliday junctions (HJ) with an array of different kinetic rates due to varying buffer concentration of $MgCl_2$ [39–42].

3 Forward Model and Inverse Strategy

In this section, we first briefly illustrate the adaptation of the general formalism described in our first companion manuscript [24] to the pulsed illumination case. Next, we present a specialized inference procedure. The details of the framework not provided herein can be found in the Supplementary Information.

As before, we consider a molecular complex labeled with a donor-acceptor FRET pair. As the molecular complex transitions through its M_σ system states indexed by $\sigma_{1:M_\sigma}$, laser pulses (optimized to excite the donor) separated by time τ may excite either the donor or acceptor to drive transitions among the photophysical states, $\psi_{1:M_\psi}$, as defined in the first companion manuscript [24]. Such photophysical transitions lead to photon emissions that may be detected in either donor or acceptor channels. The set of N observations, *e.g.*, photon arrival times, from N pulses are recorded as

$$\boldsymbol{w} = \{w_1, w_2, \ldots, w_N\}. \tag{1}$$

Here, each individual measurement is a pair $w_n = (\mu_n^d, \mu_n^a)$, where μ_n^d and μ_n^a are the recorded arrival times (also known as microtimes) after the n-th pulse in both donor and acceptor channels, respectively. In cases where there is no photon detection, we denote the absent microtimes with $\mu_n^d = \varnothing$ and $\mu_n^a = \varnothing$ for donor and acceptor channels, respectively.

As is clear from Fig. 1, smFRET traces are inherently stochastic due to the nature of photon excitation, emission, and noise introduced by detector electronics. To analyze such stochastic systems, we begin with the most generic likelihood derived in Eq. 51 of the first companion manuscript [24]

$$L \propto \boldsymbol{\rho}_{start}\ \mathbf{Q}_1\ \ldots\ \mathbf{Q}_n\ \ldots \mathbf{Q}_N\ \boldsymbol{\rho}_{norm}^T, \tag{2}$$

where $\boldsymbol{\rho}_{start}$ is the initial probability vector for the system-FRET composite to be in one of M $(= M_\psi \times M_\sigma)$ superstates, and $\boldsymbol{\rho}_{norm}$ is a vector that sums the elements of the propagated probability vector. Here, we recall that $\mathbf{Q}_n$ is the transition probability matrix between pulses n and $n+1$, characterizing system-FRET composite transitions among superstates.

The propagators $\mathbf{Q}_n$ above adopt different forms depending on whether a photon is detected or not during the associated period. Their most general forms are derived in Sec. 2.5.1 of the first companion manuscript [24]. However, these propagators involve computationally expensive integrals and thus we make a few approximations here as follows: 1) we assume that the system state remains the same over an interpulse period since typical system kinetic timescales (typically 1 ms or more) are much longer than interpulse periods ($\approx$ 100 ns) [41, 43]; 2) the interpulse period ($\approx$ 100 ns) is longer than the donor and acceptor lifetimes ($\approx$ a few ns) [41, 43] such that they relax to the ground state before the next pulse. Furthermore, we will demonstrate a specialized sampling scheme under these physically motivated approximations.

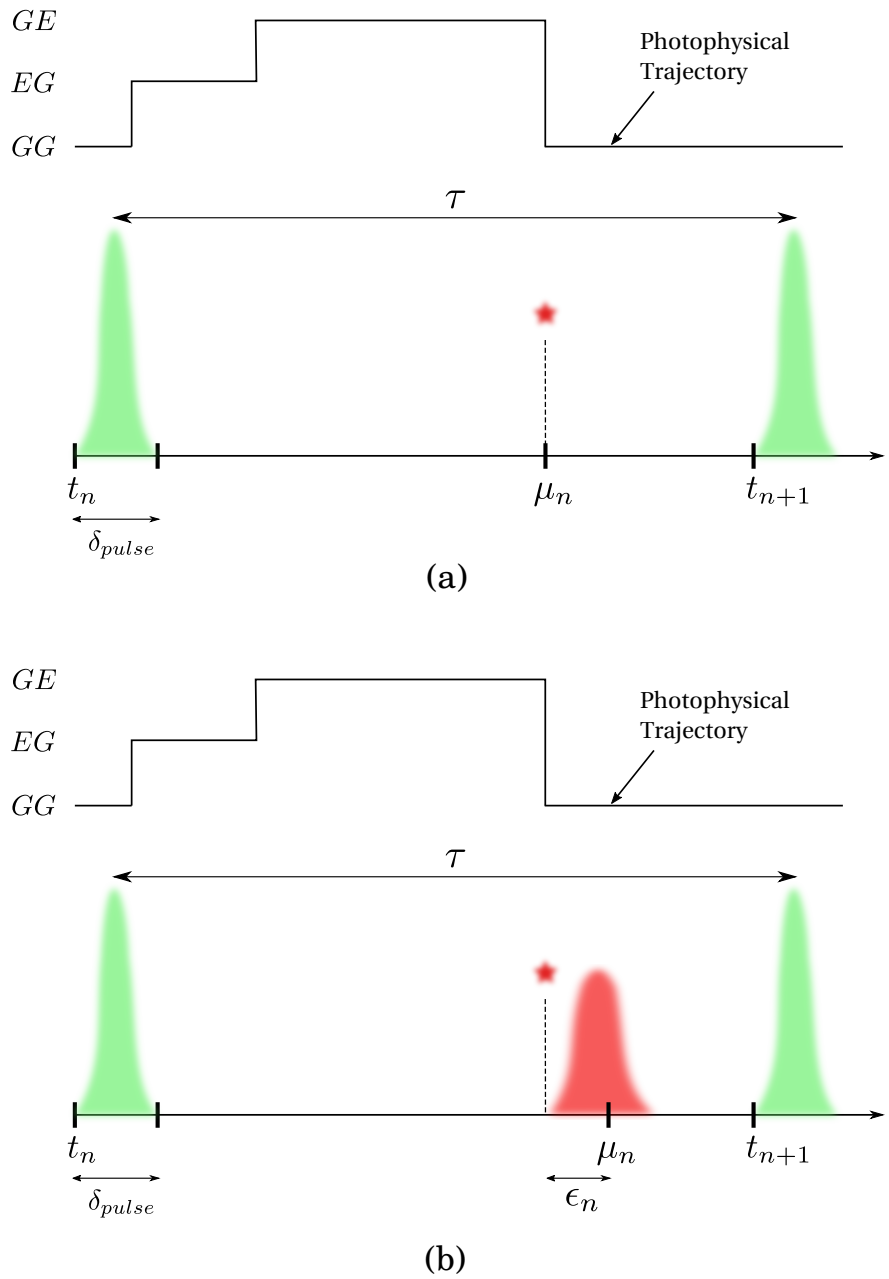

Figure 1: **Events over a pulsed illumination experiment pulse window**. Here, the beginning of the n-th interpulse window of size τ is marked by time t_n. The FRET labels originally in state GG (donor and acceptor, respectively, in ground states) are excited by a high intensity burst (shown by the green) to the state EG (only donor excited) for a very short time δ_{pulse}. If FRET occurs, the donor transfers its energy to the acceptor and resides in the ground state leaving the FRET labels in the GE state (only acceptor excited). The acceptor then emits a photon to be registered by the detector at microtime μ_n. When using ideal detectors, the microtime is the same as the photon emission time as shown in panel (a). However, when the timing hardware has jitter (shown in red), a small delay ϵ_n is added to the microtime as shown in panel (b). For convenience, we have reproduced this figure from our first companion manuscript [24].

The immediate implications of assumption (1) are that the system transitions may now, to a good approximation, only occur at the beginning of each pulse. Consequently, the evolution of the FRET pair between two consecutive pulses is now exclusively photophysical as the system state remains the same during interpulse times. As such, the system now evolves in equally spaced discrete time steps of size τ where the system state trajectory can be written as

$$s_{1:N} = \{s_1, s_2, \ldots, s_n, \ldots, s_{N-1}, s_N\},$$

where s_n is the system state between pulses n and $n+1$. The stochastic evolution of the system states in such discrete steps is then determined by the transition probability matrix designated by $\mathbf{\Pi}_\sigma$. For example, in the simplest case of a molecular complex with two system states $\sigma_{1:2}$, this matrix is computed as follows

$$\mathbf{\Pi}_\sigma = \mathbf{exp}\left(\tau \begin{bmatrix} * & \lambda_{\sigma_1\to\sigma_2} \\ \lambda_{\sigma_2\to\sigma_2} & * \end{bmatrix}\right) = \begin{bmatrix} \pi_{\sigma_1\to\sigma_1} & \pi_{\sigma_1\to\sigma_2} \\ \pi_{\sigma_2\to\sigma_1} & \pi_{\sigma_2\to\sigma_2} \end{bmatrix}, \tag{3}$$

where the matrix in the exponential contains transition rates among the system states and the $*$ represents the negative row sum.

Next, by assumption 2, we can further suppose that the fluorophores always start in the ground state at the beginning of every pulse. As a result, we treat pulses independently and write the probability of observation w_n as

$$p(w_n|s_n, \mathbf{G}_\psi) = \boldsymbol{\rho}_{ground}\mathbf{Q}_n^\psi(s_n)\boldsymbol{\rho}_{norm}^T, \tag{4}$$

where $\boldsymbol{\rho}_{ground}$ denotes the probability vector when the FRET pair is in the ground state at the beginning of each pulse, $\mathbf{G}_\psi$ is the generator matrix with only photophysical transition rates, and $\mathbf{Q}_n^\psi(s_n)$ is the photophysical propagator for the n-th interpulse period.

We further organize the observation probabilities of Eq. 4 into a newly defined detection matrix $\mathbf{D}_n^\sigma$ with its elements given by $(\mathbf{D}_n^\sigma)_{s_n\to\sigma_j} = p(w_n|s_n, \mathbf{G}_\psi)$. Here, we note that the index j does not appear on the right hand side because the system state does not change during an interpulse window resulting in the independence of observation probability from the next system state s_{n+1}. The explicit formulas for the observation probabilities are provided in the Supplementary Information.

Now, using the matrix $\mathbf{D}_n^\sigma$, we define the reduced propagators for each interpulse period as

$$\mathbf{\Pi}_n^\sigma = \mathbf{\Pi}_\sigma \odot \mathbf{D}_n^\sigma, \tag{5}$$

where $\odot$ denotes element-by-element product.

Finally, using these simplified propagators, we can write the likelihood for a an smFRET trace under pulsed illumination as

$$L = p(\boldsymbol{w}|\boldsymbol{\rho}_{start}, \mathbf{\Pi}_\sigma, \mathbf{G}_\psi) \propto \boldsymbol{\rho}_{start}\mathbf{\Pi}_1^\sigma\mathbf{\Pi}_2^\sigma \ldots \mathbf{\Pi}_N^\sigma\boldsymbol{\rho}_{norm}^T, \tag{6}$$

as also introduced in the Sec. 2.5.1 of the first companion manuscript [24]. This form of the likelihood is advantageous in that it allows empty pulses to be computed as a simple product, greatly reducing computational cost.

In the following, we first illustrate a parametric inference procedure assuming a given number of system states. We next generalize the procedure developed to the nonparametric case to deduce the number of system states along with the rest of parameters.

3.1 Inference Procedure: Parametric Sampler

Now, with the likelihood at hand, we construct the posterior as follows

$$p(\boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi | \boldsymbol{w}) \propto p(\boldsymbol{w} | \boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi) p(\boldsymbol{\rho}_{start}) p(\mathbf{G}_\psi) p(\boldsymbol{\Pi}_\sigma), \tag{7}$$

where we assume in the prior that the unknown parameters, including the initial probability vector, $\boldsymbol{\rho}_{start}$, the photophysical transition rates in the generator matrix $\mathbf{G}_\psi$, and the transition probabilities among system states in propagator $\boldsymbol{\Pi}_\sigma$, are independent. Here, we can sample the set of unknowns using the above posterior with the Gibbs sampling procedure as described in the first companion manuscript Sec. 3.2.2. However, a computationally more convenient inference procedure that allows direct sampling is accomplished by writing the posterior of Eq. 8 as a marginalization (sum) over state trajectories as follows

$$\begin{aligned} p(\boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi | \boldsymbol{w}) &= \sum_{s_{1:N}} p(\boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi, s_{1:N} | \boldsymbol{w}) \\ &\propto \sum_{s_{1:N}} p(\boldsymbol{w} | \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi, s_{1:N}) p(\boldsymbol{\rho}_{start}) p(\mathbf{G}_\psi) p(\boldsymbol{\Pi}_\sigma) p(s_{1:N} | \boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma), \end{aligned} \tag{8}$$

where $s_{1:N} = \{s_1, s_2 \ldots, s_N\}$ denotes a system state trajectory. Now, we can use the non-marginal posterior

$$p(\boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi, s_{1:N} | \boldsymbol{w}) \propto p(\boldsymbol{w} | \boldsymbol{\Pi}_\sigma, \mathbf{G}_\psi, s_{1:N}) p(\boldsymbol{\rho}_{start}) p(\mathbf{G}_\psi) p(\boldsymbol{\Pi}_\sigma) p(s_{1:N} | \boldsymbol{\rho}_{start}, \boldsymbol{\Pi}_\sigma) \tag{9}$$

to sample the trajectory $s_{1:N}$ which, in turn, allows direct sampling of the elements of propagator $\boldsymbol{\Pi}_\sigma$ described shortly. For priors on $\boldsymbol{\rho}_{start}$ and rates in $\mathbf{G}_\psi$, we, respectively, use a Dirichlet and Gamma distributions similar to Eq. 65-66 of the first companion manuscript [24]. We sample the system state trajectory $s_{1:N}$ by recursively sampling the states using a forward filtering backward sampling algorithm described in the Supplementary Information Sec. S4.3.

Finally, for each row in the propagator $\boldsymbol{\Pi}_\sigma$, we use a Dirichlet prior

$$\boldsymbol{\pi}_m \sim \textbf{Dirichlet}(\alpha \boldsymbol{\beta}), \qquad m = 1, 2, \ldots, M_\sigma, \tag{10}$$

where M_σ is the number of system states and $\boldsymbol{\pi}_m$ denotes the m-th row of the propagator. Here, the hyperparameters α and $\boldsymbol{\beta}$ are, respectively, the concentration parameter and a vector of length M_σ described in the first companion manuscript [24] Sec. 3.2.2. We can now directly generate samples for the transition probability vectors $\boldsymbol{\pi}_m$ of length M_σ via prior-likelihood conjugacy as (see Supplementary Information Sec. S4.3)

$$\boldsymbol{\pi}_m \sim \textbf{Dirichlet}(\mathbf{n}_m + \alpha \boldsymbol{\beta}), \qquad m = 1, 2, \ldots, M_\sigma^{max},$$

where the vector $\mathbf{n}_m$ collects the number of times each transition out of system state σ_m occurs obtained using the system state trajectory.

After constructing the posterior, we can make inferences on the parameters by drawing samples from the posterior. However, as the resulting posterior has a nonanalytical form, it cannot be directly sampled. Therefore, we develop a Markov chain Monte Carlo sampling (MCMC) procedure [38, 44–48] to draw samples from the posterior.

Our MCMC scheme follows a Gibbs sampling technique sweeping through updates of the set of parameters in the following order: 1) photophysical transition rates including donor relaxation rates λ_d (inverse of donor lifetime), acceptor relaxation rate λ_a (inverse of acceptor lifetime), FRET rates $\lambda^{FRET}_{\sigma_{1:M_\sigma}}$ for each system state, and excitation rate (inverse of excitation probability π_{ex}) using MH; 2) transition probabilities between system states, $\pi_{1:M_\sigma}$ by directly drawing samples from the posterior; 3) the system states trajectory, $\mathcal{S}$, using forward backward sampling procedure [49]; and 4) the initial probabilities $\boldsymbol{\rho}_{start}$ by taking direct samples. In the end, the chains of samples drawn can be used for subsequent numerical analysis.

3.2 Inference Procedure: Nonparametrics Sampler

The smFRET data analysis method illustrated above assumes a given number of system states, M_σ. However, in many applications the number of system states is not specified *a priori*. Here, we describe a generalization of our parametric method to address this shortcoming and estimate the number of system states simultaneously along with other unknown parameters.

We accomplish this by modifying our previously introduced parametric posterior as follows. First, we suppose an infinite number of system states ($M_\sigma \to \infty$) for the likelihood introduced previously and learn the transition matrix $\mathbf{\Pi}_\sigma$. The number of system states can then be interpreted as those appreciably visited over the course of the trajectory.

To incorporate this infinite system state space into our inference strategy, we leverage the iHMM [25, 26, 28–30] from the BNP repertoire, placing a hierarchical Dirichlet process prior over the infinite set of system states as described in the first companion manuscript, Sec. 3.2.2 [24]. However, as detailed in the first companion manuscript [24] Sec. 3.2.2 dealing with an infinite number of random variables, though feasible, is not computationally efficient and we approximate this infinite value with a large number M_σ^{max}, reducing our hierarchical Dirichlet process prior to

$$\begin{aligned} \boldsymbol{\beta} &\sim \textbf{Dirichlet}\left(\frac{\gamma}{M_\sigma^{max}}, \ldots, \frac{\gamma}{M_\sigma^{max}}\right), \\ \boldsymbol{\pi}_m &\sim \textbf{Dirichlet}(\alpha\boldsymbol{\beta}) \ , \qquad\qquad m = 1, ..., M_\sigma^{max}. \end{aligned}$$

Here, $\boldsymbol{\beta}$ denotes the base probability vector of length M_σ^{max} serving itself as a prior on the probability transition matrix $\mathbf{\Pi}_\sigma$, and $\boldsymbol{\pi}_m$ is the m-th row of $\mathbf{\Pi}_\sigma$. Moreover, γ is a positive scalar hyperparameter of the Dirichlet process prior often chosen to be one. As such, we ascribe identical weights across the state space *a priori* for computational convenience [28, 29, 50].

Now, equipped with the nonparametric posterior, we proceed to simultaneously make inferences on transition probabilities, excited state escape rates, and the remaining parameters. To do so, we employ the Gibbs sampling scheme detailed in Sec. 3.2.2 on the first companion manuscript [24], except that we must now also sample the system state trajectory $s_{1:N}$. More details on the overall sampling scheme are found in the Supplementary Information in section S4.

4 Results

The main objective of our method is to learn full distributions over: 1) transition probabilities among M_σ^{max} system states determining, in turn, the corresponding system transition rates and the effective number of system states; 2) photophysical transition rates including FRET rates $\lambda_{1:M}^{FRET}$, and fluorophores' relaxation rates (inverse of lifetimes) λ_a and λ_d.

To sample from distributions over these parameters, the BNP-FRET sampler requires input data comprised of photon arrival time traces from both donor and acceptor channels as well as a set of precalibrated input parameters including: camera effects such as crosstalk matrix and detection efficiency (see Sec. 2.4 and Example V of the first companion manuscript [24]); background emission (see Sec. 2.6 of the first companion manuscript and Supplementary Information Sec. S2.4); and the IRF (see Sec. 2.5 of the first companion manuscript [24] and Supplementary Information Sec. S2.3).

Here, we first show that our method samples posteriors over a set of parameters employing realistic synthetic data generated using the Gillespie algorithm [51] to simulate system and photophysical transitions while incorporating detector artefacts such as crosstalk (see Sec. 2.8 of the first companion manuscript [24]). The list of parameters used in data generation for all the figures is provided in Supplementary Information Sec. S6. Furthermore, prior hyperparameters used in the analysis of synthetic and experimental data are listed in the Supplementary Information Sec. S3.

We first show that our method works for the simplest case of slow transitions compared to the interpulse period (25 ns) with two system states using synthetic data, see Fig. 2. Next, we proceed to tackle more challenging synthetic data with three system states and higher transition rates Fig. 3. We show that our nonparametric algorithm correctly infers system transition probabilities and thus the number of system states; see Fig. 3.

After demonstrating the performance of our method using synthetic data, we use experimental data to investigate the kinetics of HJs under different $MgCl_2$ concentrations in buffer; see Fig. 4.

4.1 Simulated Data Analysis

To help validate BNPs on smFRET single photon data, we start with a simple case of a two state system and select kinetics similar to those of the experimental data sets, *c.f.*, the HJ in 10 mM $MgCl_2$, with escape rates of $40\,\mathrm{s}^{-1}$ for both system states [52]. The generated system state trajectory and photon traces over a period of 500 ms from both channels are shown in Fig. 2 (a).

Fig. 2 (b) shows the bivariate posterior distribution over FRET efficiencies, ε_{FRET}, defined as $\varepsilon_{FRET} = \lambda_{FRET}/(\lambda_{FRET} + \lambda_d)$, and system escape rates, *i.e.*, obtained by computing logarithm of the propagator matrix, with two peaks corresponding to two system states most visited by the sampler. Furthermore, the ground truths, designated by red dots, fall within the posterior with a relative error of less than 3% from the posterior modes. The results for the remaining parameters, including donor and acceptor transition rates, FRET transition rates and system transition probabilities, are presented in Supplementary Information Sec. S7.

To showcase the critical role played by BNPs, we also consider the more difficult case of a

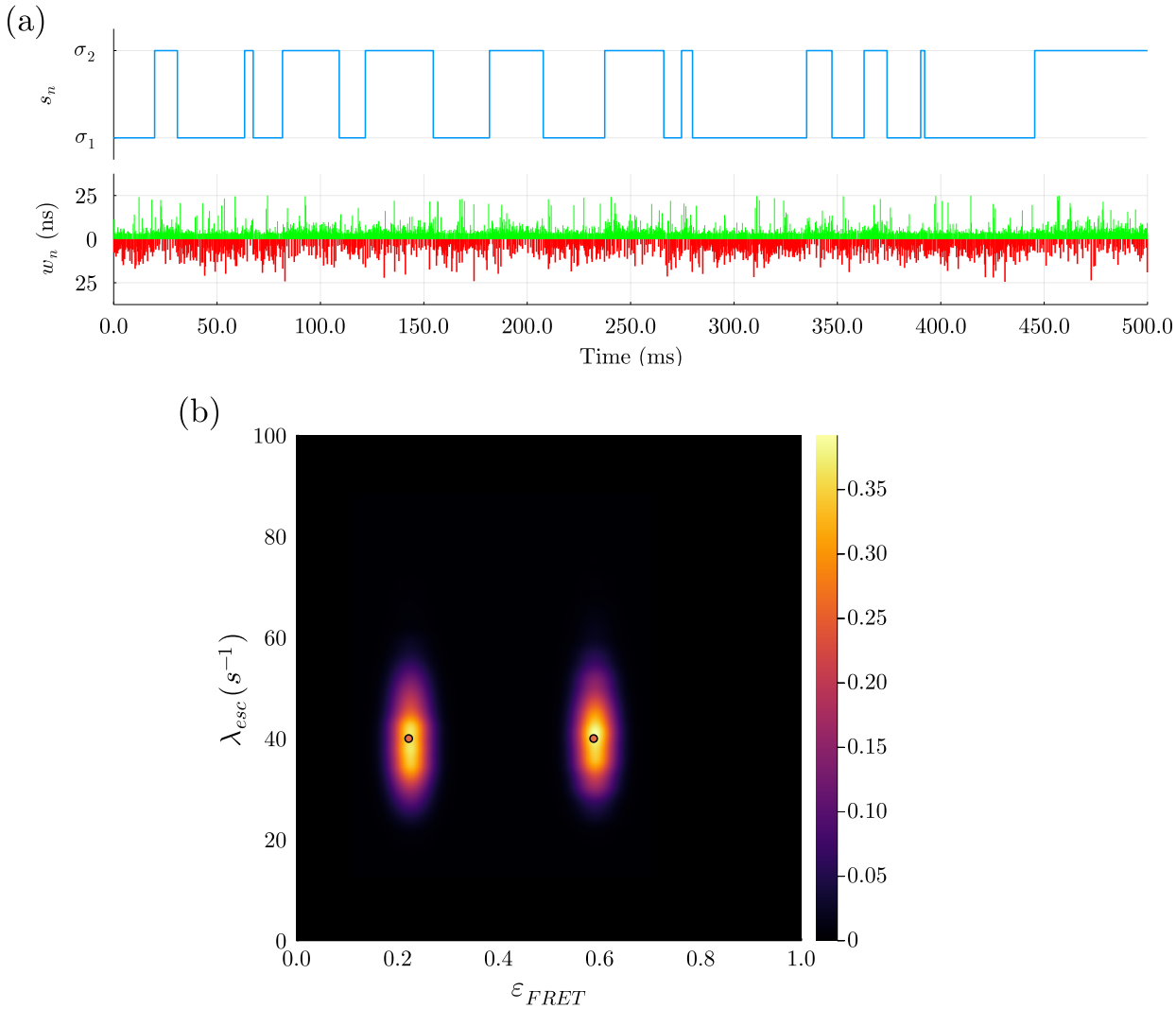

Figure 2: **Analysis on synthetic data for a system with two system states**. In panel (a), we show a section of synthetic data produced with the values in Supplementary Information Table S2. Furthermore, the system state trajectory is shown in blue. Below this, the arrival times of donor and acceptor photons μ_n^d and μ_n^a are shown in green and red, respectively. In panel (b), the ground truth is shown with red dots corresponding to escape rate of 40 s^{-1} and FRET efficiencies of 0.22 and 0.59. From our MAP estimate, we plot the bivariate distribution over escape rates λ_{esc} and FRET efficiencies ε_{FRET}. As seen, the BNP-FRET sampler clearly distinguishes two system states and locates ground truth values for the associated escape rates with values $\approx 38^{+7}_{-7}$ s^{-1} and $\approx 40^{+7}_{-7}$ s^{-1}, and FRET efficiencies $\approx 0.21^{+0.03}_{-0.03}$ and $\approx 0.59^{+0.03}_{-0.03}$. We have smoothed the distributions using kernel density estimation (KDE) for illustration purposes only.

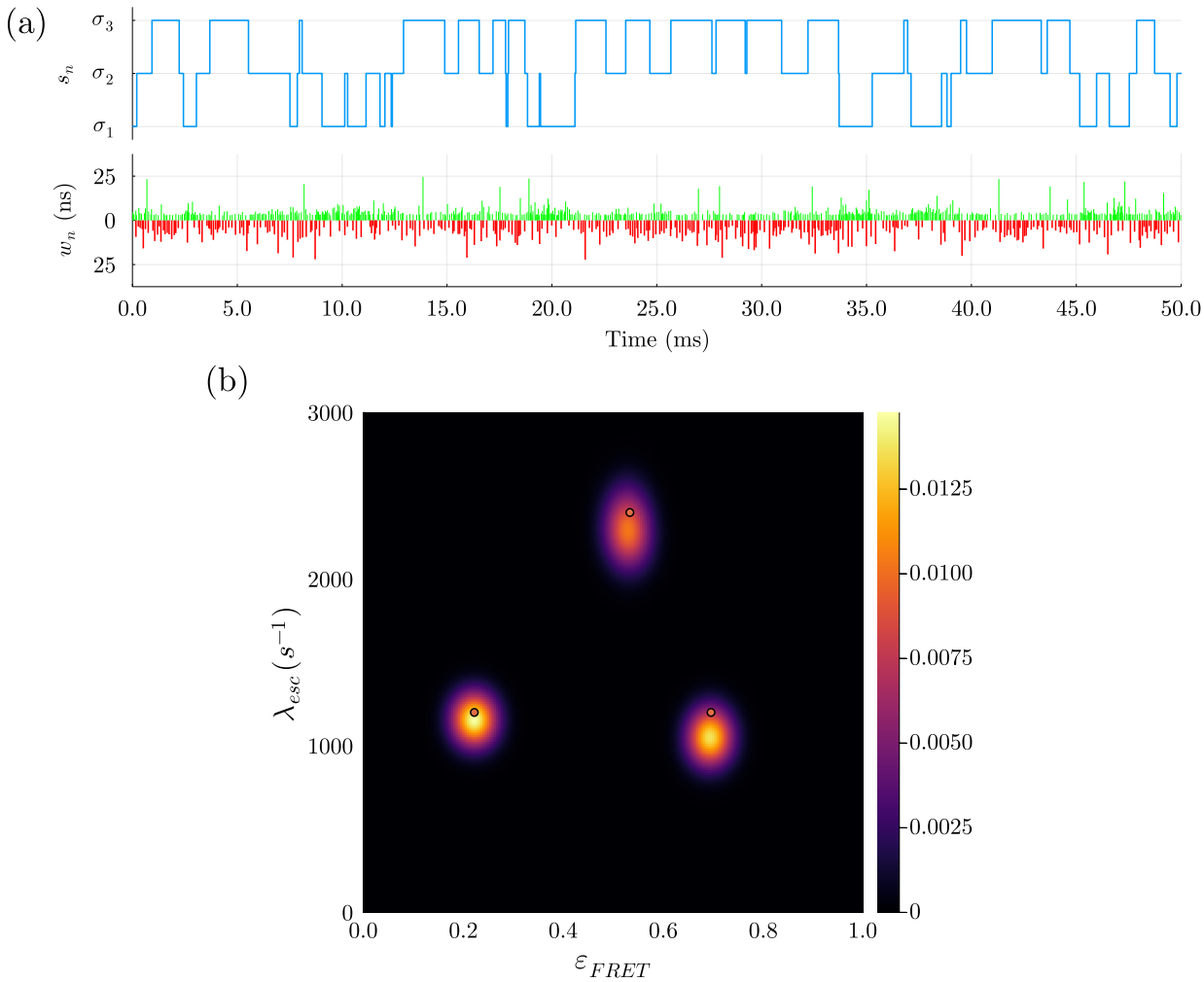

Figure 3: **Analysis on synthetic data for three system states**. In panel (a) we have a section of synthetic data produced with the values from Supplementary Information Table S3. The system state trajectory is seen in blue. Below this, the arrival times of donor and acceptor photons μ_n^d and μ_n^a are shown in green and red, respectively. In panel (b), the red dots show ground truths corresponding to escape rates of 1200 s^{-1}, 2400 s^{-1}, and 1200 s^{-1} and FRET efficiencies of 0.22, 0.53 and 0.7. From our MAP estimate, we plot the distribution over escape rates λ_{esc} and FRET efficiencies ε_{FRET}. The MAP estimate clearly shows three system states with escape rates of 1100^{+60}_{-60} s^{-1}, 2300^{+131}_{-128} s^{-1} 1050^{+80}_{-80} s^{-1}.

sample with three system states and faster system state kinetics ranging over 1200-2500 s^{-1}. We do so by simulating photon traces in both donor and acceptor channels over a period of ∼150 ms. A 50 ms section of the synthetic photon trace is shown in Fig 3(a).

Using direct photon arrivals from the generated photon trace, we find that the most probable system state trajectories sampled by BNP-FRET visit the correct number of system states, as shown in Fig 3(b), while inferring all other parameters. Furthermore, the BNP-

FRET sampler estimates the system transition rates and thus the escape rates (*i.e.*, sum of transition rates out of a given state) where the ground truth escape rates differ from the posterior peaks by a relative average error of less than 8%. The results for the remaining parameters are provided in the Supplementary Information Sec. S7.

4.2 Experimental Data Analysis: Holliday Junction (HJ)

In this section, we benchmark our method over a wide range of kinetic rates employing experimental data acquired using HJ under varying buffer $MgCl_2$ concentrations [15, 52].

HJs are four-way double-helical DNA junctions existing in various structural configurations [41, 53, 54]. When not interacting with multivalent metal ions, electrostatic repulsion between negatively charged phosphate groups of the four helical arms forces HJs to assume a wide configuration where the arms lie along the two diagonals of a square. However, in the presence of ions, such as Mg^{2+}, interaction with the phosphate groups results in electrostatic screening. This reduced repulsion induces transitions to what is believed to be primarily two compact stacked configurations/conformations. The transitions between both conformations necessitates passing through the intermediate open configuration. Since, at high ion concentrations, displacing ions away from the phosphate group becomes increasingly difficult, in this scenario we anticipate smaller transition rates between both conformations.

The HJ kinetic rates have been studied using both fluorescence lifetime correlation spectroscopy (FLCS) [15] and HMM analysis [55] on diffusing HJs assuming *a priori* a pair of high and low FRET system states. As expected, these previous studies show kinetic rates decreasing with increasing $MgCl_2$ concentrations [41, 43] and correspondingly longer dwells.

Here our method, free from averaging and binning otherwise common in HMM analysis, is particularly well-suited to learn the rapid kinetics at low Mg^{2+} concentrations. We apply our BNP-FRET to data acquired from HJs at 1, 3, 5, and 10 mM $MgCl_2$ concentrations, and sample the photophysical transition rates and the system transition probabilities.

The acquired bivariate posterior distributions over the FRET efficiencies and escape rates (computed via the logarithm of the system transition probability matrix $\mathbf{\Pi}_\sigma$) are presented in Fig. 4. Moreover, estimates for the other parameters can be found in the Supplementary Information Section S7. We note that our results are obtained on a single molecule basis with a photon budget of $10^4 - 10^5$ photons.

For all four concentrations (see Fig. 4), our BNP-FRET sampler most frequently visited only two system states, while this was given as an input to the other analysis methods [15, 55]. Moreover, both escape rates are found to have similar values with an average of approximately $1400\,\mathrm{s}^{-1}$ (1 mM $MgCl_2$), $140\,\mathrm{s}^{-1}$ (3 mM $MgCl_2$), $72\,\mathrm{s}^{-1}$ (5 mM $MgCl_2$), and $41\,\mathrm{s}^{-1}$ (10 mM $MgCl_2$). These escape rates are in close agreement with values reported by FLCS and H2MM methods [15, 55] $\approx 1300\ \mathrm{s}^{-1}$ (1 mM $MgCl_2$), $\approx 170\ \mathrm{s}^{-1}$ (3 mM $MgCl_2$), $\approx 100\ \mathrm{s}^{-1}$ (5 mM $MgCl_2$), $\approx 60\ \mathrm{s}^{-1}$ (10 mM $MgCl_2$), which lie well within the bounds of our posteriors shown in Fig. 4 while simultaneously, and self-consistently, learning number of system states.

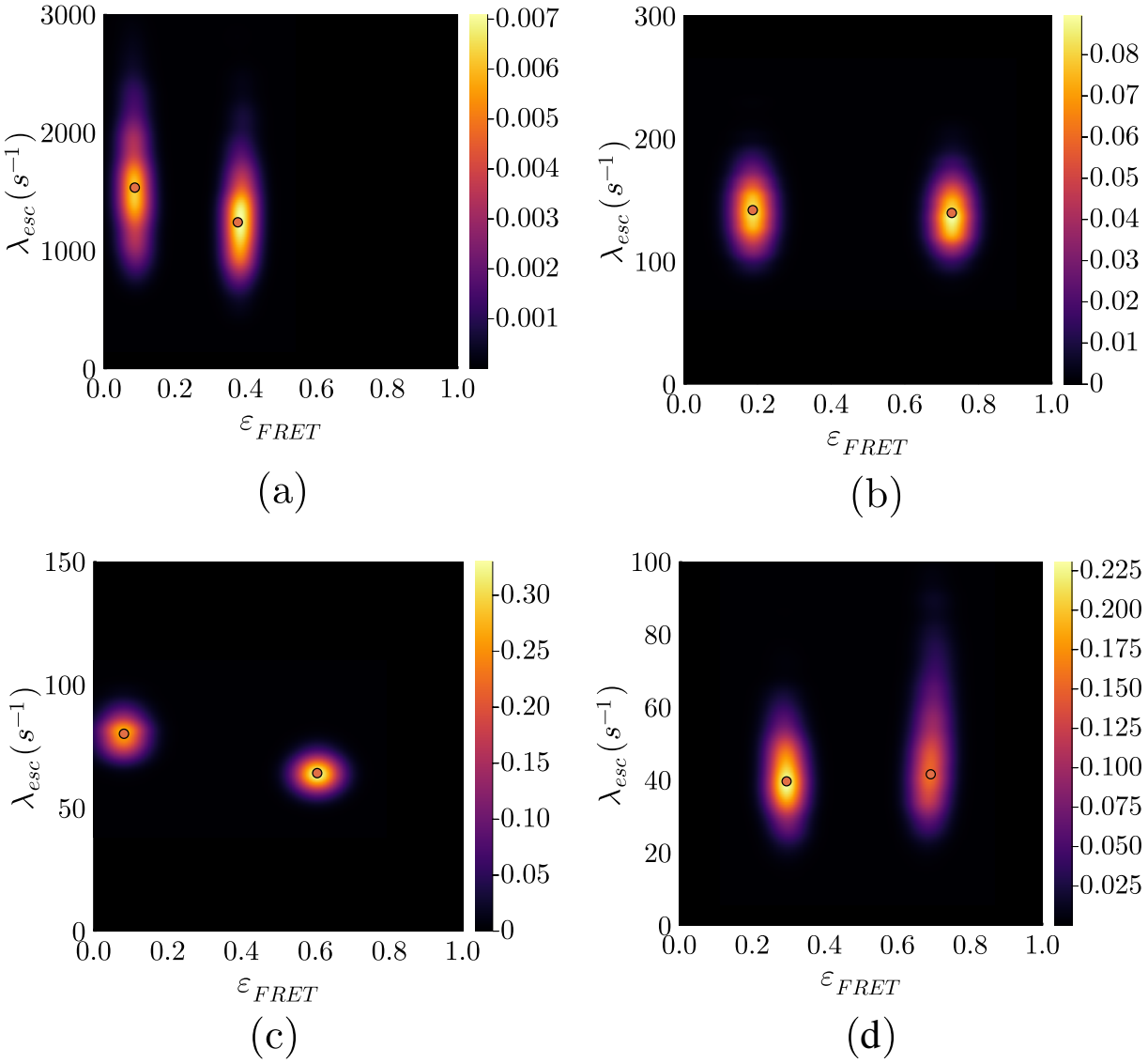

Figure 4: **The bivariate posterior for the conformational transition rates λ_{esc} and FRET efficiencies ε_{FRET} for experimental data acquired in the presence of different HJ concentrations**. Here, we show our bivariate posteriors where red dots show MAP estimates. In panel (a), we show the posterior for a sample with 1 mM $MgCl_2$. We report escape rates of 1530^{+500}_{-550} s^{-1} and 1240^{+420}_{-420} s^{-1} in this case. The posterior for a sample with 3 mM $MgCl_2$ is shown in panel (b). We report escape rates of 140^{+38}_{-38} s^{-1} and 142^{+32}_{-32} s^{-1} for this case. In panel (c), we show our posterior for a sample with 5 mM $MgCl_2$. Here, we report escape rates of 64^{+9}_{-9} s^{-1} and 80^{+10}_{-10} s^{-1}. The posterior in panel (d) is for a sample with 10 mM $MgCl_2$. We report escape rates of 39^{+17}_{-12} s^{-1} and 41^{+23}_{-12} s^{-1}.

4.3 Experimental Data Acquisition

In this section, we describe the protocol for preparing the surface immobilized HJ sample labeled with a FRET pair and the procedure for recording smFRET traces from individual immobilized molecules. The sample preparation and recording of data follow previous work [56].

Sample preparation: The HJ used in this work consists of four DNA strands whose sequences are as follows

R-strand: 5'-CGA TGA GCA CCG CTC GGC TCA ACT GGC AGT CG-3'
H-strand: 5'-CAT CT<u>T</u> 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 CC<u>T</u> CCT AGC AAG CCG CTG CTA CTA AGA TG-3'.

For surface immobilization, the X-strand was labeled with biotin at the 5'-end. For FRET measurements, the donor (ATTO-532) and acceptor (ATTO-647N) dyes were introduced into the H- and B-strands, respectively. In both cases, the dyes were labeled to thymine nucleotide at the 6th position from the 5'-ends of respective strands (shown as <u>T</u>). All DNA samples (labeled or unlabeled) were purchased from JBioS (Japan) in the HPLC purified form and were used without any further purification.

The HJ complex was prepared by mixing 1 mM solutions of R-, H-, B-, and X-strands in TN buffer (10 mM Tris-HCl with 50 mM NaCl, pH 8.0) at 3:2:3:3 molar ratio, annealing the mixture at 94 °C for 4 minutes, and gradually cooling it down (2-3 °C min^{-1}) to room temperature (25 °C). For smFRET measurements, we used a sample chamber (Grace Bio-Labs SecureSeal, GBL621502) with biotin-PEG-SVA (biotin-poly(ethylene glycol)-succinimidyl valerate) coated coverslip. The chamber was first incubated with streptavidin (0.1 mg mL^{-1} in TN buffer) for 20 min. This was followed by washing the chamber with TN buffer (3 times) and injection of 1 nM HJ solution (with respect to its H-strand) for 3-10 seconds. After this incubation period, the chamber was rinsed with TN buffer (3 times) to remove unbound DNA and it was filled with TN buffer containing 1 mM (or 5 mM) $MgCl_2$ and 2 mM Trolox for smFRET measurements.

smFRET measurements: The smFRET traces from individual HJs were recorded using a custom built confocal microscope (Nikon Eclipse Ti) equipped with the Perfect Focus System (PFS), a sample scanning piezo stage (Nano control B16-055), and a time correlated single photon counting (TCSPC) module (Becker and Hickl SPC-130EM).

The broadband light generated by a supercontinuum laser operating at 40 MHz (Fianium SC-400-4) was filtered with a bandpass filter (Semrock FF01-525/30) for exciting the donor dye, ATTO-532. This excitation light was introduced to the microscope using a single-mode optical fiber (Thorlabs P5-460B-PCAPC-1), and directed onto the sample using a dichroic mirror (Chroma ZT532/640rpc) and a water immersion objective lens (Nikon Plan Apo IR 60x, numerical aperture = 1.27).

The excitation light was focused onto the top surface of the coverslip and, during measurements, the focusing condition was maintained using the PFS. The fluorescence signals were collected by the same objective, passed through the dichroic mirror, and guided to the detection assembly (Thorlabs DFM1/M) using a multimode fiber (Thorlabs M50L02S-A).

Note that this multimode fiber (core diameter: 50 µm) also acts as the confocal pinhole. In the detection assembly, the fluorescence signals from the donor and acceptor dyes were separated using a dichroic mirror (Chroma Technology ZT633rdc), filtered using bandpass filters (Chroma ET585/65m for donor, and Semrock FF02-685/40 for acceptor), and detected using separate hybrid detectors (Becker & Hickl HPM-100-40-C).

For each detected photon, its macrotime (absolute arrival time from the start of the measurement) was recorded with 25.2 ns resolution and its microtime (relative delay from the excitation pulse) was recorded with 6.1 ps resolution using the TCSPC module operating in time-tagging mode. A router (Becker and Hickl HRT-41) was used to process the signals from the donor and acceptor detectors.

For recording smFRET traces from individual HJs, we first imaged a 10 µm×3 µm area of the sample using the piezo stage by scanning it linearly at a speed of 1 $\mu m\, s^{-1}$ in the X-direction and with an increment of 0.1 µm in the Y-direction. Individual HJs appeared as isolated bright spots in the image.

Next, we fitted the obtained donor and acceptor intensity images with multiple 2D Gaussian functions to determine the precise locations of individual HJs. Note that, during this image acquisition, the laser excitation power was kept to a minimum (∼1 µW at the back aperture of the objective lens) to avoid photobleaching the dyes. In addition, we also employed an electronic shutter (Suruga Seiki, Japan) in the laser excitation path to control the sample excitation as required.

Using the precise locations of individual HJs obtained, we recorded 30 s long smFRET traces for each molecule by moving them to the center of the excitation beam using the piezo stage. For each trace, the laser excitation was blocked (using the shutter) for the first 5 seconds and was allowed to excite the sample for the remaining 25 seconds. Note that the smFRET traces were recorded using 40 µW laser excitation (at the back aperture of the objective lens) to maximize the fluorescence photons emitted from the dyes. We automated the process of acquiring smFRET traces from different molecules sequentially and executed it using a program written in-house on Igor Pro (Wavemetrics).

5 Discussion

The sensitivity of smFRET under pulsed illumination has been exploited to investigate many different molecular interactions and geometries [8–11, 57]. However, quantitative interpretation of smFRET data faces serious challenges including unknown number of system states and robust propagation of uncertainty from noise sources such as detectors and background. These challenges ultimately mitigate our ability to determine full distributions over all relevant unknowns and, traditionally, have resulted in data pre- or post-processing compromising the information otherwise encoded in the rawest form of data: single photon arrivals.

Here, we provide a general BNP framework for smFRET data analysis starting from single photon arrivals under a pulsed illumination setting. We simultaneously learn transition probabilities among system states as well as determine photophysical rates by incorporating existing sources of uncertainty such as background and crosstalk.

We benchmark our method using both experimental and simulated data. That is, we first show that our method correctly learns parameters for the simplest case with two system

states and slow system transition rates. Moreover, we test our method on more challenging cases with more than two states using synthetic data and obtain correct estimations for the system state transition probabilities and thus the number of system states along with the remaining parameters of interest. To further assess our method's performance, we analyzed experimental data from HJs suspended in solutions with a range of $MgCl_2$ concentrations. These data were previously processed using other techniques assuming a fixed number of system states by binning photon arrival times [15].

Despite multiple advantages mentioned above for BNP-FRET, BNPs always come with an added computational cost as they take full advantage of information from single photon arrival times and all existing sources of uncertainty. For this version of our general BNP method simplified for pulsed illumination, we further reduced the computational complexity by grouping empty pulses together. Therefore, the computational complexity increased only linearly with the number of input photons as the photons are treated independently.

The method described in this paper assumes a Gaussian IRF. However, the developed framework is not limited to a specific form for the IRF and can be used for data collected using any type of IRF by modifying Eq. 4. Furthermore, the framework is flexible in accommodating different illumination techniques such as alternating color pulses, typically used to directly excite the acceptor fluorophores. This can be achieved by simple modification of the propagator $\mathbf{Q}_n^{\psi}$ in Eq. 4. A future extension of this method could relax the assumption of a static sample by adding spatial dependence to the excitation rate as we explored in previous works [35, 48, 58]. This would allow our method to learn the dynamics of diffusing molecules, as well as their photophysical and system state transition rates.

6 Code Availability

The BNP-FRET software package is available on Github at
`https://github.com/LabPresse/BNP-FRET`

7 Acknowledgments

We thank Dr. Zeliha Kilic, Weiqing Xu, and J Shephard Bryan IV for their contributions and insight into the project. We also thank Dr. Douglas Shepherd for providing insight into the workings of detectors and other experimental equipment and Irina Gopich for discussions on FRET likelihoods. T.T. acknowledges support from the JSPS KAKENHI (Grant No. 18H05265). S. P. acknowledges support from the NIH NIGMS (R01GM130745) for supporting early efforts in nonparametrics and NIH NIGMS (R01GM134426) for supporting single photon efforts.

Bibliography

[1] Shimon Weiss. Fluorescence spectroscopy of single biomolecules. *Science*, 283(5408):1676, 1999.

[2] Jennifer Lippincott-Schwartz, Erik Snapp, and Anne Kenworthy. Studying protein dynamics in living cells. *Nature Reviews Molecular Cell Biology*, 2(6):444, 2001.

[3] Bo Huang, Mark Bates, and Xiaowei Zhuang. Super-resolution fluorescence microscopy. *Annual Review of Biochemistry*, 78:993, 2009.

[4] Mickaël Lelek, Melina T Gyparaki, Gerti Beliu, Florian Schueder, Juliette Griffié, Suliana Manley, Ralf Jungmann, Markus Sauer, Melike Lakadamyali, and Christophe Zimmer. Single-molecule localization microscopy. *Nature Reviews Methods Primers*, 1(1):1, 2021.

[5] Mohamadreza Fazel and Michael J Wester. Analysis of super-resolution single molecule localization microscopy data: A tutorial. *AIP Advances*, 12(1):010701, 2022.

[6] Rupsa Datta, Tiffany M Heaster, Joe T Sharick, Amani A Gillette, and Melissa C Skala. Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications. *Journal of Biomedical Optics*, 25(7):071203, 2020.

[7] Yuval Garini, Ian T Young, and George McNamara. Spectral imaging: principles and applications. *Cytometry Part A: The Journal of the International Society for Analytical Cytology*, 69(8):735, 2006.

[8] Rahul Roy, Sungchul Hohng, and Taekjip Ha. A practical guide to single-molecule FRET. *Nature Methods*, 5(6):507, 2008.

[9] Hisham Mazal and Gilad Haran. Single-molecule FRET methods to study the dynamics of proteins at work. *Current Opinion in Biomedical Engineering*, 12:8, 2019.

[10] Benjamin Schuler. Single-molecule FRET of protein structure and dynamics - a primer. *Journal of Nanobiotechnology*, 11(1):S2, 2013.

[11] Maolin Lu, Xiaochu Ma, Luis R. Castillo-Menendez, Jason Gorman, Nirmin Alsahafi, Utz Ermel, Daniel S. Terry, Michael Chambers, Dongjun Peng, Baoshan Zhang, Tongqing Zhou, Nick Reichard, Kevin Wang, Jonathan R. Grover, Brennan P. Carman, Matthew R. Gardner, Ivana Nikić-Spiegel, Akihiro Sugawara, James Arthos, Edward A. Lemke, Amos B. Smith, Michael Farzan, Cameron Abrams, James B. Munro, Adrian B. McDermott, Andrés Finzi, Peter D. Kwong, Scott C. Blanchard, Joseph G. Sodroski, and Walther Mothes. Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET. *Nature*, 568(7752):415, 2019.

[12] Steven M. Mooney, Ruoyi Qiu, John J. Kim, Elizabeth J. Sacho, Krithika Rajagopalan, Dorhyun Johng, Takumi Shiraishi, Prakash Kulkarni, and Keith R. Weninger. Cancer/testis antigen PAGE4, a regulator of c-Jun transactivation, is phosphorylated by

homeodomain-interacting protein kinase 1, a component of the stress-response pathway. *Biochemistry*, 53(10):1670, 2014.

[13] Anna K Woźniak, Gunnar F Schröder, Helmut Grubmüller, Claus AM Seidel, and Filipp Oesterhelt. Single-molecule FRET measures bends and kinks in DNA. *Proceedings of the National Academy of Sciences*, 105(47):18337, 2008.

[14] Hoi Sung Chung, John M. Louis, and Irina V. Gopich. Analysis of fluorescence lifetime and energy transfer efficiency in single-molecule photon trajectories of fast-folding proteins. *The Journal of Physical Chemistry B*, 120(4):680, 2016.

[15] Wooseok Heo, Kazuto Hasegawa, Kenji Okamoto, Yasushi Sako, Kunihiko Ishii, and Tahei Tahara. Scanning two-dimensional fluorescence lifetime correlation spectroscopy: Conformational dynamics of DNA Holliday junction from microsecond to subsecond. *The Journal of Physical Chemistry Letters*, 13(5):1249, 2022.

[16] André Zeug, Andrew Woehler, Erwin Neher, and Evgeni G Ponimaskin. Quantitative intensity-based fret approaches—a comparative snapshot. *Biophysical journal*, 103(9):1821–1827, 2012.

[17] Sahiti Kuppa, Jaigeeth Deveryshetty, Rahul Chadda, Jenna R Mattice, Nilisha Pokhrel, Vikas Kaushik, Angela Patterson, Nalini Dhingra, Sushil Pangeni, Marisa K Sadauskas, et al. Rtt105 regulates rpa function by configurationally stapling the flexible domains. *Nature communications*, 13(1):1–16, 2022.

[18] Zeliha Kilic, Ioannis Sgouralis, and Steve Pressé. Generalizing hmms to continuous time for fast kinetics: hidden markov jump processes. *Biophysical journal*, 120(3):409–423, 2021.

[19] Peter Kapusta, Michael Wahl, Aleš Benda, Martin Hof, and Jörg Enderlein. Fluorescence lifetime correlation spectroscopy. *Journal of Fluorescence*, 17(1):43, 2007.

[20] Kunihiko Ishii and Tahei Tahara. Two-dimensional fluorescence lifetime correlation spectroscopy. 1. principle. *The Journal of Physical Chemistry B*, 117(39):11414, 2013.

[21] Takuhiro Otosu, Kunihiko Ishii, and Tahei Tahara. Microsecond protein dynamics observed at the single-molecule level. *Nature Communications*, 6(1), 2015.

[22] Janghyun Yoo, Jae-Yeol Kim, John M. Louis, Irina V. Gopich, and Hoi Sung Chung. Fast three-color single-molecule FRET using statistical inference. *Nature Communications*, 11(1), 2020.

[23] Eitan Lerner, Antonino Ingargiola, and Shimon Weiss. Characterizing highly dynamic conformational states: The transcription bubble in RNAP-promoter open complex as an example. *The Journal of Chemical Physics*, 148(12):123315, 2018.

[24] Ayush Saurabh, Matthew Safar, Ioannis Sgouralis, Mohamadreza Fazel, and Steve Pressé. Single photon smFRET. I. theory and conceptual basis. *bioRxiv*, 2022.

[25] Ioannis Sgouralis and Steve Pressé. An introduction to infinite HMMs for single-molecule data analysis. *Biophysical Journal*, 112(10):2021, 2017.

[26] Ioannis Sgouralis, Shreya Madaan, Franky Djutanta, Rachael Kha, Rizal F. Hariadi, and Steve Pressé. A Bayesian nonparametric approach to single molecule Förster resonance energy transfer. *The Journal of Physical Chemistry B*, 123(3):675, 2019.

[27] Emily B. Fox, Erik B. Sudderth, Michael I. Jordan, and Alan S. Willsky. A sticky HDP-HMM with application to speaker diarization. *The Annals of Applied Statistics*, 5(2A):1020, 2011.

[28] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476):1566, 2006.

[29] Jayaram Sethuraman. A constructive definition of Dirichlet priors. *Statistica Sinica*, 4(2):639, 1994.

[30] Jim Pitman. Poisson–Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition. *Combinatorics, Probability and Computing*, 11(5):501, 2002.

[31] Thomas S. Ferguson. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209, 1973.

[32] Samuel J. Gershman and David M. Blei. A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, 56(1):1, 2012.

[33] Ioannis Sgouralis, Miles Whitmore, Lisa Lapidus, Matthew J. Comstock, and Steve Pressé. Single molecule force spectroscopy at high data acquisition: A Bayesian nonparametric analysis. *The Journal of Chemical Physics*, 148(12):123320, 2018.

[34] Meysam Tavakoli, J Nicholas Taylor, Chun-Biu Li, Tamiki Komatsuzaki, and Steve Pressé. Single molecule data analysis: An introduction. *Advances in Chemical Physics*, 162:205–305, 2017.

[35] Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, Omer M. Shafraz, Sanjeevi Sivasankar, Bryan Donaphon, Marcia Levitus, and Steve Pressé. Pitching single-focus confocal data analysis one photon at a time with Bayesian nonparametrics. *Physical Review X*, 10(1):011021, 2020.

[36] Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, Wooseok Heo, Kunihiko Ishii, Tahei Tahara, and Steve Pressé. Direct photon-by-photon analysis of time-resolved pulsed excitation data using Bayesian nonparametrics. *Cell Reports Physical Science*, 1(11):100234, 2020.

[37] J. Shepard Bryan IV, Ioannis Sgouralis, and Steve Pressé. Diffraction-limited molecular cluster quantification with Bayesian nonparametrics. *Nature Computational Science*, 2(2):102, 2022.

[38] Mohamadreza Fazel, Sina Jazani, Lorenzo Scipioni, Alexander Vallmitjana, Enrico Gratton, Michelle A. Digman, and Steve Pressé. High resolution fluorescence lifetime maps from minimal photon counts. *ACS Photonics*, 9(3):1015, 2022.

[39] Kenji Okamoto and Yasushi Sako. State transition analysis of spontaneous branch migration of the Holliday junction by photon-based single-molecule fluorescence resonance energy transfer. *Biophysical Chemistry*, 209:21, 2016.

[40] Sungchul Hohng, Chirlmin Joo, and Taekjip Ha. Single-molecule three-color FRET. *Biophysical Journal*, 87(2):1328, 2004.

[41] Sean A. McKinney, Anne-Cécile Déclais, David M.J. Lilley, and Taekjip Ha. Structural dynamics of individual Holliday junctions. *Nature Structural Biology*, 10(2):93, 2003.

[42] S. A. McKinney, A. D. J. Freeman, D. M. J. Lilley, and T. Ha. Observing spontaneous branch migration of Holliday junctions one step at a time. *Proceedings of the National Academy of Sciences*, 102(16):5715, 2005.

[43] I.G. Panyutin, I. Biswas, and P. Hsieh. A pivotal role for the structure of the Holliday junction in DNA branch migration. *The EMBO Journal*, 14(8):1819, 1995.

[44] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087, 1953.

[45] W Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97, 1970.

[46] J Shepard Bryan IV, Ioannis Sgouralis, and Steve Pressé. Diffraction-limited molecular cluster quantification with Bayesian nonparametrics. *Nature Computational Science*, 2(2):102, 2022.

[47] Mohamadreza Fazel, Michael J Wester, Hanieh Mazloom-Farsibaf, Marjolein Meddens, Alexandra S Eklund, Thomas Schlichthaerle, Florian Schueder, Ralf Jungmann, and Keith A Lidke. Bayesian multiple emitter fitting using reversible jump Markov chain Monte Carlo. *Scientific Reports*, 9(1):1, 2019.

[48] Sina Jazani, Ioannis Sgouralis, Omer M Shafraz, Marcia Levitus, Sanjeevi Sivasankar, and Steve Pressé. An alternative framework for fluorescence correlation spectroscopy. *Nature Communications*, 10(1):1, 2019.

[49] L.R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257, 1989.

[50] Mohamadreza Fazel, Alexander Vallmitjana, Lorenzo Scipioni, Enrico Gratton, Michelle A Digman, and Steve Presse. Fluorescence lifetime: Beating the IRF and interpulse window. *bioRxiv*, 2022.

[51] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403, 1976.

[52] Zeliha Kilic, Ioannis Sgouralis, Wooseok Heo, Kunihiko Ishii, Tahei Tahara, and Steve Pressé. Extraction of rapid kinetics from smFRET measurements using integrative detectors. *Cell Reports Physical Science*, 2(5):100409, 2021.

[53] Mikhail Karymov, Douglas Daniel, Otto F Sankey, and Yuri L Lyubchenko. Holliday junction dynamics and branch migration: single-molecule analysis. *Proceedings of the National Academy of Sciences*, 102(23):8186, 2005.

[54] EE Ferapontova, CP Mountford, J Crain, AH Buck, P Dickinson, JS Beattie, P Ghazal, JG Terry, AJ Walton, and AR Mount. Electrochemical control of a DNA Holliday Junction nanoswitch by Mg2+ ions. *Biosensors and Bioelectronics*, 24(3):422, 2008.

[55] Menahem Pirchi, Roman Tsukanov, Rashid Khamis, Toma E. Tomov, Yaron Berger, Dinesh C. Khara, Hadas Volkov, Gilad Haran, and Eyal Nir. Photon-by-photon hidden Markov model analysis for microsecond single-molecule FRET kinetics. *The Journal of Physical Chemistry B*, 120(51):13065, 2016.

[56] Zeliha Kilic, Ioannis Sgouralis, and Steve Pressé. Generalizing HMMs to continuous time for fast kinetics: Hidden Markov jump processes. *Biophysical Journal*, 120(3):409, 2021.

[57] Judith S Sebolt-Leopold and Jessie M English. Mechanisms of drug inhibition of signalling molecules. *Nature*, 441(7092):457, 2006.

[58] Sina Jazani, Ioannis Sgouralis, and Steve Pressé. A method for single molecule tracking using a conventional single-focus confocal setup. *The Journal of Chemical Physics*, 150(11):114108, 2019.