Single Photon smFRET. II. Application to Continuous Illumination

Ayush Saurabh[1,2], Matthew Safar[1,3], Mohamadreza Fazel[1,2], Ioannis Sgouralis[4], and Steve Pressé[1,2,5]

[1]Center for Biological Physics,
Arizona State University, Tempe, AZ, USA
[2]Department of Physics,
Arizona State University, Tempe, AZ, USA
[3]Department of Mathematics and Statistical Science,
Arizona State University, Tempe, AZ, USA
[4]Department of Mathematics, University of Tennessee Knoxville,
Knoxville, TN, USA
[5]School of Molecular Sciences, Arizona State University,
Tempe, AZ, USA

November 8, 2022

Contents

Abstract

Here we adapt the Bayesian nonparametrics (BNP) framework presented in the first companion manuscript to analyze kinetics from single photon, single molecule Förster Resonance Energy Transfer (smFRET) traces generated under continuous illumination. Using our sampler, BNP-FRET, we learn the escape rates and the number of system states given a photon trace. We benchmark our method by analyzing a range of synthetic and experimental data. Particularly, we apply our method to simultaneously learn the number of system states and the corresponding kinetics for intrinsically disordered proteins (IDPs) using two-color FRET under varying chemical conditions. Moreover, using synthetic data, we show that our method can deduce the number of system states even when kinetics occur at timescales of interphoton intervals.

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 time-invariant laser intensities. In particular, we study interactions of IDPs.

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

Single molecule Förster Resonance Energy Transfer (smFRET) experiments are widely used [1] to study molecular kinetics across timescales on both stationary [2–5] and freely diffusing molecules [6]. These timescales include faster events, below the micro- to millisecond timescales, including domain rotations, configurational kinetics of disordered proteins, protein folding, protein-protein interactions, all the way to slower events, such as misfolding and refolding events, occurring on minute and even hour long timescales [7].

In a typical experiment we consider herein, a continuous wave (CW) laser illuminates a sample with a beam of constant intensity and power over a period of time. CW sources are common as they are both cheaper and technically simpler to implement in an experimental setup than their pulsed counterparts [8, 9] that we explore in our third companion manuscript [10]. However, as compared to pulsed sources, a disadvantage lies in the increased photon flux through the sample which can accelerate photodamage [11].

While pulsed illumination can significantly reduce sample photobleaching and phototoxicity [12] and more readily reveals excited state lifetimes of fluorophores, in practice it is restricted to analyzing one (time-stamped) photon per interpulse period. This in turn limits the data acquisition rate and sets a bound on the temporal resolution of the kinetics we may deduce from pulsed single photon arrival.

By contrast, continuous illumination avoids this problem, by allowing a larger number of photons to be detected in the time that would normally be considered an interpulse period in pulsed illumination [13]. The cost then comes at the loss of direct knowledge of excited state lifetime which can, with difficulty and high uncertainty, then be decoded from photon-antibunching statistics if required [14] as shown in the first companion manuscript [15].

It is common practice to analyze photon arrival data to extract kinetics under continuous illumination by binning the data and subsequently using hidden Markov models (HMMs) [16–19]. As noise distributions are better characterized in unprocessed data, it remains conceptually preferred, though more computationally costly, to use photon-by-photon methods [13, 14, 20–24]. Indeed, photon-by-photon methods can be used to learn both photophysical and system transition rates directly from the detected photon colors and interphoton arrival times. Additionally, this has the benefit of avoiding averaging kinetics that may occur when binning data [17].

Currently available methods to analyze smFRET data in a photon-by-photon manner [13, 20] rely on the foundational works of Gopich and Szabo [13, 14, 25], where the likelihood is taken as the product of as many generator matrix exponentials as there are photons in a FRET trace. Such a generator matrix constitutes transition rates encoding the kinetics of the system-FRET composite [15].

When analyzing smFRET data, of particular interest is the dimensionality of this generator matrix determined by the number of system states. In all existing analyses, the dimensionality is fixed by hand *a priori* and the transition rates are then learned as point estimates using maximum likelihood methods.

Yet point estimates can be biased. In fact, limited data, lack of temporal resolution to estimate very fast kinetics [15], and noise, all contribute to bias [26] in addition to a flattening of possibly multimodal likelihoods [27, 28]. This motivates why we wish to operate in a Bayesian setting to learn distributions over the number of system states and transition rates, while incorporating unavoidable noise sources such as detector electronics and background.

For this reason, we developed a complete Bayesian nonparametric (BNP) framework in the first companion manuscript [15]. This framework incorporates many key complexities of a typical smFRET experimental setup, including background emissions, fluorophore photophysics (blinking, photobleaching, and direct acceptor excitation), instrument response function (IRF), detector dead time, and crosstalk.

Here, we delve deeper into this framework for the case of continuous illumination by exploring its utility in cases where the number of system states is unknown.

We first test the robustness of our nonparametric method and its software implementation BNP-FRET by analyzing synthetically generated data for kinetics varying from very slow to timescales as fast as the interphoton arrival times. We then apply our method to experimental smFRET data capturing interactions between intrinsically disordered protein (IDP) fragments [29, 30] relevant to signaling and regulation.

IDPs are of particular interest to nonparametric analyses as IDP's lack of order and stability results in broader spectra of dominant FRET pair distances sensitive to their chemical environment. In particular, we study interactions between the nuclear-coactivator binding domain (NCBD) of a CBP/p300, *i.e.*, transcription coactivator and the activation domain of SRC -3 (ACTR) under varying chemical conditions affecting their coupled folding and binding reaction rates [29–31]. We use a single FRET pair under continuous illumination to observe the possible physical configurations (system states) of the NCBD-ACTR complex. Further, we report new bound/transient system states for the NCBD P20A mutation, not observed using previous point estimation techniques [30].

3 Forward Model and Inference Strategy

For the sake of completeness, we begin with relevant aspects of the methods presented in the first companion manuscript [15], including the likelihood needed in Bayesian inference, and our parametric and nonparametric Markov Chain Monte Carlo (MCMC) samplers.

An smFRET experiment involves at least two single photon detectors collecting information on stochastic arrival times. We denote these arrival times with

$$\{T_{start}, T_1, T_2, T_3, \ldots, T_K, T_{end}\},$$

in detection channels

$$\{c_1, c_2, c_3, \ldots, c_K\},$$

for a total number of K photons. In this representation above, T_{start} and T_{end} are the experiment's start and end times, respectively.

Using this dataset, we would like to infer parameters governing a system's kinetics. That is, the number of system states M_σ and the associated transition rates $\lambda_{\sigma_i \to \sigma_j}$, as well as

M_ψ photophysical transition rates $\lambda_{\sigma_i,\psi_l\to\psi_m}$ corresponding to each system state σ_i. Here, $\sigma_i \in \{\sigma_1,\ldots,\sigma_{M_\sigma}\}$ and $\psi_l \in \{\psi_1,\ldots,\psi_{M_\psi}\}$ are the system states and photophysical states, respectively. These rates populate a generator matrix $\mathbf{G}$ of dimension $M_\phi = M_\sigma \times M_\psi$ now representing transitions among composite superstates, $\phi_i \equiv (\sigma_j,\psi_k)$ where $i=(j-1)M_\psi+k$ (see the first companion manuscript for details [15] on the structure of such a matrix). This matrix governs the evolution of the system-FRET composite via the master equation

$$\frac{d\boldsymbol{\rho}(t)}{dt} = \boldsymbol{\rho}(t)\mathbf{G}, \tag{1}$$

as described in Sec. 2.3 of the first companion manuscript [15]. Here, $\boldsymbol{\rho}(t)$ is a row vector populated by probabilities for finding the composite in a given superstate at time t.

In estimating these parameters, we must account for all sources of uncertainty present in the experiment, such as shot noise and detector electronics. Therefore, we naturally work within the Bayesian paradigm where the parameters are learned by sampling from probability distributions over these parameters termed posteriors. Such posteriors are proportional to the product of the likelihood, which is the probability of the collected data $\boldsymbol{w}$ given the physical model, and prior distributions over the parameters as follows

$$p(\mathbf{G}|\boldsymbol{w}) \propto L(\boldsymbol{w}|\mathbf{G})p(\mathbf{G}), \tag{2}$$

where $\boldsymbol{w}$ constitutes the set of all observations, including photon arrival times and detection channels.

To construct the posterior, we begin with the likelihood

$$L(\boldsymbol{w}|\mathbf{G}) \propto \boldsymbol{\rho}_{start}\, \mathbf{\Pi}_1^{non}\mathbf{G}_1^{rad}\, \mathbf{\Pi}_2^{non}\mathbf{G}_2^{rad}\, \ldots\, \mathbf{\Pi}_{K-1}^{non}\mathbf{G}_{K-1}^{rad}\, \mathbf{\Pi}_K^{non}\mathbf{G}_K^{rad}\, \mathbf{\Pi}_{end}^{non}\, \boldsymbol{\rho}_{norm}^T, \tag{3}$$

derived in Sec. 2.3 of the first companion manuscript. Here, $\mathbf{\Pi}_k^{non}$ and $\mathbf{G}_k^{rad}$ are the non-radiative and radiative propagators, respectively. Furthermore, $\boldsymbol{\rho}_{start}$ is computed by solving the master equation assuming the system was at steady-state immediately preceding the time at which the experiment began. That is, we solve

$$\boldsymbol{\rho}_{start}\mathbf{G} = 0.$$

Next, assuming that the transition rates are independent of each other, we can write the associated prior as

$$p(\mathbf{G}) = \prod_{i,j} p(\lambda_{\phi_i\to\phi_j}),$$

where we choose Gamma prior distributions over individual rates. That is,

$$p(\lambda_{\phi_i\to\phi_j}) = \mathbf{Gamma}\left(\lambda_{\phi_i\to\phi_j};\alpha,\frac{\lambda_{ref}}{\alpha}\right),$$

to guarantee positive values. Here, ϕ_i represents one of the M_ϕ superstates of the system-FRET composite collecting both the system and photophysical states as described in Sec. 2.2. Furthermore, α and λ_{ref} are parameters of the Gamma prior.

In what follows, we first assume that the number of system states are known and will describe an inverse strategy that uses the posterior above to learn only transition rates. Next, we generalize our model to a nonparametric case accommodating more practical situations with unknown system state numbers. We do so by assuming an infinite dimensional system state space and making the existence of each system state itself a random variable.

3.1 Inference Procedure: Parametric Sampler

Now, with the posterior defined, we prescribe a sampling scheme to learn distributions over all parameters of interest, namely, transitions rates populating $\mathbf{G}$ and the number of system states. However, our posterior in Eq. 2 does not assume a form amenable to analytical calculations. Therefore, we employ Markov Chain Monte Carlo (MCMC) techniques to draw numerical samples.

Particularly convenient here is the Gibbs algorithm that sequentially and separately generates samples for individual transition rates in each MCMC iteration. This requires us to first write the posterior in Eq. 2 using the chain rule as follows

$$p(\mathbf{G}|\boldsymbol{w}) = p(\lambda_{\phi_i\to\phi_j}|\mathbf{G}\backslash\lambda_{\phi_i\to\phi_j}, \boldsymbol{w})p(\mathbf{G}\backslash\lambda_{\phi_i\to\phi_j}|\boldsymbol{w}), \tag{4}$$

where the backslash after $\mathbf{G}$ indicates exclusion of the subsequent rate parameter. Furthermore, the first term on the right hand side is the conditional posterior for the individual rate $\lambda_{\phi_i\to\phi_j}$. The second term in the product is a constant in the corresponding Gibbs step as it is independent of $\lambda_{\phi_i\to\phi_j}$. Similarly, the priors $p(\mathbf{G}\backslash\lambda_{\phi_i\to\phi_j})$ for the rest of the rate parameters on the right hand side of Eq. 2 are also considered constant. Equating the right hand sides of Eqs. 2 & 4 then allows us to write the following conditional posterior for $\lambda_{\phi_i\to\phi_j}$ as

$$p(\lambda_{\phi_i\to\phi_j}|\mathbf{G}\backslash\lambda_{\phi_i\to\phi_j}, \boldsymbol{w}) \propto L(\boldsymbol{w}|\mathbf{G}) \;\; \textbf{Gamma}\left(\lambda_{\phi_i\to\phi_j}; \alpha, \frac{\lambda_{ref}}{\alpha}\right). \tag{5}$$

Since the conditional posterior above does not take a closed form that allows for direct sampling, we use the Metropolis-Hastings (MH) step [32–34] where new samples are drawn from a proposal distribution q and accepted with probability

$$\alpha(\lambda^*_{\phi_i\to\phi_j}, \lambda_{\phi_i\to\phi_j}) = \textbf{min}\left\{1, \frac{p(\lambda^*_{\phi_i\to\phi_j}|\boldsymbol{w}, \mathbf{G}\backslash\lambda_{\phi_i\to\phi_j}) \; q(\lambda_{\phi_i\to\phi_j}|\lambda^*_{\phi_i\to\phi_j})}{p(\lambda_{\phi_i\to\phi_j}|\boldsymbol{w}, \mathbf{G}\backslash\lambda_{\phi_i\to\phi_j}) \; q(\lambda^*_{\phi_i\to\phi_j}|\lambda_{\phi_i\to\phi_j})}\right\}, \tag{6}$$

where the asterisk denotes proposed rate values from the proposal distribution q.

Now, to generate an MCMC chain of samples, we first initialize the chains for all transition rates $\lambda_{\phi_i\to\phi_j}$, by randomly drawing values from their corresponding prior distributions. We then successively iterate across each transition rate in each new MCMC step and draw new samples from the corresponding conditional posterior using the MH criterion.

In the MH step, a convenient choice for the proposal is a Normal distribution leading to a simpler formula for the acceptance probability in Eq. 6. This is due to its symmetry resulting in $q(\lambda_{\phi_i\to\phi_j}|\lambda^*_{\phi_i\to\phi_j}) = q(\lambda^*_{\phi_i\to\phi_j}|\lambda_{\phi_i\to\phi_j})$. However, a Normal proposal distribution would allow forbidden negative transition rates, leading to automatic rejection in the MH step and thus inefficient sampling. Therefore, it is more convenient to propose new samples using a Normal distribution in logarithmic space to allow exploration along the full real line as follows

$$\log(\lambda^*_{\phi_i\to\phi_j}/\kappa) \,\big|\, \log(\lambda_{\phi_i\to\phi_j}/\kappa), \sigma^2 \sim \textbf{Normal}\left(\log(\lambda_{\phi_i\to\phi_j}/\kappa), \sigma^2\right),$$

where $\kappa = 1$ is an auxiliary parameter in the same units as $\lambda_{\phi_i\to\phi_j}$ introduced to obtain a dimensionless quantity within the logarithm.

The transformation above requires introduction of Jacobian factors in the acceptance probability as follows

$$\alpha(\lambda^*_{\phi_i\to\phi_j}, \lambda_{\phi_i\to\phi_j}) = \mathbf{min}\left\{1, \frac{p(\lambda^*_{\phi_i\to\phi_j}|\boldsymbol{w}, \mathbf{G}\backslash\lambda_{\phi_i\to\phi_j})}{p(\lambda_{\phi_i\to\phi_j}|\boldsymbol{w}, \mathbf{G}\backslash\lambda_{\phi_i\to\phi_j})}\frac{(\partial\log(\lambda_{\phi_i\to\phi_j}/\kappa)/\partial\lambda_{\phi_i\to\phi_j})}{(\partial\log(\lambda_{\phi_i\to\phi_j}/\kappa)/\partial\lambda_{\phi_i\to\phi_j})^*}\right\},$$

where the derivatives represent the Jacobian and the proposal distributions are canceled by virtue of using a Normal distribution.

The acceptance probability above depends on the difference of the current and proposed values for a given transition rate. This difference is determined by the covariance of the Normal proposal distribution σ^2 which needs to be tuned for each rate individually to achieve an optimum performance of the BNP-FRET sampler, or equivalently approximately one-third acceptance rate for the proposals [35].

In our case, where the smFRET traces analyzed contain about 10^5 photons, we found it prudent to make the sampler alternate between two sets of variances at every MCMC iteration, $\{\sigma^2_{ex} = 10^{-5}, \sigma^2_{FRET} = 0.01, \sigma^2_{sys} = 0.1\}$ and $\{\sigma^2_{ex} = 10^{-5}, \sigma^2_{FRET} = 0.5, \sigma^2_{sys} = 5.0\}$, for the excitation rates, FRET rates, and system transition rates. This ensure that the sampler is quickly able to explore values at different orders of magnitude.

Intuitively, these covariance values in the proposal distributions above would ideally scale with the relative widths of the conditional posteriors for these parameters (in log-space) if the approximate width could be estimated. Since posterior widths depend on the amount of data used, an increase in the number of photons available in the analysis would require a correspondingly smaller variance.

3.2 Inference Procedure: Nonparametric BNP-FRET Sampler

Here, we first, briefly summarize our inference procedure described in Sec. 3.1 & 3.2.1 of the first companion manuscript [15] for ease of reference.

In realistic situations, the system state space's dimensionality is usually unknown as molecules under study may exhibit complex and unexpected behaviors across conditions and timescales. Consequently, the dimensionality M_ϕ of the generator matrix $\mathbf{G}$ is also unknown, and must be determined by adopting a BNP framework.

In such a framework, we assume an infinite set of system states and place a binary weight, termed load, on each system state such that if it is warranted by the data, the value of the load is realized to one. Put differently, we must place a Bernoulli prior on each candidate state (of which there are formally an infinite number) [36, 37]. In practice, we learn distributions over Bernoulli random variables b_i that activate/deactivate different portions of the full generator matrix as (see Sec. 3.2.1 of the first companion manuscript

[15])

$$\mathbf{G}=\begin{bmatrix} * & b_1^2\lambda_{\psi_1\to\psi_2} & b_1^2\lambda_{\psi_1\to\psi_3} & b_1b_2\lambda_{\sigma_1\to\sigma_2} & 0 & 0 & \dots \\ b_1^2\lambda_{\psi_2\to\psi_1} & * & b_1^2\lambda_{\sigma_1,\psi_2\to\psi_3} & 0 & b_1b_2\lambda_{\sigma_1\to\sigma_2} & 0 & \dots \\ b_1^2\lambda_{\psi_3\to\psi_1} & b_1^2\lambda_{\psi_3\to\psi_2} & * & 0 & 0 & b_1b_2\lambda_{\sigma_1\to\sigma_2} & \dots \\ b_1b_2\lambda_{\sigma_2\to\sigma_1} & 0 & 0 & * & b_2^2\lambda_{\psi_1\to\psi_2} & b_2^2\lambda_{\psi_1\to\psi_3} & \dots \\ 0 & b_1b_2\lambda_{\sigma_2\to\sigma_1} & 0 & b_2^2\lambda_{\psi_2\to\psi_1} & * & b_2^2\lambda_{\sigma_2,\psi_2\to\psi_3} & \dots \\ 0 & 0 & b_1b_2\lambda_{\sigma_2\to\sigma_1} & b_2^2\lambda_{\psi_3\to\psi_1} & b_2^2\lambda_{\psi_3\to\psi_2} & * & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where active loads are set to 1, while inactive loads are set to 0. Furthermore, $*$ represents negative row-sums. Finally, the number of active loads provides an estimate of the number of system states warranted by a given dataset.

As we have introduced new variables we wish to learn, we upgrade the posterior of Eq. 2 to incorporate the full set of loads, $\mathbf{b}=\{b_1,b_2,\dots,b_\infty\}$, as follows

$$p(\mathbf{b},\mathbf{G}|\boldsymbol{w})\propto L(\boldsymbol{w}|\mathbf{b},\mathbf{G})\,p(\mathbf{b})p(\mathbf{G}),$$

where we assume that all parameters of interest are independent of each other.

As in the parametric sampler presented in the previous subsection, we generate samples from the nonparametric posterior above using Gibbs sampling. That is, we first initialize the MCMC chains for loads and rates by drawing random samples from their priors. Next, to construct the chains, we iteratively draw samples from the posterior in two steps: 1) sequentially sample all rates using the MH procedure; then 2) loads by direct sampling, from their corresponding conditional posteriors (as described in Sec. 3.2.1 of the first companion manuscript [15]). Since step (1) is similar to the parametric case, we only focus on the second step in what follows.

To generates samples for load b_i, the corresponding conditional posterior is given by [38]

$$p(b_i|\mathbf{b}\backslash b_i,\mathbf{G},\boldsymbol{w})\propto L(\boldsymbol{w}|\mathbf{b},\mathbf{G})\ \textbf{Bernoulli}\left(b_i;\frac{1}{1+\frac{M_\sigma^{max}-1}{\gamma}}\right),$$

where the backslash after $\mathbf{b}$ indicates exclusion of the following load. We may set the hyperparameters M_σ^{max}, the maximum allowed number of system states used in computations, and γ, the expected number of system states based on simple visual inspection of the smFRET traces.

Now, the conditional posterior in the equation above is discrete and describes the probability for the load to be either active or inactive, that is, it is itself a Bernoulli distribution as follows

$$p(b_i|\mathbf{b}\backslash b_i,\mathbf{G},\boldsymbol{w})=\textbf{Bernoulli}(b_i;q_i),$$

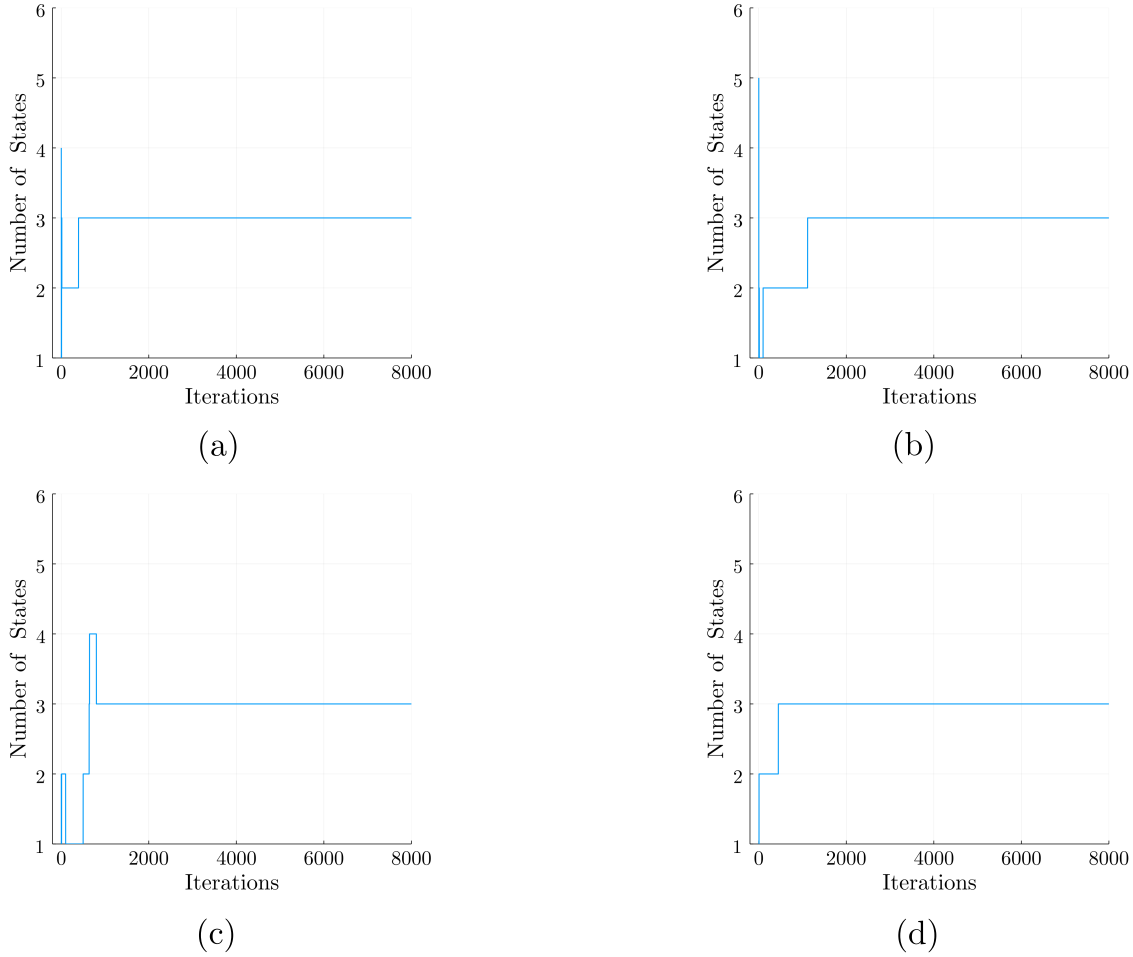
where

$$q_i=\frac{L(\boldsymbol{w}|b_i=1,\mathbf{b}\backslash b_i,\mathbf{G},\boldsymbol{\rho}_{start})}{L(\boldsymbol{w}|b_i=1,\mathbf{b}\backslash b_i,\mathbf{G})+L(\boldsymbol{w}|b_i=0,\mathbf{b}\backslash b_i,\mathbf{G})}.$$

The simple form of this posterior is amenable to direct sampling. In the end, the chain of generated samples can be used for subsequent statistical analysis.

4 Results

In this section, we first demonstrate the robustness of our BNP-FRET sampler by investigating the effects of excitation rate on the distributions over transitions rates and system state numbers. Once we have illustrated the BNP-FRET sampler's performance on synthetic data, we apply it to estimate the number of system states along with associated escape rates from publicly available experimental data for a complex involving instrinsically disorder proteins (ACTR-NCBD). We compare our results with reported literature values [29, 30].



(a)

(b)

(c)

(d)

Figure 1: **MCMC chains generated by the BNP-FRET sampler for the number of system states**. The synthetic smFRET datasets used to generate these chains assume uniform excitation rate of 10 ms^{-1} and FRET efficiencies of 0.09, 0.5, and 0.9, for a three state system. However, the system's escape rates for all three states become faster by a factor of 10 as we move from panel (a) to (d). That is, in the slowest case, we use escape rates of 0.01, 0.02, and 0.03 ms^{-1} for the three system states, while in the fastest case kinetics are as fast as the excitation rate itself. Our method converges to the correct number of system states for each dataset. As we will see later, the rates become more difficult to estimate for panel (d) which we consider to be the point at which the method breaks down.

4.1 Resolution of Timescales Given Excitation Rate: Nonparametrics

To demonstrate the performance of our BNP-FRET sampler over a range of timescales given a fixed excitation rate, we follow the same approach as presented in the first companion manuscript (see Sec. 4.1) [15]. That is, we generate four synthetic smFRET traces containing $K = 2$ million photons each for a biomolecular complex with three system states, $\{\sigma_1, \sigma_2, \sigma_3\}$. The kinetic scheme for this system is a generalization of the example presented in the first companion manuscript [15] (brown boxes) with two system states.

Now, to synthesize smFRET traces, we fix the excitation rate to $\lambda_{ex} = 10\,\text{ms}^{-1}$ and FRET efficiencies ε_{FRET} to 0.09, 0.5, and 0.9 for the three system states, respectively, motivated by experiments in [30]. The remaining parameters are the system transition rates $\lambda_{\sigma_i \to \sigma_j}$, varied across datasets to test our BNP-FRET sampler over a wide range of timescales ranging from a thousand times longer than the average interphoton arrival time $(1/\lambda_{ex})$ to as short as the average interphoton arrival time itself (representing an extreme case). We do not probe kinetics any faster because the excitation rate does not provide enough temporal resolution for resolving system transitions in this regime, as demonstrated in the first manuscript (see Sec. 4.1).

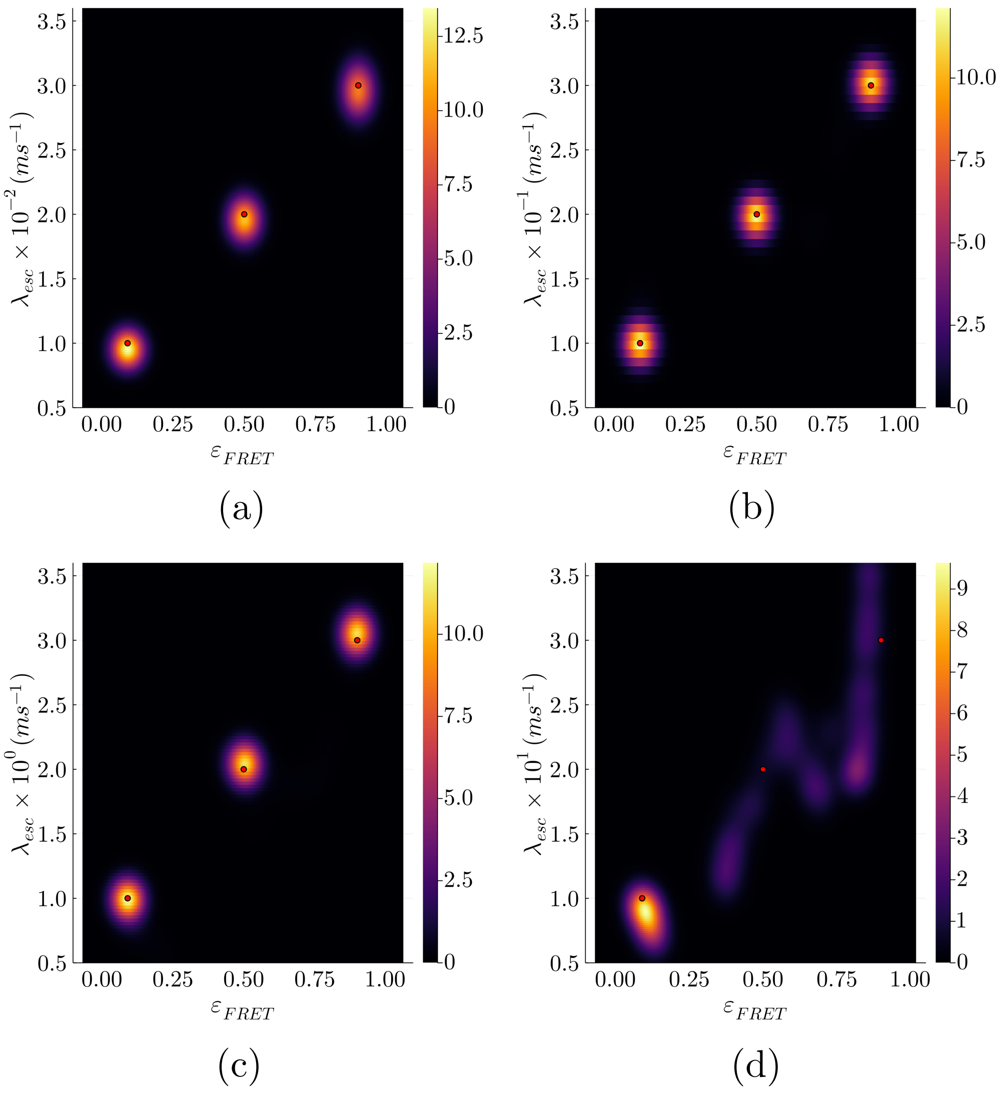
We start the analysis by applying our BNP-FRET sampler to learn the number of system states for the case with slowest escape rates, *i.e.*, the sum of all transition rates out of a given system state. These escape rates are $\lambda_{esc} = 0.01$, 0.02, and 0.03ms^{-1}. We show that our BNP-FRET sampler can correctly learn the number of system states and the associated escape rates and FRET efficiencies; see Fig. 1(a) and Fig. 2(a).

Next we analyze, one-by-one, datasets generated using escape rates that are 10 times faster in each subsequent dataset. BNP-FRET deduces the correct number of system states in all cases (see Fig. 2a-c), however the determination of the rates begins to fail in panel (d) of Fig. 2.

The failure to estimate escape rates approximating the excitation rate can also be predicted using a "photon budget index" defined in the first companion manuscript [15] Sec. 4.1 as

$$s = \frac{K\lambda_{ex}}{\lambda_{probe} M_\sigma} \tag{7}$$

where K and λ_{probe} are, respectively, the photon counts and the escape rate to be probed. Plugging the parameter values associated to the dataset shown in both Fig. 1(d) and 2(d) with three escape rates, *i.e.*, $K = 2 \times 10^6$, $M_\sigma = 3$, $\lambda_{ex} = 10\,\text{ms}^{-1}$ and $\lambda_{probe} = \lambda_{esc} = 10 - 30\,\text{ms}^{-1}$, into the above equation, we obtain $s = 2/3 \times 10^6$, $2/6 \times 10^6$ and $2/9 \times 10^6$. The index obtained for $\lambda_{probe} = 10\text{ms}^{-1}$ is on par with the threshold of $s_{\text{thresh}} = 10^6$ derived in the first companion manuscript in [15] Sec. 4.1 where the sampler had available sufficient information to drawn an accurate inference. By contrast, moving to the larger escape rates of $\lambda_{esc} = 20$, 30ms^{-1} the photon budget indices obtained are much smaller than the threshold and the sampler starts failing due to lack of information. To be more precise, our sampler is capable of learning any escape rates, even those larger than excitation rate, given sufficient photons. As this is counter intuitive, we note that the excitation rate is an average value and there are often photons detected with interphoton intervals much smaller than $1/\lambda_{ex}$. As such, given long photon traces, there are always enough photons with small interphoton



(a)

(b)

(c)

(d)

Figure 2: **Learned bivariate posterior for the escape rates λ_{esc} and FRET efficiencies ε_{FRET} from synthetic data also used in Fig. 1**. Going from panels (a) to (d), we speed up the kinetics (escape rates) by a factor of 10 each time leading to a gradual loss of temporal resolution needed to identify system transitions. The ground truth is shown with the red dots. The estimates for escape rates and FRET rates in panels (a) to (c), have less than 10% errors. However, as seen in panel (d), the excitation rate does not provide enough temporal resolution to resolve system transitions occurring at interphoton arrival time-scales, resulting in large errors in the parameter estimates. The estimated escape rates in panel (d) are $0.8^{+0.1}_{-0.4}$ s^{-1}, $1.4^{+1.0}_{-0.2}$ s^{-1}, and, $2.0^{+1.8}_{-0.3}$ s^{-1}with very large uncertainties (95% confidence intervals). We have smoothed the posterior distributions here using KDE for visualization purposes only.

intervals to learn faster escape rates (and indeed to learn excited state lifetimes as we show in the first companion manuscript [15]) that would otherwise evade binned photon analysis methods [39].

4.2 Analysis of Experimental Data: NCBD-ACTR Interactions

Here, we apply our BNP-FRET sampler to two datasets probing the interactions between partner IDPs, NCBD and ACTR, under different conditions [29, 30]. Precise knowledge of binding and unbinding reactions of such proteins is of fundamental importance toward understanding how they regulate expression of their target genes.

Methods that have been used in the past [29, 30] to analyze smFRET traces from experiments on NCBD-ACTR interaction assumed a fixed number of system states to obtain maximum likelihood point estimates for transition rates. In addition, these methods bin photons to mitigate computational expense. However, given the inherently unstructured and flexible nature of IDPs, fixing the dimensionality of the model *a priori* can be limiting and, as we will see, may bias analysis. Therefore, our nonparametric method which places no constraints on the number of system states while incorporating all major noise sources, is naturally suited.

In the following subsections, we first analyze data for a system where an immobilized ACTR labeled with a Cy3B donor interacts with an NCBD labeled with a CF680R acceptor in the presence of ethylene glycol (EG), 36% by volume, in order to more closely mimic cellular viscosity [29]. Here, the binding of NCBD to ACTR is monitored in smFRET experiments using a confocal microscope setup. Next, we analyze data for a system in a buffer without EG, and therefore with faster kinetics. Here an immobilized ACTR interacts with a freely-diffusing mutated NCBD (P20A) [30].

To acquire both experimental FRET datasets containing about 200000 photons each, laser powers of 0.5 μW and 0.3 μW were used leading to excitation rates varying from 3000 to 11000 s^{-1} in the confocal region depending on where the immobilized sample lies with respect to the center of the excitation laser beam.

Moreover, we are provided a calibrated route correction matrix (RCM) by the authors of [29, 30] to account for spectral crosstalk, and relative detection efficiencies of donor and acceptor channels. We defined such an RCM in Sec. 2.4.1 of the first companion manuscript [15] and specify it for each dataset separately in the following subsections.

Finally, by contrast to the first companion manuscript [15] we ignore the IRF. The latter typically acts over a period of hundreds of picoseconds. As such, it is immaterial on the seconds timescale over which system transitions occur. Moreover, the background values vary for each dataset, and are therefore precalibrated, independently, for each dataset in the corresponding sections.

Now, with all experimental details at hand, we proceed to analyze the experimental data using our BNP-FRET sampler.

4.2.1 Immobilized ACTR in 36% EG

Binding of NCBD to ACTR leads to the formation of a stable and ordered complex in the presence of EG. In addition, when two fluorescent dyes labeling the IDPs come in close proximity, we expect FRET interactions. Therefore, bound and unbound system states of the NCBD-ACTR complex correspond to high and low FRET efficiency signals, respectively.

For the analysis of the collected smFRET data from such a complex, we must take into account all sources of noise such as crosstalk and background. The crosstalk/detection

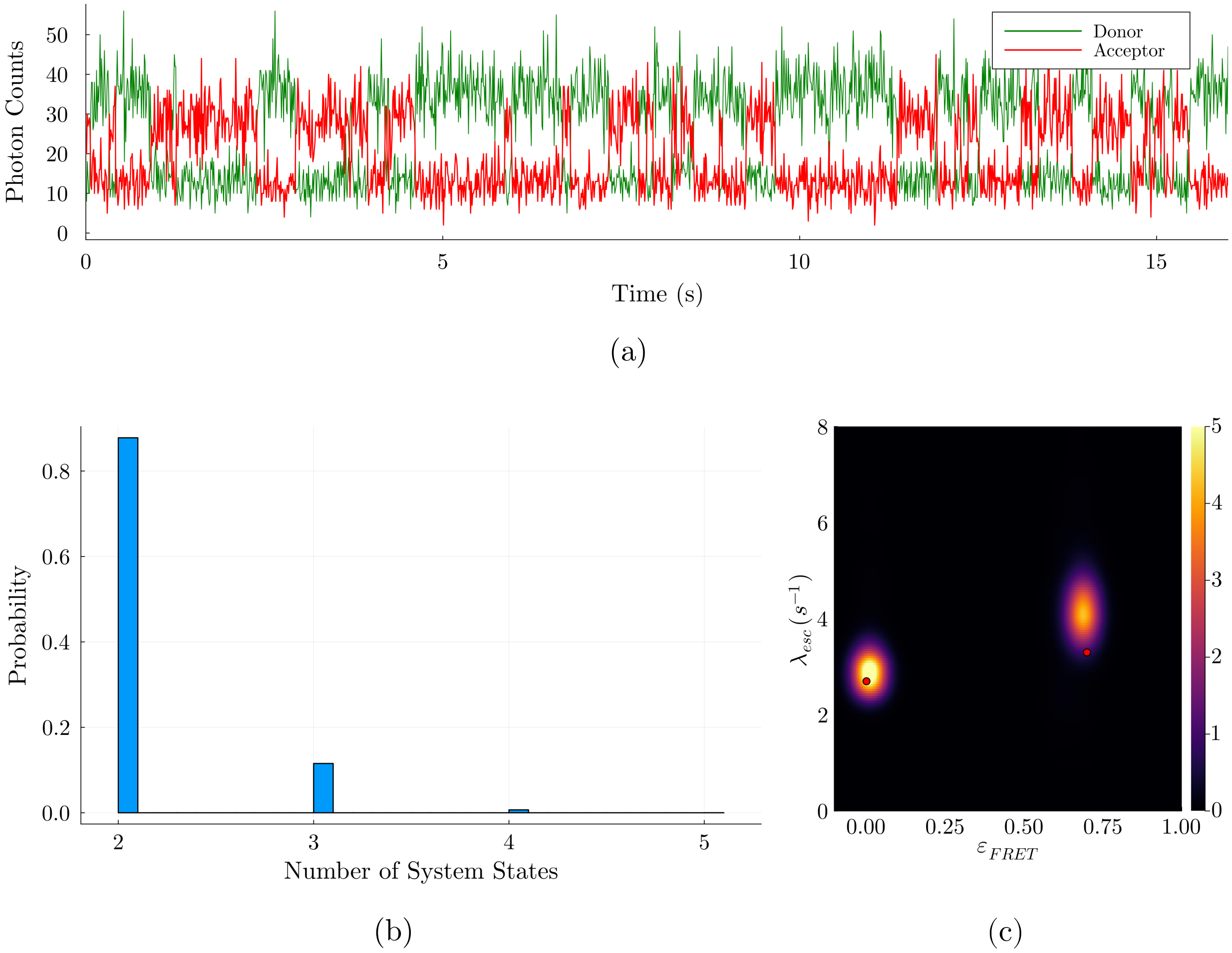


Figure 3: **Results for NCBD-ACTR interactions in the presence of ethylene glycol (EG)**. Panel (a) shows the raw photon counts (bin width of 0.01s) recorded by the two detection channels during the experiment. In panel (b), we show a probability distribution for the number of system states estimated by the BNP-FRET sampler. The sampler spends a majority of its time in two system states with only a small relative probability ascribed to more states. In the posterior distribution for the escape rates and FRET efficiencies in panel (c), two distinct FRET efficiencies are evident with values of about $0.003^{+0.020}_{-0.002}$ (unbound) and $0.70^{+0.02}_{-0.02}$ (bound), and corresponding escape rates of about $2.9^{+0.3}_{-0.3}$ $\mathrm{s}^{-1}$and $4.1^{+0.5}_{-0.4}$ s^{-1}. The red dots show results reported by [29] using maximum likelihood method. We have smoothed the distribution for demonstrative purposes only.

efficiency values are computed from the RCM given by the authors of [29] as

$$\mathbf{RCM} \propto \begin{bmatrix} \phi_{d2} & -\phi_{d1} \\ -\phi_{a2} & \phi_{a1} \end{bmatrix} \propto \begin{bmatrix} 1.0 & -0.22 \\ 0.0 & 1.02 \end{bmatrix},$$

where channels 1 & 2 are, respectively, designed to receive acceptor and donor photons. Furthermore, ϕ_{ai} and ϕ_{di}, respectively, denote probabilities of acceptor and donor photons being registered by channel i. Adopting the same normalization convention for the RCM as in the first companion manuscript [15] (see Example V) gives the following values for the effective crosstalk factors as

$$\phi_{a1} = 0.84,\ \phi_{a2} = 0.0,\ \phi_{d1} = 0.18,\ \text{and}\ \phi_{d2} = 0.82.$$

As such, these values imply that approximately 18% of the emitted donor photons are detected in the acceptor channel due to crosstalk. Furthermore, only 84% of emitted acceptor photons are detected in the acceptor channel, and acceptor photons do not suffer any crosstalk.

We must also incorporate precalibrated background rates for donor and acceptor channels given as 0.283 s^{-1} and 0.467 s^{-1}, respectively [29].

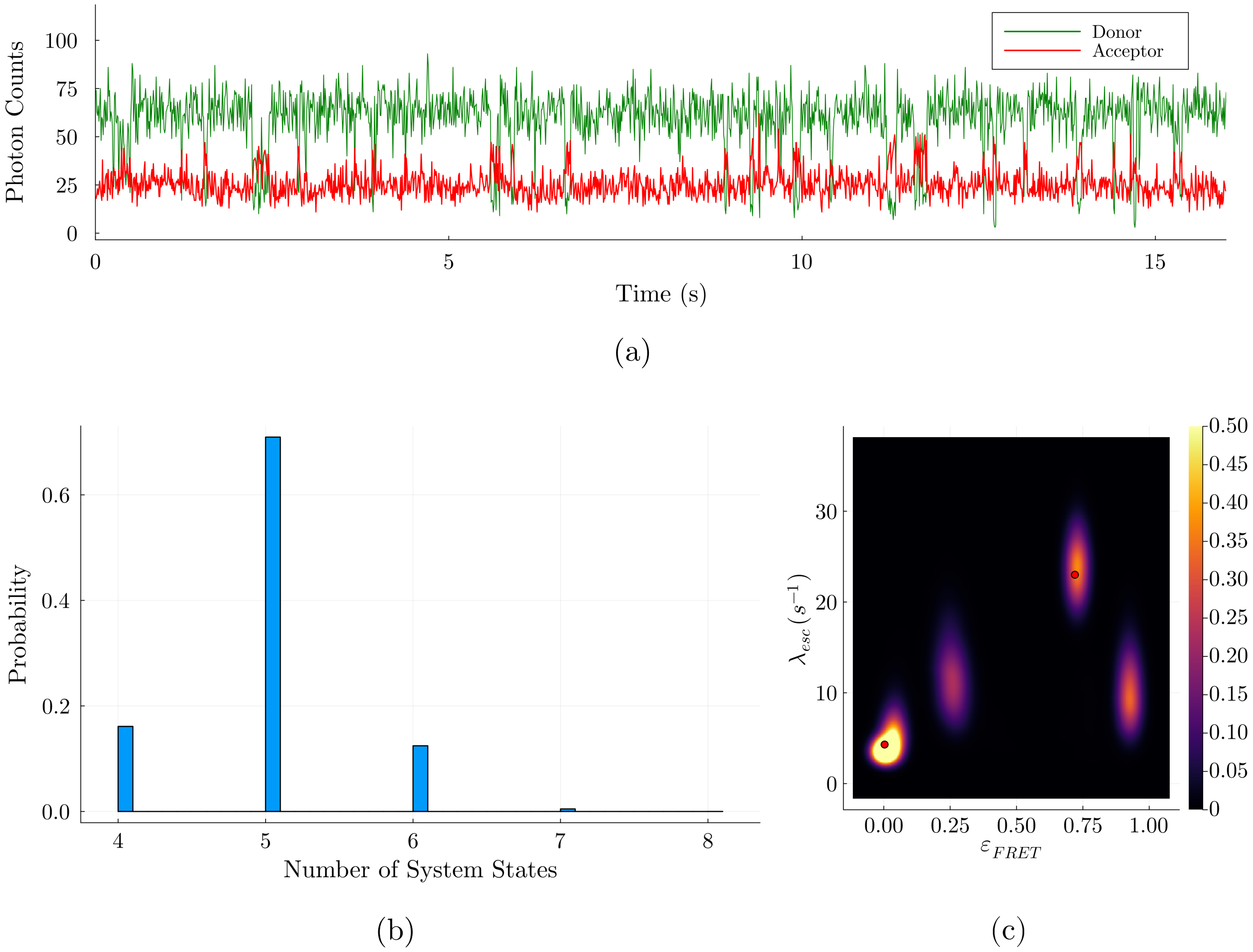
With all such corrections applied, our BNP-FRET sampler now predicts two system states; see Fig. 3. The system state with the lowest FRET efficiency of 0.0 corresponds to the unbound NCBD. The remaining system state with higher FRET efficiency of ≈ 0.7 coincides with the bound NCBD-ACTR complex configuration. The associated escape rates we obtain from our method for both of the system states are approximately 2.9 s^{-1} and 4.1 s^{-1} as seen in Fig. 3b. These results are consistent with results reported in supplementary table S1 of [29] with an average relative difference of $\approx 15\%$.

4.2.2 Immobilized ACTR in buffer

Here, in the absence of EG, the viscosity of the solution is lowered [29], leading to faster system transitions representing a unique analysis challenge.

As in the previous subsection, from the RCM provided by the authors of [30] for the current dataset, we found crosstalk factors of $\phi_{a1} = 0.72$, $\phi_{a2} = 0.0$, $\phi_{d1} = 0.10$, and $\phi_{d2} = 0.90$. After correcting for these crosstalk/detection efficiency values and background rates of 0.312 s^{-1} and 1.561 s^{-1} for the donor and acceptor channels, respectively, our BNP-FRET sampler now predicts five system states (see Fig. 4(a)&(b)) with FRET efficiencies of 0.0, 0.72, 0.03, 0.28, and 0.92 approximately. Here, the first two system states with vanishingly small estimated FRET efficiencies, namely 0.0 and 0.03, most likely represent the same configuration where NCBD is diffusing freely away from the immobilized ACTR, leading to no FRET interactions. Various sources of noise in the dataset may have resulted in this splitting of the unbound system state. Furthermore, the system state with the FRET efficiency and escape rate of approximately 0.72 and 25.0 s^{-1}, respectively, coincides with the previously predicted bound configuration found using a maximum likelihood method with a fixed number of system states [30]. We have compiled the learned transition rates (median values) in the generator matrix below (in s^{-1} units)

$$\mathbf{G}_\sigma = \begin{bmatrix} * & \lambda_{\sigma_1\to\sigma_2} & \lambda_{\sigma_1\to\sigma_3} & \lambda_{\sigma_1\to\sigma_4} & \lambda_{\sigma_1\to\sigma_5} \\ \lambda_{\sigma_2\to\sigma_1} & * & \lambda_{\sigma_2\to\sigma_3} & \lambda_{\sigma_2\to\sigma_4} & \lambda_{\sigma_2\to\sigma_5} \\ \lambda_{\sigma_3\to\sigma_1} & \lambda_{\sigma_3\to\sigma_2} & * & \lambda_{\sigma_3\to\sigma_4} & \lambda_{\sigma_3\to\sigma_5} \\ \lambda_{\sigma_4\to\sigma_1} & \lambda_{\sigma_4\to\sigma_2} & \lambda_{\sigma_4\to\sigma_3} & * & \lambda_{\sigma_4\to\sigma_5} \\ \lambda_{\sigma_5\to\sigma_1} & \lambda_{\sigma_5\to\sigma_2} & \lambda_{\sigma_5\to\sigma_3} & \lambda_{\sigma_5\to\sigma_4} & * \end{bmatrix} = \begin{bmatrix} -4.31 & 3.44 & 0.70 & 0.15 & 0.02 \\ 18.0 & -24.97 & 3.48 & 0.99 & 2.5 \\ 0.21 & 3.98 & -5.10 & 0.91 & 0.003 \\ 1.85 & 0.013 & 7.0 & -8.87 & 0.007 \\ 0.08 & 6.23 & 0.12 & 0.71 & -6.6 \end{bmatrix}, \tag{8}$$



(a)

(b)

(c)

Figure 4: **Results for NCBD-ACTR interactions in buffer, without EG**. Panel (a) shows the raw photon counts (bin width of 0.01s) recorded by the two detection channels during the experiment. In panel (b), we show a probability distribution produced by the BNP-FRET sampler for the number of system states. Models with less than four system states in the histogram are not shown as we ascribe to them zero probability. Indeed, the most probable model contains five system states. Next, in panel (c) depicting the posterior distribution for the escape rates and FRET efficiencies, five distinct FRET efficiencies are evident with values of $0.002^{+0.03}_{-0.001}$, $0.72^{+0.02}_{-0.02}$, $0.03^{+0.02}_{-0.02}$, $0.28^{+0.02}_{-0.02}$, and $0.92^{+0.02}_{-0.01}$ with corresponding escape rates of about $4.3^{+1.9}_{-1.8}$, $25.0^{+2.1}_{-2.9}$, $5.1^{+1.8}_{-1.9}$, $8.9^{+3.5}_{-0.8}$, and $6.6^{+4.0}_{-1.0}$ s^{-1}. The first two system states with almost vanishing FRET efficiencies may represent the same unbound configuration with the small splitting likely arising from various sources of noise present in the dataset. The red dots show the results reported in [30] using maximum likelihood method.

where the diagonal elements correspond to negative of the escape rate values. Furthermore, the steady-state populations/probabilities for these system states can be computed by solving $\boldsymbol{\rho}_{steady}\mathbf{G}_\sigma = 0$, resulting in

$$\boldsymbol{\rho}_{steady} = \begin{bmatrix} 0.55 & 0.12 & 0.23 & 0.05 & 0.05 \end{bmatrix}. \tag{9}$$

Here, the two newly observed system states, with FRET efficiencies of 0.28 and 0.92 and corresponding escape rates of approximately 8.87 s^{-1} and 6.6 s^{-1}, are bound configurations not previously detected [30] and deserve further attention. For instance, lower viscosity buffer (as compared to cases in the presence of EG) may allow the system to visit transient system states more readily under observation timescales [40, 41]. Additionally, steady-state probabilities for these new transient system states that we recover are indeed expectedly low (0.05 and 0.05) as compared to other system states of the NCBD-ACTR complex. Furthermore, IDPs interact in a complex manner with high possibility for residual secondary structures [42]. Competing parametric methods would need to posit a high number of system states *a priori* in order for their kinetics to be quantifiable. Finally, despite a difference in the estimate of the number of system states, our slower kinetics in the presence of EG are consistent with those of Ref. [29]. Direct comparison of escape rates across system states recovered by BNP-FRET versus Ref. [29] however is questionable on account of having recovered a different number of system states.

One way by which we may assure ourselves that these system states are not artefactually added by our computational algorithm (overfitting), is to analyze synthetic data generated under the same conditions (excitation rate, crosstalk, and background) as the experiment but with a ground truth of two system states. We can then ask whether the noise properties force our method to introduce artefactual states. Thus, we simulate a two system state model with the previously reported escape rates [30]) of 4.3 s^{-1} and 23.0 s^{-1} with corresponding FRET efficiencies of 0.0 and 0.8, and the same photon budget of 200000 photons. The results for the analysis of this synthetic dataset in Fig. 5(a)&(b) show no additional system states introduced by our method under this parameter regime suggesting the robustness of our findings for the experimental data.

Another way by which we may assure ourselves is by analyzing synthetically generated data for the four most distinct system states (on the basis of FRET efficiency) predicted by the BNP-FRET sampler for the experimental dataset. These system states correspond to FRET efficiencies of approximately 0.0, 0.72, 0.28, and 0.92 with associated escape rates of 4.31, 24.97, 8.87, and 6.6 s^{-1} as computed from the matrix in Eq. 8. We tested whether our sampler BNP-FRET underfits or overfits with regards to the estimated number of system states. As shown in Fig. 6, the most probable model predicted by the sampler has four system states, again verifying the robustness of our method.

5 Discussion

FRET techniques have been essential in investigating molecular interactions on nanometer scales, for instance, most recently in directly monitoring interaction of the SARS-COV2 virus spike protein with host receptors [43, 44]. Yet, the quantitative interpretation of smFRET data suffers from several issues including difficulties in estimating the number of system states, dealing with fast transition rates and providing uncertainties over estimates, particularly uncertainties over the number of system states [45, 46] originating from multiple noise sources.

Here, we implemented a general nonparametric smFRET data analysis framework presented in the first companion manuscript [15] to address the issues associated with smFRET

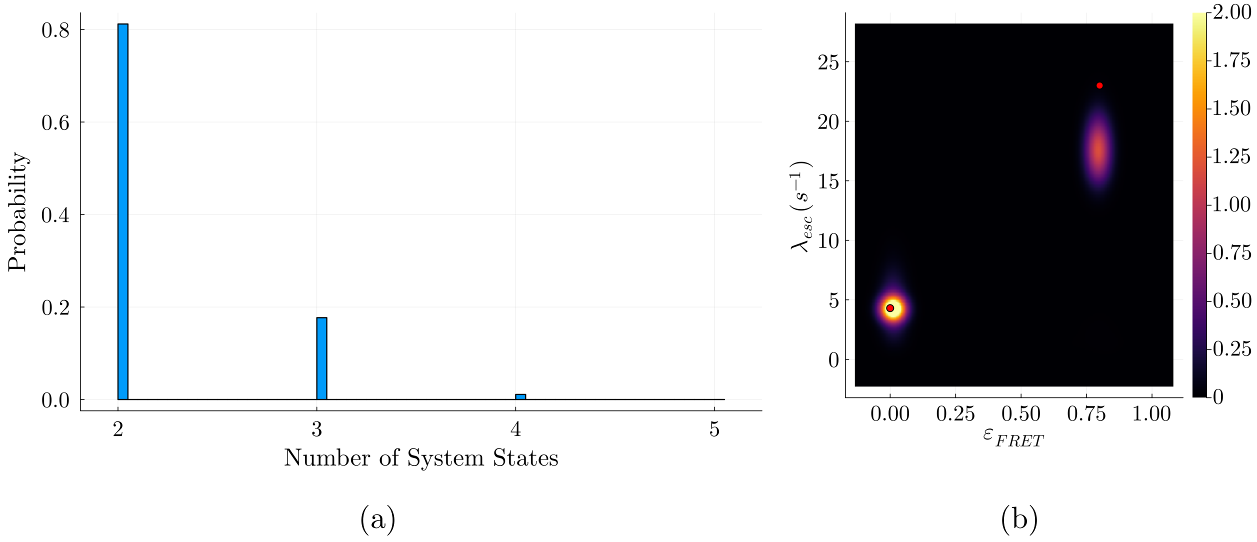
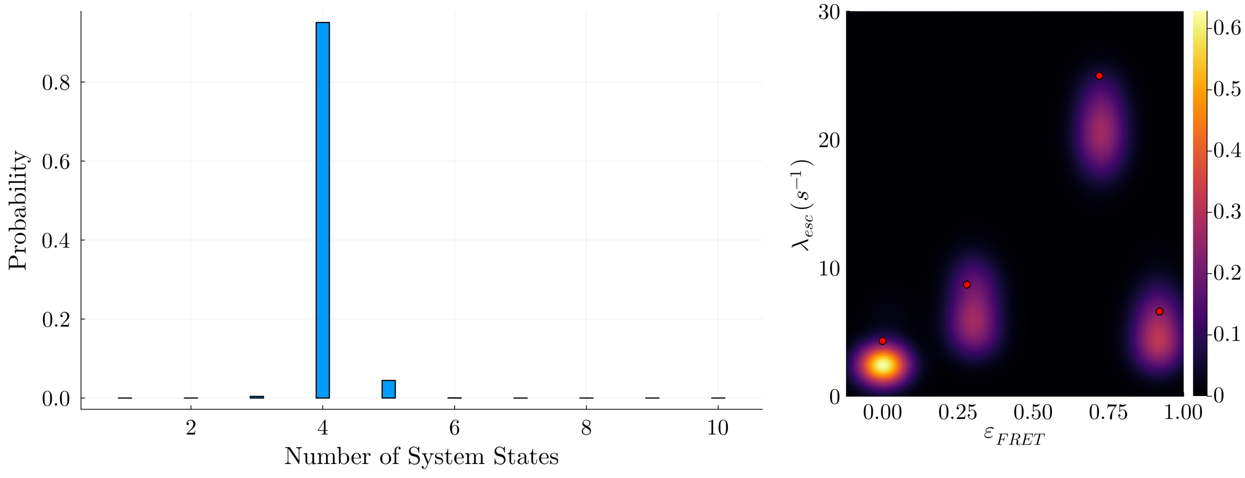


Figure 5: **Robustness test using synthetic data with realistic noise parameters**. The synthetic data here is generated under the same conditions (excitation rate, crosstalk, background, and photon budget) as the experiment whose results are shown in Fig. 4 with only two states as ground truth to see whether the multiple noise sources are likely to result in our method introducing spurious states (such as five states as seen in Fig. 4 using previously reported transitions rates [30]). In panel (a), we show the posterior produced by the BNP-FRET sampler for the number of system states. Fortunately, the most sampled model contains two system states, showing that noise sources do not introduce spurious states in this case. Small relative probability is ascribed to higher dimensional models. In the joint posterior distribution over the escape rates and FRET efficiencies in panel (b), two distinct FRET efficiencies are evident with values of $0.003^{+0.010}_{-0.002}$, $0.8^{+0.013}_{-0.012}$ with corresponding escape rates of about $4.4^{+0.9}_{-0.8}$ and $17.6^{+2.0}_{-1.9}$ s^{-1}. The red dots show the ground truth. The slight bias away from the ground truth results from high noise (background) in the data. The absence of additional system states suggests that the additional system states encountered in the experimental results are not artefactual.

data analysis acquired under continuous illumination. The framework developed can learn posterior distributions over the number of system states as well as the corresponding kinetics ranging from slow values all the way up to kinetic of events occurring on timescales approaching excitation rates. That is, our method propagates uncertainty over not only kinetic parameters but their associated models as well. This is especially significant in avoiding over-commitment to any one model when multiple models are almost equally probable given the data.

We benchmarked our method starting from synthetic data with three system states with a range of different timescales. We challenged our method by simulating data with kinetics as fast as the interphoton arrival times and correctly deduced the system state numbers even under such extreme conditions. We further assessed our method using experimental data acquired observing NCBD interacting with ACTR under different ethylene glycol (EG) concentrations that may impact the timescales at which the binding/unbinding reactions



(a)

(b)

Figure 6: **Second robustness test using synthetic data with realistic noise parameters**. The synthetic data here is generated under the same conditions (excitation rate, crosstalk, background, and photon budget) as the experiment whose results are shown in Fig. 4 with four distinct system states (on the basis of FRET efficiency) as ground truth to see whether our sampler overfits or underfits with regards to the number of system states. These system states correspond to FRET efficiencies of 0.0, 0.72, 0.28, and 0.92 with associated escape rates of 4.31, 24.97, 8.87, and 6.6 s^{-1} as computed from the matrix in Eq. 8. In panel (a), we show the posterior produced by the BNP-FRET sampler for the number of system states. Fortunately, the most sampled model contains four system states, verifying the robustness of our method. Small probabilities are also ascribed to models with different numbers of system states. In the posterior distribution over the escape rates and FRET efficiencies in panel (b), four distinct FRET efficiencies are evident with values of $0.002^{+0.03}_{-0.001}$, $0.72^{+0.03}_{-0.03}$, $0.28^{+0.04}_{-0.03}$, and $0.92^{+0.03}_{-0.03}$ and corresponding escape rates of $3.9^{+2.0}_{-1.5}$, $23.8^{+2.2}_{-1.5}$, $7.1^{+1.9}_{-1.5}$, and $5.9^{+1.7}_{-1.5}$ s^{-1}. Here, the ground truth is shown with red dots. High noise from background results in the underestimates seen here.

occur. In the previous point estimate methods [29, 30], two system states were assumed *a priori* for 0 and 36% EG concentrations. However, our nonparametric method predicts the number of system states and obtains two additional system states in the absence of EG (fast kinetics). This observation may be tied to the inherently unstable nature of the two IDPs under investigation [40].

A careful treatment of how experimental noise propagates into uncertainties over the number of system states and rates does come with associated computational cost. Other methods have managed to mitigate these costs by making approximations including: 1) assuming kinetics much slower than fluorophore excitation and relaxation rates [13, 14]; 2) assuming fast dye photophysics is completely irrelevant to the system transition rate and that FRET efficiency sufficiently identifies transitions between system states [14]; 3) ignoring detector effects and relegating other noise sources, such as background, to post-processing steps [13]; and, most popularly, 4) binning data [45, 47, 48]. For the general case without such approximations, however, the primary computation—the likelihood—remains expensive due to the required evaluation of many matrix exponentials. This cost can be mitigated in a number of ways by, for instance, computing likelihoods for several data traces in parallel. The scaling of the method is provided in the first companion manuscript [15].

The method described in this paper was developed for cases with discrete system state spaces. For continuous state spaces, both the likelihood and priors would require major modification in the spirit of Refs. [49, 50].

Our framework can accommodate different illumination modalities such as alternating laser excitation (ALEX) [51] to directly excited both donor and acceptor dyes by assuming nonzero direct excitation rates in the generator matrix. Indeed, direct excitation of the acceptor would further help in the simultaneous determination of crosstalk factors, detection efficiencies, and quantum yield of the dyes alongside kinetics.

6 Code availability

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

7 Declaration of Interest

The authors declare no competing interests.

8 Acknowledgments

We thank Weiqing Xu and Dr. Zeliha Kilic for regular feedback and help, especially during the development of the nonparametrics samplers. We also thank Prof. Benjamin Schuler, Dr. Daniel Nettels, and Oliver Stach for regular feedback and providing the necessary experimental data. S.P. acknowledges support from the NIH NIGMS (R01GM130745) for supporting early efforts in nonparametrics and NIH NIGMS (R01GM134426) for supporting single-photon efforts. The majority of the computations were performed on the Agave and Sol supercomputers at ASU.

Bibliography

[1] Eitan Lerner, Anders Barth, Jelle Hendrix, Benjamin Ambrose, Victoria Birkedal, Scott C Blanchard, Richard Börner, Hoi Sung Chung, Thorben Cordes, Timothy D Craggs, Ashok A Deniz, Jiajie Diao, Jingyi Fei, Ruben L Gonzalez, Irina V Gopich, Taekjip Ha, Christian A Hanke, Gilad Haran, Nikos S Hatzakis, Sungchul Hohng, Seok-Cheol Hong, Thorsten Hugel, Antonino Ingargiola, Chirlmin Joo, Achillefs N Kapanidis, Harold D Kim, Ted Laurence, Nam Ki Lee, Tae-Hee Lee, Edward A Lemke, Emmanuel Margeat, Jens Michaelis, Xavier Michalet, Sua Myong, Daniel Nettels, Thomas-Otavio Peulen, Evelyn Ploetz, Yair Razvag, Nicole C Robb, Benjamin Schuler, Hamid Soleimaninejad, Chun Tang, Reza Vafabakhsh, Don C Lamb, Claus AM Seidel, and Shimon Weiss. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife*, 10:e60416, 2021.

[2] James J McCann, Ucheor B Choi, Liqiang Zheng, Keith Weninger, and Mark E Bowen. Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. *Biophysical Journal*, 99(3):961, 2010.

[3] Steve Pressé, Julian Lee, and Ken A Dill. Extracting conformational memory from single-molecule kinetic data. *The Journal of Physical Chemistry B*, 117(2):495, 2013.

[4] Steve Pressé, Jack Peterson, Julian Lee, Phillip Elms, Justin L MacCallum, Susan Marqusee, Carlos Bustamante, and Ken Dill. Single molecule conformational memory extraction: P5ab RNA hairpin. *The Journal of Physical Chemistry B*, 118(24):6597, 2014.

[5] 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, 2018.

[6] Jae-Yeol Kim, Cheolhee Kim, and Nam Ki Lee. Real-time submillisecond single-molecule FRET dynamics of freely diffusing molecules with liposome tethering. *Nature Communications*, 6(1):6992, 2015.

[7] Benjamin Schuler and Hagen Hofmann. Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Current Opinion in Structural Biology*, 23(1):36, 2013.

[8] Y. Wu, X. Wu, R. Lu, M. Li, L. Toro, and E. Stefani. Super-resolution light microscopy: Stimulated emission depletion and ground-state depletion. In Ralph A. Bradshaw and Philip D. Stahl, editors, *Encyclopedia of Cell Biology*, page 76. Waltham, 2016.

[9] 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.

[10] Matthew Safar, Ayush Saurabh, Bidyut Sarkar, Mohamadreza Fazel, Kunihiko Ishii, Tahei Tahara, Ioannis Sgouralis, and Steve Pressé. Single photon smFRET. III. application to pulsed illumination. *bioRxiv*, 2022.

[11] Christian Eggeling, Jerker Widengren, Leif Brand, Jörg Schaffer, Suren Felekyan, and Claus A. M. Seidel. Analysis of photobleaching in single-molecule multicolor excitation and Förster resonance energy transfer measurements. *The Journal of Physical Chemistry A*, 110(9):2979, 2006.

[12] Colton Boudreau, Tse-Luen (Erika) Wee, Yan-Rung (Silvia) Duh, Melissa P. Couto, Kimya H. Ardakani, and Claire M. Brown. Excitation light dose engineering to reduce photo-bleaching and photo-toxicity. *Scientific Reports*, 6(1):30892, 2016.

[13] Irina V. Gopich and Attila Szabo. Decoding the pattern of photon colors in single-molecule FRET. *The Journal of Physical Chemistry B*, 113(31):10965, 2009.

[14] Irina Gopich and Attila Szabo. Theory of photon statistics in single-molecule Förster resonance energy transfer. *The Journal of Chemical Physics*, 122(1):014707, 2005.

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

[16] Sean A. McKinney, Chirlmin Joo, and Taekjip Ha. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical Journal*, 91(5):1941, 2006.

[17] 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.

[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, 2021.

[19] Zeliha Kilic, Ioannis Sgouralis, Wooseok Heo, Kunihiko Ishii, Tahei Tahara, and Steve Pressé. A continuous time representation of smFRET for the extraction of rapid kinetics. *Biophysical Journal*, 120(3):186a, 2021.

[20] 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.

[21] Daniel Nettels, Irina V. Gopich, Armin Hoffmann, and Benjamin Schuler. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proceedings of the National Academy of Sciences*, 104(8):2655, 2007.

[22] Irina V. Gopich and Attila Szabo. Single-molecule FRET with diffusion and conformational dynamics. *The Journal of Physical Chemistry B*, 111(44):12925, 2007.

[23] 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):3336, 2020.

[24] Janghyun Yoo, John M. Louis, Irina V. Gopich, and Hoi Sung Chung. Three-color single-molecule FRET and fluorescence lifetime analysis of fast protein folding. *The Journal of Physical Chemistry B*, 122(49):11702, 2018.

[25] Irina V. Gopich and Attila Szabo. Single-macromolecule fluorescence resonance energy transfer and free-energy profiles. *The Journal of Physical Chemistry B*, 107(21):5058, 2003.

[26] K.V. Mardia, H.R. Southworth, and C.C. Taylor. On bias in maximum likelihood estimators. *Journal of Statistical Planning and Inference*, 76(1):31, 1999.

[27] Rolf Sundberg. Flat and multimodal likelihoods and model lack of fit in curved exponential families. *Scandinavian Journal of Statistics*, 37(4):632, 2010.

[28] Malcolm Sambridge. A parallel tempering algorithm for probabilistic sampling and multimodal optimization. *Geophysical Journal International*, 196(1):357, 2013.

[29] Franziska Zosel, Andrea Soranno, Karin J. Buholzer, Daniel Nettels, and Benjamin Schuler. Depletion interactions modulate the binding between disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences*, 117(24):13480, 2020.

[30] Franziska Zosel, Davide Mercadante, Daniel Nettels, and Benjamin Schuler. A proline switch explains kinetic heterogeneity in a coupled folding and binding reaction. *Nature Communications*, 9(1):3332, 2018.

[31] Narayanan Gopalakrishna Iyer, Hilal Özdag, and Carlos Caldas. p300/CBP and cancer. *Oncogene*, 23(24):4225, 2004.

[32] 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.

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

[34] Christopher M. Bishop. Pattern recognition. *Machine Learning*, 128(9), 2006.

[35] Andrew Gelman, Walter R Gilks, and Gareth O Roberts. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1):110, 1997.

[36] 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.

[37] 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.

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

[39] Mark Jäger, Alexander Kiel, Dirk-Peter Herten, and Fred A Hamprecht. Analysis of single-molecule fluorescence spectroscopic data with a Markov-modulated Poisson process. *ChemPhysChem*, 10(14):2486, 2009.

[40] Andrea Soranno, Iwo Koenig, Madeleine B Borgia, Hagen Hofmann, Franziska Zosel, Daniel Nettels, and Benjamin Schuler. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. *Proceedings of the National Academy of Sciences*, 111(13):4874, 2014.

[41] Daniel Johansen, Cy MJ Jeffries, Boualem Hammouda, Jill Trewhella, and David P Goldenberg. Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophysical Journal*, 100(4):1120, 2011.

[42] Anthony Banks, Sanbo Qin, Kevin L Weiss, Christopher B Stanley, and Huan-Xiang Zhou. Intrinsically disordered protein exhibits both compaction and expansion under macromolecular crowding. *Biophysical Journal*, 114(5):1067, 2018.

[43] Maolin Lu. Single-molecule FRET imaging of virus spike–host interactions. *Viruses*, 13(2):332, 2021.

[44] Byunghoon Kang, Youngjin Lee, Jaewoo Lim, Dongeun Yong, Young Ki Choi, Sun Woo Yoon, Seungbeom Seo, Soojin Jang, Seong Uk Son, Taejoon Kang, et al. FRET-based hACE2 receptor mimic peptide conjugated nanoprobe for simple detection of SARS-CoV-2. *Chemical Engineering Journal*, 442:136143, 2022.

[45] Sean A. McKinney, Chirlmin Joo, and Taekjip Ha. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical Journal*, 91(5):1941, 2006.

[46] Jonathan E. Bronson, Jingyi Fei, Jake M. Hofman, Ruben L. Gonzalez, Jr., and Chris H. Wiggins. Learning rates and states from biophysical time series: A Bayesian approach to model selection and single-molecule FRET data. *Biophysical Journal*, 97(12):3196, 2009.

[47] Benjamin Schuler. Single-molecule fluorescence spectroscopy of protein folding. *ChemPhysChem*, 6(7):1206, 2005.

[48] 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.

[49] J Shepard Bryan and Steve Presse. Learning continuous potentials from smfret. *bioRxiv*, 2022.

[50] Irina V. Gopich and Attila Szabo. Theory of the statistics of kinetic transitions with application to single-molecule enzyme catalysis. *The Journal of Chemical Physics*, 124(15):154712, 2006.

[51] Achillefs N Kapanidis, Ted A Laurence, Nam Ki Lee, Emmanuel Margeat, Xiangxu Kong, and Shimon Weiss. Alternating-laser excitation of single molecules. *Accounts of Chemical Research*, 38(7):523, 2005.