Supporting Information
A Bayesian Nonparametric Approach to Single Molecule Förster Resonance Energy Transfer

Ioannis Sgouralis,[1] Shreya Madaan,[2] Franky Djutanta,[3] Rachael Kha,[4] Rizal F. Hariadi,[3,1] and Steve Pressé[1,5]

[1)] *Center for Biological Physics, Department of Physics, Arizona State University, Tempe, AZ 85287, USA*

[2)] *School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287, USA*

[3)] *Biodesign Center for Molecular Design and Biomimetics, Biodesign Institute, Arizona State University, Tempe, AZ 85287, USA*

[4)] *School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA*

[5)] *School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA*[a)]

Appendix A: Notation

Table S1 summarizes the notation used throughout this study. Overall, we use $n = 0, 1, \dots, N$ to index time levels (timestamp) and $m = 1, 2, \dots, M$ to index molecular states. For clarity, we note that t_0 is not associated with any of the intensities in $\vec{I}^D, \vec{I}^A$ and that our analysis employs $M = \infty$.

In grouping similar variables, we use: arrows to denote lists (traces) over time e.g. $\vec{s} = (s_1, s_2, \dots, s_N)$; tildes to denote lists over the molecule state-space e.g. $\tilde{\lambda}^D = (\lambda^D_{\sigma_1}, \lambda^D_{\sigma_2}, \dots, \lambda^D_{\sigma_M})$; double tildes to denote lists of lists over the molecular state space e.g. $\tilde{\tilde{\pi}} = (\tilde{\pi}_{\sigma_1}, \tilde{\pi}_{\sigma_2}, \dots, \tilde{\pi}_{\sigma_M}, \tilde{\pi}_*)$; and bars to gather photo-physical parameters e.g. $\bar{\omega}^D = (\omega^D_0, \omega^D_1, \omega^D_*)$.

The sttistical notation "$X \sim F$", where X is a random variable and F is a probability distribution, indicates[1–3] that X is sampled from F or simply that X follows the probability measure (i.e. probability density or probability mass function) associated with F.

TABLE S1. Summary of notation used in this study.

Variable	Description
δt	data acquisition time
$\delta \tau$	exposure time
t_n	time at step n
I^D_n	donor intensity measured at step n
I^A_n	acceptor intensity measured at step n
f^D_n	donor photo-state at step n
f^A_n	acceptor photo-state at step n
s_n	molecule state at step n
σ_m	distinct molecule state
$\pi_{\sigma_m \to \sigma_{m'}}$	transition probability from σ_m to $\sigma_{m'}$ in one step
$\pi_{* \to \sigma_m}$	initial probability for σ_m
ξ^D	background donor photo-emission rate
ξ^A	background acceptor photo-emission rate
λ^S	donor photo-emission rate without FRET
$\lambda^D_{\sigma_m}$	donor photo-emission rate of σ_m with FRET
$\lambda^A_{\sigma_m}$	acceptor photo-emission rate of σ_m with FRET
ϵ_{σ_m}	FRET efficiency of σ_m
λ^{D*}_n	apparent donor photo-emission rate at step n
λ^{A*}_n	apparent donor photo-emission rate at step n
ϵ^*_n	apparent FRET efficiency at step n
ω^D_0	donor photo-activation, after dark step
ω^D_1	donor photo-activation, after bright step
ω^D_*	donor photo-activation, at initial step
ω^A_0	acceptor photo-activation, after dark step
ω^A_1	acceptor photo-activation, after bright step
ω^A_*	acceptor photo-activation, at initial step
$c^{D \to D}$	donor-to-donor cross-talk coefficient
$c^{D \to A}$	donor-to-acceptor cross-talk coefficient
$c^{A \to A}$	acceptor-to-acceptor cross-talk coefficient
$c^{A \to D}$	acceptor-to-donor cross-talk coefficient
q^D	quantum efficiency in the donor's wavelength
q^A	quantum efficiency in the acceptor's wavelength

Appendix B: Summary of Equations

For completeness, below we summarize the entire set of equations used in the formulation and analysis of smFRET measurements. In generative form and full statis-

[a)]Electronic mail: spresse@asu.edu

tical notation, these are

$$\tilde{\beta} \sim \mathbf{GEM}_{\sigma_1,\sigma_2,\ldots}(\gamma) \quad \text{(B1)}$$
$$\tilde{\pi}_*|\tilde{\beta} \sim \mathbf{DP}_{\sigma_1,\sigma_2,\ldots}(\alpha\tilde{\beta}) \quad \text{(B2)}$$
$$\tilde{\pi}_{\sigma_m}|\tilde{\beta} \sim \mathbf{DP}_{\sigma_1,\sigma_2,\ldots}(\alpha\tilde{\beta}) \quad \text{(B3)}$$

$$\kappa^D_{\sigma_m} \sim \mathbf{Gamma}\left(\phi^D_\kappa, \psi^D_\kappa/\phi^D_\kappa\right) \quad \text{(B4)}$$
$$\kappa^A_{\sigma_m} \sim \mathbf{Gamma}\left(\phi^A_\kappa, \psi^A_\kappa/\phi^A_\kappa\right) \quad \text{(B5)}$$
$$\kappa^S \sim \mathbf{Gamma}\left(\phi^S_\kappa, \psi^S_\kappa/\phi^S_\kappa\right) \quad \text{(B6)}$$

$$\theta \sim \mathbf{Gamma}\left(\phi_\theta, \psi_\theta/\phi_\theta\right) \quad \text{(B7)}$$
$$\rho^D \sim \mathbf{Gamma}\left(\phi^D_\rho, \psi^D_\rho/\phi^D_\rho\right) \quad \text{(B8)}$$
$$\rho^A \sim \mathbf{Gamma}\left(\phi^A_\rho, \psi^A_\rho/\phi^A_\rho\right) \quad \text{(B9)}$$

$$\omega^D_* \sim \mathbf{Beta}\left(\eta^D_*, \zeta^D_*\right) \quad \text{(B10)}$$
$$\omega^D_0 \sim \mathbf{Beta}\left(\eta^D_0, \zeta^D_0\right) \quad \text{(B11)}$$
$$\omega^D_1 \sim \mathbf{Beta}\left(\eta^D_1, \zeta^D_1\right) \quad \text{(B12)}$$

$$\omega^A_* \sim \mathbf{Beta}\left(\eta^A_*, \zeta^A_*\right) \quad \text{(B13)}$$
$$\omega^A_0 \sim \mathbf{Beta}\left(\eta^A_0, \zeta^A_0\right) \quad \text{(B14)}$$
$$\omega^A_1 \sim \mathbf{Beta}\left(\eta^A_1, \zeta^A_1\right) \quad \text{(B15)}$$

$$s_1|\tilde{\tilde{\pi}} \sim \mathbf{Categorical}_{\sigma_1,\sigma_2,\ldots}\left(\tilde{\pi}_*\right) \quad \text{(B16)}$$
$$f^D_1|\bar{\omega}^D \sim \mathbf{Bernoulli}(\omega^D_*) \quad \text{(B17)}$$
$$f^A_1|\bar{\omega}^A \sim \mathbf{Bernoulli}(\omega^A_*) \quad \text{(B18)}$$

$$s_n|s_{n-1},\tilde{\tilde{\pi}} \sim \mathbf{Categorical}_{\sigma_1,\sigma_2,\ldots}\left(\tilde{\pi}_{s_{n-1}}\right) \quad \text{(B19)}$$
$$f^D_n|f^D_{n-1},\bar{\omega}^D \sim \mathbf{Bernoulli}(\omega^D_{f^D_{n-1}}) \quad \text{(B20)}$$
$$f^A_n|f^A_{n-1},\bar{\omega}^A \sim \mathbf{Bernoulli}(\omega^A_{f^A_{n-1}}) \quad \text{(B21)}$$

$$I^D_n|\mu^D_n \sim \mathbf{Poisson}\left(\mu^D_n\delta\tau\right) \quad \text{(B22)}$$
$$I^A_n|\mu^A_n \sim \mathbf{Poisson}\left(\mu^A_n\delta\tau\right) \quad \text{(B23)}$$

From those: Eqs. (B3)–(B5) apply for $m = 1, 2, \ldots$, Eqs. (B19)–(B21) apply for $n = 2, \ldots, N$, and Eqs. (B21)–(B22) apply for $n = 1, \ldots, N$.

According to the parametrization of the $\mathbf{Gamma}(\phi, \psi/\phi)$ probability distribution we employ, ϕ is the usual shape parameter, and ψ is the mean value.

Appendix C: Induced Priors on FRET Efficiency and Emission Rates

Under the Gamma priors employed on θ and $\kappa^S, \kappa^D_{\sigma_m}, \kappa^A_{\sigma_m}, \rho^D, \rho^A$ (see App. B), the induced (i.e. marginal) priors on FRET efficiency and photo-emission rates can be obtained by appropriate transformations of random variables[4]. Specifically, the associated prior probability densities, after the involved calculus, become

$$\mathcal{P}\left(\epsilon_{\sigma_m}\right) \propto \frac{\left(\frac{1}{\epsilon_{\sigma_m}} - 1\right)^{\phi^D_\kappa - 1}}{\left(1 + \frac{\psi^A_\kappa \phi^A_\kappa}{\psi^D_\kappa \phi^D_\kappa}\left(\frac{1}{\epsilon_{\sigma_m}} - 1\right)\right)^{\phi^D_\kappa + \phi^A_\kappa} \epsilon^2_{\sigma_m}} \quad \text{(C1)}$$

$$\mathcal{P}\left(\lambda^D_{\sigma_m}\right) \propto \frac{K_{\phi_\theta - \phi^D_\kappa}\left(2\sqrt{\frac{\phi_\theta \phi^D_\kappa}{\psi_\theta \psi^D_\kappa}\lambda^D_{\sigma_m}}\right)}{\left(\lambda^D_{\sigma_m}\right)^{1 - \frac{\phi_\theta + \phi^D_\kappa}{2}}} \quad \text{(C2)}$$

$$\mathcal{P}\left(\lambda^A_{\sigma_m}\right) \propto \frac{K_{\phi_\theta - \phi^A_\kappa}\left(2\sqrt{\frac{\phi_\theta \phi^A_\kappa}{\psi_\theta \psi^A_\kappa}\lambda^A_{\sigma_m}}\right)}{\left(\lambda^A_{\sigma_m}\right)^{1 - \frac{\phi_\theta + \phi^A_\kappa}{2}}} \quad \text{(C3)}$$

$$\mathcal{P}\left(\lambda^S\right) \propto \frac{K_{\phi_\theta - \phi^S_\kappa}\left(2\sqrt{\frac{\phi_\theta \phi^S_\kappa}{\psi_\theta \psi^S_\kappa}\lambda^D_S}\right)}{\left(\lambda^S\right)^{1 - \frac{\phi_\theta + \phi^S_\kappa}{2}}} \quad \text{(C4)}$$

$$\mathcal{P}\left(\xi^D\right) \propto \frac{K_{\phi_\theta - \phi^D_\rho}\left(2\sqrt{\frac{\phi_\theta \phi^D_\rho}{\psi_\theta \psi^D_\rho}\xi^D}\right)}{\left(\xi^D\right)^{1 - \frac{\phi_\theta + \phi^D_\rho}{2}}} \quad \text{(C5)}$$

$$\mathcal{P}\left(\xi^A\right) \propto \frac{K_{\phi_\theta - \phi^A_\rho}\left(2\sqrt{\frac{\phi_\theta \phi^A_\rho}{\psi_\theta \psi^A_\rho}\xi^A}\right)}{\left(\xi^A\right)^{1 - \frac{\phi_\theta + \phi^A_\rho}{2}}} \quad \text{(C6)}$$

where K_ϕ denotes the modified Bessel function of the second kind[5] and, due to clarity, normalization constants in all densities are omitted.

In the case of $\phi^D_\kappa = \phi^A_\kappa$ and $\psi^D_\kappa = \psi^A_\kappa$, the prior on ϵ_{σ_m} takes the form of a Beta probability distribution, which becomes non-informative (uniform) under $\phi^D_\kappa = \phi^A_\kappa = 1$. The prior on the photo-emission rates are already in the form of K probability distributions[6].

Appendix D: Computational Scheme

A fully working version of the method described in this study, termed **bl-ICON**, in Matlab source code and GUI formats, is available in the SUPPLEMENTAL MATERIAL.

As mentioned in Sec. , the implementation generates samples from the posterior probability distribution $\mathcal{P}\left(\theta, \rho^D, \rho^A, \tilde{\kappa}^D, \tilde{\kappa}^A, \tilde{\tilde{\pi}}, \bar{\omega}^D, \bar{\omega}^A, \vec{s}, \vec{f}^D, \vec{f}^A | \vec{I}^D, \vec{I}^A\right)$ through an MCMC scheme[7,8] that relies on Gibbs sampling.

Specifically, for the nonparametric variables, as in previous studies[9–11], we use a weak limit approximation that replaces Eq. (13) with:

$$\tilde{\beta} \sim \mathbf{Dir}_{\sigma_1,\sigma_2,\ldots,\sigma_M}\left(\frac{\gamma}{M}, \frac{\gamma}{M}, \cdots, \frac{\gamma}{M}\right). \quad \text{(D1)}$$

The existing statistical theory[9,10] ensures that, for sufficiently large M, Eq. (D1) yields estimates that are indistinguishable from Eq. (13).

Under Eq. (D1), our implementation generates posterior samples by iterating the following Gibbs steps:

Step 1: Update fluorophore and molecule states $(\vec{f}^D, \vec{f}^A, \vec{s})$ jointly using forward filtering-backward sampling[12?] .

Step 2: Update the non-parametric base $\tilde{\beta}$ as described previously[9].

Step 3: Update transition probabilities $\tilde{\tilde{\pi}}$, $\tilde{\pi}_*$ and photo-switching probabilities $\vec{\omega}^D$, $\vec{\omega}^A$.

Step 4: Update the reference photo-emission rate θ.

Step 5: Update photo-emission background ρ^D, ρ^A and fluorophore $\tilde{\kappa}^D, \tilde{\kappa}^A, \kappa^S$ multipliers jointly using a Hamiltonian Monte Carlo[13,14] or Metropolis-Hastings-within-Gibbs[7,8] steps.

Due to the conjugacies[2] Categorical-Dirichlet and Poisson-Gamma, sampling of all variables in steps 3 and 4 is achieved directly.

Fine details of the outlined implementation can be found in the **bl-ICON**'s source code, available in the Supplemental Material

[1] U. Von Toussaint, "Bayesian inference in physics," Reviews of Modern Physics **83**, 943 (2011).
[2] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, Vol. 2 (CRC press Boca Raton, FL, 2014).
[3] D. Sivia and J. Skilling, *Data analysis: a Bayesian tutorial* (OUP Oxford, 2006).
[4] P. M. Lee, *Bayesian statistics: an introduction* (John Wiley & Sons, 2012).
[5] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Vol. 55 (Courier Corporation, 1964).
[6] E. Jakeman, "On the statistics of k-distributed noise," Journal of Physics A: Mathematical and General **13**, 31 (1980).
[7] C. Robert and G. Casella, *Monte Carlo statistical methods* (Springer Science & Business Media, 2013).
[8] C. Robert and G. Casella, *Introducing Monte Carlo Methods with R* (Springer Science & Business Media, 2009).
[9] E. Fox, E. Sudderth, M. Jordan, and A. Willsky, "A sticky HDP-HMM with application to speaker diarization," The Annals of Applied Statistics , 1020–1056 (2011).
[10] Y. Teh, M. Jordan, M. Beal, and D. Blei, "Hierarchical Dirichlet processes," Journal of the American Statistical Association (2012).
[11] I. Sgouralis, M. Whitmore, L. Lapidus, M. J. Comstock, and S. Pressé, "Single molecule force spectroscopy at high data acquisition: A bayesian nonparametric analysis," The Journal of chemical physics **148**, 123320 (2018).
[12] C. Bishop, "Pattern recognition and machine learning (information science and statistics), 1st edn. 2006. corr. 2nd printing edn," (2007).
[13] R. M. Neal *et al.*, "Mcmc using hamiltonian dynamics," Handbook of Markov Chain Monte Carlo **2**, 2 (2011).
[14] M. Betancourt, "A conceptual introduction to hamiltonian monte carlo," arXiv preprint arXiv:1701.02434 (2017).