Article

Fast and robust two-dimensional inverse Laplace transformation of single-molecule fluorescence lifetime data

Saurabh Talele[1,2] and John T. King[1,*]

[1]Center for Soft and Living Matter, Institute for Basic Science, Ulsan, Republic of Korea and [2]Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea

ABSTRACT   Fluorescence spectroscopy at the single-molecule scale has been indispensable for studying conformational dynamics and rare states of biological macromolecules. Single-molecule two-dimensional (2D) fluorescence lifetime correlation spectroscopy is an emerging technique that holds promise for the study of protein and nucleic acid dynamics, as the technique is 1) capable of resolving conformational dynamics using a single chromophore, 2) resolves forward and reverse transitions independently, and 3) has a dynamic window ranging from microseconds to seconds. However, the calculation of a 2D fluorescence relaxation spectrum requires an inverse Laplace transform (ILT), which is an ill-conditioned inversion that must be estimated numerically through a regularized minimization. Current methods for performing ILTs of fluorescence relaxation can be computationally inefficient, sensitive to noise corruption, and difficult to implement. Here, we adopt an approach developed for NMR spectroscopy (T1-T2 relaxometry) to perform one-dimensional (1D) and 2D-ILTs on single-molecule fluorescence spectroscopy data using singular-valued decomposition and Tikhonov regularization. This approach provides fast, robust, and easy to implement Laplace inversions of single-molecule fluorescence data. We compare this approach to the widely used maximal entropy method.

SIGNIFICANCE   Inverse Laplace transformations are a powerful approach for analyzing relaxation data. The inversion computes a relaxation rate spectrum from experimentally measured temporal relaxation, circumventing the need to choose appropriate fitting functions. They are routinely performed in NMR spectroscopy and are becoming increasingly used in single-molecule fluorescence experiments. However, as Laplace inversions are ill-conditioned transformations, they must be estimated from regularization algorithms that are often computationally costly and difficult to implement. In this work, we adopt an algorithm first developed for NMR relaxometry to provide fast, robust, and easy to implement one-dimensional and two-dimensional inverse Laplace transformations on single-molecule fluorescence data.

INTRODUCTION

Single-molecule fluorescence spectroscopy has provided unparalleled access into dynamics of biological macromolecules and the mechanism of biochemical processes (1). Typical experimental techniques rely on measuring time-dependent fluctuations in the fluorescence emission wavelength (2–8), intensity (9–11), or lifetime (12–14) from a single chromophore or chromophore pair. The dynamics of biochemical processes of interest are then inferred from the analysis of the photon stream. The most widespread

technique is Single Molecule Förster Resonance Energy Transfer (sm-FRET) (1,15–18), which leverages the highly sensitive distance dependence for dipole-dipole coupling between two chromophores to monitor nanometer scale motions of a biomolecule. Conformations are distinguished by the emission wavelength, which is either from the acceptor or donor chromophore depending on the separation distance. Transitions between two conformations are observed as abrupt changes in emission intensity at both the acceptor and donor emission wavelength. This approach has several advantages, including a clear interpretation of emission wavelength transitions observed in the photon stream and the ability to study highlighted conformational motions through the selective placement of the chromophores (17). However, the need for site-specific labeling of two

chromophores on a single biomolecule makes its application to proteins challenging, in particular when compared with nucleic acids (18), which are relatively easy to modify.

Single-molecule two-dimensional (2D) fluorescence lifetime correlation spectroscopy (sm-2D-FLCS), first developed by Tahara and co-workers (19–21) and later extended into the single-molecule regime by Kondo and workers (22), provides an alternative analysis of single-molecule data that does not sacrifice chemical selectivity or temporal resolution. In this approach, time-correlated single-photon counting (TCSPC) is used to detect fluorescence intensity and emission delay time from a single chromophore, either freely diffusing in dilute solution (19–21) or surface immobilized (22). The data are recorded in the Time Tagged Time-Resolved (TTTR) mode which generates a real-time photon stream (characterized by a global macrotime) with a recorded emission delay time (characterized by a microtime) for each registered photon (23). Distinct chemical species in the system are distinguished by their fluorescence lifetime. Thus, chemical exchange between two states can be kinetically resolved on the condition that they have different fluorescence lifetimes.

A series of 2D photon histograms is generated by cataloging all photon pairs separated by a systematically varying waiting time ΔT from an experimentally recorded photon stream (19,20). A 2D inverse Laplace transform (2D-ILT) of the lifetime histogram generates a 2D-FCLS spectrum at the given ΔT. Analogous to 2D-NMR (24), species that do not undergo any form of exchange during ΔT appear as diagonal peaks in the 2D spectrum, whereas species that exchange during ΔT appear as cross-peaks. Measuring 2D spectra for a series of ΔT allows the overall correlation function to be effectively split into its components comprising autocorrelations for the diagonal components and cross correlations for the off-diagonal components. Thus, the chemical exchange kinetics among the components can be measured directly through time correlation functions reflecting chemical exchange between two states (19–22). As this technique relies on a single chromophore to resolve conformational dynamics, it can be applied to systems in which fluorescence labeling is challenging or to native proteins with an endogenous chromophore. Additionally, for a conformational transition of interest, the forward and reverse kinetics are resolved as they are recorded in different regions of the 2D spectrum. The ability to measure asymmetric forward and reverse kinetics is relevant for studying systems operating out of equilibrium. Lastly, as the technique requires no additional optical setup or heavy computing hardware, it can be implemented on any setup already used for Fluorescence Correlation Spectroscopy (FCS) or FRET.

The challenge of this experimental approach is the need to perform a 2D-ILT, which is an ill-conditioned problem and is, therefore, numerically unstable. This results in multiple solutions satisfying the same problem when solved with traditional least-square analysis (25). Instead, this class of problems must be solved via regularized least-square analysis, which imposes a penalty on solutions with undesired features and hence promotes smoother solutions (26). To date, the calculation of the 2D spectra is traditionally performed using the maximum entropy method (MEM) (19–22), which produces a solution using Bayesian inference. The application of MEM has two severe limitations. First, it involves a constrained optimization because entropy cannot be defined for negative values, so one must check for the positivity constraint during each iteration and modify the fit accordingly. This makes the approach computationally inefficient. Second, the 2D spectra is fit as a lexicographically ordered one-dimensional (1D) vector, resulting in a fitting kernel matrix whose size is of the order of the fourth power of the number of points in the lifetime spectra, leading to a tremendous computational cost. To decrease the computational cost, the data are often binned nonuniformly before fitting, which results in a loss of temporal information in the input data (20).

The challenges of performing 2D-ILTs (i.e., converting the time-dependent relaxation data to a relaxation rate spectrum) are not unique to fluorescence spectroscopy. Indeed, multidimensional Laplace inversions are commonly employed in NMR spectroscopy, for instance, in the computation of T1-T2 correlation spectrum from relaxometry data. Work by Venkataramanan and co-workers (27) outlined an efficient algorithm for Laplace inversion in NMR relaxometry data that leveraged singular-valued decomposition (SVD) and Tikhonov regularization (TK). Subsequent improvements of this approach have been reported (28,29). In this study, we adopt this general approach for application to single-molecule fluorescence spectroscopy analysis. This algorithm reduces the computation time to merely a few seconds per 2D spectrum and, hence, enables the analysis of large data sets with high resolution.

MATERIALS AND METHODS

Generation of 2D photon histograms

A 2D-FLCS spectrum is generated from a 2D-ILT of a lifetime correlation histogram, easily measured using standard TCSPC techniques (19,20). We start with a familiar 1D picture to introduce the concept of inverse Laplace transform (ILT). Given a time series of recorded photons, the emission delays are distributed exponentially and can be represented as the following equation:

$$I(t) = \sum_{j=1}^{n} a_j exp^{\left(-t/\tau_j\right)}, \tag{1}$$

for n independent components with amplitudes a_j and respective fluorescence lifetimes τ_j. This can be represented in matrix form and generalized for the sake of broader application as the following equation:

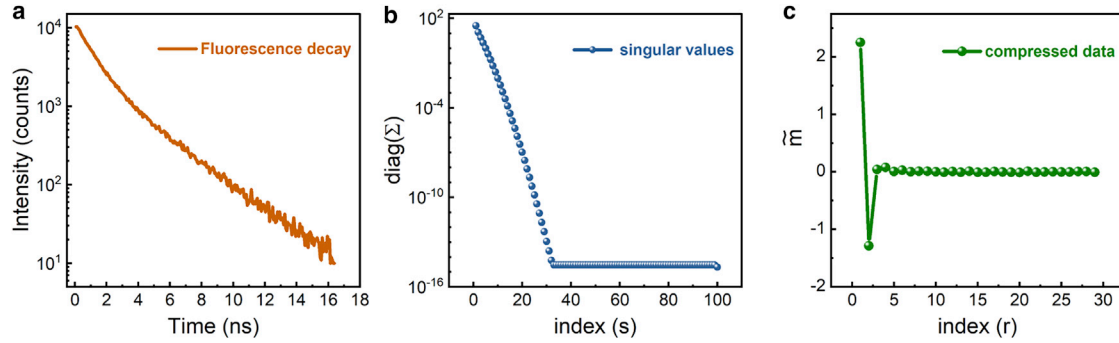$$I_i = \sum_{j=1}^{n} K_{ij} \cdot A_j, \tag{2}$$

FIGURE 1 (*a*) An example fluorescence decay data exhibiting a biexponential decay. (*b*) Singular values of the kernel matrix K in Eq. 10 comprising of exponential decays with 100 distinct lifetimes in range (0.1 ns–10 ns). The singular values decrease rapidly, and K can be represented accurately by ∼30 highest singular values. (*c*) Compressed data $\bar{m}$ obtained by Eq. 13 represented using 30 points.

where K is a kernel with predefined basis of fluorescence lifetime decays $K_{ij} = (1/\tau_j)exp^{(-t_i/\tau_j)}$, and A_j is a column vector representing the amplitudes corresponding the species τ_j. Kernel K can be defined such that the experimental lifetimes are included in $[t_{min}, t_{max}]$ with sufficient resolution such that any measured decay can be represented by a precalculated kernel and an amplitude vector. Under this construction, $A(\tau_j)$ represents a 1D-ILT of $I(t)$, or the lifetime spectrum corresponding to $I(t)$, as we convert the time-dependent decay into its lifetime components and their probability distribution amplitudes.

This can be easily extended to 2D. Cataloging all the photon pairs separated by a global lag time ΔT in the experimentally recorded time series and having emission delays t_1 and t_2, we can generate a 2D histogram with axes as t_1 and t_2, where each point $M(t_1, t_2, \Delta T)$ represents the number of coincidences when a photon with delay t_2 is detected ΔT after having detected a photon with delay t_1:

$$M(t_{1,i}, t_{2,j}, \Delta T) = \sum_{k,l} F(\tau_{1,k}, \tau_{2,l}, \Delta T) \cdot exp^{(-t_{1,i}/\tau_{1,k})} \cdot exp^{(-t_{2,j}/\tau_{2,l})}. \quad (3)$$

Using the kernel structure mentioned above, we can write in matrix notation,

$$M(t_1, t_2, \Delta T) = K_1(t_1, \tau_1) \, F(\tau_1, \tau_2, \Delta T) \, K_2(t_2, \tau_2)^T, \quad (4)$$

where $F(\tau_1, \tau_2, \Delta T)$ is the joint probability distribution of occurrence of a photon with lifetime τ_2 occurring ΔT after registering a photon with lifetime τ_1. Each point on $F(\tau_1, \tau_2, \Delta T)$ represents the amplitude of correlation between the components at (τ_1, τ_2) separated by lag time ΔT, the autocorrelations appear along the diagonal of $F(\tau_1 = \tau_2)$ and the cross correlations appear as off-diagonal peaks of $F(\tau_1 \neq \tau_2)$. By varying ΔT, one can determine the separated time correlation functions of the constituent components in any fluorescence emission time series.

Determining $F(\tau_1, \tau_2, \Delta T)$ is equivalent to performing a 2D-ILT on M. For efficient handling, we can convert the 2D form of Eq. 4 to 1D by lexicographically ordering the matrices $m = vec[M], f = vec[F]$ and can write the kernel operations K_1 and K_2 by a single operator K_0 given by the Kronecker product:

$$K_0 = K_1 \times K_2. \quad (5)$$

And we represent the equivalent 1D problem as the following equation:

$$m = K_0 f. \quad (6)$$

Now, the only task at hand is to solve for f given m and K_0. As previously mentioned, this equation cannot be solved analytically because it is an ill-conditioned inversion resulting in a solution that is not stable or unique. Instead, regularized least-squared technique must be used to approximate the inversion. In regularized least-square minimization, the least square minimizes the difference between the input data and the fit, whereas the regularization imposes a penalty on undesired features of the fitted solution. The regularization prevents overfitting or numerical instability. To avoid any bias, we start with a uniform before the guess of f and iteratively find the solution by minimizing the objective function given by the following equation:

$$Q(f) = \|K_0 f - m\|^2 + \alpha R(f), \quad (7)$$

where the first term on RHS represents the least-square term, $R(f)$ is the regularization function of choice, and α is a regularization constant that weighs the importance of the least-square fitting versus regularization. The choice of α is critical to appropriate fitting: too small, and the minimization remains unstable (overfitting); too large, and fit may not reflect the underlying experimental data (underfitting). Two regularization methods discussed here include the commonly employed MEM (30–32) and TK (33,34). Ultimately, we highlight the strength of TK at providing a fast and robust method for performing 2D-ILT.

2D-ILT by Maximum Entropy Method

The MEM is a commonly employed fitting technique used for a number of applications, including image reconstruction (30,31) and spectroscopy (32). It was also the approach applied by Tahara and co-workers in their original development of 2D-FLCS spectroscopy (19,20). MEM uses a regularization function based on Shannon entropy penalization $\phi(s) = -\sum s \log s$. The regularization function can be written as follows,

$$R(f) = \sum f - x - f \, ln\left(\frac{f}{x}\right), \quad (8)$$

where x represents the estimated prior fit of the experimental data. The algorithm is initiated by taking f to be a flat distribution and then optimized according to Eq. 7. Fitting is typically initiated with a large α-parameter, which is iteratively decreased until the classical MEM condition is satisfied (35). This method ensures that of all the possible solutions, the solution with the maximal information entropy is chosen. In our results, we follow the MEM approach described by Tahara and co-workers (20).

For typical 2D-FLCS spectra, the data sets are too large to compute ILT using MEM efficiently. Instead, the data are binned logarithmically, which decreases the data size at the cost of resolution. Furthermore, to obtain a

TABLE 1 Computation Times in Seconds for Methods or 1D ILT of Same Set of Data Comprising 2×10^6 Total Photons

Regularizer	No Compression	Compression Only	BRD	BRD Only
MEM	32 ± 3	9.1 ± 0.3	–	–
Tikhonov	29 ± 3	8.90 ± 0.25	0.17 ± 0.02	0.22 ± 0.02

The corresponding fits are shown in Figure 7, a–f.

meaningful ILT using MEM, one must impose a nonnegativity constraint on f after every iteration because the entropy cannot be defined for negative values. This significantly decreases the efficiency of the calculation.

2D-ILT by Tikhonov Regularization

Regularization function

An alternative approach has been developed in NMR spectroscopy, which uses SVD-based data compression and TK (27–29). This approach differs from MEM in several ways. First, the size of the numerical calculation is greatly reduced by compressing the kernels using SVD instead of nonuniform binning of the data. Because the kernels are smooth functions, the elementwise data are vastly redundant and SVD can reduce the data size by roughly 100-fold without compromising the quality of the fit or the spectral resolution. Second, it uses TK on the compressed data, which ensures a unique solution to the optimization due to the quadratic nature of the terms. Third, it employs Butler-Reeds-Dawson (BRD) method to transform the constrained optimization to an unconstrained optimization which is computationally efficient to implement. Lastly, it provides a method for choosing an appropriate regularization constant proportional to variance in the data. The outline of the method follows what is presented in (27) and (29). Using TK, the objective function to be minimized from Eq. 7 becomes the following:

$$\widehat{f} = \arg \min_{f > 0} Q(f), \tag{9}$$

where,

$$Q(f) = \|\tilde{m} - \tilde{K}_0 f\|^2 + \alpha \|f\|^2. \tag{10}$$

The quadratic nature of the terms ensures a unique solution. The second term penalizes solutions that have large norms, which is characteristic of functions with sharp features. Solutions with smoothly varying features are therefore promoted, and α scales the desired smoothness with respect to the least-squares difference. However, because f is the probability distribution, it cannot take negative values, which makes this a constrained optimization problem. To convert the inversion to an unconstrained problem, the BRD algorithm is employed (27). We do not anticipate significant speed improvement by implementing Tikhonov regularizer alone. Instead, its purpose is to enable the use of the BRD algorithm, which is incompatible with maximal entropy regularizer.

TABLE 2 Computation Times in Seconds for Methods or 2D ILT of Same Set of Data with the 2D Emission Delay Histograms Evaluated at $\Delta T = 2$ ms

Regularizer	No Compression	Compression Only	BRD	BRD Only
MEM	9×10^3	4×10^3	–	–
Tikhonov	8×10^3	1.4×10^3	20	500

The corresponding fits are shown in Figure 7, g–l.

Kernel compression

Here, we demonstrate efficient reduction of computational size of the problem in Eq. 7 without significant loss of temporal information of the data. To get a sense of magnitude of the computational size, TCSPC experiment performed at the resolution of 64 ps has 256 data points Fig. 1 a. The 2D emission delay histogram generated from the time trace are thus of size 256×256, if we desire to perform 2D ILT with a basis size of 100 lifetime constants, we have the estimated spectrum of the size 100×100 and hence the kernels of size 256×100. Converting the problem to 1D by lexicographic ordering (Eq. 6) gives us $m_{65,536 \times 1} = K_{65,536 \times 10,000} f_{10,000 \times 1}$, which is a tremendously large problem. Traditional solutions include nonuniform binning of the data and using smaller basis of lifetimes. However, maximal storage space is occupied by the kernel, which is precalculated and free of noise.

Here, we show how the size of the problem can be reduced by compressing the kernel using SVD as demonstrated in (27). For simplicity, we illustrate the implementation for 1D ILT (Eq. 2). For the implementation in the case of 2D ILT, the compression is performed before converting the problem to 1D as shown in the Supporting materials and methods, Section 4. The 1D problem at hand is as follows,

$$m_{d \times 1} = K_{d \times s} f_{s \times 1}, \tag{11}$$

where the subscripts denote the dimensions ($d \approx 256$, $s = 100$). We represent the kernel K by its SVD, $K = U\Sigma V^T$, where U and V are unitary matrices ($U^T U = V^T V = I$) and Σ is a diagonal matrix with the singular values of K arranged in descending order:

$$m_{d \times 1} = U_{d \times s} \Sigma_{s \times s} V_{s \times s}^T f_{s \times 1}. \tag{12}$$

Fig. 1 b shows the singular values of a kernel formed using basis of 100 decays within range of lifetimes [0.1–10 ns] and for time [0–16 ns]. We see that the singular values become negligible after ~30 values ($\sigma_i < 10^{-15}$; i > 30). This is attributed to the smooth nature of the kernels, which implies that we can represent the kernel matrix by a compact SVD where we only keep the r highest singular values greater than 10^{-15} (first 30 values in this case). Thus, the matrix $\Sigma_{s \times s}$ can be compressed to $\tilde{\Sigma}_{r \times r}$ without significant loss of information. $\tilde{U}_{d \times r}$ and $\tilde{V}_{s \times r}$ are formed by selecting the first r columns of U and first r rows of V. The resultant compact kernel $\tilde{U}_{d \times r}\tilde{\Sigma}_{r \times r}\tilde{V}_{r \times s}^T$ is an approximation of the original kernel and has the same dimensions as $K_{d \times s}$. The difference between the original and the compact kernel is shown in Fig. S2. We can observe that there is negligible loss of information and that $\tilde{U}_{d \times r}\tilde{\Sigma}_{r \times r}\tilde{V}_{r \times s}^T$ approximates K well.

Using the compact SVD, Eq. 12 now becomes the following,

$$m_{d \times 1} = \tilde{U}_{d \times r} \tilde{\Sigma}_{r \times r} \tilde{V}_{r \times s}^T f_{s \times 1}, \tag{13}$$

$$\tilde{U}_{r \times d}^T m_{d \times 1} = \tilde{\Sigma}_{r \times r} V_{r \times s}^T f_{s \times 1}, \tag{14}$$

$$\tilde{m}_{r \times 1} = \tilde{K}_{r \times s} f_{s \times 1}, \text{and} \tag{15}$$

where $\tilde{m} = \tilde{U}^T m$ is referred to as compressed data and $\tilde{K} = \tilde{\Sigma}\tilde{V}^T$. Note that the size of the problem is reduced using the SVD of kernels and not the experimental data m. The fitting is now performed on $\tilde{m}$ that has dimensions of $r \times 1$ Fig. 1 c. For the typical case of $r = 30$,

$$\tilde{m}_{30 \times 1} = \tilde{K}_{30 \times 100} f_{100 \times 1}. \tag{16}$$

This makes it different than the traditional nonuniform binning approach. Venkataramanan et al. have shown that the structure of the problem Eq. 7 remains unchanged after this transformation (27). That is, the solution f

Basis Size	1D ILT Time (s)	2D ILT Time (s)
1000	0.40	1000
100	0.17	20
50	0.05	1

that satisfies Eq. 11 is the same one that satisfies Eq. 16 and again, the solution is unique because of the quadratic nature of the problem.

Comparing the size of the problem in Eq. 11 to that of Eq. 16, we observe that the problem size is reduced by a factor of 10. The reduction factor is even greater in the case of 2D problem (250 times, see Supporting materials and methods, Section 4). This compression is a general technique and can be used alongside the existing MEM approach (Figure 7, b and h; Tables 1 and 2) or broadly, with any other problem involving smooth large kernel functions. The projection of the original data on the range space is given by the following,

$$\tilde{\tilde{m}} = \tilde{U}\tilde{U}^T m = \tilde{U}\tilde{m}, \tag{17}$$

where $\tilde{\tilde{m}}$ resembles a smoother version or a low pass filtered version of original data m (Fig. 4 a).

Tikhonov Regularization with BRD

A brief overview of the BRD approach (27,36) summarizing is given below. In Eq. 10, the gradient of Q with respect to f must be zero at its minimum,

$$\frac{\partial Q}{\partial f_i} = -\tilde{K}_{0,i}^T(\tilde{m} - \tilde{K}_0 f) + \alpha f_i = 0, \tag{18}$$

where,

$$\alpha f_i = \tilde{K}_{0,i}^T(\tilde{m} - \tilde{K}_0 f), \tag{19}$$

and,

$$f_i = \tilde{K}_{0,i}^T c. \tag{20}$$

We invoke a vector c which maps f as,

$$c = \frac{\tilde{m} - \tilde{K}_0 f}{\alpha}, \tag{21}$$

where c has the same dimensions as the compressed data $\tilde{m}$. The unconstrained problem then becomes (29),

$$\hat{c} = arg \min \chi(c), \tag{22}$$

where,

$$\chi(c) = \frac{1}{2}c^T[G(c) + \alpha I]c - c^T \tilde{m}, \tag{23}$$

and,

$$G(c) = \tilde{K}_0 \, diag\big(H(\tilde{K}_0^T c)\big)\tilde{K}_0^T. \tag{24}$$

Here, $H(\cdot)$ denotes the Heaviside function, which ensures positive semidefiniteness. The minimization of $\chi(c)$ can be carried out via standard unconstrained inverse Newton minimization routines such as *fminunc* in MATLAB (The MathWorks, Natick, MA). The required gradients and Hessian of Eq. 22 are easily computed as follows,

$$\nabla\chi(c) = (G(c) + \alpha I)c - \tilde{m}, \tag{25}$$

and,

$$\nabla\nabla\chi(c) = G(c) + \alpha I. \tag{26}$$

Using the optimized vector c, the ordered vector f is calculated as follows

$$f = \max\big(0, \tilde{K}_0^T c\big). \tag{27}$$

The 2D-ILT spectrum F is then given by reshaping the $f_{1\times kl}$ vector to a matrix $F_{k\times l}$. We start with a large value of α to estimate c. The recommended optimum α for the following iterations is given by the following,

$$\alpha_{opt} = \frac{\sqrt{k \times l} \cdot \hat{\sigma}}{\|c\|_F}, \tag{28}$$

with,

$$\hat{\sigma} = std\big(\sqrt{\tilde{\tilde{m}}} - \sqrt{m}\big). \tag{29}$$

Equation 29 represents the approximation of noise level in the data obtained by first normalizing the data by square root of the photon counts and then comparing it with a smoothened version of itself (Fig. 4 a). The optimized vector c is iteratively evaluated using this α_{opt}. A suitable stopping criterion is achieved when

$$\|c\|_F > 1/\hat{\sigma}, \tag{30}$$

or when the relative difference between consecutive α to be less than 0.1%.

$$\frac{|\alpha_i - \alpha_{i+1}|}{\alpha_i} < 10^{-3}. \tag{31}$$

Typically, convergence is achieved within 10 iterations.

RESULTS AND DISCUSSION

We demonstrate the performance of both MEM and TK using Markovian Monte Carlo simulations to generate artificial photon time series for a two-state system with user defined inputs for fluorescence lifetime, emission intensity and transition rate matrix (see Supporting materials and methods, Section 2 for details) We compare the compression efficiency, tolerance to noise, and the timing of the algorithm for various relevant parameters.

Implementation

We simulate a two-state system with fluorescence lifetimes of $\tau_1 = 1.0$ ns and $\tau_2 = 3.0$ ns, and exchange rates of $k_f = k_r = 1 \times 10^3 s^{-1}$ corresponding to an exchange time of 1 ms. The number of photons analyzed is roughly 2.5×10^5. The corresponding 1D and 2D lifetime spectra
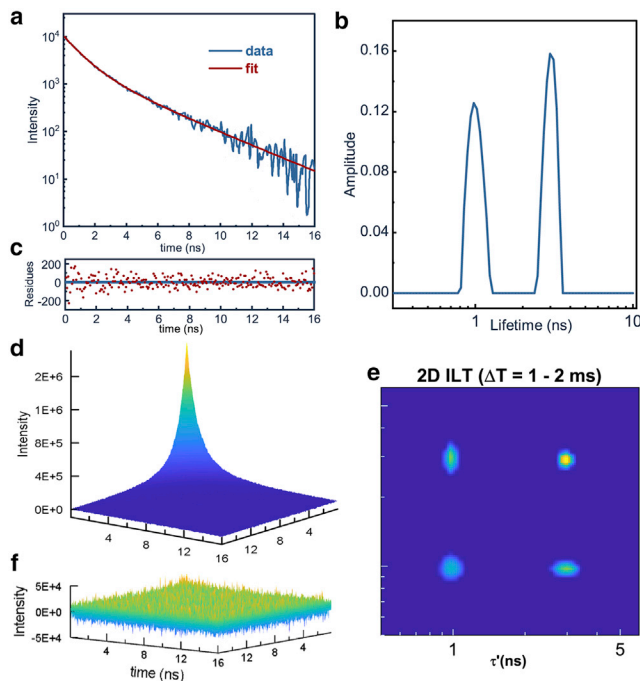
FIGURE 2 Laplace inversions by Tikhonov Regularization . Monte Carlo simulations generate an artificial photon stream of a two-component system with fluorescence lifetimes of $\tau_1 = 1$ ns and $\tau_2 = 3$ ns undergoing equilibrium chemical exchange at a rate of 1×10^3 s^{-1}. (*a*) Raw 1D-photon histogram (*blue*) and the obtained fit (*red*) from 1D-ILT. (*b*) The 1D lifetime spectrum shows two peaks with relaxation rates of 1 and 3 ns, consistent with fluorescence lifetime of the two components in the system. (*c*) Residual between the 1D photon histogram and the fit obtained by 1D-ILT. (*d*) 2D photon correlation histogram and (*e*) 2D relaxation rate spectrum computed at a waiting time of $\Delta T = 2$ ms obtained from a 2D-ILT. The spectrum consists of diagonal peaks at 1 and 3 ns, as well as cross-peaks indicating chemical exchange between the two species. (*f*) Residual between the 2D photon correlation histogram and 2D-ILT fit.

were obtained by ILT by TK as described above, using a basis set of size $L = 100$ and with an SVD using 30 greatest values. The 1D photon histogram, resulting fit computed by a 1D-ILT, and residual are shown in Fig. 2, *a* and *c*. The 1D relaxation rate spectrum shows two peaks located at $\tau_1 = 1.0$ ns and $\tau_2 = 3.0$ ns, consistent with the input parameters of the simulation Fig. 2 *b*.

The 2D photon correlation histogram evaluated at $\Delta T = 2$ ms, resulting fit computed by a 2D-ILT, and residual are shown in Fig. 2, *d* and *f*. A 2D-FLCS spectrum shows two diagonal peaks at $\tau_1 = 1.0$ ns and $\tau_2 = 3.0$ ns, and two off-diagonal cross-peaks between these transitions Fig. 2 *e*. The cross-peaks indicate that chemical exchange has occurred within the timescale $\Delta T = 2$ ms, which is consistent with the input transition rate matrix.

Computing the 2D-FLCS spectrum for a series of ΔT delay times allows the chemical exchange kinetics to be directly measured. Fig. 3, *a–c* shows the 2D-FLCS spectrum measured for delay times ranging from 10 μs to 2 s. At early times $\Delta T < \tau_{exchange}$, only diagonal peaks are observed as no chemical exchange had occurred Fig. 3 *a*. As ΔT is increased, cross-peaks

emerge as chemical exchange occurs Fig. 3 *b*. The exchange cross-peaks reach a steady state as $\Delta T > \tau_{exchange}$ Fig. 3 *c*. Monitoring the amplitude of these diagonal and cross-peaks provides direct information regarding the chemical exchange process. The amplitude of the peaks can be expressed as,

$$C_{ij}(\Delta T) = \langle S_i(t)S_j(t+\Delta T)\rangle, \qquad (32)$$

where i and j represent the state of the system with lifetimes τ_i and τ_j. $C_{ij}(\Delta T)$ thus represents the probability of the system in state i at time t to be observed in state j at later time $t + \Delta T$. The diagonal is given by the condition $i = j$, whereas cross-peaks are given by the condition $i \neq j$. The two autocorrelation and cross-correlation functions obtained from the 2D-FLCS spectra are shown in Fig. 3 *d*. As expected, the autocorrelation functions decay on the timescale of $\tau_{exchange}$, whereas the cross-correlation functions grow on the timescale $\tau_{exchange}$. Fitting the data provides a direct measure of $\tau_{exchange}$ for the system.

In practice, lifetime histograms are convoluted by a systematic instrument response function (IRF) and do not manifest as pure exponential decays. We can account for this by fitting the data with kernels convoluted by a known IRF that can be measured experimentally from scattering of the excitation pulse from a colloidal medium. The experimentally observed TCSPC histogram is modeled from Eq. 1 as

$$I^{obs}(t,\tau) = \sum_i \text{IRF}(t_i - t_0)\exp\left(-\frac{t - t_i}{\tau}\right), \qquad (33)$$

where, t_0 is the unavoidable zero-time shift for the IRF. It is important so estimate the correct value of t_0 as it can lead to undesirable errors in the ILT spectrum. To circumvent this, we calculate the 1D-ILT spectrum at various values of zero-time shifts and use the range over which the fitting least-square error χ^2 is minimal. The ILT obtained from this range is then averaged to obtain the final 1D-ILT spectrum to be used for further analysis. 2D spectrum is calculated using the same range of zero-time-shifts. The kernels to be fitted are also convoluted with the IRF similarly (see Supporting materials and methods, Section 3) and Fig. S2.

Effect of compression

Kernel compression (KC) increases the computational speed by excluding redundant or insignificant information and thus reducing the size of the problem. However, there is a tradeoff between the magnitude of compression and loss of information. The measure of compression in our case is given by the number of singular values chosen to represent the kernel of exponential. Fig. 1 *b* shows the singular values for a kernel comprising of 100 basis exponentials ranging from lifetimes [0.1–10 ns]. It is observed that only the greatest 30 singular
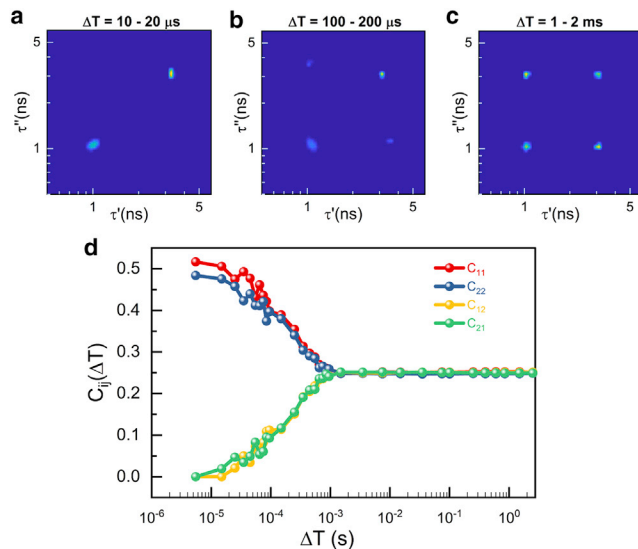
FIGURE 3 Chemical Exchange kinetics of a two-state system. 2D-FLCS spectra computed at (*a*) $\Delta T = 10$–20 μs, (*b*) $\Delta T = 100$–200 μs, and (*c*) $\Delta T = 1$–2 ms. At $\Delta T < \tau_{exchange}$, no cross-peaks are observed between the two states, indicating no exchange has occurred. As ΔT approaches $\tau_{exchange}$ cross-peaks emerge indicating both forward and reverse transitions between the two states. (*d*) The kinetics of the exchange process are reflected in the autocorrelation and cross-correlation functions of the diagonal and off-diagonal peaks, respectively. Fitting either the autocorrelation or cross-correlation functions provide a direct measure of the exchange kinetics in the system.

values are enough to sufficiently describe the kernel. This is demonstrated in Fig. S1 *d* showing the difference between the true kernel and the approximated kernel constructed by choosing only the 30 greatest singular values and Fig. S1 *c* showing the least-square error in the kernel approximation $lsq = \Sigma(K - K_{compact})^2$.

Finally, transforming the problem by operation of U^T, Eq. 14 reduces the computational size by 10 times for 1D-ILT and 250 times for 2D-ILT. The number of significant singular values depends largely on the range of lifetimes used to generate the basis and not the number of basis vectors chosen for fitting (Fig. S1, *a* and *b*).

Dependence on total photon counts

A robust algorithm should produce consistent results for various realizations of data obtained from same source. In case of regularized optimization, this is reflected in the choice of the regularization parameter α (Eq. 7). α represents the bias toward the smoothness with respect to the least-square errors, which manifests as the width of the peaks in case of the ILT. We start with a sufficiently high value of α and gradually decrease it following Eq. 28 as suggested in (36), where $\widehat{\sigma}^2$ is a measure of the variance in data.

Determining $\widehat{\sigma}$ is straightforward in experiments influenced by uniform gaussian noise however, for the fluorescence lifetime histogram, the photon counts follow Poisson
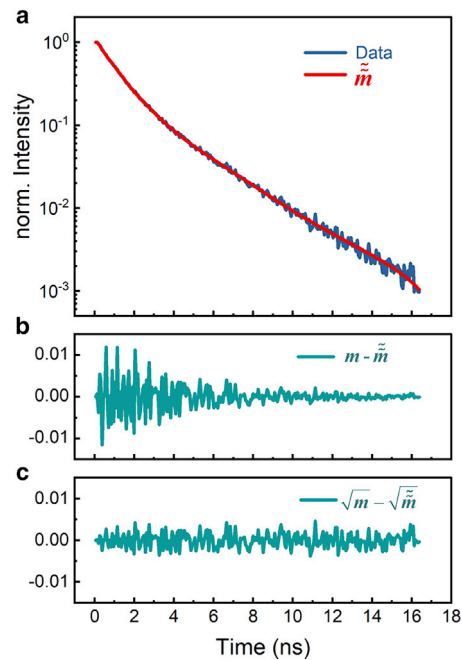

FIGURE 4 Noise estimation for stopping criterion. (*a*) Normalized fluorescence data superimposed with $\tilde{m}$, projection of compressed data onto the range space of kernel given by Eq. 17. It is observed that $\tilde{m}$ serves as a smooth approximation to input data referring to which we infer the corresponding noise level given by Eq. 29. (*b*) Shows the difference $\tilde{m} - m$, which is dependent on the photon counts. We normalize both traces using square root of photon counts which eliminates the intensity dependence. (*c*) Shows the difference after normalization given by $\sqrt{\tilde{m}} - \sqrt{m}$.

distribution at each point individually. As a result, the variance at each point is different and is proportional to the total accumulated photons. The best case is to determine the point-wise variance experimentally however, the data used for determination can be simply added together into a fluorescence decay with more counts and hence less relative noise. Thus, we chose to approximate the point-wise variance by comparing the data with a smoothened version of itself. Luckily, $\tilde{m} = \tilde{U}^T \tilde{U} m$ which is the projection of compressed data to the range space of the kernels serves as an approximate smoothened curve corresponding to the data given that the total number of photons is sufficiently high (above $\sim 10^4$ total photons) Fig. 4 *a*. Then we estimate $\widehat{\sigma}$ using Eq. 29, where the traces m and $\tilde{m}$ are normalized by their square root which effectively removes the intensity dependence of each data point and produces uniformly distributed residues Fig. 4, *b* and *c*. Decrease in photon counts leads to higher $\widehat{\sigma}$, which in turn leads to a higher α_{opt} and hence smoother fits. Fig. 5, *a* and *b* shows dependence of $\widehat{\sigma}$ and α_{opt} on the total photon counts. Fig. 5, *c*–*f* also shows the normalized 1D and 2D-ILT fits for trajectories with various total photon counts. A similar approach of using preliminary fits has been shown to improve the accuracy and precision of estimated parameters in fitting of single-molecule fluorescence decays (37). A qualitative analysis of effect of noise on the resulting ILT
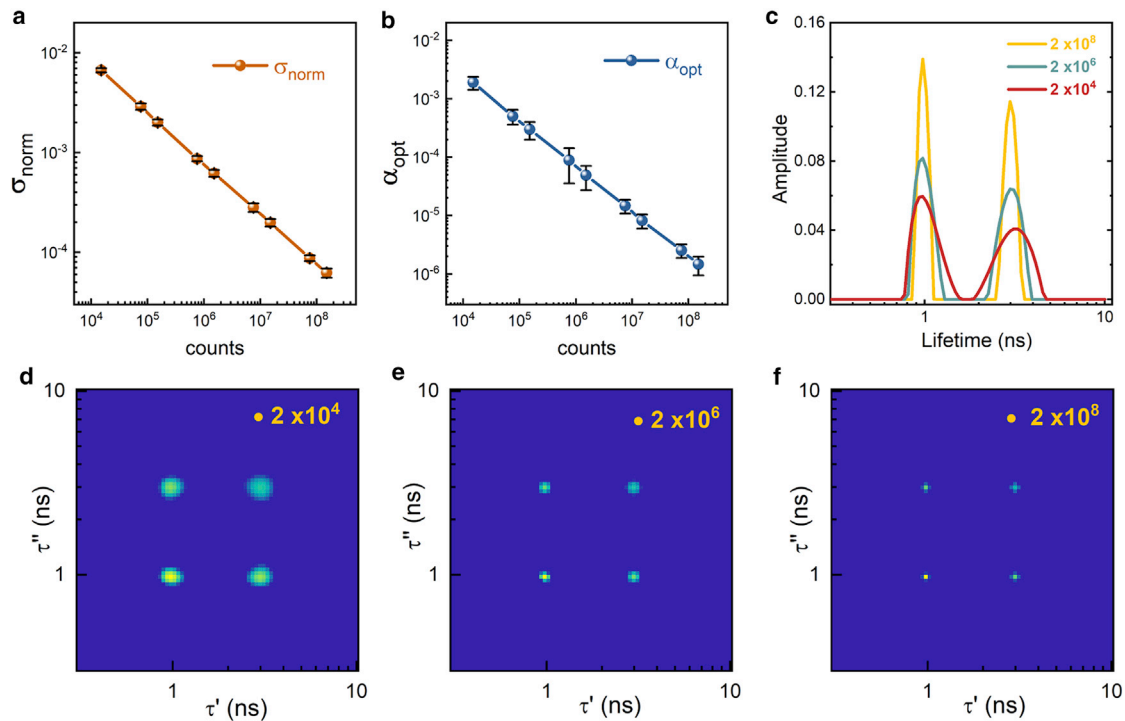
FIGURE 5 Dependence of 2D spectrum on total photon counts. (*a*) The estimated $\hat{\sigma}$ with respect to total photon counts in the fluorescence data. (*b*) Shows the optimum regularization parameter obtained with respect to the total photon counts in the fluorescence data. Error bars denote the standard deviation using 30 instances of simulated data. Higher α_{opt} due to lower photon counts results in smoother fits and vice versa as demonstrated in (*c*), which showing 1D-ILT of three trajectories with varying total photons. (*d–f*) show the normalized 2D-ILT obtained at $\Delta T = 2\ ms$ using the simulated trajectories from (*c*).

fits is detailed in Supporting materials and methods, Section 5 and Figs. S3 and S4.

Application to experimental data

In this section, we demonstrate the utility of the outlined 2D-ILT algorithm on experimental data. We measure 2D-FLCS for bacteriorhodopsin (bR), a transmembrane proton pump protein with an endogenous retinal chromophore. Photon absorption by the retinal chromophore initiates a complex reaction cycle that involves a series of conforma-tional transitions that occur on timescales ranging from microseconds to milliseconds (38). The dynamics of confor-mational transitions involved in the biochemical cycle can be mapped through changes in the retinal chromophore fluo-rescence lifetime (39). Note, here, we are interested only in demonstrating the use of the outlined approach to obtaining reliable 2D-FLCS spectra for an experimental system. An analysis of the dynamics of bR reaction cycle are reported elsewhere (39).

The crystal structure and absorption spectra of bR is shown in Fig. 6, *a* and *b*. The experiments were carried out
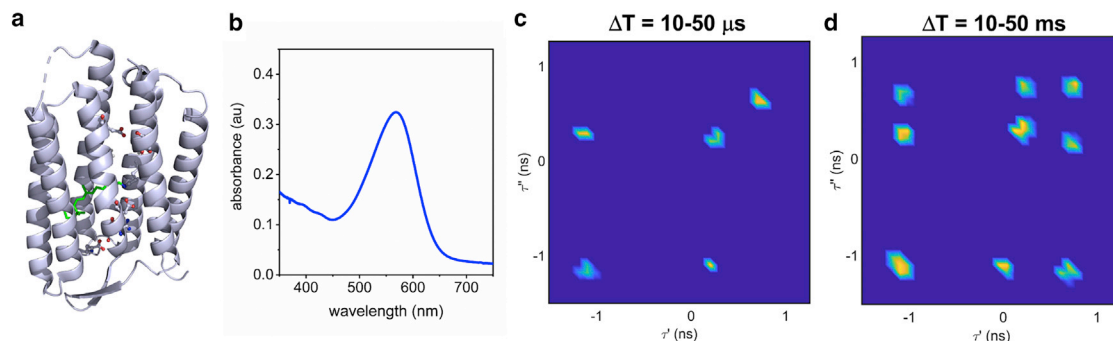

FIGURE 6 2D-ILT on bR single-molecule fluorescence data. (*a*) Crystal structure of bR, with the retinal chromophore highlighted in green. (*b*) Absorption spectrum of the bR in the ground state, with $\lambda_{\max} = 560$ nm. (*c* and *d*) 2D-FLCS spectra computed for $\Delta T = 10–50\ \mu s$ and $\Delta T = 10–50$ ms. The three diagonal peaks can be assigned to the K, L, and N intermediate formed during the photoinduced reaction cycle of bR. Cross-peaks evolve because of chemical exchange between the intermediate species.
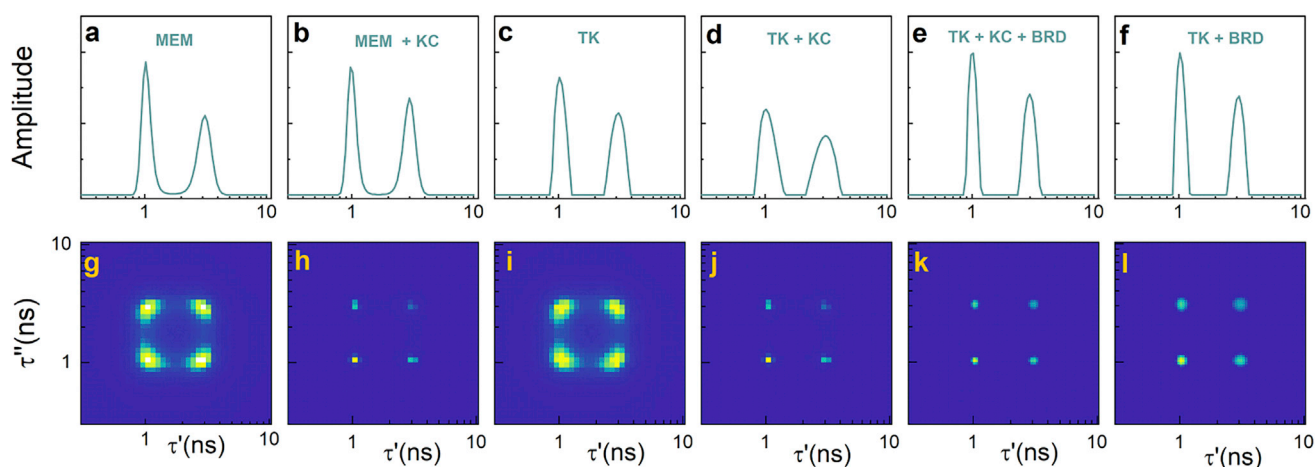
FIGURE 7 Comparison of TK and MEM. (*a–f*) Show 1D-ILT of same simulated data set comprising 2×10^6 total photons with lifetimes 1 and 3 ns. (*g–l*) show the 2D-ILT of the same data set evaluated at $\Delta T = 2$ms. The computational times for the simulations are given in Tables 1 and 2.

with 532-nm excitation. 2D-FLCS spectra computed at $\Delta T = 10$–50 μs and $\Delta T = 10$–50 ms are shown in Fig. 6, *c* and *d* - (single-molecule fluorescence traces contained $\sim 5 \times 10^6$ photons). At $\Delta T = 10$–50 μs, the spectrum contains three diagonal peaks, which can be assigned to the K, L, and N intermediates of the reaction cycle (Fig. 6 *c*) (38–40). Cross-peaks are observed between the K and L intermediates indicating exchange, consistent with the millisecond exchange rates measured from bulk spectroscopy. At later times, cross-peaks between the L and N state, as well as the K and N state, appear, consistent with a millisecond exchange time (41).

Comparison to existing methods

We tested the fitting of same sets of simulated data and quantified the performance of each step individually. The key improvements suggested are kernel compression (KC), Tikhonov Regularization (TK), and implementation of the BRD algorithm (BRD). Tables 1 and 2 show the computational time in seconds required to evaluate the 1D and 2D-ILT fits using MEM alone, MEM with KC, TK alone, TK and KC, TK and BRD, and finally TK with KC and BRD. The corresponding 1D and 2D-ILT fits are shown in Fig. 7. We observe that all the above techniques give accurate results. For cases involving constrained optimization without compression, Fig. 7, *g* and *i* (MEM and TK) the 2D-ILTs appear smoother as compared to the fits obtained using constrained optimization along with KC. However, considering computational efficiency and memory requirements (Tables 1 and 2), we observe that the best and results are obtained when the techniques (TK, KC, BRD) are used in conjunction Fig. 7, *e* and *k*.

Table 3 shows typical computational times required to evaluate 1D and 2D-ILT fits for various sizes of lifetime basis. The fittings were performed on an Intel-Core i7-4790 CPU and using home-built code written in MATLAB 2020a software. The code is publicly available at (https://github.com/saurabhtauke/2D-FLCS_King_lab).

CONCLUSIONS

sm-2D-FLCS is a powerful tool for studying dynamics of biological macromolecules, though the difficulty of computing ILTs may be prohibitive to its widespread application. Here, we outline a fast and robust method of computing 2D-ILTs. The method, based on SVD and TK, is adopted from NMR spectroscopy for application to single-molecule fluorescence spectroscopy. The approach allows for stable inversions of large, noisy data sets, common in single-molecule spectroscopy, to be carried out efficiently, without sacrificing the resolution of the spectra. Furthermore, using Monte Carlo simulations to generate artificial photon streams, we demonstrate that this technique is robust in terms of the spectral resolution, noise tolerance, and computational efficiency. This provides an alternative method (beyond typical MEMs) for performing Laplace inversions of single-molecule fluorescence data that is easily implemented.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.08.031.

SUPPORTING CITATION

Reference (42) appears in the Supporting materials and methods.

REFERENCES

1. Weiss, S. 1999. Fluorescence spectroscopy of single biomolecules. *Science.* 283:1676–1683.

2. Ha, T., T. Enderle, …, S. Weiss. 1996. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. USA.* 93:6264–6268.

3. Deniz, A. A., T. A. Laurence, …, S. Weiss. 2000. Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2. *Proc. Natl. Acad. Sci. USA.* 97:5179–5184.

4. Deniz, A. A., M. Dahan, …, P. G. Schultz. 1999. Single-pair fluorescence resonance energy transfer on freely diffusing molecules: observation of Förster distance dependence and subpopulations. *Proc. Natl. Acad. Sci. USA.* 96:3670–3675.

5. Nettels, D., I. V. Gopich, …, B. Schuler. 2007. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl. Acad. Sci. USA.* 104:2655–2660.

6. Prabhakar, A., E. V. Puglisi, and J. D. Puglisi. 2019. Single-molecule fluorescence applied to translation. *Cold Spring Harb. Perspect. Biol.* 11:a032714.

7. Schuler, B., E. A. Lipman, and W. A. Eaton. 2002. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature.* 419:743–747.

8. Rueda, D., G. Bokinsky, …, N. G. Walter. 2004. Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance. *Proc. Natl. Acad. Sci. USA.* 101:10066–10071.

9. Bonnet, G., O. Krichevsky, and A. Libchaber. 1998. Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc. Natl. Acad. Sci. USA.* 95:8602–8606.

10. Lu, H. P., L. Xun, and X. S. Xie. 1998. Single-molecule enzymatic dynamics. *Science.* 282:1877–1882.

11. Gunn, K. H., J. F. Marko, and A. Mondragón. 2017. An orthogonal single-molecule experiment reveals multiple-attempt dynamics of type IA topoisomerases. *Nat. Struct. Mol. Biol.* 24:484–490.

12. Yang, H., G. Luo, …, X. S. Xie. 2003. Protein conformational dynamics probed by single-molecule electron transfer. *Science.* 302:262–266.

13. Goldsmith, R. H., and W. E. Moerner. 2010. Watching conformational- and photo-dynamics of single fluorescent proteins in solution. *Nat. Chem.* 2:179–186.

14. Schlau-Cohen, G. S., Q. Wang, …, W. E. Moerner. 2013. Single-molecule spectroscopy reveals photosynthetic LH2 complexes switch between emissive states. *Proc. Natl. Acad. Sci. USA.* 110:10899–10903.

15. Schuler, B., and W. A. Eaton. 2008. Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* 18:16–26.

16. Joo, C., H. Balci, …, T. Ha. 2008. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.* 77:51–76.

17. Roy, R., S. Hohng, and T. Ha. 2008. A practical guide to single-molecule FRET. *Nat. Methods.* 5:507–516.

18. Ha, T. 2001. Single-molecule fluorescence methods for the study of nucleic acids. *Curr. Opin. Struct. Biol.* 11:287–292.

19. Ishii, K., and T. Tahara. 2013. Two-dimensional fluorescence lifetime correlation spectroscopy. 1. Principle. *J. Phys. Chem. B.* 117:11414–11422.

20. Ishii, K., and T. Tahara. 2013. Two-dimensional fluorescence lifetime correlation spectroscopy. 2. Application. *J. Phys. Chem. B.* 117:11423–11432.

21. Otosu, T., K. Ishii, and T. Tahara. 2015. Microsecond protein dynamics observed at the single-molecule level. *Nat. Commun.* 6:7685.

22. Kondo, T., J. B. Gordon, …, G. S. Schlau-Cohen. 2019. Microsecond and millisecond dynamics in the photosynthetic protein LHCSR1 observed by single-molecule correlation spectroscopy. *Proc. Natl. Acad. Sci. USA.* 116:11247–11252.

23. Kapusta, P., M. Wahl, …, J. Enderlein. 2007. Fluorescence lifetime correlation spectroscopy. *J. Fluoresc.* 17:43–48.

24. Jeener, J., B. H. Meier, …, R. R. Ernst. 1979. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* 71:4546–4553.

25. Tikhonov, A. N. 1977. Solutions of Ill-Posed Problems. Winston, New York.

26. Tikhonov, A. N. 1963. Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady.* 4:1035–1038.

27. Venkataramanan, L., Y. Q. Song, and M. D. Hurlimann. 2002. Solving Fredholm integrals of the first kind with tensor product structure in 2 and 2.5 dimensions. *IEEE Trans. Signal Process.* 50:1017–1026.

28. Chouzenoux, E., S. Moussaoui, …, F. Mariette. 2010. Efficient maximum entropy reconstruction of nuclear magnetic resonance T1–T2 spectra. *IEEE Trans. Signal Process.* 58:6040–6051.

29. Su, G., X. Zhou, Y. Zhang…, 2019. Improved Butler–Reeds–Dawson algorithm for the inversion of two-dimensional NMR relaxometry data. *Math. Probl. Eng.* 2019:1–12.

30. Skilling, J., and R. Bryan. 1984. Maximum entropy image reconstruction-general algorithm. *Mon. Not. R. Astron. Soc.* 211:111–124.

31. Narayan, R., and R. Nityananda. 1986. Maximum-Entropy image-restoration in astronomy. *Annu. Rev. Astron. Astrophys.* 24:127–170.

32. Livesey, A. K., and J. C. Brochon. 1987. Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. *Biophys. J.* 52:693–706.

33. Tikhonov, A. N. 1963. Solution of incorrectly formulated problems and the regularization method. *Dokl. Akad. Nauk. SSSR.* 151:501–504.

34. Tikhonov, A. N. 1963. On the solution of ill-posed problems and the method of regularization. *Dokl. Akad. Nauk. Russian Academy of Sciences.* 151:501–504.

35. Gull, S. F., and J. Skilling. 1999. Quantified maximum entropy MemSys5 users' manual. Maximum Entropy Data Consultants Ltd., Suffolk, UK, pp. 1–108.

36. Butler, J., J. Reeds, and S. Dawson. 1981. Estimating solutions of first kind integral equations with nonnegative constraints and optimal smoothing. *SIAM J. Numer. Anal.* 18:381–397.

37. Turton, D. A., G. D. Reid, and G. S. Beddard. 2003. Accurate analysis of fluorescence decays from single molecules in photon counting experiments. *Anal. Chem.* 75:4182–4187.

38. Lanyi, J. K. 2004. Bacteriorhodopsin. *Annu. Rev. Physiol.* 66:665–688.

39. Talele, S., and J. T. King. 2021. Stability of a nonequilibrium biochemical cycle revealed by single-molecule spectroscopy. *bioRxiv* https://doi.org/10.1101/2021.05.24.445545.

40. Lanyi, J. K. 1993. Proton translocation mechanism and energetics in the light-driven pump bacteriorhodopsin. *Biochim. Biophys. Acta.* 1183:241–261.

41. Váró, G., and J. K. Lanyi. 1991. Kinetic and spectroscopic evidence for an irreversible step between deprotonation and reprotonation of the Schiff base in the bacteriorhodopsin photocycle. *Biochemistry.* 30:5008–5015.

42. Marcus Sackrow, P. K. 2016. Read_PTU.m, [Online]. Available: https://github.com/PicoQuant/PicoQuant-TimeTagged-File-Format-Demos/blob/master/PTU/Matlab/Read_PTU.m.