A Bayesian method for single molecule, fluorescence burst analysis

P. R. Barber,[1,2,*] S. M. Ameer-Beg,[2] S. Pathmananthan,[2] M. Rowley,[2] and A. C. C. Coolen[2]

[1] *University of Oxford, Gray Institute for Radiation Oncology and Biology, Old Road Campus, Oxford OX3 7DQ, UK*
[2] *King's College London, Division of Cancer Research & Randall Division of Cell and Molecular Biophysics, Richard Dimbleby Department of Cancer Research, Guy's Campus, London SE1 1UL, UK*
*Paul.Barber@rob.ox.ac.uk

Abstract: There is currently great interest in determining physical parameters, e.g. fluorescence lifetime, of individual molecules that inform on environmental conditions, whilst avoiding the artefacts of ensemble averaging. Protein interactions, molecular dynamics and sub-species can all be studied. In a burst integrated fluorescence lifetime (BIFL) experiment, identification of fluorescent bursts from single molecules above background detection is a problem. This paper presents a Bayesian method for burst identification based on model selection and demonstrates the detection of bursts consisting of 10% signal amplitude. The method also estimates the fluorescence lifetime (and its error) from the burst data.

References and links

1. M. Böhmer, and J. Enderlein, "Fluorescence spectroscopy of single molecules under ambient conditions: methodology and technology," ChemPhysChem **4**(8), 792–808 (2003).
2. C. Eggeling, J. R. Fries, L. Brand, R. Günther, and C. A. Seidel, "Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy," Proc. Natl. Acad. Sci. U.S.A. **95**(4), 1556–1561 (1998).
3. J. R. Fries, L. Brand, C. Eggeling, M. Kollner, and C. A. M. Seidel, "Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy," J. Phys. Chem. A **102**(33), 6601–6613 (1998).
4. R. Kühnemuth, and C. A. M. Seidel, "Principles of Single Molecule Multiparameter Fluorescence Spectroscopy," Single Mol. **2**(4), 251–254 (2001).
5. W. Becker, A. Bergmann, E. Haustein, Z. Petrasek, P. Schwille, C. Biskup, L. Kelbauskas, K. Benndorf, N. Klöcker, T. Anhut, I. Riemann, and K. König, "Fluorescence lifetime images and correlation spectra obtained by multidimensional time-correlated single photon counting," Microsc. Res. Tech. **69**(3), 186–195 (2006).
6. J. Enderlein, D. L. Robbins, W. P. Ambrose, and R. A. Keller, "Molecular Shot Noise, Burst Size Distribution, and Single-Molecule Detection in Fluid Flow: Effects of Multiple Occupancy," J. Phys. Chem. A **102**(30), 6089–6094 (1998).
7. W. Grange, P. Haas, A. Wild, M. A. Lieb, M. Calame, M. Hegner, and B. Hecht, "Detection of transient events in the presence of background noise," J. Phys. Chem. B **112**(23), 7140–7144 (2008).
8. J. Enderlein, P. M. Goodwin, A. Van Orden, W. Patrick Ambrose, R. Erdmann, and R. A. Keller, "A maximum likelihood estimator to distinguish single molecules by their fluorescence decays," Chem. Phys. Lett. **270**(5-6), 464–470 (1997).
9. E. Brooks Shera, N. K. Seitzinger, L. M. Davis, R. A. Keller, and S. A. Soper, "Detection of single fluorescent molecules," Chem. Phys. Lett. **174**(6), 553–557 (1990).
10. M. Prummer, C. G. Hubner, B. Sick, B. Hecht, A. Renn, and U. P. Wild, "Single-molecule identification by spectrally and time-resolved fluorescence detection," Anal. Chem. **72**(3), 443–447 (2000).
11. M. Prummer, B. Sick, A. Renn, and U. P. Wild, "Multiparameter microscopy and spectroscopy for single-molecule analytics," Anal. Chem. **76**(6), 1633–1640 (2004).
12. S. H. Lee, M. Jae, and R. P. Gardner, "Non-Poisson counting statistics of a hybrid G-M counter dead time model," Nucl. Instrum. Methods Phys. Res. B **263**, 46–49 (2007).
13. T. C. Voss, I. A. Demarco, and R. N. Day, "Quantitative imaging of protein interactions in the cell nucleus," Biotechniques **38**(3), 413–424 (2005).
14. H. Wallrabe, and A. Periasamy, "Imaging protein molecules using FRET and FLIM microscopy," Curr. Opin. Biotechnol. **16**(1), 19–27 (2005).

15. C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).
16. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992).
17. M. Parsons, J. Monypenny, S. M. Ameer-Beg, T. H. Millard, L. M. Machesky, M. Peter, M. D. Keppler, G. Schiavo, R. Watson, J. Chernoff, D. Zicha, B. Vojnovic, and T. Ng, "Spatially distinct binding of Cdc42 to PAK1 and N-WASP in breast carcinoma cells," Mol. Cell. Biol. **25**(5), 1680–1695 (2005).
18. M. Peter, S. M. Ameer-Beg, M. K. Y. Hughes, M. D. Keppler, S. Prag, M. Marsh, B. Vojnovic, and T. Ng, "Multiphoton-FLIM quantification of the EGFP-mRFP1 FRET pair for localization of membrane receptor-kinase interactions," Biophys. J. **88**(2), 1224–1237 (2005).
19. P. R. Barber, S. M. Ameer-Beg, J. Gilbey, L. M. Carlin, M. Keppler, T. C. Ng, and B. Vojnovic, "Multiphoton time-domain fluorescence lifetime imaging microscopy: practical application to protein–protein interactions using global analysis," J. R. Soc. Interface **6**, S93–S105 (2009).

1. Introduction

Single molecule fluorescence techniques are currently of great interest in biology for their ability to remove the inherent loss of information when ensemble averaging over many molecules [1]. Multiparameter fluorescence detection (MFD) and burst integrated fluorescence lifetime (BIFL) experiments can determine the characteristic fluorescence decay time of single fluorophores as they traverse a focal volume illuminated within the sample [2–4]. Two challenges that accompany such experiments are the identification of the bursts of photons from the individual fluorophores and the determination of the lifetime from the data integrated over each burst.

BIFL experiments are often performed with time correlated single photon counting (TCSPC) [5] in which two arrival times of each photon are recorded; one relative to the start of the recording (the 'macrotime') and the other relative to the excitation pulse train (the 'microtime'). For most organic dyes, the characteristic lifetime of the excited state is in the order of nanoseconds [1] and as such a macrotime resolution in the order of 50 ns and a microtime resolution in the order of picoseconds are suitable and achievable with modern electronics [5].

Burst identification is traditionally achieved by filtering (e.g. Lee filter [6]) and thresholding the detected photon flux based on the macrotime data (at, say, 50% above background) or through the use of statistical analysis and an iterative optimisation of an intensity threshold [7]. A sliding or stepped time window, of predetermined size, is used to isolate portions of the macrotime data. Lifetime can be determined with least-squares fitting of the histogram of microtimes; this reveals the characteristic exponential decay. However, it is often the case that maximum likelihood techniques should be used to overcome bias when dealing with Poissonian noise limited data, as routinely arises from bursts of limited photon counts [8].

Time gating has previously been used to discriminate the fluorescence emission from the Raman signal that may occur in some experiments. This occurs on a very short time scale almost coincident with the excitation pulse, and Shera *et al.* [9] showed a significant signal to noise improvement using a time-gated technique. Time gating in TCSPC can also easily be achieved, during or post acquisition and so additional and existing time-gating techniques can also be applied in parallel to the technique presented in this paper. In our simulations the Raman signal has been excluded in order that we can concentrate on the more difficult problem of identifying signal in the presence of a uniform background signal, which will always be the fundamental noise floor in any experiment, however carefully designed, and after temporal filtering by time gating has been applied.

Bayesian techniques have been used to combine evidence in MFD experiments in order to classify burst data [10,11] using information from reference samples for each class. We offer an alternative application of probabilistic Bayesian techniques in order to identify the bursts themselves in the presence of only approximate prior knowledge about the fluorophore lifetime.

In this paper we present a method of burst detection through Bayesian inference which makes use of the microtime data to select between background and signal models. The macrotime data is used purely to order the photon arrival microtimes on a real-time line and to

select, by means of a sliding time window, consecutive sections of the photon arrival data as previous authors have done. The algorithm identifies bursts based on the distribution of microtimes, as described below, and can also measure the fluorophore lifetime from the burst isolated data. The key to this work is the way that evidence from individual photon counts is combined, working directly with the arrival microtimes and not their accumulation histogram, thus removing the need to make *any* assumptions about the noise in the accumulated data, be it Poissonian or Gaussian, as a dead-time in TCSPC detection can cause counting statistics to deviate from these models [12]. Also key is the selection of background and signal models and the use of Bayesian inference to choose between them. This method allows bursts to be detected when the signal-to-background ratio is much smaller than 50%, and we demonstrate, with simulated data, detection of bursts from a 10% signal component contaminated by a uniform random background.

Experimental data were also collected and used to test the algorithms. A two-photon microscope was used for this purpose because of its inherently small focal volume. The target fluorophores in this case were quantum dots in a very dilute solution which are similar in size and fluorescence lifetime to a typical fluorescent protein. They can also be used as a donor for Förster Resonance Energy Transfer (FRET) [13,14] and this experiment may be the first step towards detecting single molecule FRET via BIFL when single or multiple acceptor fluorophores are in close proximity to the donor.

2. Methods

Bayes' theorem offers a robust and consistent way of combining probabilistic evidence from different sources. If we have a model of a physical phenomenon that contains undetermined parameters, we can use Bayes' theorem to determine likely values for these parameters in the light of given evidence. Also, a significant aspect of Bayesian analysis is its ability to select between different physical models. For our purposes, the selection is between two models: a signal model and a background noise model. Using the Bayes' equation, we may write the probability of a signal being present within the current section of data, $P(H_{sig}|D)$, as [15]

$$P\left(H_{sig}|D\right) = \frac{P\left(D|H_{sig}\right)P\left(H_{sig}\right)}{P\left(H_{sig}\right)P\left(D|H_{sig}\right) + P\left(H_{bg}\right)P\left(D|H_{bg}\right)}, \qquad (1)$$

where $P(A)$ indicates the probability of some event A, and $P(A|B)$ the conditional probability of A given an event B has occurred, H_{sig} represents the signal model, H_{bg} the background model and D represents a set of data. Equation (1) expresses the resulting *posterior* conditional probability of H_{sig} being the appropriate model given the data D, $P(H_{sig}|D)$, in terms of *prior* knowledge about the likelihood of a particular model expressed as a probability $P(H_{sig})$ and the probability of obtaining the data from a probabilistic signal model $P(D|H_{sig})$, and a normalising denominator. It should be emphasised that D represents a section of data of many photon arrival times and the result of applying this equation will be the probability of this data set containing meaningful information about a fluorescence event that matches our model H_{sig}. This analysis does not distinguish between single molecule events and those emanating from multiple molecules and the usual experimental precautions must be taken to ensure single molecule detection [1].

We can formalise suitable models for a random background and an exponential signal plus background in the following way. Our model of random background, H_{bg}, can be simply a uniform probability of a photon detection within the TCSPC measurement period, occurring between the periodic excitation pulses, allowing us to write the probability density function (PDF) of recording a photon as a function of the microtime Δt; i.e. $p(\Delta t|H_{bg})$ represents the probability of recording a microtime within the time period Δt to $\Delta t + d\Delta t$, as

$$p\left(\Delta t \middle| H_{bg}\right) = \frac{1}{T}, \tag{2}$$

where T is the total TCSPC measurement period, also noting $\Delta t \in [0,T]$, and that $\int_0^T p\left(\Delta t \middle| H_{bg}\right) d\Delta t = 1$. The PDF for the signal model, H_{sig}, in the simple case where the excitation function is an assumed delta pulse and the instrument response is assumed to be perfect (i.e. introduces no temporal spreading), can be written as

$$p\left(\Delta t \middle| H_{sig}, \mathbf{w}\right) = (1 - w_0) F\left(\Delta t \middle| w_1\right) + \frac{w_0}{T}, \tag{3}$$

where $\mathbf{w}$ is the vector of model parameters ($w_0 \in [0,1]$, $w_1 \geq 0$) that represents the proportional background signal (w_0) and the fluorescence lifetime (w_1), and $F\left(\Delta t \middle| w_1\right)$ describes the natural decay of the fluorophore as a function of the microtime Δt. If the microtime data are taken in sections of real macrotime by the sliding time window, the assumption can be made that the model parameters, $\mathbf{w}$, are constant for the real-time duration of the current section of the data. We write $F\left(\Delta t \middle| w_1\right)$ as a mono-exponential function (but more complex functions such as multiple or stretched exponentials can be used):

$$F\left(\Delta t \middle| w_1\right) = \frac{e^{-\Delta t / w_1}}{w_1 (1 - e^{-T/w_1})}, \tag{4}$$

ensuring that $\int_0^T p\left(\Delta t \middle| H_{sig}, \mathbf{w}\right) d\Delta t = 1$.

The use of a mono-exponential model is justified since single fluorophores passing through the focal volume may reasonably be assumed to consist of a single species of a particular lifetime in many experiments.

In the presence of very limited reliable prior knowledge, we use a prior PDF for $\mathbf{w}$, $p\left(\mathbf{w}\right)$, that has maximum entropy within the allowed parameter ranges ($w_0 \in [0,1]$, $w_1 \geq 0$), i.e.

$$p\left(\mathbf{w}\right) = \alpha e^{-\alpha w_1}, \tag{5}$$

(note that dependence on w_0 is not necessary). This introduces one hyper-parameter, α, which constrains the average w_1 ($<w_1> = 1/\alpha$) in the prior, and can be set according to the likely order of magnitude of the parameter w_1.

The probability of recording a particular data set, D, consisting of p photons, from the two models, is denoted $P\left(D \middle| H_{sig}\right)$ and $P\left(D \middle| H_{bg}\right)$, can be written as a joint probability as follows,

$$P\left(D \middle| H_{sig}\right) = \int d\mathbf{w} \, p(\mathbf{w}) . \prod_{\mu=1}^{p} \left[p(\Delta t_\mu \mid H_{sig}, \mathbf{w}) \right],$$

$$P\left(D \middle| H_{sig}\right) = \int d\mathbf{w} \, \alpha \, e^{-\alpha w_1} \prod_{\mu=1}^{p} \left[(1 - w_0) \, F\left(\Delta t_\mu \middle| w_1\right) + \frac{w_0}{T} \right], \tag{6}$$

and similarly for the background model, noting that here we can make use of the fact that $\int p\left(\mathbf{w}\right) d\mathbf{w} = 1$.

$$P\left(D|H_{bg}\right) = \prod_{\mu=1}^{p}\left[\frac{1}{T}\right] = \frac{1}{T^{p}}. \tag{7}$$

We can now write down the probability of the signal model, H_{sig}, by substituting (6) and (7) into Eq. (1) and rearranging, as

$$P\left(H_{sig}|D\right) = \frac{P\left(H_{sig}\right).\int d\mathbf{w}\,\alpha\,e^{-\alpha w_1}\prod_{\mu=1}^{p}\left[\left(1-w_0\right)T\,F\left(\Delta t_\mu|w_1\right)+w_0\right]}{\left[1-P\left(H_{sig}\right)\right]+P\left(H_{sig}\right).\int d\mathbf{w}\,\alpha\,e^{-\alpha w_1}\prod_{\mu=1}^{p}\left[\left(1-w_0\right)T\,F\left(\Delta t_\mu|w_1\right)+w_0\right]}, \tag{8}$$

noting that the prior probabilities of each model sum to unity and can be estimated empirically. The parameter vector, $\mathbf{w}$, for the signal model can also be inferred as the PDF, $p\left(\mathbf{w}|D,H_{sig}\right)$, can be written using the Bayesian equation as

$$p\left(\mathbf{w}|D,H_{sig}\right) = \frac{e^{-\alpha w_1}\prod_{\mu=1}^{p}\left[\left(1-w_0\right)T\,F\left(\Delta t_\mu|w_1\right)+w_0\right]}{\int d\mathbf{w}\,e^{-\alpha w_0'}\prod_{\mu=1}^{p}\left[\left(1-w_0'\right)T\,F\left(\Delta t_\mu|w_1'\right)+w_0'\right]}. \tag{9}$$

We may, should it be necessary, include the effect of an instrument response function (IRF) at this point. An IRF, estimated by a Gaussian function, may account for excitation pulse width and detection response time by replacing $F\left(\Delta t|w_1\right)$ by $\overline{F}\left(\Delta t|w_1\right)$:

$$\overline{F}\left(\Delta t|w_1\right) = \frac{\int ds\,\Gamma\left(\Delta t-s\right)F\left(s|w_1\right)}{\int_0^T dt\int ds\,\Gamma\left(t-s\right)F\left(s|w_1\right)}, \tag{10}$$

$$\Gamma\left(u\right) = \frac{e^{-\frac{1}{2}\left(u-u_c\right)^2/\sigma^2}}{\sigma\sqrt{2\pi}}\frac{2\theta[u]}{1+\text{erf}\left(u_c/\sigma\sqrt{2}\right)}. \tag{11}$$

$\Gamma(u)$ is the likelihood that the IRF delays any photon detection by a time u, $\theta[u]$ denotes the step function ($\theta[u > 0] = 1$ and $\theta[u < 0] = 0$), and erf(x) denotes the error function. In reality the excitation pulse will not occur at $\Delta t = 0$ but will be slightly delayed; this delay is accounted for by u_c and the width of the IRF is estimated by σ.

Equation (9) describes how each photon arrival microtime contributes some evidence towards determining the most likely values in the parameter vector and this is illustrated in Fig. 1.

Most likely values of the parameter vector can therefore be determined from the maximum in $p\left(\mathbf{w}|D,H_{sig}\right)$, or from the minimum in $-\log p\left(\mathbf{w}|D,H_{sig}\right)$. We use the *amoeba* routine [16], based on the downhill simplex method, for minimisation of this function. The average values and error bars can also be determined from the first and second moment of $p\left(\mathbf{w}|D,H_{sig}\right)$, where the most likely values serve as starting points for determining these moments. Where lifetimes were determined, we report the most likely value with an error interval equal to the second moment.
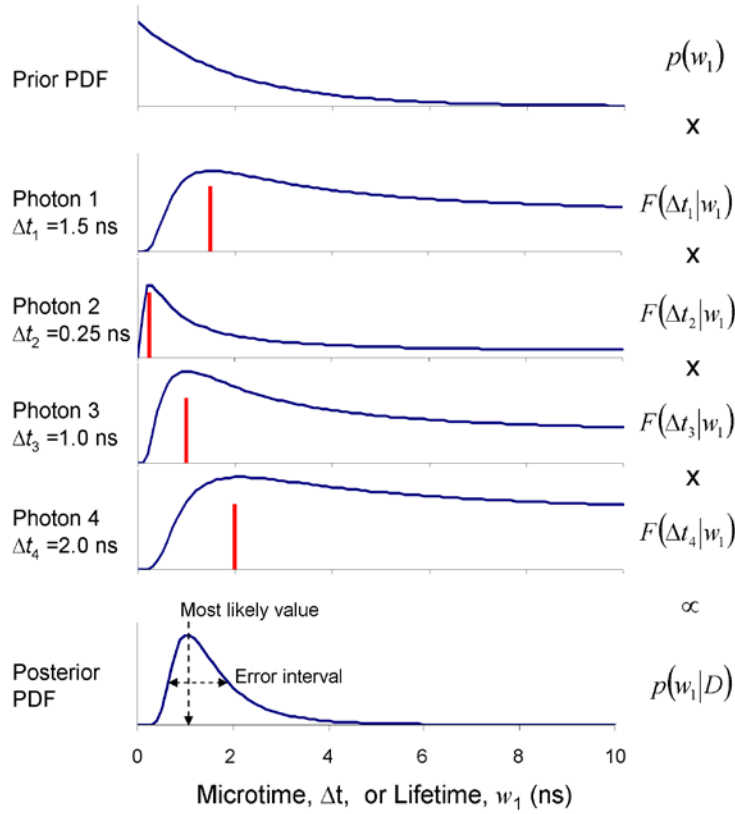
Fig. 1. A simplified schematic illustrating the mathematical process of forming the posterior probability density function (PDF, blue lines) versus lifetime, from which the most likely lifetime and an associated error interval can be calculated. Starting from the prior probability density function, successive photon arrival times (red bars, shown at the appropriate microtime) contribute some evidence under this Bayesian framework, as calculated by the function F (see text). As more photons are measured the posterior PDF may become a single well-defined peak indicating the most likely lifetime value, w_1.

We have thus far allowed the recorded photon arrival times, Δt, to occupy any value in the continuous range [0,T], thus we denote the above as the "continuous algorithm". In reality Δt will be binned and therefore occupy a finite number of discrete values. If, in addition, the values of **w** are also restricted to discrete values within a pre-defined parameter range, the PDF can be calculated on a discrete grid of values and we can now have a very efficient algorithm for deriving the PDF and $P(H_{sig})$ numerically where a trade-off between computational speed and accuracy can be made. This method forms the "discretized algorithm", but does not necessarily limit w_0 and w_1 to the grid values as we can use the first two moments to calculate the means and standard deviations of the marginalised PDF. Thus a minimisation routine, such as *amoeba,* need not be used in this discrete framework as an exhaustive search of the grid to locate the peak posterior likelihood can be performed quickly.

Care should be taken in the design of the grid. The variable range covered must sufficiently encompass the PDF in all dimensions, and the separation of grid points must be sufficiently small to prevent the entire PDF from falling between any two neighbours. This grid technique is thus highly suited to these experiments as the limited number of photon counts results in a smooth and broad PDF from which most likely values and moments are easily calculated. Many more photon counts may result in a sharply peaked PDF and the continuous formalism should be used in this case.

A computer program was written in C to perform both the continuous and discretized algorithms. The discrete algorithm acts by pre-calculating a 3-dimensional 'multiplier grid' of values equal to $\left[(1-w_0)\,T\,F\left(\Delta t\,|w_1\right)+w_0\right]$ for all possible allowed values of w_0, w_1 and Δt. The 2-dimensional PDF (Eq. (9) can then be calculated very speedily by multiplying together the appropriate planes of the multiplier grid that correspond to the microtimes of the detected photons within the sliding macrotime window. Thus a discrete 2D map of the numerator of Eq. (9) is calculated and can be used for lifetime estimates (Eq. (9) and for the signal model probability (Eq. (8).

In Section 3 we describe how the algorithm was tested with simulated and experimental data. Firstly the experimental methods are presented followed by the results for simulated and then experimental data.

2.1 Two-photon Single molecule spectroscopy

Single molecule spectroscopy was performed using a custom-built two-photon excitation system constructed around an upright 90i fluorescence microscope (Nikon, Tokyo, Japan) and similar to that described elsewhere [17,18]. In brief, excitation was provided by a titanium-sapphire laser (Mira 900F, Coherent Inc, Santa Clara, CA, USA), operating at 890 nm and scanned into the image plane using an afocal NIR achromatic scanner and reprojection optics. Time-resolved detection was provided by a non-descanned detection channel with a hybrid photomultiplier (HPM-100-40, Becker and Hickl GmbH, Berlin, Germany) placed in the re-projected pupil plane and a TCSPC personal computer plug-in board (SPC830, Becker and Hickl GmbH, Berlin, Germany). For single-molecule studies, the scanning system parks the excitation beam in the centre of the field of view while data are acquired. The laser power was adjusted during serial dilution of the samples under test to adjust the signal-background ratio while avoiding potential pulse-pile up due to excessive burst counting rates. (typically powers would not exceed 10 mW at the objective pupil). The First-in First-out (FIFO) mode of data acquisition was used for data collection with an ADC resolution of 4096 over a 10 ns measurement period. Data were collected using a bandpass interference filter at centre wavelength 520 nm (FF01-520/15-25: Semrock, Rochester, NY, USA). A Nikon 1.30 NA Plan Fluor, 40x oil immersion objective was used to focus the laser beam into the sample solution.

2.2 Sample Preparation

For single molecule spectroscopy, a 1 nM solution of Streptavidin conjugated quantum dots (Qdots, Invitrogen Corporation UK, Q10141MP) with peak emission wavelength of 525 nm was prepared by serial dilution of the stock solution in spectrophotometric grade, ultrapure water. The sample was pipetted into a clean, single depression cavity slide (Agar Scientific, L4090) and secured with a type 1.5 coverslip.

3. Results

3.1 Analysis of simulated data

Photon arrival time data were generated in a Monte Carlo fashion to simulate a TCSPC experiment with a certain proportion of background and a given mono-exponential lifetime signal in 10 bursts (Fig. 2). Photon arrival macrotimes were randomised and microtimes were generated from random numbers weighted according to the form and lifetime of the fluorescence signal being simulated. During signal periods the background count rate was maintained and the lifetime was set to be 2 ns in all cases. A macrotime window was moved, in a sliding fashion, along the data set to extract a portion of the data to be analysed. The discretized Bayesian methods were applied and if the probability of the signal model being appropriate exceeded a threshold (P_{Hsig}) then a burst was indicated. Table 1 presents the input parameters to the algorithm used to produce Fig. 2. A sliding window of 50 ms traversed the macrotime data in 10 ms steps, and so 50 ms sections of data were independently analysed (the fact that the simulated data are highly periodic is irrelevant to the algorithm). The size of

the sliding window was chosen to achieve a high temporal resolution whilst maintaining sufficient photons for analysis (and therefore a sufficient lifetime precision). The PDF was calculated on a grid where w_0 was allowed to vary between 0 and 1 in 20 steps, w_1 was allowed to vary between 0 and 10 ns in 50 steps and with 4096 Δt time bins, each separated by 2.44 ps. The choice of $P(H_{sig})$ reflects the prior probability of a burst and should be estimated for the system under study. It can be tuned to balance true and false positive detections. In practice one would start with a low value and increase until bursts are detected; more detailed examination of the detected bursts should reveal if true bursts are being detected.

Table 1. Values used by the discrete algorithm when analysing the simulated experiments

Parameter	Value	Parameter	Value	Parameter	Value
Window size	50 ms	$P(H_{sig})$	0.3	Microtime unit	2.44 ps
Window step	10 ms	Min w_1	0.0 ns	Macrotime clock	50 ns
Signal probability Threshold (p_{Hsig})	0.5	Max w_1	10.0 ns	Microtime interval (T)	10 ns
		# w_1 bins	50	# microtime bins	4096
		# w_0 bins	20		
		α	0.5		

We can see in Fig. 2 a and b that the red line of P(signal) = $P(H_{sig}|D)$ peaks when a burst is detected. The 'background' (dotted line) trace is of w_0 which is near 1 in the presence of no signal. Where a burst of signal occurs, this trace decreases indicating the proportion of signal contributing to the total intensity. Note that the background intensity remains constant throughout but the proportion of background to signal changes when bursts occur.

For demonstration purposes, several sections of data were chosen for analysis with the continuous Bayesian method. These are also presented in Fig. 2. The discrete method provided estimates of lifetime of 3.2 ± 0.9 ns and 3.6 ± 1.8 ns for the data in Fig. 2 c and d respectively, whereas the continuous method provides estimates of 1.8 ± 0.3 ns and 1.9 ± 1.4 ns (photon counts 1467 and 1078). We must bear in mind that the photon count that contributes to the signal is approximately 500 and 100 counts respectively. We can see that the discrete method is good at identifying bursts but the analysis of these bursts for lifetime should be performed with a more thorough analysis such as the continuous algorithm.

The performance of the burst detection algorithm with decreasing signal proportion is presented in Fig. 2h, showing significant performance degradation below 10% signal. The parameters have been arranged such that there were minimal false positive detections; Fig. 2 h indicates the true and false positive detection rates.

The total processing time for the discrete analysis was 12.7 seconds, on a standard office PC, to process the full photon train of 40,395 counts in the 50% signal case, including reading the data, setting up the multiplier grid and performing Bayesian analysis on each macrotime slice of 50 ms, stepping through 1.6 seconds of data in 10 ms steps.
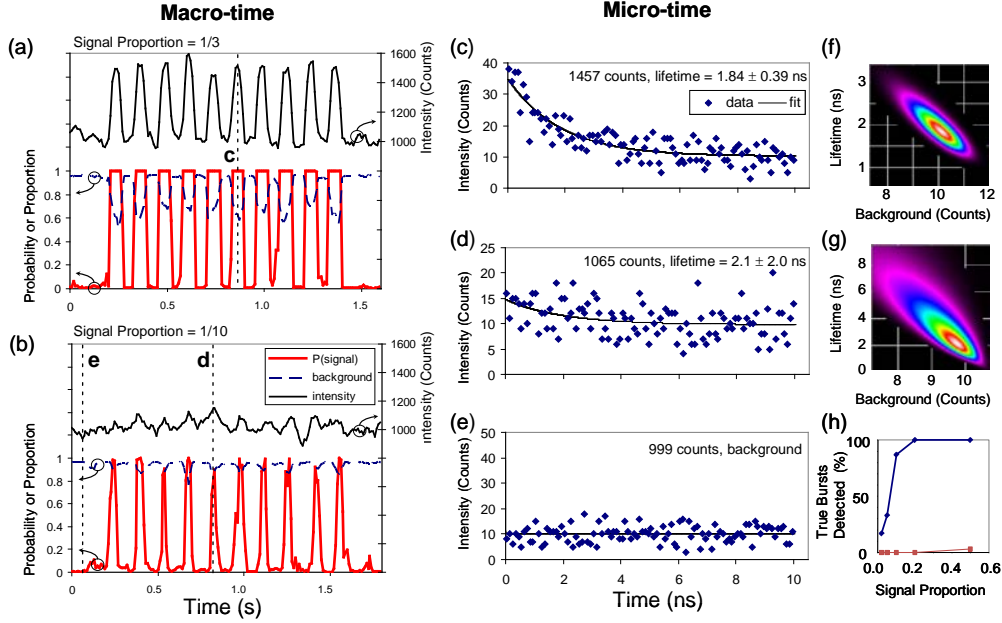
Fig. 2. Bayesian burst detection from simulated data. Signal fluorescence lifetime is set to 2 ns in all cases. (a) Bursts where the signal proportion of 1/3 (bg: 2 x 10^4 s^{-1} sig: 1 x 10^4 s^{-1}) are easily detected (and would equally be by simple thresholding of the macrotime trace). (b) Bursts with a signal proportion of 1/10 (bg: 2 x 10^4 s^{-1} sig: 2 x 10^3 s^{-1}) cannot be detected from the macrotime trace alone. The Bayesian technique detects these bursts as shown by the resulting probability of a signal, P(signal) = $P(H_{sig}|D)$ (red line). (c) The microtime trace with 1/3 signal, extracted from the photon train near 0.8 seconds (see dotted line labelled 'c' in (a)). (d and e) Microtime traces representing 1/10 signal and pure background noise (see dotted lines in (b)). Microtime traces were formed from data re-binned into 100 bins for plotting clarity. Where bursts were detected, lifetime fitting was performed by the full continuous Bayes algorithm on the re-binned data. The fitted curves are shown (solid line) and on the right hand side (f and g) are plots of the 2D PDF around the most likely value, where white indicates high values, decreasing through a spectrum to black which indicate low values. (h) The percentage of true detected bursts (blue line) and false positive detections (red line) as a function of signal proportion. ('counts' = number of photon counts in trace, 'lifetime' = determined lifetime).

3.2 Analysis of experimental data

The data acquired with the experimental setup from the dilute quantum dot suspension are presented in Fig. 3. Several clearly identified bursts were detected (where P(signal) is above the P_{Hsig} threshold) by the discrete Bayesian algorithm, using the parameters shown in Table 2, one of which could be easily identified because of its high peak intensity but many of which would have been missed by simple thresholding of the macrotime trace. Again, 2 bursts and a section of background were selected for lifetime analysis with the continuous Bayesian method and these are shown in Fig. 3 b, c and d. The discrete method provided estimates of lifetime of 6.21 ± 1.25 ns and 4.55 ± 1.64 ns for the data in Fig. 3 b and d respectively, whereas the continuous method provides estimates of 6.20 ± 1.44 ns and 4.88 ± 2.39 ns, respectively (photon counts 1015 and 84). An IRF was used in this analysis with empirically determined values of u_c = 1.40 ns and σ = 0.1 ns. These data were also analysed with a traditional Levenberg-Marquardt (LM) non-linear least squares algorithm [19] which failed to find a suitable solution in the first case but gave an estimate of 2.6 ± 6.1 ns in the second case with 84 photons counted. This demonstrates the robustness of Bayesian techniques.

Table 2. Values used by the discrete algorithm when analysing the experimental data

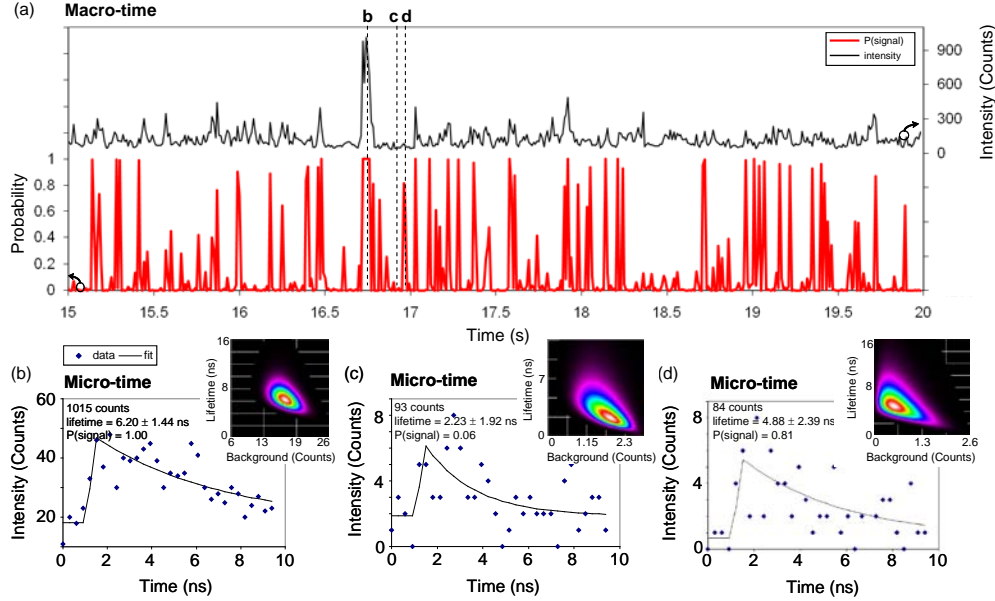| Parameter	Value	Parameter	Value	Parameter	Value
Window size	10 ms	$P(H_{sig})$	0.001	Microtime unit	2.44 ps
Window step	10 ms	Min w_1	0.0 ns	Macrotime clock	50 ns
Signal probability Threshold (p_{Hsig})	0.5	Max w_1	12.0 ns	Microtime interval (T)	10 ns
		# w_1 bins	50	# microtime bins	4096
		# w_0 bins	20		
		α	0.5		

Fig. 3. Bayesian burst detection from experimental data of a dilute solution of quantum dots. (a) Macrotime intensity trace and corresponding signal probability. The obvious high-intensity burst is detected along with several low intensity bursts that are not apparent through analysis of the macrotime trace alone. (b, c and d) Microtime traces extracted from the photon train at the positions indicated by the dotted lines in (a). Lifetime fitting was performed on the microtime data and the fitted curve is shown (solid line, re-binned into 32 bins for clarity) and inset is a plot of the PDF around the most likely value. Data in graph (c) was rejected as background since P(signal) is below 0.5. Lifetimes of 6.20 ns and 4.88 ns were extracted by a full, continuous, Bayesian analysis of the high- and low-intensity bursts, respectively.

A larger section of data was processed (approximately 65,000 photons counted) as a verification of the expected lifetime. Both Bayesian and LM techniques indicated an average lifetime of 4.5 ns when fitted with a mono-exponential model. Further analysis with LM and bi- and tri-exponential models indicated that these data for the ensemble are bi-exponential with components near 5.5 ns and 0.4 ns, in a 57:43 amplitude ratio.

4. Conclusion

In conclusion our fast discrete algorithm for burst identification in BIFL experiments uses the combination of evidence from the microtime data within the macrotime-based sliding window to determine the most likely model for that section of data. With this Bayesian technique we demonstrate the detection of bursts with 10% signal on a 90% background from simulated data. The more general continuous Bayesian algorithm can also be used for lifetime estimation with improved accuracy over a traditional Levenberg-Marquardt non-linear least squares algorithm with low photon count data. A direct comparison of the fully continuous

Bayes method, working directly on the photon arrival times, with a previously published maximum likelihood method, working on the accumulation histogram, will be the subject of a future paper.

Acknowledgments