Bayesian analysis of fluorescence lifetime imaging data

Mark I Rowley[a], Paul R Barber[a,b], Anthony CC Coolen[a] and Borivoj Vojnovic[a,b]

[a]King's College London, Randall Division of Cell and Molecular Biophysics, Guy's Campus, London, SE1 1UL, UK;

[b]University of Oxford, Gray Institute for Radiation Oncology and Biology, Old Road Campus, Oxford, OX3 7DQ, UK

ABSTRACT

Fluorescence Lifetime Imaging (FLIM) is an intensity independent and sensitive optical technique for studying the cellular environment but its accuracy is often compromised when low photon counts are available for analysis. We have developed a photon-by-photon Bayesian analysis method targeted at the accurate analysis of low photon count time-domain FLIM data collected using Time Correlated Single Photon Counting (TCSPC). Parameter estimates obtained with our mono-exponential Bayesian analysis compare favorably with those using maximum likelihood, least squares, and phasor analysis, offering robust estimation with greater precision at very low total photon counts, particularly in the presence of significant background levels. Details of the Bayesian implementation are presented alongside results of mono-exponential analysis of both real and synthetic data. We demonstrate that for low photon count data, obtained by imaging human epithelial *carcinoma* cells expressing cdc42-GFP, Bayesian analysis estimates the green fluorescent protein (GFP) lifetime to a level of accuracy not obtained using maximum likelihood estimation or other techniques. These results are echoed by the analysis of synthetic decay data incorporating a 10% uniform background, with our Bayesian analysis routines yielding lifetime estimates within an accuracy of 20% with about 50 counts. This level of precision is not achieved with maximum likelihood nor phasor analysis techniques with fewer than 100 counts.

Keywords: Time-resolved microscopy, TCSPC, FLIM, Bayesian analysis

1. INTRODUCTION

Time-Correlated Single Photon Counting (TCSPC) can be used for data acquisition in experiments that use Fluorescence Lifetime Imaging Microscopy (FLIM), permitting the accurate measurement of the detection time of single photons after periodic pulse excitation. In this type of time-domain FLIM a pulsed laser provides periodic excitory illumination of the sample causing subsequent photon emission; the detection time (relative to the time of excitation) of single photon pulses is then determined with high precision and recorded. Over time, after many periods of excitation, a data set that reflects the characteristics of the fluorescence decay of the sample accumulates.

The effective analysis of fluorescence decay data is a statistical problem and is not a simple task when photons are in short supply. Numerous methods of analysis to estimate, with the greatest accuracy possible, parameter values that best fit the experimentally measured distribution have been previously studied, each having their own advantages and disadvantages.[1] When significant counts are available, where Gaussian statistics closely approximate the Poissonian statistics associated with photon count data, data fitting is relatively straightforward and both least squares (LS) and maximum likelihood estimation (MLE) have been shown to perform well.[2–7] The collection of significant counts is however inevitably linked to longer experimental acquisition times. When only a low number of counts is available, resulting from either low fluorophore concentrations, or their propensity to fade or when dynamic processes need to be studied, difficulties in determining the correct lifetime, or establishing the error associated with lifetime estimation inevitably increase. Under such conditions, standard LS has been shown to yield poor estimates,[2,3,5,7] though data manipulation prior to analysis[4,5] and a modification of the LS weighting[8] have both been shown to improve LS estimates. The use of MLE methods for the analysis of exponential decays has also been thoroughly studied[3,5] and although it does yield parameter predictions that are more accurate than those of LS[3] and offer good resilience to changes in bin width and photon count,[5] the critical dependence of MLE algorithms on their starting location (in the parameter space) lead to the suggestion of a hybrid LS-MLE

Further author information:
Mark Rowley: E-mail: mark.rowley@kcl.ac.uk; Paul Barber: E-mail: Paul.Barber@rob.ox.ac.uk

analysis.[5] Further practical limitations arise due to non-ideal data acquisition systems and detector quantum efficiency limitations. In this paper, we do not consider these limitations and the terms "photon" and "count" are assumed to have similar meanings. We also assume that the non-linearity of the time axis is adequately small, a situation that is indeed the case with most of the currently available TCSPC data acquisition systems.

The justification for even the most well-established, and a commonly used analysis method, LS, is theoretically unsound when only a meagre number of photons are detected.[2] Though MLE and LS do perform optimally for large sample sizes,[9] it has been suggested that when the large sample assumptions are not met the Bayesian framework should be a preferred alternative analysis technique.[9] The effectiveness of such FLIM techniques to study the cellular environment is, therefore, critically dependent on the accuracy of the data analysis methods. The aim of this work was to develop a technique that extracts the maximum amount of information from available data and to be confident that uncertainties associated with the derived parameters can be quantified.

The analysis presented here avoids the fitting approach altogether, offering a new approach utilizing the Bayesian method to relate the parameters of a fluorescence decay model, developed in a continuous time domain and adapted to accommodate the discrete time nature of TCSPC data, to a set of binned photon arrival times. This type of Bayesian framework was introduced in Ref. 10 for model selection of burst integrated fluorescence lifetime studies. This work now takes the framework forward to determine accurate lifetime parameter estimates and to include the repetitive excitation of TCSPC and seeks to establish performance limitations of both this method and other more established techniques.

2. METHODS

In formalizing our Bayesian approach to the analysis of fluorescence lifetime imaging data, we aim to describe the likelihood of a photon being detected at a particular measured time given a purely exponential decay process while also accounting for any significant effects introduced by the data acquisition. The model is then applied in the Bayesian framework, photon-by-photon, to a set of p photon arrival times that constitute the data $D = \{\Delta t_1, \ldots, \Delta t_p\}$, each arrival time being recorded during a measurement interval $[0, T]$ following pulsed excitation. The assumed mono-exponential fluorescence decay model is of the form:

$$p(\Delta t | w_0, w_1) = \frac{w_0}{T} + (1 - w_0)\overline{F}(\Delta t | w_1), \tag{1}$$

where w_0 is the fraction of detected photons due to a uniform background signal (with $w_0 \in [0, 1]$), w_1 is the relaxation time of the photon-generating emission process, and $\overline{F}(\Delta t | w_1)$ describes the likelihood of a photon due to the emission process being detected at time Δt following some stimulation event and is normalized over the measurement interval $[0, T]$. In this section, we define a fluorescence decay model in continuous time that reflects the nature of the observed data, then discuss the compensations necessary for the discrete time and repetitive nature of TCSPC data, before applying the Bayesian framework to extract the statistics of the model parameters of interest in Sec. 3.

2.1 Mono-exponential fluorescence decay model in continuous time

In arriving at a realistic and generally useful expression for $\overline{F}(\Delta t | w_1)$ it is sensible to consider the impact of the data acquisition process on the recorded data, considering especially the influence that repetitive pulsed excitation has on the physical photon generating process, $F(\Delta t | w_1)$, and also the effects of the instrument on the photon arrival times that are eventually recorded. The photon recording system includes the contributions of the tails from any previous decays. Indeed, kinetic information is now available for use from these contributions. The effects of this process on the resultant waveform are illustrated in Fig. 1 (a). Suppose that we observe a portion of this waveform immediately following the most recent pulse, the likelihood of a photon being detected at a time Δt immediately following the most recent pulse, conditioned on a purely exponential decay of lifetime w_1 and incorporating the 'history' of all previous decays, is described by:

$$F(\Delta t | w_1) \propto \sum_{\ell=0}^{k} e^{-(\Delta t + \ell T_m)/w_1}, \tag{2}$$

where $k = 0, 1, 2, 3, \ldots$ is the number of excitation periods T_m that have preceded the most recent pulse. For simplicity, we now only consider photons that are counted once an 'equilibrium' state has been reached, (2) thereby simplifies as follows:

$$F(\Delta t|w_1) \quad = \quad \left(1 + \frac{1}{e^{T_m/w_1} - 1}\right) e^{-\Delta t/w_1}, \tag{3}$$

a situation which, in practice, is reached after a few pulses in typical conditions where the excitation period is significantly longer than the lifetime.

The effects of the instrument response on the measured photon arrival times are now incorporated, the instrument being assumed to introduce a delay u_c and an uncertainty, approximated by a Gaussian of width σ, into the precise time between the physical photon generation and photon detection, according to:

$$\Gamma(u) = \frac{e^{-\frac{1}{2}(u-u_c)^2/\sigma^2}}{\sigma\sqrt{2\pi}} \frac{2\theta[u]}{1 + \operatorname{erf}(u_c/\sigma\sqrt{2})}, \tag{4}$$

where $\operatorname{erf}(z)$ denotes the error integral $\operatorname{erf}(z) = (2/\sqrt{\pi})\int_0^z dx\ e^{-x^2}$ and $\theta[u]$ denotes the Heaviside step function (i.e. $\theta[u \geq 0] = 1$, $\theta[u < 0] = 0$). In order that the effects of previous pulses and the effects of the detection instrument are incorporated into $\overline{F}(\Delta t|w_1)$, the likelihood of a photon being detected at arrival time Δt given exponential decay time w_1 and including, Eqn. (3) is adapted to include a 'pre-pulse' term:

$$F(\Delta t|w_1) = \left(1 + \frac{1}{e^{T_m/w_1} - 1}\right) \left(\underbrace{\theta\left[-\Delta t\right].e^{-(\Delta t + T_m)/w_1}}_{\text{``Pre-pulse''component}} + \theta\left[\Delta t\right].e^{-\Delta t/w_1}\right), \tag{5}$$

and convolved with the approximated instrument response (Eqn. (4)):

$$\overline{F}(\Delta t|w_1) = \frac{\int ds\ \Gamma(\Delta t - s)F(s|w_1)}{\int_0^T dt\ \int ds\ \Gamma(t - s)F(s|w_1)}. \tag{6}$$

2.2 Compensating for discrete time TCSPC data

We now consider the discrete nature of TCSPC FLIM data (i.e. photon arrival times are collected in discrete bins) and adapt the above model to account for such data. The measurement interval is partitioned into m non-overlapping intervals or bins of finite width, with p_j being the number of detected photon arrival times falling in the jth bin, that is, the interval between the bin walls b_j^{L} and b_j^{H}. Returning to Eqn. (6), rather than our model operating directly on a measured photon arrival time Δt, we now need to work with the average likelihood $\widehat{F}(\Delta t \in b_j|w_1) = (b_j^H - b_j^L)^{-1} \int_{b_j} d\Delta t\ \overline{F}(\Delta t|w_1)$ of a photon being detected in the bin b_j, arriving at the following expression:

$$\widehat{F}(\Delta t \in b_j|w_1) = \frac{1}{b_j^{\text{H}} - b_j^{\text{L}}} \cdot \frac{\chi(b_j^{\text{H}}, w_1, u_c, \sigma) - \chi(b_j^{\text{L}}, w_1, u_c, \sigma)}{\chi(T, w_1, u_c, \sigma) - \chi(0, w_1, u_c, \sigma)}, \tag{7}$$

where the quantity $\chi(x, w_1, u_c, \sigma)$ is defined as:

$$\begin{aligned}
\chi(x, w_1, u_c, \sigma) \quad = \quad & e^{-(x-u_c)/w_1}\left[\operatorname{erf}\left(\frac{u_c}{\sigma\sqrt{2}} + \frac{\sigma}{w_1\sqrt{2}}\right) - (1 - e^{-T_m/w_1})\operatorname{erf}\left(\frac{u_c - x}{\sigma\sqrt{2}} + \frac{\sigma}{w_1\sqrt{2}}\right) + e^{-T_m/w_1}\right] \\
& + (1 - e^{-T_m/w_1})e^{-\frac{1}{2}\sigma^2/w_1^2}\operatorname{erf}\left(\frac{u_c - x}{\sigma\sqrt{2}}\right). \tag{8}
\end{aligned}$$

The bin-averaged arrival time likelihood, for a selection of lifetimes, is displayed in Fig. 1 (b).
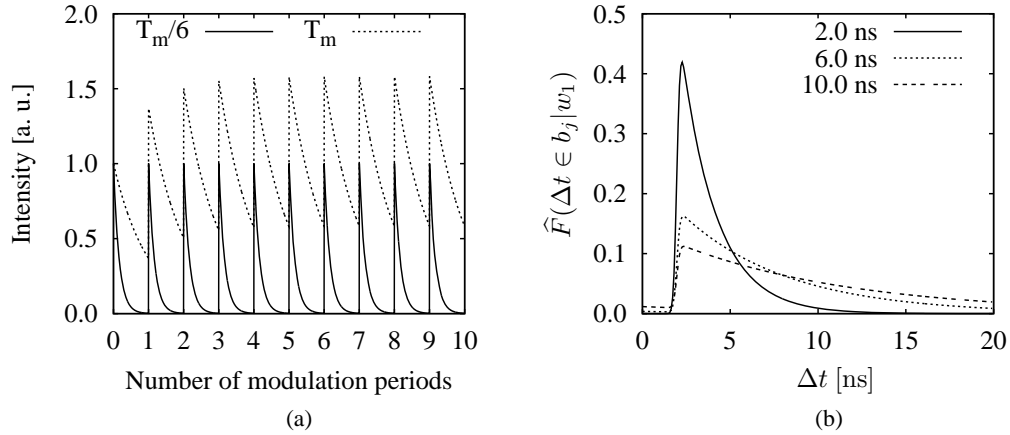
Figure 1. *Modeling repetitive excitation:* (a) Influence of repetitive excitation on the photon generating process emission waveform, observe that for a decay lifetime of the order of the period between successive excitations (i.e. lifetime T_m) the waveform is elevated and differs significantly from the case where the lifetime is considerably shorter than the modulation period (i.e. lifetime $T_m/6$). (b) The quantity $\widehat{F}(\Delta t \in b_j | w_1)$ computed for three different lifetimes, for a system having modulation period 25.0 ns, delay $u_c = 2.0$ ns, convolution width $\sigma = 0.15$ ns and a measurement interval of 20.0 ns partitioned into 256 bins of equal width.

2.3 Bayesian analysis

Having defined our fluorescence decay model we are now able to apply the Bayesian formalism and determine expressions to extract the statistics of the model parameters $\boldsymbol{w} = (w_0, w_1)$, in a similar way to Ref. 10. The likelihood of a photon being detected at arrival time Δt falling in the jth bin b_j, conditioned on a uniform background proportion w_0 and a fluorescence decay of lifetime w_1, is given by:

$$p(\Delta t \in b_j | \boldsymbol{w}) = \left\{ \begin{array}{ll} \frac{w_0}{T} + (1 - w_0)\widehat{F}(\Delta t \in b_j | w_1) & \text{if} \quad \Delta t \in [0, T] \\ 0 & \text{otherwise} \end{array} \right. . \tag{9}$$

The prior distribution $P(\boldsymbol{w})$ of the model parameters is determined from maximum entropy (under constraints $w_0 \in [0, 1]$, $w_1 \geq 0$), and involves one non-negative hyper-parameter α, which constrains the average $\langle w_1 \rangle = 1/\alpha$ in the prior:

$$p(\boldsymbol{w}) = \left\{ \begin{array}{ll} \alpha e^{-\alpha w_1} & \text{if} \quad w_0 \in [0, 1], w_1 \geq 0 \\ 0 & \text{otherwise} \end{array} \right. . \tag{10}$$

The posterior parameter distribution $P(\boldsymbol{w}|D)$ is given by:

$$\begin{aligned} P(\boldsymbol{w}|D) &= \frac{P(\boldsymbol{w})P(D|\boldsymbol{w})}{\int d\boldsymbol{w}' \, P(\boldsymbol{w}')P(D|\boldsymbol{w}')} \\ &= \frac{e^{-\alpha w_1} \prod_{j=1}^{m} \left[w_0 + (1 - w_0)T\widehat{F}(\Delta t \in b_j | w_1) \right]^{p_j}}{\int dw_0' \int dw_1' e^{-\alpha w_1'} \prod_{j=1}^{m} \left[w_0 + (1 - w_0)T\widehat{F}(\Delta t \in b_j | w_1') \right]^{p_j}}, \end{aligned} \tag{11}$$

where the quantity $\widehat{F}(b_j | \tau_\ell)$ denotes the averaged likelihood of photon arrival times in the interval of the bin b_j, and the data is comprised of a set of m bin counts where p_j is the number of times a photon has been detected in the time interval b_j. Defining the quantity $S(\boldsymbol{w}, D) = -\ln P(\boldsymbol{w}|D)$,

$$S(\boldsymbol{w}, D) = \alpha w_1 - \sum_{j=1}^{m} p_j \ln\left[w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)\right], \tag{12}$$

the most likely parameter values $\boldsymbol{w}^\star = (w_0^\star, w_1^\star)$, are defined as $S(\boldsymbol{w}^\star, D) = \min_{\boldsymbol{w}} S(\boldsymbol{w}, D)$ and can be located numerically by minimization of $S(\boldsymbol{w}, D)$ with respect to $\boldsymbol{w} = (w_0, w_1)$. Alternatively, or additionally, in locating the extrema of $S(\boldsymbol{w}, D)$ its partial derivatives may be utilized. The derivative with respect to w_0,

$$\frac{\partial}{\partial w_0} S(\boldsymbol{w}, D) = \sum_{j=1}^{m} p_j \frac{T\widehat{F}(\Delta t \in b_j | w_1) - 1}{w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)}, \tag{13}$$

offers the following relation for the optimal w_0 as a function of w_1,

$$\frac{1}{p} \sum_{j=1}^{m} p_j \frac{T\widehat{F}(\Delta t \in b_j | w_1) - 1}{w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)} = 0 \quad \Rightarrow \quad \frac{1}{p} \sum_{j=1}^{m} p_j \frac{1}{w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)} = 1$$

$$\Rightarrow \quad w_0 = \frac{1}{p} \sum_{j=1}^{m} p_j \left[1 + (w_0^{-1} - 1)T\widehat{F}(\Delta t \in b_j | w_1)\right]^{-1}, \tag{14}$$

effectively reducing any numerical search for the most likely $\boldsymbol{w}$ to a search involving variation of w_1 only. It is also possible to extremize analytically with respect to w_1; this, however, is not considered in this current work.

The marginal posterior distributions are used to obtain the statistics of w_0 and w_1, and yield the predictions $w_0 = \overline{w}_0 \pm \Delta w_0$ and $w_1 = \overline{w}_1 \pm \Delta w_1$, where the posterior average value of the lifetime parameter w_1 (which is usually of greatest interest) is given by:

$$\overline{w}_1 = \frac{\int_0^\infty dw_1\, w_1 e^{-\alpha w_1} \int_0^1 dw_0 \prod_{j=1}^{m} p_j \left[w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)\right]}{\int_0^\infty dw_1' \int_0^1 dw_0'\, e^{-\alpha w_1'} \prod_{j=1}^{m} p_j \left[w_0' + (1-w_0')T\widehat{F}(\Delta t \in b_j | w_1')\right]}, \tag{15}$$

and the posterior width is given by:

$$\Delta w_1 = \left\{ \frac{\int_0^\infty dw_1\, w_1^2 e^{-\alpha w_1} \int_0^1 dw_0 \prod_{j=1}^{m} p_j \left[w_0 + (1-w_0)T\widehat{F}(\Delta t \in b_j | w_1)\right]}{\int_0^\infty dw_1' \int_0^1 dw_0'\, e^{-\alpha w_1'} \prod_{j=1}^{m} p_j \left[w_0' + (1-w_0')T\widehat{F}(\Delta t \in b_j | w_1')\right]} - \overline{w}_1^2 \right\}^{\frac{1}{2}}. \tag{16}$$

Routines have been implemented in C and incorporated into the fitting program TRI2 [10] to compute the model parameter statistics and also enable the posterior distribution $P(\boldsymbol{w}|D)$ to be captured and explored graphically. A minimization routine based on the downhill simplex method is employed to determine the mode of the posterior distribution; this serves as the starting point in the exploration of the posterior distribution and the computation of its statistics. In Sec. 3 the results of the Bayesian analysis algorithms are presented and compared to existing parameter estimation methods for both synthetic and real experimental data. The Bayesian framework can also be used for the determination of the optimal hyper-parameter value given the photon arrival time data, however, in this work (to avoid a significant increase in computation), the hyper-parameter α value is determined from the average arrival time value.

2.4 Time-domain FLIM

Time domain FLIM was performed with an in-house system, excitation being provided by a supercontinuum 40MHz source (SC450-4, Fianium, UK), pulse width ~ 4 ps, and photon counting performed using a SPC830 TCSPC board (Becker & Hickl, Berlin, Germany). De-scanned detection was afforded by the use of a fast single-photon response, photomultiplier tube (PMH-100-0, Becker & Hickl, Berlin, Germany) placed behind a pinhole, parfocal with the image plane. The following filters were used in data acquisition; 470 nm excitation filter (Semrock FF01-470/22-25, Laser 2000, UK), 495 nm dichromatic reflector (Semrock FF 495-Di02-25x36, Laser 2000, UK), and 510 nm emission filter (Semrock FF01-510/20-25, Laser 2000, UK) and an IR cut filter. To avoid pulse pile-up, peak photon counting rates remained well below the maximum counting rate offered by the TCSPC electronics, [11] with average photon counting rates of the order $10^4 - 10^5$ photons/second. The photon arrival times, with respect to the approximately 40 MHz repetitive laser pulses, were binned into 256 time windows over a total measurement period of 15 ns. Images were captured with a 0.75 NA objective lens (S Fluor 20x/0.75 air, Nikon, UK) at 256×256 pixels corresponding to 334×334 μm at the sample.

2.5 Sample preparation

Human epithelial *carcinoma* cells (A431) stably expressing cdc42-GFP [12] were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum. Cells were seeded on a glass coverslip (50000 cells/ml) and 48 hours later were fixed with 4% paraformaldehyde, treated with 1mg/ml NaBH4, then mounted with Mowiol (VWR, UK) containing antifade agent and kept at -20 degC. Data acquisition was performed at room temperature.

3. RESULTS

The photon-by-photon Bayesian parameter estimation routines were tested against LS and MLE fitting routines as well as phasor analysis. The latter established techniques operate directly on the accumulated histogram. The LS implementation is based upon the modified Levenberg-Marquardt (MLM) algorithm and is described in Refs. 6 and 7. Maximum likelihood estimation routines were implemented as described in Ref. 13, also based on the MLM algorithm. The LS, MLE, and phasor analysis routines all consider only a window of the collected data that excludes time points before the rise of the transient has occurred. Automated selection of the valid data window is discussed in Ref. 6. In contrast, the Bayesian analysis routines operate on all of the available data, except for small portions at the start and end of the transient, potentially corrupted by consequences of dithering associated with the time-amplitude converter in the TCSPC electronics. [11] In the Bayesian analysis, the hyper-parameter α acts as a regularizer (e.g. Ref. 14) and is configured separately for each analyzed data set, its value being determined solely from the binned arrival times of that data set. As such, the Bayesian analysis presented in this paper does not benefit from any additional information over the other methods.

Phasor analysis (e.g. Ref. 15, 16) was implemented incorporating automated background level estimation at each pixel (using data before the rise of the transient), with Fourier components computed using modulation frequency $\omega = 2\pi/m$ set according to the number of time bins m that form the window of data to be analyzed, and with the estimated background level being subtracted from the photon count at each bin of the data window. The average of the phase and modulation lifetime values is reported.

The fitting algorithms were tested with simulated data of the expected signal from a model of the TCSPC system which included a fixed background and the effect of repetitive excitation [6] at 40 MHz; simulated Poisson noise was added to the convolution of a raw fluorescence signal and a Gaussian excitation pulse to simulate realistic photon count transients.

3.1 Analysis of simulated data

A comparison of the performance of the estimation techniques at low total photon counts is illustrated in Fig. 2 for synthetic data comprising a decay of lifetime $\tau = 2.0$ ns, a uniform background of 10%, and incorporating an instrument response having a convolution width 0.15 ns at FWHM; from 2^{16} transients analyzed it can be observed that the Bayesian estimates converge towards the true lifetime value much more rapidly than the those of the other estimation methods. Lifetime estimates with an accuracy (measured according to the standard deviation of the lifetime distribution) of about 10% are offered by our Bayesian routines with a total photon count of about 200; MLE and phasor analysis offer lifetime estimates to an accuracy of only about 15% at such a total photon count, with LS estimates having an uncertainty of almost 20%. Lifetime estimates with an accuracy of about 10% are obtained with about 400 total counts per pixel using MLE, about 450 total counts with phasor analysis, and greater than 500 total counts using LS. Under such conditions, to achieve lifetime

estimates to an accuracy of about 10%, the Bayesian analysis would enable a decrease in the imaging duration of about a factor of two over that required for MLE or phasor analyses. The F value, as introduced in Ref. 17, is here used to compare the sensitivity of the lifetime estimation techniques. The lifetime estimates presented in Fig. 2, offer an F value of about 1.43 for our Bayesian analysis, and about 1.90, 2.63 and 2.24 for the MLE, LS and phasor analysis estimates respectively. Our Bayesian analysis lifetime estimates are more narrowly distributed than those from the other estimation techniques, as shown in Fig. 2 (b), and this is reflected in a smaller F number which indicates the greater sensitivity of our technique.
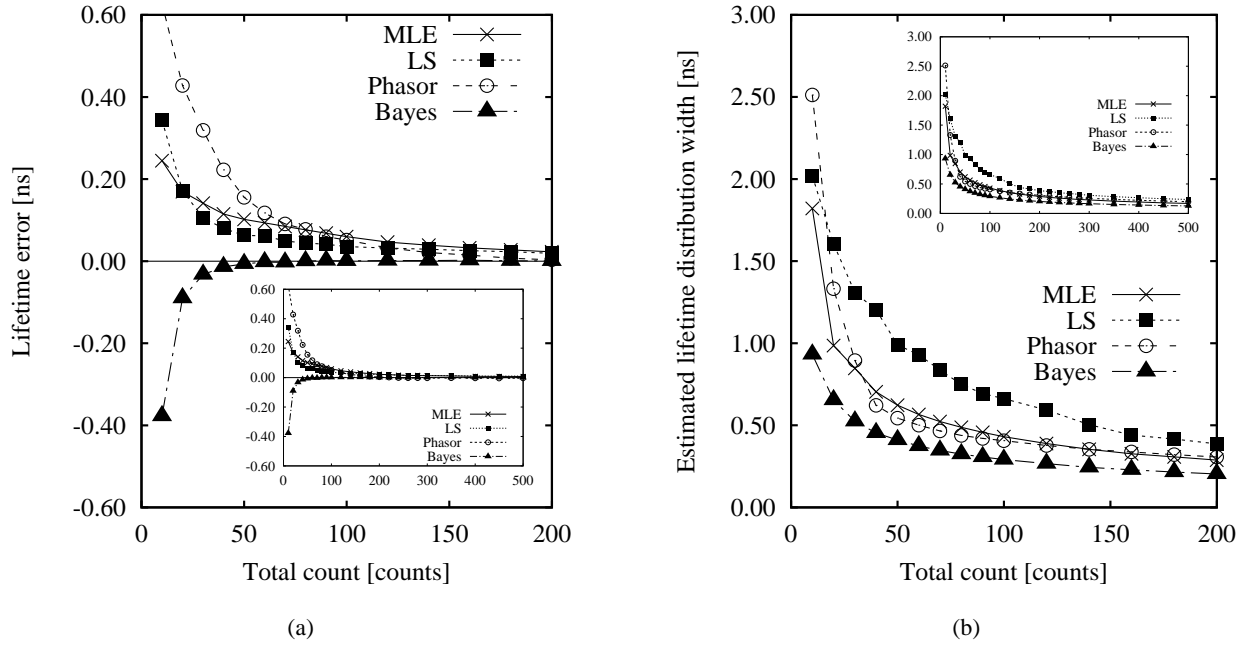
Figure 2. *Increasing average total photon count:* The estimated lifetime distribution statistics for increasing total count, for the different analysis methods applied to synthetic data having a decay lifetime $\tau = 2.0$ ns, generated to yield an average total count of between 10 and 500 counts at each pixel with a background of 10%, and incorporating an instrument response with delay 2.0 ns and FWHM width 0.15 ns. In (a) the lifetime error (deviation of average estimated lifetime value from the true value), and in (b) the standard deviation of the estimated lifetime distribution. The analysis was performed on synthetic 256×256 pixel images such that each data point represents the results from 2^{16} transients.

The influence of background on lifetime estimates is shown in Fig. 3 for an average total count of 50 counts and 1000 counts (insets); observe that the Bayesian analysis lifetime estimates do not deviate significantly from the actual lifetime value even when counts due to background account for about half of the total counts. Notice also that the accuracy of phasor analysis lifetime estimates degrades rapidly in the presence of background counts, most likely a symptom of a background offset being determined at each pixel from the (very limited) data that precedes the rise of the transient.

The sensitivity of our Bayesian technique is highlighted once more in Fig. 4, the lifetime estimates of images comprising three panels, with each panel separated in lifetime by 0.2 ns from its neighbor, are compared. At an average total count of $p = 100$ counts (Fig. 4 (a, e)) the three vertical panels are very nearly discernable from each other by eye on inspecting the Bayesian analysis lifetime image, although the estimated lifetime distribution does not reflect such structure. At the higher average total count of $p = 1000$ (Fig. 4 (b, e)), the panels are clearly discernable by eye on inspection of the estimated lifetime image for each of the methods, and additionally the Bayesian analysis estimated lifetime distribution is clearly tri-modal. The contrast metric $(I_{\max} - I_{\min})/(I_{\max} + I_{\min})$ is used to quantifying the ease with which the modes of the lifetime distribution are resolved, where $I_{\max}$ is the frequency at the dominant mode and $I_{\min}$ is the lowest frequency between the modes. The contrast achieved with an average total count using our Bayesian analysis is 0.26 between the
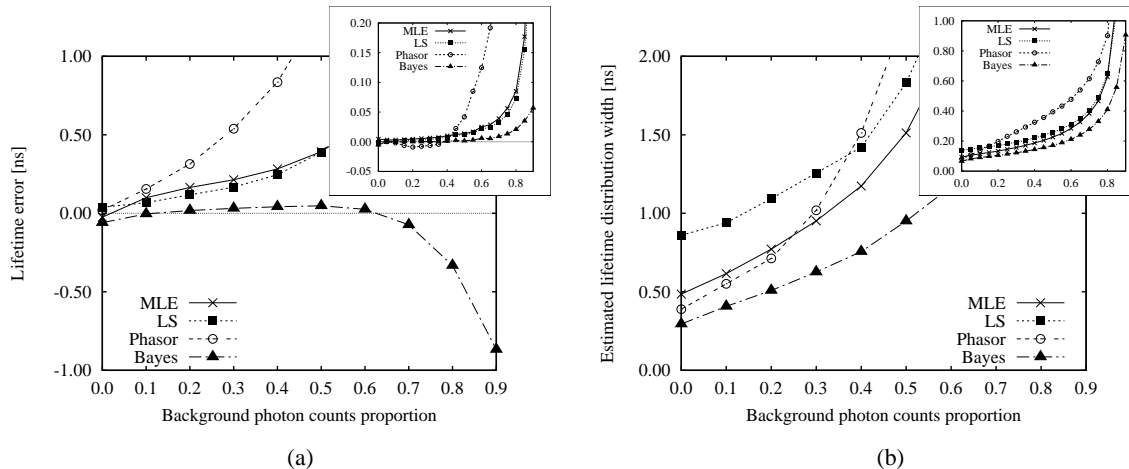
Figure 3. *Influence of background on lifetime estimates:* Statistics of the estimated lifetime distribution for increasing background proportion for an average total photon count of 50 counts and 1000 counts (insets), for the different analysis methods. In (a) the lifetime error (deviation of average estimated lifetime value from the true value), and in (b) the standard deviation of the estimated lifetime distribution. In all cases the analysis was performed on synthetic 256×256 pixel images, generated to yield data corresponding to a lifetime $\tau = 2.0$ ns decay at each pixel, and incorporating an instrument response with delay 2.0 ns and FWHM width 0.15 ns.

1.8 ns and 2.0 ns panels and 0.15 between the 2.0 ns and 2.2 ns panels; the reduction in contrast between the panels with greater lifetime is due to increasing accuracy in the estimates with increasing measurement interval. At this total count it is not yet possible to quantify the contrast of the LS estimates as their distribution does not exhibit prominent separate modes. On increasing the average total count further it becomes possible to quantify the contrast between adjacent panels, for the sake of comparison, of all methods; at an average total count of 2500 (Fig. 4 (c, f)), the contrast achieved between the 1.8 ns and 2.0 ns panels using Bayesian analysis is 0.83, a figure that compares favorably with those of MLE, LS and phasor analysis estimates which yield contrast measures of 0.48, 0.18 and 0.61 respectively. It is also noticeable that the width of the estimated lifetime distributions increases with increasing lifetime; a consequence of the ratio of the decay lifetime to the measurement period.[1,5] The lifetime images and distributions in Fig. 4 (d, g) result from the analysis of decays having the same properties as those of Fig. 4 (c, f), though with a modeled repetition rate of 120 MHz rather than 40 MHz.

3.2 Analysis of experimental data

We now present improvements in lifetime estimates made possible by our method on experimental biological data. For this purpose we have used a well-established biological model system using cells which express GFP; the fluorescence emission from this closely follows a mono-exponential decay with a lifetime of ~ 2.2 ns. Inevitably some deviations from this value will be present and the effects of this biological noise are included in the lifetime distributions. A431 human epithelial *carcinoma* cells expressing cdc42-GFP (Sec. 2.5) were imaged for short durations using the system described in Sec. 2.4. Fluorophore bleaching is negligible, under our experimental protocol, at exposure times well beyond those presented here (data not shown). The estimated lifetime *vs.* pixel photon count are shown for each analysis method in Fig. 5 (a). The diffuse nature of the LS cloud reinforces the fact that lifetime estimates should be interpreted with caution. All analysis methods ultimately converge to a lifetime value of around 2.18 ns as the average total photon count increases. The Bayesian approach does however always lead to a tighter distribution (Fig. 5 (d, e)), as evidenced by a cleaner lifetime image (Fig. 5 (b, c)).

4. DISCUSSION

The benefit of our photon-by-photon Bayesian analysis is most apparent at low total counts. This in turn leads to a faster acquisition or an improved spatial resolution or both. In time-resolved fluorescence lifetime work, some form of spatial binning is commonly used to increase the photon counts available and our approach reduces the need for this. It is apparent that 'cleaner' images are obtained with our Bayesian approach on both synthetic and real image data. Furthermore, in practice, some background signal is inevitable either due to detector noise or stray photons entering the imaging chain
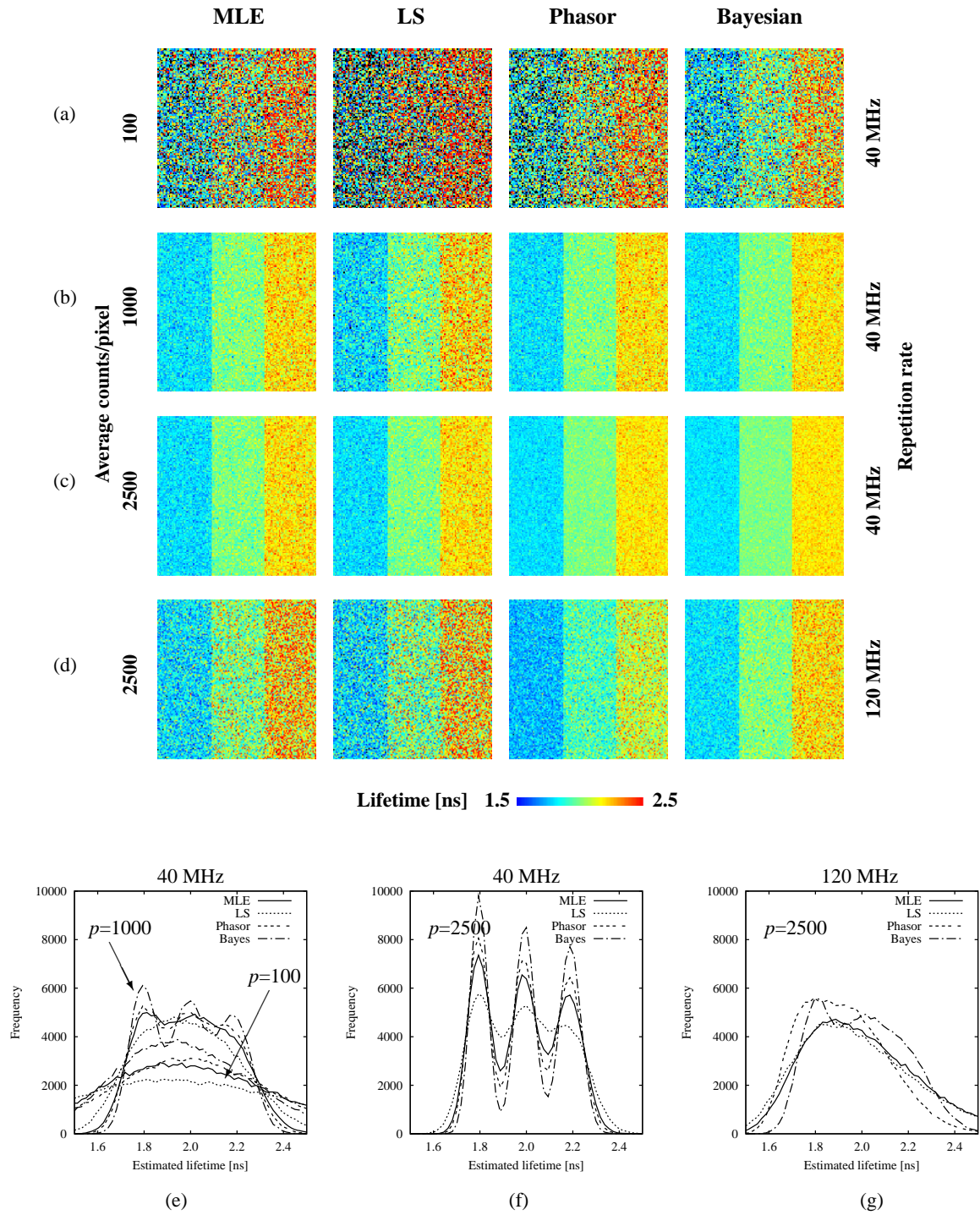
Figure 4. *Lifetime separation:* Estimated lifetime distributions for the different estimation methods applied for the analysis of synthetic images generated with lifetimes $\tau = 1.8$ ns (left panel), $\tau = 2.0$ ns (central panel), $\tau = 2.2$ ns (right panel), a uniform background averaging about 0.02 counts/bin being present in all transients (i.e. background proportion decreases as the photon count is increased). All transients incorporate an instrument response delayed by 10% of the measurement interval and FWHM width of 0.15 ns with (a,b,c) generated to model a repetition rate of 40 MHz and (d) a repetition rate of 120 MHz; in (d) data generated with bin widths of 29 ps and in (a,b,c) with 78 ps bin widths. Average total count of 100 in (a,e), 1000 in (b,e), and 2500 in (c,d,f,g).
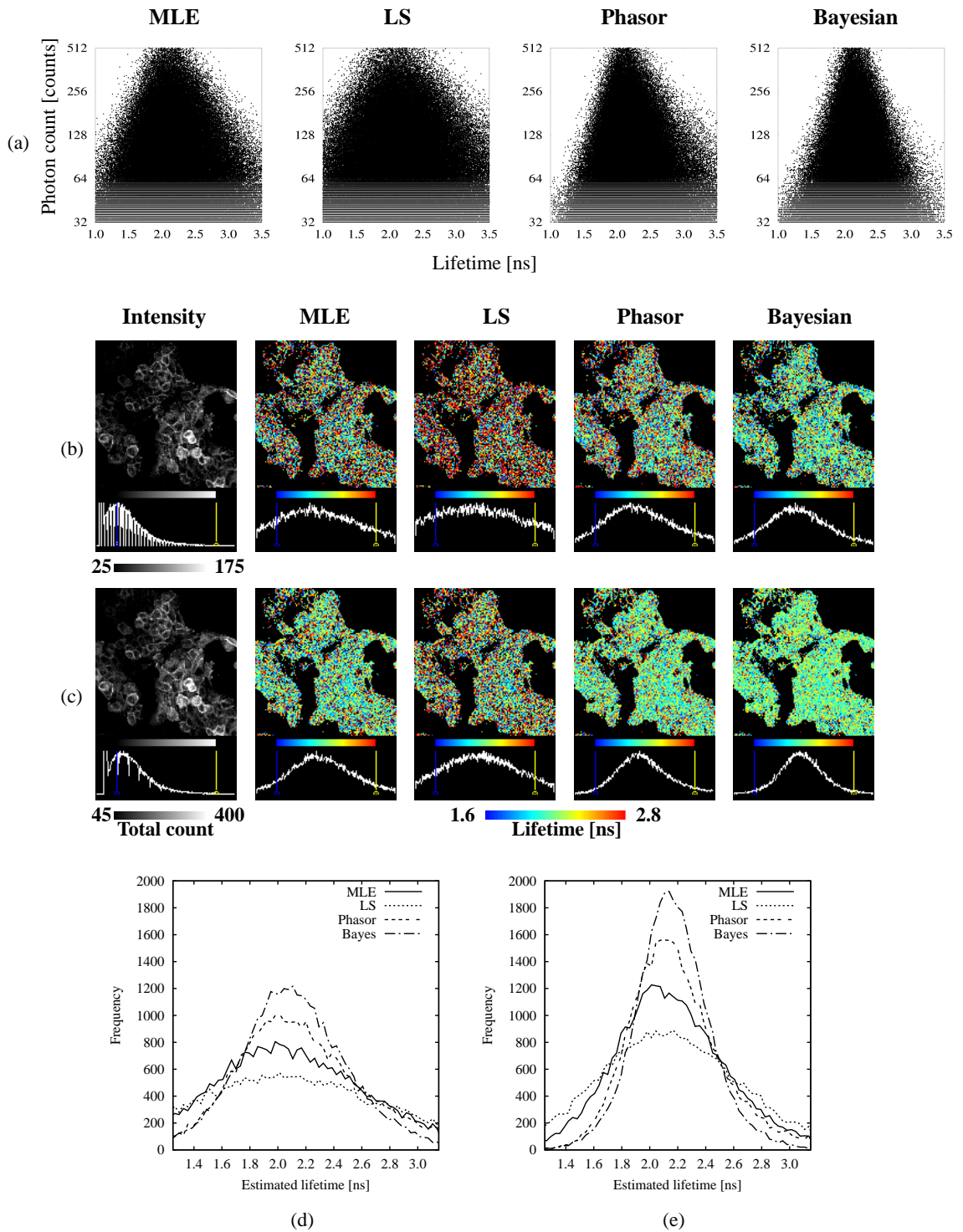
Figure 5. *Mono-exponential analysis of GFP data with low total count:* The estimated lifetime distributions for the different estimation methods applied for the analysis of GFP fluorescence decay data (2.18 ns); in (a) pixel-by-pixel lifetime *vs.* photon count estimates accumulated over ten low count images (per pixel count range between about 20 and 1300), in (b) lifetime images for analysis of data where most of the analyzed pixels having a total count of between 20 and 50 at each analyzed pixel, in (c) the same field is imaged longer such that the majority of analyzed pixels have a total count between 45 and 150 counts, and in (d) and (e) the estimated lifetime distributions corresponding to (b) and (c) respectively. Pixels having a total count of less than 20 in (b) and less than 44 in (c) have been masked and excluded from analysis. The images panels in (b) and (c) correspond to a field of view of $334 \times 334 \ \mu$m.

and our approach accounts for this unwanted signal. Indeed, in some cases, such as during in-vivo imaging, stray light is inevitable as it is in non-microscopy applications (clinical FLIM etc.). The data presented in Fig. 3 suggest that a similar approach may be useful in absorption measurements where a high background is inherent.

Since most biologically useful fluorophore lifetimes are short (< 3.5 ns), the majority of time-resolved imaging instruments operate at repetition rates above 20 MHz. Higher repetition rates however are desirable in order to reduce data collection times, but repetition periods may then not be long enough to allow full decay of the transient. The Bayesian approach, which takes account of the consequences of high repetition rates allows somewhat higher repetition rate laser sources to be used since triggering rates of TCSPC instruments are high (the Becker & Hickl SPC830 may be triggered at up to 200 MHz). This advantage is exemplified in the lower image panels of Fig. 4. Our approach provides more accurate average lifetime values; more conventional approaches bias the results towards a lower lifetime. The distribution is necessarily wider due to the increase in bandwidth and the fact that less information is available since the transient has not decayed fully by the time the next excitation pulse arrives.

At no point during the Bayesian analysis presented in this paper was there a need to consider a specific noise model. We do not try to fit the histogram and nowhere in our analysis is it necessary to compare the fit with the data, and thus the probability of one matching the other, in the form of a χ^2 value, is not required. However, signal corruption by the detector (afterpulsing) and pulse walk may be present and should of course be minimized. Nevertheless these effects vary in proportion to the signal and at least in the mono-exponential case, influence the recorded signal to a small extent. Other corruptions of the recorded data, such as excitation light feedthrough into the detection channel cannot be readily accounted for, but in most cases assuming careful instrument design and use of appropriate filters can be rendered negligible. This is exemplified by the clear improvement in lifetime distributions in real experimental data presented here over those associated with more established methods.

In our modeling of the fluorescence excitation and decay processes (Sec. 2.1) we have assumed that a Gaussian response is a reasonable approximation to the actual excitation response of the detection apparatus. Symmetrical responses, such as the Gaussian, are indeed obtained from detection electronics as used in most modern instruments including ours. However, non-symmetrical instrument responses e.g. with other detection systems which may incorporate some form of low-pass filtering could be handled by suitable modifications to Eqn. (4).

Our Bayesian analysis has been shown to consistently provide lifetime estimates with a greater accuracy than those of MLE, LS, and phasor analysis. Bayesian analysis estimates exhibit a more rapid convergence to the true lifetime as counts are increased and with noticeably less uncertainty than those of the other methods, as shown in Fig. 2. An advantage of the greater accuracy of Bayesian analysis lifetime estimates was highlighted in Fig. 4, it being possible to resolve (either qualitatively on visual inspection or quantitatively from the lifetime distribution) three closely separated lifetimes at a lower count than with the other analysis methods. On analyzing real data, less noisy lifetime maps are produced from the accurate Bayesian analysis estimates with a smaller uncertainty than that of the other methods. A useful feature of our Bayesian analysis is the provision, for each parameter estimate, of a measure of its uncertainty (e.g. Eqn. (16)). The Bayesian estimates of uncertainty have been found to correspond well with the empirically determined spread of lifetime distributions (data not shown). The availability of such reliable uncertainty estimates could have a bearing on the data acquisition process, enabling imaging to be performed for only the duration sufficient for a prescribed degree of uncertainty in lifetime (or other parameter) estimates.

ACKNOWLEDGMENTS

REFERENCES

1. M. Kollner and J. Wolfrum, "How many photons are necessary for fluorescence-lifetime measurements?," *Chemical Physics Letters* **200 (1–2)**, pp. 199–204, 1992.

2. P. Hall and B. Sellinger, "Better estimates of exponential decay parameters," *Journal of Physical Chemistry* **85 (20)**, pp. 2941–2946, 1981.

3. Z. Bajzer, T. M. Therneau, J. C. Sharp, and F. G. Prendergast, "Maximum likelihood method for the analysis of time-resolved fluorescence decay curves," *European Biophysics Journal* **20 (5)**, pp. 247–262, 1991.

4. J. Tellinghuisen, "Bias and precision in the estimation of exponential decay parameters from sparse data," *Analytical Chemistry* **65 (9)**, pp. 1240–1246, 1993.

5. G. Nishimura and M. Tamura, "Artefacts in the analysis of temporal response functions measured by photon counting," *Physics in Medicine and Biology* **50**, pp. 1327–1342, 2005.

6. P. R. Barber, S. M. Ammer-Beg, J. Gilbey, R. J. Edens, I. Ezike, and B. Vojnovic, "Global and pixel kinetic data analysis for FRET detection by multi-photon time-domain FLIM," *Proc. SPIE* **5700**, pp. 171–181, 2005.

7. P. R. Barber, S. M. Ammer-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," *Journal of The Royal Society Interface* **6**, pp. S93–S105, 2009.

8. D. A. Turton, G. D. Reid, and G. S. Beddard, "Accurate analysis of fluorescence decays from single molecules in photon counting experiments," *Analytical Chemistry* **75 (16)**, pp. 4182–4187, 2003.

9. E. N. Brown, Z. Zhang, and A. D. McCollom, "On the uncertainty in single molecule fluorescent lifetime and enery emission measurements," tech. rep., United States. National Aeronautics and Space Administration, 1994.

10. P. R. Barber, S. M. Ameer-Beg, S. Pathmananthan, M. Rowley, and A. C. C. Coolen, "A Bayesian method for single molecule, fluorescence burst analysis," *Biomedical Optics Express* **1**(4), pp. 1148–1158, 2010.

11. W. Becker, *The bh TCSPC Handbook*. Becker & Hickl GmbH, Third ed., 2008.

12. K. Makrogianneli, L. M. Carlin, M. D. Keppler, D. R. Matthews, E. Ofo, A. Coolen, S. M. Ameer-Beg, P. R. Barber, B. Vojnovic, and T. Ng, "Integrating Receptor Signal Inputs That Influence Small Rho GTPase Activation Dynamics at the Immunological Synapse," *Molecular and cellular biology* **29**, pp. 2997–3006, 2009.

13. T. A. Laurence and B. A. Chromy, "Efficient maximum likelihood estimator fitting of histograms," *Nature Methods* **7**(5), pp. 338–339, 2010.

14. C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

15. A. H. A. Clayton, Q. S. Hanley, and P. J. Verveer, "Graphical representation and multicomponent analysis of single-frequency fluorescence lifetime imaging microscopy data," *Journal of Microscopy* **213**, pp. 1–5, 2004.

16. A. Leray, C. Spriet, D. Trinel, and L. Héliot, "Three-Dimensional Polar Representation for Multispectral Fluorescence Lifetime Imaging Microscopy," *Cytometry Part A* **75A**, pp. 1007–1014, 2009.

17. H. C. Gerritsen, M. A. H. Asselbergs, A. V. Agronskaia, and W. G. J. H. M. V. Sark, "Fluorescence lifetime imaging in scanning microscopes: acquisition speed, photon economy and lifetime resolution," *Journal of Microscopy* **206**, pp. 218–224, 2002.