pawFLIM: reducing bias and uncertainty to enable lower photon count in FLIM experiments

2017 Methods Appl. Fluoresc. 5 024016

(http://iopscience.iop.org/2050-6120/5/2/024016)

Methods and Applications in Fluorescence

PAPER

pawFLIM: reducing bias and uncertainty to enable lower photon count in FLIM experiments

RECEIVED
4 November 2016

REVISED
21 April 2017

ACCEPTED FOR PUBLICATION
12 May 2017

PUBLISHED
26 June 2017

Mauro Silberberg and Hernán E Grecco

Departamento de Física, FCEyN, UBA and IFIBA, CONICET, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina

E-mail: hgrecco@df.uba.ar

Keywords: denoising, Förster resonant energy transfer (FRET), fluorescence lifetime imaging microscopy (FLIM), wavelets, protein-protein interaction

Supplementary material for this article is available online

Abstract

Förster resonant energy transfer measured by fluorescence lifetime imaging microscopy (FRET-FLIM) is the method of choice for monitoring the spatio-temporal dynamics of protein interactions in living cells. To obtain an accurate estimate of the molecular fraction of interacting proteins requires a large number of photons, which usually precludes the observation of a fast process, particularly with time correlated single photon counting (TCSPC) based FLIM. In this work, we propose a novel method named pawFLIM (phasor analysis via wavelets) that allows the denoising of FLIM datasets by adaptively and selectively adjusting the desired compromise between spatial and molecular resolution. The method operates by applying a weighted translational-invariant Haar-wavelet transform denoising algorithm to phasor images. This results in significantly less bias and mean square error than other existing methods. We also present a new lifetime estimator (named normal lifetime) with a smaller mean squared error and overall bias as compared to frequency domain phase and modulation lifetimes. Overall, we present an approach that will enable the observation of the dynamics of biological processes at the molecular level with better temporal and spatial resolution.

1. Introduction

Through vision, light has always been the preferred phenomenon to learn from nature. Telescopes and microscopes expanded the limit of our senses, allowing us to discover the immensity of the Universe and the micro-cosmos that lies in every cell. Since such developments, using light to interrogate and manipulate the world has been a cornerstone of modern technology, from medical imaging to data storage, from optical communications to harnessing the energy of the Sun. At the core of all these examples is our understanding of the interaction between light and matter.

Fluorescence microscopy, empowered by the development of the confocal microscope (Minsky 1961, Conchello and Lichtman 2005) and the retooling of fluorescent proteins (Shimomura *et al* 1962, Tsien 1998), revolutionized cell biology as it allowed us to observe the localization and dynamics of proteins in living cells (Amos and White 2003). For example, in the study of mitosis we have come a long way from the outstanding work of the German biologist W Flemming, who used handmade drawings in the XIXth century (Paweletz 2001), to the recent super-resolution movies revealing membrane dynamics during the process (Aguet *et al* 2016).

But fluorescence is more than just an excellent way to achieve contrast in microscopy. Molecules are sensitive to their environment, a feature (not a bug) which can be exploited as a reporter. Fluorescence lifetime, a measure of the time a molecule spends in the excited state before returning to the ground state by emitting a photon, is a particularly good readout of the environment, as was first demonstrated by Enrique Gaviola. While working in Berlin during the 1920's (Bernaola 2001), the Argentinean physicist accurately measured for the first time the fluorescence lifetimeof fluorescein and rhodamine in solution, both in the nanosecond range. He also demonstrated that it was affected by temperature, solvent and viscosity (Gaviola 1926). He achieved this remarkable

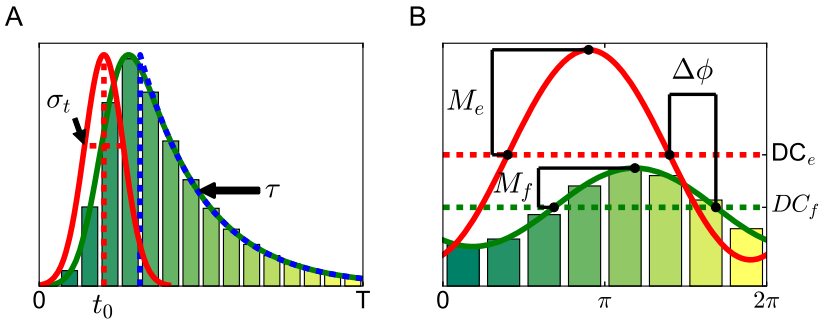

Figure 1. Depiction of time and frequency domain FLIM techniques. (a) In the time domain, the sample is illuminated with a train of pulses with period T. Each pulse (red) excites the fluorophore and its emission (green) is measured as a function of time. The detected emission is the convolution of the fluorescence emission (in this case, an exponential decay with parameter τ) (blue) and the pulse (in this case, a Gaussian distribution with parameters t_0 and σ_t). (b) In the frequency domain, the sample is sinusoidally excited (red), and the emission (green) measured with a phase sensitive detection scheme. The parameters of the exponential decay can be obtained from the phase shift ($\Delta\phi$) and the demodulation between excitation and fluorescence emission (M_f/M_e). DC_e and DC_f correspond to the average intensities of the excitation and fluorescence emission, respectively.

observation by building a phase fluorometer (we also owe him this word) with a sub-nanosecond resolution using Kerr cells (Gaviola and Pringsheim 1927).

Since the pioneering work by Gaviola, it has been found that fluorescence lifetime is affected by several additional factors. Among these is the presence of other molecules in close vicinity due to a dipole–dipole interaction known as Förster resonant energy transfer (FRET) (Förster 1948). FRET has been of major importance in understanding how function emerges from protein networks. Indeed, biological function arises from the interaction of nanometer-sized molecules which operate in a context dependent manner (Barabasi and Oltvai 2004, Grecco *et al* 2011). Finding those interactions requires co-localizing proteins in a volume which is seven orders of magnitude smaller than the resolvable volume in standard fluorescence microscopy (Jares-Erijman and Jovin 2003, Grecco and Bastiaens 2008, Margineanu *et al* 2016). As an analogy, it is like assessing whether two people are holding hands in a hotel which occupies an area of 100 m by 100 m and is 100 stories tall. FRET allows us to assess interaction without providing any information about which room they are in.

Methods for measuring lifetime are generally divided into time domain and frequency domain techniques (Lakowicz 2006). In the former, a pulsed excitation is used and the emission is collected in a time resolved manner (figure 1(a)). In the latter, a sinusoidal excitation combined with a phase sensitive detection of the emission is used (figure 1(b)). Fluorescence lifetime imaging microscopy (FLIM) allows us to obtain spatially resolved maps of decay rates in fluorescent samples by combining such techniques with a wide-field or confocal microscopy.

FRET measured by FLIM (FRET-FLIM) is the technique of choice for quantifying the interaction dynamics of proteins in living systems (Wouters *et al* 2001). However, accurate quantification of the fraction of interacting proteins requires a compromise. The number of photons per image is limited either by the acquisition device, the brightness of the sample or by the photostability of the fluorophore. In certain cases, the photon count can be increased by integrating over a longer time or in a larger space but consequently yields a lower temporal or spatial resolution. The same is true for any other FLIM based quantification, such as the fluorescence lifetime of an environmentally sensitive probe or an intramolecular activity FRET sensor.

Until recently, little attention had been given to applying image analysis techniques that take advantage of the spatial correlations present in FLIM datasets. Most work has been focused on accurately extracting the apparent fluorescence lifetimes from decay curves (i.e. single-pixel analysis). One of the first approaches to benefit from spatial information was global analysis, which assumes that some parameters are spatially fixed (usually lifetimes) and simultaneously fits every pixel with that restriction (Beechem *et al* 1983, Knutson *et al* 1983, Verveer *et al* 2000, Barber *et al* 2005). However, image analysis techniques allow us to adaptively exploit local spatial information. They are not only valuable for improving intensity images, but can also assist in the lifetime analysis of FLIM images (Buranachai *et al* 2008, Spring and Clegg 2009, Chang and Mycek 2009, Lin 2012, Roudot *et al* 2013).

In this work, we propose a novel method named pawFLIM (phasor analysis via wavelets of FLIM) that allows the denoising of FLIM datasets by adaptively adjusting the desired compromise between spatial and molecular resolution. After presenting the theory and previous work on this topic, we show through numerical simulations that pawFLIM outperforms existing methods, resulting in a reduced bias and mean square error. As it operates at the phasor space, it also improves the determination global lifetimes, relative intensities and molecular fractions. It is applicable to both time- and frequency-domain data and not

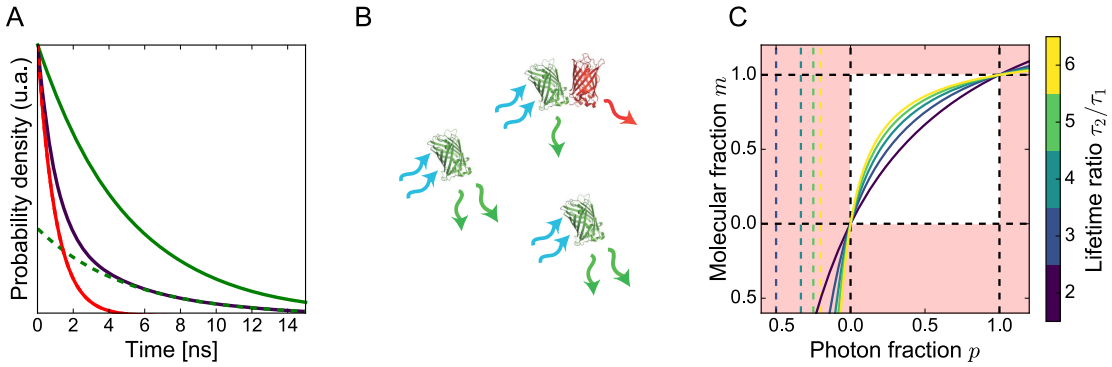

Figure 2. From photons to molecules in FRET-FLIM. (a) Monoexponential decays (normalized to $t = 0$) of *long* (green) and *short* (red) lifetimes; the unnormalized long decay curve is also shown (dashed green). In a FRET-FLIM experiment with two distinct lifetime populations, these correspond to populations of donor alone and donor in complex with acceptor. A biexponential mixture of them (violet) results when molecules in both states are found in the same pixel. (b) From the biexponential decay, the fraction of photons (p) detected from donor in complex can be derived. Since FRET quenches the donor emission (green arrows), the molecular fraction (m) will differ from p. As an example, for these three molecules $p = 1/5$ but the fraction of donor in complex molecules is larger ($m = 1/3$). (c) Photon to molecular fraction conversion, for different lifetime ratios (color coded). Although values outside 0–1 (limited by dashed black lines) are unphysical (red-shaded region), they can't be omitted in the analysis as this will introduce a bias in most statistical estimators. For $p < 0$, this functional relation has an asymptotic value (vertical colored dashed lines corresponding to the different lifetime ratios).

limited to binary systems as it can improve the estimation of average lifetimes in any complex mixture. We demonstrate the performance on a real image by applying pawFLIM to a FRET-FLIM experimental dataset.

2. Theory

2.1. Fluorescence lifetime

Fluorescence emission can be modeled as a first order kinetic process (Pelet *et al* 2004). The solution is usually described as a continuous intensity decay, but a single molecule emits one photon at a time. From a probabilistic point of view, the photon emission time from a fluorophore excited at $t = 0$ follows an exponential distribution of parameter τ, whose probability density function (PDF) is:

$$D(t|\tau) = \begin{cases} 0 & \text{if } t < 0 \\ \dfrac{\exp(-t/\tau)}{\tau} & \text{if } t \geqslant 0 \end{cases} \tag{1}$$

where τ corresponds to the fluorophore's lifetime.

Originally, FLIM was demonstrated for single lifetime estimates. However, a mixture of different molecules or molecular species in different biochemical states may be present in each pixel. In such cases, its decay kinetics are not accurately described by a single exponential. Therefore, a single lifetime estimation is only a semiquantitative estimate and, instead, the lifetimes and populations of each species should be resolved (Verveer *et al* 2000).

If several fluorophores are considered, the weighted sum of its individual PDF for the photon emission time yields the overall PDF:

$$D(t|p_i, \tau_i) = \sum_i \frac{p_i}{\tau_i} \exp(-t/\tau_i) \tag{2}$$

where $\sum_i p_i = 1$. This results in a mixture distribution, where p_i is the probability that the emitted photon comes from the fluorophore with lifetime τ_i (figure 2(a)). That is, if N photons are detected, Np_i are expected to come from fluorophore i.

It is important to differentiate between p_i, the photon fraction, and m_i, the molecular fraction (figure 2(b)). The photon emission of each molecular species is proportional to its detected brightness, an expression to encompass from fluorophore properties such as quantum yield to experimental setup conditions such as detector response. To obtain the molecular fraction, each photon fraction has to be divided by its corresponding detected brightness, and then the sum renormalized to one (Verveer *et al* 2000).

In most FRET-FLIM experiments only the donor emission is observed. Therefore any change in brightness is due to a reduction of its quantum yield when a complex with an acceptor is formed. Assuming that only two states are present in the sample, just the ratio between their respective quantum yields is needed to transform from photon (p) to molecular (m) fraction. For FRET, this is equivalent to the lifetime ratio (Lakowicz 2006) and therefore:

$$m = \frac{1}{1 + \dfrac{\tau_D}{\tau_{FRET}} \dfrac{(1-p)}{p}} \tag{3}$$

where we have defined $m_{FRET} = m$ and $m_D = 1 - m$. It's important to note that, since this is a nonlinear relationship (figure 2(c)), it is not equivalent to average in p or m.

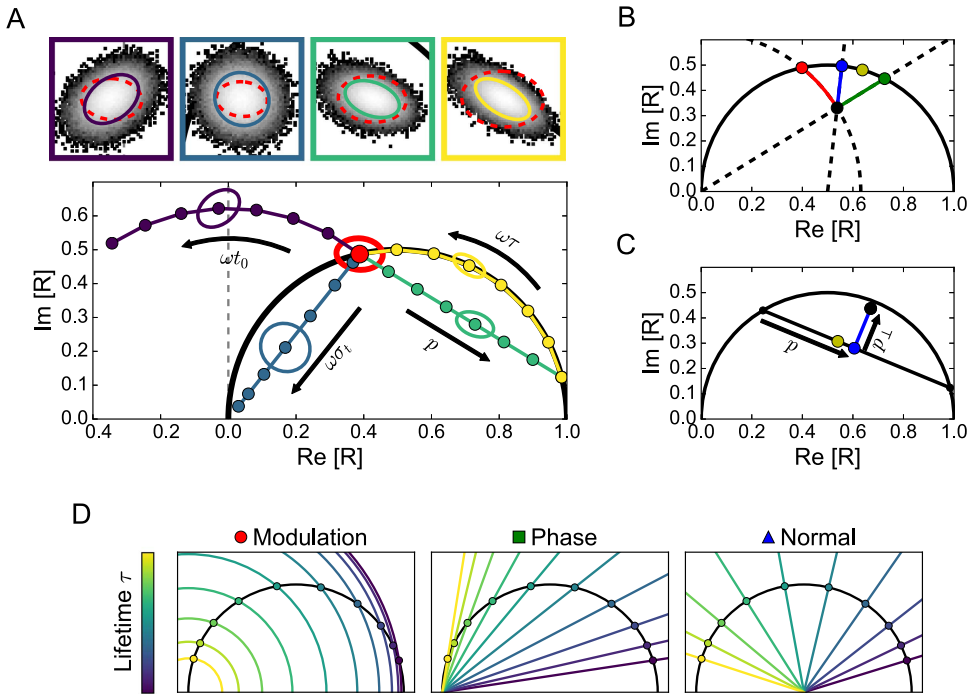

Figure 3. FLIM analysis using the phasor approach. (a) A phasor corresponding to a monoexponential decay moves counterclockwise along the semicircle as $\omega\tau$ increases (yellow). For biexponential decays, it lies along a line joining the monoexponential phasors (green). We define $p = 0$ at the red extreme, corresponding to the long lifetime. Due to the IRF, it moves towards (blue) or rotates around (violet) the origin, where the parameters correspond to a Gaussian IRF of center t_0 and width σ_t. To demonstrate the effect of noise on each line, the phasor distribution is shown above in four boxes labeled with the same color. The colored ellipses represent the 60% confidence area. Notice that, compared to the phasor distribution in the semicircle (red dashed line), the others are skewed. (b) Geometrical interpretation of lifetime estimators for a phasor (yellow) to which noise has been added (black). Red and green correspond to historically used modulation and phase lifetimes, respectively. Blue corresponds to newly introduced normal lifetime (see text). (c) The photon fraction p is estimated from the distance of the scalar projection (blue) of a phasor (yellow) to which noise has been added (black) over the biexponential line to the long lifetime phasor, normalized by the length of the biexponential line. Also shown is the perpendicular distance to the line, $p_\perp$, a useful measure for global analysis. (d) Contour curves of single lifetime estimators, chosen to be logarithmically spaced to consider the relative error in lifetime.

2.2. Phasor approach: analyzing FLIM datasets in the complex plane

The phasor approach consists of analyzing the data in the frequency domain (Clayton *et al* 2004, Digman *et al* 2008, Grecco *et al* 2009). A phasor R is defined as the Fourier transform of the normalized decay, which for a single exponential decay, (1), is:

$$R = \int_{-\infty}^{\infty} D(t)\, \mathrm{e}^{\mathrm{i}\omega t}\, \mathrm{d}t = \frac{1}{1 - \mathrm{i}\omega\tau}. \qquad (4)$$

This value moves counterclockwise in the complex plane, along a semicircle, as $\omega\tau$ increases (figure 3(a), yellow line).

From a given phasor R located on the semicircle, the corresponding lifetime can be obtained. Nevertheless, a measured phasor is usually located off the semicircle due to noise. Therefore it is important to choose an appropriate estimator. For historical reasons, the most common estimators are the modulation and phase lifetimes, respectively obtained from the modulus and phase of (4). We propose an alternative estimator, the normal lifetime, which consists of projecting the measured phasor to the closest point on the semicircle and then obtaining τ from (4).

These lifetime estimators can be geometrically interpreted. In the case of the modulation lifetime, each measured modulation gets assigned to a lifetime or, equivalently, a phasor in the semicircle. But there are several possible phasors which share the same modulation. All phasors lying in a circle around the origin share the same modulus and, hence, get assigned the same lifetime. Therefore, this lifetime can be thought of as projecting a measured phasor to the monoexponential semicircle along a circle centered in the origin (figure 3(b), red line). Analogously, the phase modulation can be thought of as projecting along a line from the origin, since these are the points that share the same phase (figure 3(b), green line). Our new estimator assigns the closest point in the semicircle, which is a normal projection, hence its name (figure 3(b), blue line).

As can be seen from this geometrical representation (figure 3(d)), circumferences corresponding to short lifetimes are crammed together, as are phase lines corresponding to long lifetimes. Therefore, a small error in the phasor gets translated to a large error in lifetime. Instead, the normal estimator is more evenly distributed.

For biexponential decays, the phasor moves along a line joining the monoexponential phasors R_1 and R_2 (figure 3(a), green line):

$$R = p\, R_1 + (1 - p)\, R_2. \tag{5}$$

If these are known, we can solve (5) for p. However, in the presence of noise, a phasor R won't necessarily lie on this segment. The usual estimator is the normalized scalar projection to the biexponential line, which is equivalent to a least squares minimization (Grecco *et al* 2009). A noisy phasor is projected perpendicularly to the biexponential line and, then, p is estimated as the distance between that point and R_2, normalized by the length of the segment defined by R_1 and R_2 (figure 3(c), and supplementary information 1 and 2 available online at stacks.iop.org/MAF/5/024016/mmedia).

2.3. Effect of the instrument response function (IRF)

In the previous sections we presented the results for a delta-like pulse illuminating the sample at $t = 0$ and with electronics with an infinite bandwith. However, since the excitation pulse is broad, not all molecules are excited simultaneously but rather at a random time within the pulse width. The detector introduces an additional random time offset for each photon (see supplementary information 3). Both effects, which offset the emission time, can be grouped in a single variable described by a distribution called the instrument response function (IRF).

Therefore, the expected fluorescence F can be expressed as a convolution of the fluorescence decay D with the instrument response function *IRF*:

$$\begin{aligned} F(t|\tau,\, T) &= D(t|\tau) \,*\, \mathrm{IRF}(t|T) \\ &= \int_{-\infty}^{\infty} D(u)\, \mathrm{IRF}(t - u|T)\, \mathrm{d}u \end{aligned} \tag{6}$$

where T is the excitation period. To exemplify the effect of the IRF on the phasor we will consider a Gaussian function centered at t_0 and of σ_t width.

$$\mathrm{IRF}(t|t_0,\, \sigma_t) = \frac{1}{\sqrt{2\pi}\ \sigma_t} \exp\left(-\frac{(t - t_0)^2}{2\,\sigma_t^2}\right). \tag{7}$$

As the Fourier transform of a convolution is the product of the transforms:

$$\begin{aligned} \widehat{F}\,(\omega|T) &= \widehat{D}\,(\omega) \times \widehat{\mathrm{IRF}}(\omega|T) \\ &= \begin{cases} \widehat{D}\,(\omega)\, \widehat{\mathrm{IRF}}(\omega|T), & \text{if } \omega = \dfrac{2\pi}{T} n \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{8}$$

where the only remaining harmonics are multiples of the fundamental frequency $\omega_0 = \frac{2\pi}{T}$ of the train of excitation pulses. For the first harmonic, convolving with a monoexponential decay, we have:

$$\begin{aligned} \widehat{F}\,(\omega) &= \widehat{D}\,(\omega)\, \widehat{\mathrm{IRF}}(\omega|t_0,\, \sigma_t) \\ &= \frac{1}{1 - \mathrm{i}\omega\tau}\ \mathrm{e}^{-\frac{1}{2}(\omega\sigma_t)^2} \mathrm{e}^{\mathrm{i}\omega t_0}. \end{aligned} \tag{9}$$

In the phasor plot, the center parameter t_0 produces a rotation around the origin, while the width parameter σ_t produces a contraction towards the origin (figure 3(a), violet and blue curves, respectively). To estimate the distribution corresponding to the *unconvolved* decay D, the IRF has to be measured (or inferred). Deconvolving in the time domain is dividing in the frequency domain, which is an ill-conditioned problem as higher harmonics tend to be dominated by noise. However, in most FLIM experiments there is a separation of temporal scales between the fluorescence decay (long) and the IRF (short). Therefore, an accurate analysis of the data can be achieved using only the lower harmonics of the data as they contain most of the information (Grecco *et al* 2009). In particular, as has been shown in global analysis, only the first harmonic is necessary to resolve a bi-exponential decay in a heterogeneous population (Verveer and Bastiaens 2003).

2.4. Average and covariance matrix of phasor distributions

The equations presented in the previous sections describe the behaviour of the noise-free value, but phasor distributions do not move rigidly (figure 3(a), red ellipse versus ellipses in blue, green and yellow boxes). To analyze further the effect of noise and the IRF, it is convenient to describe the phasor distribution rather than the noise-free phasor value.

In TCSPC, each detected photon is assigned to a histogram bin at time t_k of width Δt, with probability $p_k = \int_{t_k}^{t_k+\Delta t} F(t|\tau)\, \mathrm{d}t$. If N photons have been detected, the number of photons X_k in channel k is a multinomial random variable. Then, the corresponding phasor can be calculated as:

$$\begin{aligned} R_n &= \frac{1}{N} \sum_k^N X_k\, \mathrm{e}^{\mathrm{i}n\omega t_k} = \mathrm{Re}[R_n] + \mathrm{i}\, \mathrm{Im}[R_n] \\ &= B_n + \mathrm{i}\, A_n \end{aligned} \tag{10}$$

where n is the number of harmonics considered. A straightforward calculation applying the properties of the expectation E[.] and covariance Cov[. , .] operators gives:

$$\mathrm{E}[R_n] = \frac{1}{N} \sum_k^N \mathrm{E}[X_k]\, \mathrm{e}^{\mathrm{i}n\omega t_k} = \sum_k^N p_k\, \mathrm{e}^{\mathrm{i}n\omega t_k} = b_n + \mathrm{i}\, a_n. \tag{11}$$

For a high number of bins (or equivalently small bin width Δt), equation (11) becomes (4), and the phasor estimator is unbiased. Instead, if the bin width is not negligible, a correction has to be considered (Fereidouni *et al* 2011).

Since a phasor is a complex value, it can be considered as a two-dimensional vector. Its associated covariance matrix can be used as a measure for the shape of the corresponding distribution.

$$\begin{aligned} \mathrm{Cov}[R_n] &= \begin{pmatrix} \mathrm{Var}[B_n] & \mathrm{Cov}[B_n,\, A_n] \\ \mathrm{Cov}[B_n,\, A_n] & \mathrm{Var}[A_n] \end{pmatrix} \\ &= \frac{1}{2N} \begin{pmatrix} 1 + b_{2n} - 2b_n^2 & a_{2n} - 2b_n a_n \\ a_{2n} - 2b_n a_n & 1 - b_{2n} - 2a_n^2 \end{pmatrix}. \end{aligned} \tag{12}$$

As mentioned before, to correct (deconvolve) for the effect of the IRF it is necessary to divide by the corresponding harmonic $R_n^{IRF} = |R_n^{IRF}|\, e^{i\phi_n^{IRF}}$. This propagates to the covariance matrix as a division by $|R_n^{IRF}|^2$ and a (matrix) rotation by $-\phi_n^{IRF}$.

As a measure of the similarity of the *unconvolved* distribution with the *deconvolved* one, we calculate the difference in the trace of their covariances. Applying properties of the variance and expectation value we arrive at:

$$\Delta = \mathrm{tr}\left(\mathrm{Cov}\left[\frac{R}{R_{IRF}}\right]\right) - \mathrm{tr}(\mathrm{Cov}[R_{Y=0}]) = \frac{1}{N}\frac{1-|R_{IRF}|^2}{|R_{IRF}|^2} \geqslant 0 \quad (13)$$

where R/R_{IRF} and $R_{Y=0}$ correspond to the *deconvolved* and *unconvolved* distribution. It is evident that the traces, and therefore the covariance matrices, are different if $|R_{IRF}|^2 \neq 1$. This indicates that when the IRF is not a delta-like function, it introduces not only a rotation and shift of the phasor distribution but also changes its shape. This can also be seen in simulations (supplementary figure 4). Therefore, dividing the measured phasor distribution by the IRF coefficient will rotate and shift (more correctly, contract) the distribution to the *unconvolved* location in the phasor plot, resizing but not reshaping to the *unconvolved* form.

2.5. FLIM workflow

In full-field homodyne FD-FLIM, a high-frequency repetitive voltage signal modulates the gain of the image intensifier at the same frequency as the excitation light modulation. A stack of FLIM intensity images is acquired by phase shifting the modulation of the detector gain with respect to the modulation of the excitation light. Every phase image is integrated on a CCD camera for a time to acquire sufficient signal-to-noise ratio (Spring and Clegg 2009).

The analysis of FLIM images proceeds from this stack of phase-delayed images from which a phasor image can be obtained. Time-gated and TCSPC detection generate a completely different type of dataset but can also be transformed to a phasor image (figure 4). From such a phasor image, either a fluorescence lifetime image can be directly calculated or, for binary systems, we can apply global analysis to arrive at the parameter of interest, the molecular fraction *m*.

With the exception of global analysis, the previous sections have been concerned with single pixel estimators. Global analysis does use multiple pixel information but in an uncorrelated manner, i.e. no information about pixel location is used. As pointed out in the introduction, we developed a denoising algorithm applied in the phasor image that exploits spatial correlations to improve estimations locally. Nevertheless, image analysis techniques could be applied to several instances along the FLIM analysis workflow: intensity, phasor, and lifetime or photon/molecular fraction images. However, a procedure that utilizes local averages will perform differently in each of these spaces.

While we have previously shown that the phasor estimator is unbiased, lifetime estimators are nonlinear functions of it and, therefore, biased. Hence, to estimate an average lifetime it is better to take averages upstream in phasor space than directly in lifetime space.

Analogously, the molecular fraction will also be biased since it is a nonlinear function of the photon fraction *p* (equation (3)), which in turn is unbiased, being a linear function of the phasor. Hence, it is preferable to average in *p* and then transform to *m* rather than take averages in *m* directly. Additionally, due to noise, both positive and negative values can be found for *p* when $p_{\mathrm{true}} \sim 0$. Negatives must be kept as the average would be positively biased otherwise. But *m* diverges, and therefore its average, as *p* approaches $(1 - \tau_{\mathrm{FRET}}/\tau_D)^{-1} < 0$ (in FRET-FLIM). Again, this is not a problem if averaging in *p* rather than *m*.

Finally, while averaging either in phasor or photon fraction space leads to the same result, the information to decide whether to average or not is different. In phasor space we have two components (real and imaginary parts) and we can leverage correlations between them, while the photon fraction is only one component, which is akin to discarding the imaginary part after a suitable transformation (see supplementary information 1).

Moreover, as the phasor image is upstream of any subsequent calculation it can benefit the analysis independently of the sample (single fluorophores, non-binary systems, binary systems). For instance, averaging in phasor space impacts the determination of global lifetimes through global analysis, which, in turn, impacts the determination of the photon fraction. As any dataset can be converted to a phasor image, any posterior analysis will benefit from a denoising applied in this instance, independently of the acquisition method (FD-FLIM, TD-FLIM).

2.6. Using wavelets for image denoising

There are diverse image analysis techniques to denoise an image. Among the simpler ones are linear filtering techniques, which are specified by their transfer function, as is a Gaussian filter. Others choose an appropriate representation where information is concentrated in few coefficients, i.e. it is sparse. Assuming the small coefficients correspond to noise, they are discarded and an inverse transform applied to recover an estimation of the original image. This procedure is called thresholding or shrinkage. These techniques are classified within nonlinear filtering methods, which can vastly outperform optimal linear procedures (Mallat 1999).

One of the most used representations is the frequency (or Fourier) basis. As the signal is represented

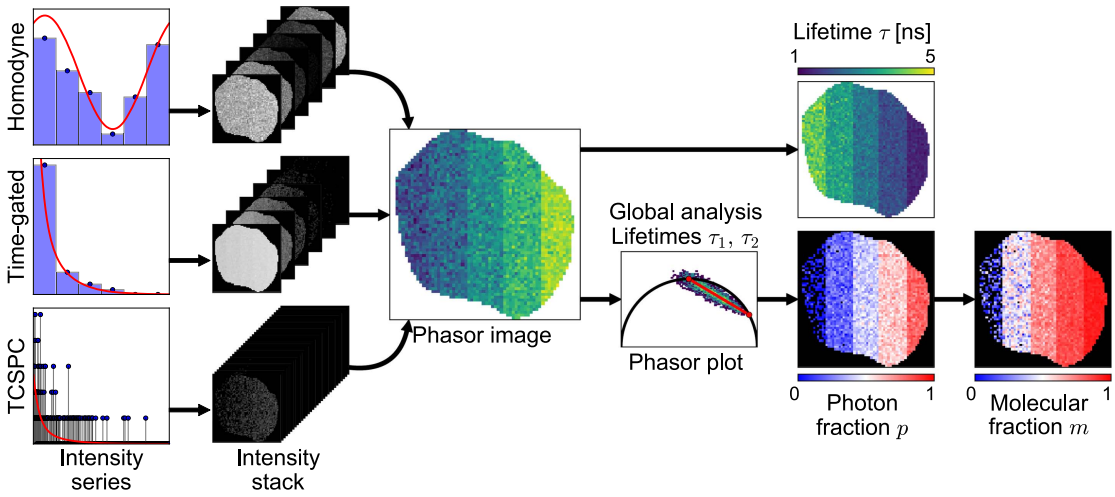

Figure 4. Phasor-approach FLIM analysis workflow. Independently of the experimental method, a phasor can be calculated for each spatial pixel from its intensity series, yielding a phasor image. Then, either an apparent lifetime can be estimated directly for each pixel or, after global analysis, a photon fraction. In turn, the photon fraction can be converted to molecular fraction normalizing by the quantum yield of each species (or lifetimes, in the case of FRET). A cell-shaped mask was placed over simulated data for illustrative purposes.

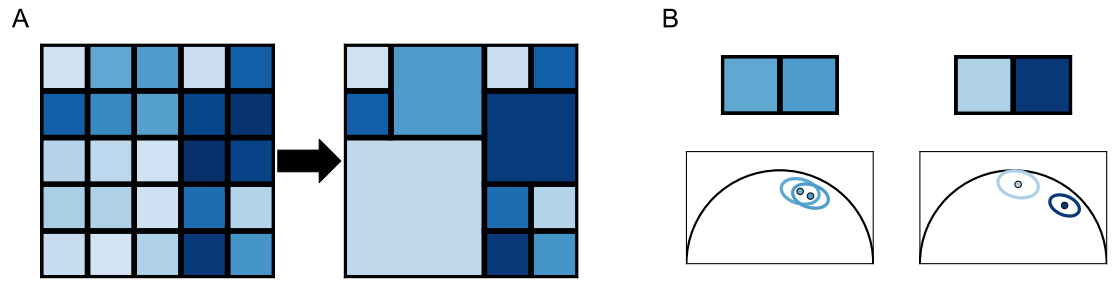

Figure 5. Spatial averaging at the phasor images. (a) We can adaptively trade spatial resolution for accuracy by averaging similar pixels. Note that image denoising methods such as wavelet thresholding may average non-locally (see supplementary figure 5). (b) A statistical test on the phasor similarity is performed to decide whether two adjacent pixels should be averaged. The desired statistical significance is a parameter that can be chosen to adjust the trade off between spatial and molecular resolution. For example, for a given significance level, the left case will be averaged but not the right one.

as a sum of sinusoidals, this basis is most adequate for periodic signals. When local features or sharp discontinuities are present, a considerable number of coefficients are necessary to accurately reconstruct the signal since they are determined globally. Several space-frequency representations, such as the wavelet transform, were developed to overcome this problem. Wavelets are useful to describe signals with local features due to their space-frequency localization. They provide a good spatial resolution for high frequency events, and a good frequency resolution while sacrificing spatial resolution for low frequency events.

From a statistical point of view, wavelet thresholding is related to hypothesis testing (Abramovich and Benjamini 1996). It can be shown that for normal random variables, their wavelet coefficients are also normal random variables. A p-value can be calculated for each coefficient and discarded according to a preset significance level (SL). On a wavelet basis, this procedure implements an adaptive smoothing (figure 5(a)), which averages the data with a kernel that depends on the underlying signal (Mallat 1999).

2.7. FLIM analysis using wavelets

The original Spring and Clegg's algorithm uses a TI-Haar wavelet transform thresholding algorithm over the intensity images to improve lifetime estimations. Yet it only uses the spatial information of each intensity image separately, without considering temporal (or phase) information. One way to integrate such information is to fit a sinusoidal function to the intensity stack in each pixel and from there recover images for the modulation and phase. In other words, obtain a phasor image. Then, the algorithm should compare if two pixels are close in phasor space. This is the idea behind pawFLIM.

We implemented a weighted translational-invariant Haar-wavelet transform denoising algorithm applied over the phasor images. The algorithm uses wavelets to adaptively average phasors. We modified

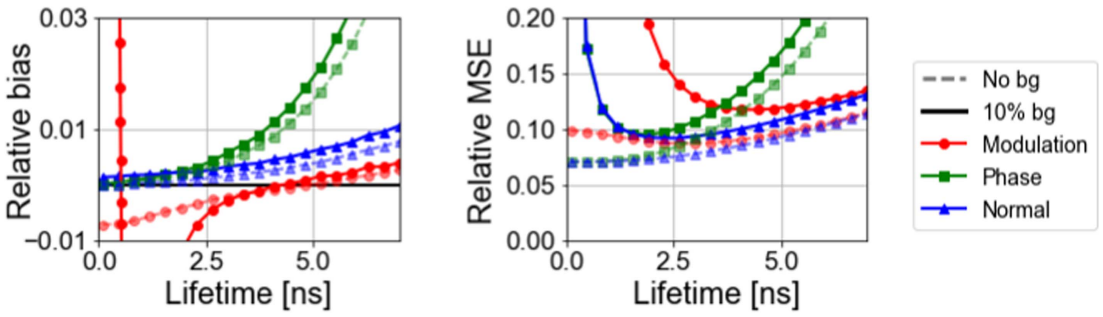

Figure 6. Performance of lifetime estimators. Bias and mean squared error (MSE) relative to the lifetime. Data from a Monte Carlo simulation of TCSPC for 200 detected photons, a frequency of 40 MHz, and 1024 bins, 1 million iterations, negligible error bars. Red, green and blue curves correspond to the modulation, phase and normal lifetime, respectively. Light and dark colors correspond to simulations without and with background.

the usual Haar approximation a and detail d coefficients as a weighted average and a difference.

$$\begin{cases} \mathbf{a} = (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1} (\boldsymbol{\Sigma}_1^{-1}\mathbf{x}_1 + \boldsymbol{\Sigma}_2^{-1}\mathbf{x}_2) \\ \mathbf{d} = \mathbf{x}_1 - \mathbf{x}_2 \end{cases} \quad (14)$$

where $\mathbf{x_1}$ and $\mathbf{x_2}$ are two neighboring complex values, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ their corresponding covariance matrices. Unlike previous denoising at the phasor level (Lin 2012), our algorithm explicitly considers the uncertainty of each phasor. To decide if two phasors should be averaged we use the distance between them in relation to their error as determined by the covariance matrix (figure 5(b)).

To resolve whether to keep the detail coefficient, a χ^2 is calculated:

$$\chi^2_{(2)} = \mathbf{d}^T(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)^{-1}\,\mathbf{d} \quad (15)$$

and the coefficient is thresholded according to a desired significance level. These equations are valid for one-dimensional complex data, but can be easily expanded for two-dimensional complex data. See supplementary information 4 for a complete description of the algorithm and the adaptation of Haar wavelets to weighted complex data.

3. Simulations

3.1. Lifetime estimators: single-pixel analysis

To compare lifetime estimators, we performed a Monte Carlo simulation of time correlated single photon counting (TCSPC) datasets and quantified the bias and mean squared error (MSE). In the general case, considering background and IRF, the simulation procedure is the following: for each detected photon, a value is drawn from a Bernoulli distribution with parameter p_{bg} deciding whether this photon comes from background noise or signal. Background photons are then drawn from a uniform distribution. Signal photons are drawn as following: a value X is drawn from an exponential distribution and another value Y is drawn from the IRF distribution, and the sum $X + Y$, corresponding to the photon arrival time, is placed in a histogram. This process is repeated until N values fill the histogram. Finally, a phasor is calculated from this histogram and, in turn, each lifetime estimator is obtained from this phasor value. This process is iterated, obtaining a distribution for each estimator, from which bias and MSE are evaluated.

The distributions for each lifetime estimator were bell shaped, approaching a normal distribution as an increasing number of photons were considered. Therefore we calculated the bias (deviation of the mean from the true value) and mean squared error (MSE), both relative to the true lifetime to describe the distribution (figure 6). For optimized systems in which $\omega\tau \leqslant 2\pi/10$ (in this case $\tau < 2.5$ ns) the improvement of the normal estimator is subtle as compared to the usual phase estimator. But when a larger lifetime range is considered, the normal lifetime estimator has an overall better performance (less bias and and MSE). This provides a practical way to extend the operation range without changing the equipment or the sample.

3.2. Lifetime estimators for spatially heterogeneous images

Having demonstrated that the normal estimator improves monoexponential lifetime determination in a single-pixel basis, we then moved to measure the performance of pawFLIM denoising on a spatially heterogeneous image. Using in each pixel the same procedure as in the previous section, we now simulated an image in which the lifetime changes between four different values (from 1 to 3 ns). The intensity for each pixel was sampled from a normal distribution with a mean of 400 and standard deviation of 40 photons, yielding a spatially unstructured intensity image (figure 7(a)).

A direct calculation of the undenoised image shows that the phasor distributions spread over the semicircle making it impossible to resolve among the underlying 4 distributions (figure 7(b)). The image reconstruction presents four clearly distinct bands (like in the template image) but with a rather large variance in each of them. Application of the Clegg

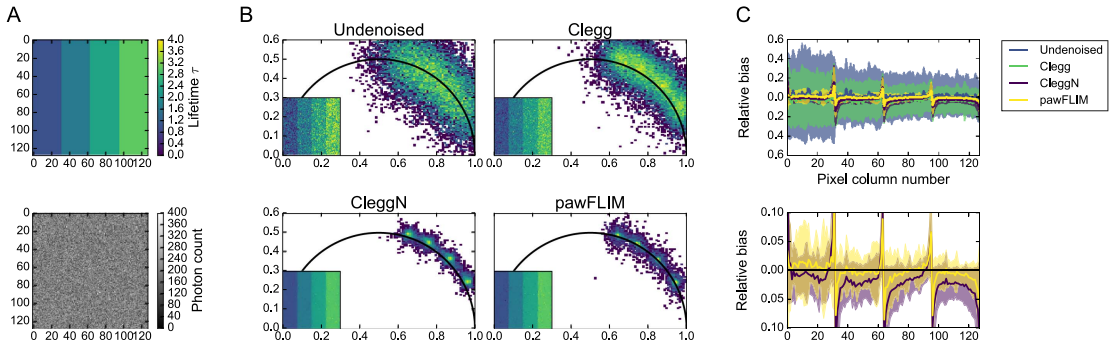

Figure 7. Performance of the different methods for intensity inhomogeneous simulated FD-FLIM data. (a) Lifetime and intensity images used as templates for the simulation, to which Poissonian noise is added. (b) Phasor plot and image reconstruction for each of the considered processing methods, with a significance level of 0.01. (c) Bias (line) and MSE (band) as a function of the horizontal coordinate of the image. pawFLIM outperforms all other methods in any case. Below is shown a zoom, to appreciate the differences between CleggN and pawFLIM.

algorithm significantly improves the result, reducing the variance in each band. Additionally, four small peaks start to appear in the distribution. The denoising is less effective than in the original work presented by Clegg (Spring and Clegg 2009). The reason is that Clegg only works well when DC levels are uniform across the image but it fails when there is a considerable variation. For instance, while two neighboring pixels may correspond to the same lifetime, their intensities could be very different due to distinct DC levels (i.e. total photons collected). Hence, those two pixels will not be averaged in intensity.

To make a fairer comparison with pawFLIM, we have included the results from our own variation of Clegg (naming the method CleggN), which solves this issue by including a normalization step. In practice, we expect different DC levels, either by photon counting fluctuations or different absolute fluorophore concentrations. Note that normalizing is a way of integrating temporal information, since we divide by the total number of photons collected. The image reconstruction using CleggN already displays a much smoother image in each band correlated with four clear peak in the phasor plot. However, a bias towards the outside of the semicircle is clearly visible. pawFLIM corrects this bias as well as narrowing down the distribution (i.e. reducing the MSE).

The comparison is more clearly visible if we calculate the average and variance along the columns of the lifetime image. pawFLIM results in a much smaller bias and MSE as compared to other methods (figure 7(c)).

3.3. Photon and molecular fractions estimators

For binary systems consisting of two monoexponential decays, as is typical of properly designed FRET-FLIM experiments, the photon and molecular fractions are the readout of choice instead of the fluorescence lifetime. Such determination is a two step process. In the first, the monoexponential lifetimes corresponding to both states are estimated (R_1 and R_2) and later the fractions are derived. Obtaining R_1 and R_2 is rather straightforward when a sample with spatially separated populations of each kind can be constructed. Even when the photon count is very low, applying the methods described in the previous sections can be used to accurately determine R_1 and R_2.

When such a sample is not available, using multiple measurements with distinct photon fraction values can be used. Using global analysis, R_1 and R_2 can be derived from the intersection of a linear fit to the phasors with the monoexponential semicircle (Clayton *et al* 2004). It has been shown that there is a requirement of a sufficient variation in p to estimate the global lifetimes (Verveer and Bastiaens 2003), which is more evident in the phasor plot representation.

However, a variation in p is not always enough to obtain an accurate determination of R_1 and R_2. Phasor distributions corresponding to a single p fraction aren't circular, but rather elongated along a direction which doesn't necessarily coincide with the biexponential line direction. In particular, for low number of photons, distributions corresponding to different p coalesce and the linear fit is dominated by the individual direction of each of them (supplementary figure 2). Therefore, we have introduced the normalized perpendicular distance to the biexponential line, $p_\perp$, which is a measure of interest in the determination of R_1 and R_2, or equivalently the global lifetimes τ_1 and τ_2 (see supplementary information 2).

We evaluated the performance of the different algorithms for an image composed of a step followed by a gradient and a sinusoidal in the p fraction (figure 8). The number of photons in each pixel was drawn from a Poisson distribution (mean value 50). While barely any structure is appreciable without denoising, it can be seen that the three algorithms recover a distribution of p similar to the true one. The second row, which displays an average in the vertical direction, shows that both Clegg and CleggN have a greater bias than pawFLIM, which explains both the observed pink areas in Clegg and CleggN as well as the

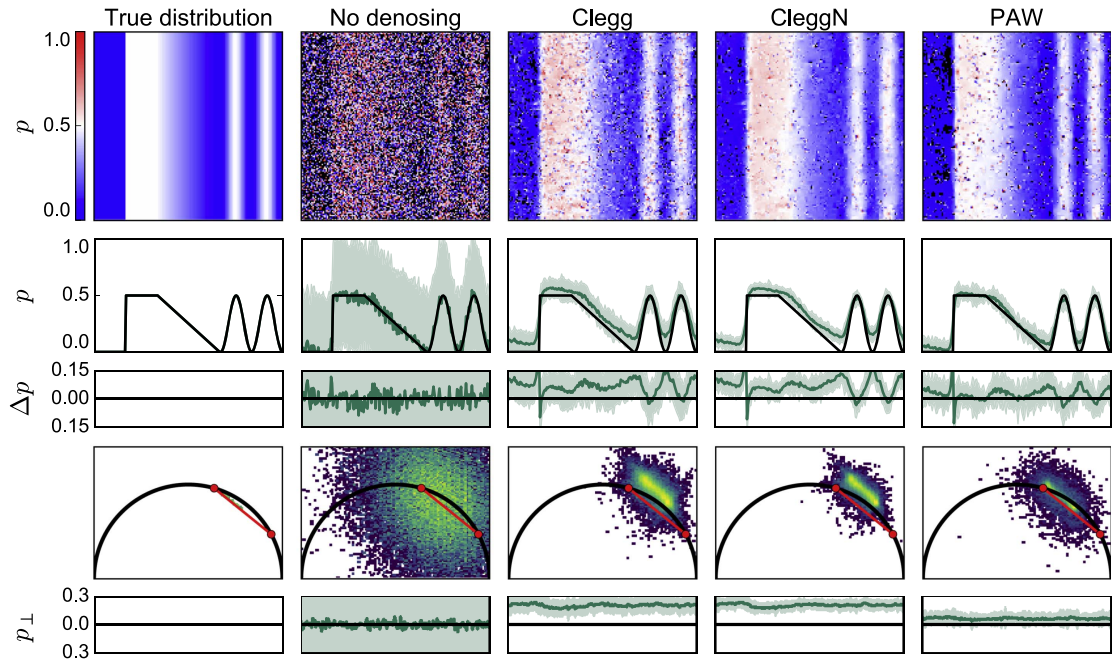

Figure 8. Performance of different methods for simulated single-frequency FD-FLIM data. Row 1: photon fraction image. Row 2: mean value (green line) and standard deviation (green band) in the vertical spatial direction, contrasted with the true value (black line). Row 3: difference from the true value. Row 4: phasor plot distribution, where the red line corresponds to the true biexponential lifetimes. Row 5: analogous to row 2, but showing the perpendicular distance to the biexponential line. (Data was simulated for a 192×192 image, with lifetimes of $\tau_1 = 1$ and $\tau_2 = 3$ ns, a frequency of $f = 40$ MHz, $N = 50$ photons per pixel, and a significance level of 0.01.)

greater absence of black spots which signal $p < 0$ (left area, which was simulated as $p = 0$). Also, pawFLIM shows a greater contrast in the sinusoidal section of the p distribution. In terms of the standard deviation, there are no great differences for this level of averaging (i.e. p-value threshold).

In this comparison, global analysis was not performed but the true lifetimes were used. While the enhancement of the phasor distribution (from circular to elongated along the biexponential line) would increase the accuracy in the linear fit (supplementary figure 3), the absence of bias in the perpendicular direction is also crucial in the determination of the global lifetimes.

In row 4, the behaviour of the phasor distribution in the phasor plot is shown. As can be easily seen, both Clegg and CleggN present a bias in the direction perpendicular to the biexponential line, while pawFLIM keeps the distribution around it. Analogously to row 2, in row 5 an average in the vertical direction of the perpendicular distance to the biexponential line is shown. It can be seen that pawFLIM also has a smaller bias in this direction.

3.4. Impact of the chosen significance level

While in a single pixel error analysis every parameter could be swept through all values of interest, there is no systematic way to analyze every possible image. Indeed, different spatial distributions, both for photon fraction and number of photons, would have to be considered. Therefore the performance comparison was computed for a few representative different N distributions (horizontal and vertical gradient and random, all subjected to Poisson noise) and gradient p fraction (figure 9).

The algorithm is controlled by a single parameter to decide when two phasors are to be averaged: the significance level (SL) of the test (also shown in terms of the standard deviation of a normal variable for comparison). The behaviour of the different algorithms was measured by the following statistics: χ^2, for the phasor R, the fraction p and the direction perpendicular to p, $p_\perp$; and bias in both parallel and perpendicular directions to p.

pawFLIM outperforms Clegg in every $\chi^2(p)$ measure. It is also superior to CleggN except in a small region of significance values. Notice that Clegg, being unnormalized, depends on the photon count distribution and therefore the relative gradient orientation (p versus N) plays a major role (figure 9, red squares versus circles in χ^2).

Using the true R_1 and R_2, the phasors corresponding to the donor only and donor in complex lifetime, pawFLIM has a lower p bias than Clegg and CleggN for all significance levels (figure 9). Additionally, pawFLIM also outperforms Clegg and CleggN for the $p_\perp$ bias. As has been described previously, $p_\perp$ has a strong influence in the determination of R_2 and R_1. This in turn affects the estimation of p as it is calculated as the fractional distance of the line connecting R_2 and R_1. Therefore, using pawFLIM improves the estimation of the global lifetimes (See supplementary

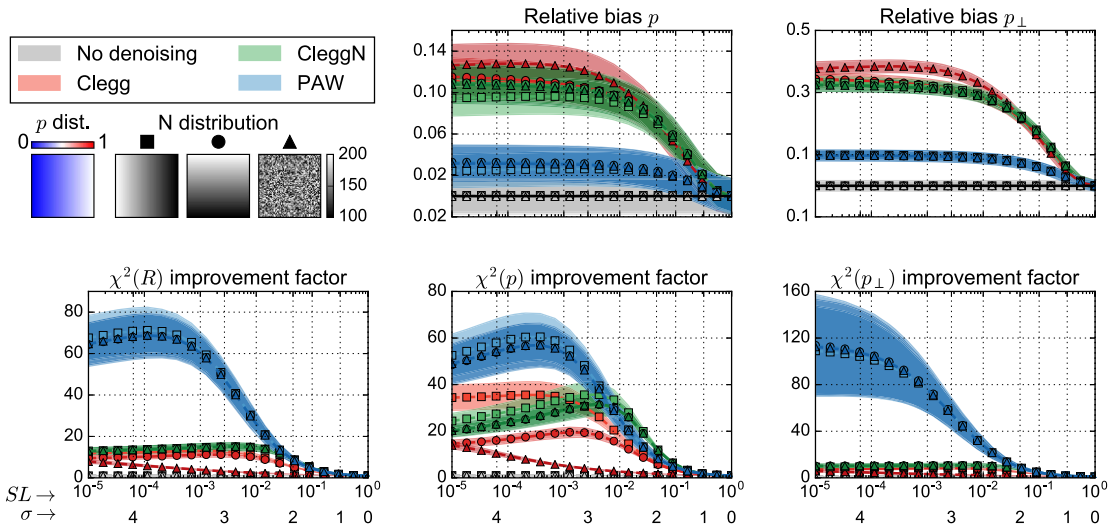

Figure 9. Statistics of the different denoising methods as a function of the significance level *SL*, for a photon fraction *p* gradient and different total photons *N* distributions. The improvement factor in χ^2 is against no applied denoising, while the relative bias is calculated as $\sum_i \frac{f_i - f_{i,0}}{\sigma_{f_i}}$, where $f_{i,0}$ and σ_{f_i} are the true value and standard deviation, respectively. The lines and bands correspond to the mean value and a standard deviation of 1000 iterations. The image size was 64 × 64, and the global lifetimes were 1 and 3 ns, with a frequency of 40 MHz.

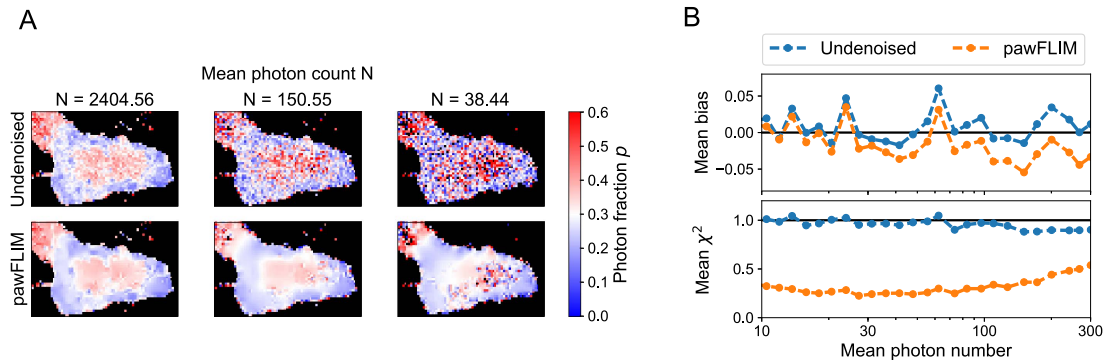

Figure 10. Application to experimental data. Quantification of the spatial distribution of EGFR phosphorylation after 30 min EGF stimulus. (a) Comparison of pawFLIM to an undenoised image for different average numbers of photons per pixel (*N*). (b) Using the undenoised image with maximum number of photons as ground-truth, the mean bias and χ^2 is shown as a function of *N* for undenoised and pawFLIM processed datasets. Image width: 80 μm.

information 2) and local photon fraction (both directly and indirectly). Overall, this provides a better estimation of the local molecular fraction *m*.

4. Application to experimental data

To demonstrate pawFLIM on experimental FRET-FLIM datasets, we image the phosphorylation of epidermal growth factor receptor (EGFR) after 30 minutes of epidermal growth factor (EGF) stimulation, as has been previously described (Grecco *et al* 2010). Briefly, we transfected MCF-7 cells with a plasmid encoding for EGFR fused to mCitrine (a yellow fluorescent protein). We stimulated the cells with EGF and fixed them after 30 min. As a FRET acceptor we used Cy3.5 conjugated to a generic anti-phosphotyrosine antibody (PY72).

To evaluate the performance of pawFLIM for different average numbers of photons per pixel (*N*), we truncated the TCSPC data series to different lengths. When the average intensity is high, both methods produce similar results (figure 10(a)). However, as the number of photons is reduced, the undenoised image quickly loses the spatial structure while pawFLIM remains mostly intact.

To quantify such a difference, we used as ground-truth the undenoised image with a maximum number of photons ($N = 2404.56$). For each pixel, the variance corresponding to the given phasor can be estimated as described in the previous sections. Comparing the ground-truth and reconstructed image, we calculated the normalized difference as the pixel-wise difference divided by the standard deviation. The mean bias can be defined as the normalized difference averaged over the whole image. In a similar way, the mean χ^2 can be

obtained by using the squared normalized difference instead. We calculated these estimators as a function of the mean photon number. It is clear from the plot that pawFLIM does not significantly affect the bias but yields a smaller mean χ^2 than the undenoised image (figure 10(b)). As an example, with as little as 30 photons on average per pixel, pawFLIM outperforms the results obtained with undenoised images by a factor of four. Even if 10 times more photons are considered, the undenoised image does not improve significantly. This means that image acquisition can be obtained 10 times faster without sacrificing molecular resolution.

5. Discussion

In this work we have demonstrated theoretically, with simulation and experimental data, a novel FLIM analysis method (named pawFLIM) that allows us to optimise the balance between molecular and spatial resolution. Such improvement is achieved by averaging neighbouring pixels with similar phasor values and thereby enhancing locally the signal to noise ratio. It outperforms other existing methods, particularly for images with very low numbers of photons per pixel. This results in a typically seven-fold reduction in bias and a typical five-fold improvement in uncertainty. This means that five times fewer photons are necessary with pawFLIM, thereby allowing for faster acquisition without sacrificing quality.

Our method combines a series of important features. First, averaging is not performed with a constant window size but rather by using an adaptive wavelet transform. This allows us to preserve high frequency features of the images such as edges. In other words, areas with similar values will be averaged with a coarser grain than others which are inhomogeneous. Second, averaging is done in the complex plane representation of the FLIM datasets (i.e. the phasor image). This allows us to properly compare neighbouring pixels in a way that is unbiased by the number of photons performing better than methods operating on intensity- or lifetime-images (Buranachai *et al* 2008, Chang and Mycek 2009, Spring and Clegg 2009). Unlike previous methods (Lin 2012), pawFLIM does not average phasors directly but rather appropriately weights the phasor values using the corresponding uncertainty when testing for similarity. The significance level of the test is a parameter of the method, allowing us to conserve spatial features of interest. It is important to notice that the performance of the method is highly dependent on the image itself. The topography of intensity and the fraction of interacting molecules defines the improvement provided by the method.

Within this work we have also presented a new estimator for the lifetime (named normal lifetime) which has a smaller MSE and bias than the usual, historically based, phase and modulation lifetimes.

Overall, this method enables us to reduce the required number of photons per pixel in FRET-FLIM datasets. In turn, movies quantifying the dynamics of protein interaction could be obtained with better temporal or spatial resolution.

Acknowledgments

The funders ANPCYT and MPG had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors also acknowledge financial support from ANPCYT PICT 2013-1301, Universidad de Buenos Aires 20020130200271BA and MPG Partner Group.

Conflicting interests

The authors have declared that no conflicting interests exist.

References

Abramovich F and Benjamini Y 1996 Adaptive thresholding of wavelet coefficients *Comput. Stat. Data Anal.* **22** 351–61
Aguet F *et al* 2016 Membrane dynamics of dividing cells imaged by lattice light-sheet microscopy *Mol. Biol. Cell.* **27** 3418–35
Amos W and White J 2003 How the confocal laser scanning microscope entered biological research *Biol. Cell* **95** 335–42
Barabasi A-L and Oltvai Z N 2004 Network biology: understanding the cell's functional organization *Nat. Rev. Genet.* **5** 101–13
Barber P R *et al* 2005 Global and pixel kinetic data analysis for FRET detection by multi-photon time-domain FLIM *Proc. SPIE* **5700** 171
Beechem J M *et al* 1983 Global resolution of heterogeneous decay by phase/modulation fluorometry: mixtures and proteins *Biochemistry* **22** 6054–8
Bernaola O 2001 *Enrique Gaviola y el Observatorio Astronómico de Córdoba. Su impacto en el desarrollo de la ciencia argentina* Saber y Tiempo
Buranachai C *et al* 2008 Rapid frequency-domain FLIM spinning disk confocal microscope: lifetime resolution, image improvement and wavelet analysis *J. Fluoresc.* **18** 929–42
Chang C-W and Mycek M-A 2009 Improving precision in time-gated FLIM for low-light live-cell imaging *Molecular Imaging II SPIE* **7370** 737009
Clayton A H A *et al* 2004 Graphical representation and multicomponent analysis of single-frequency fluorescence lifetime imaging microscopy data *J. Microsc.* **213** 1–5
Conchello J-A and Lichtman J W 2005 Optical sectioning microscopy *Nat. Methods* **2** 920–31
Digman M A *et al* 2008 The phasor approach to fluorescence lifetime imaging analysis *Biophys. J.* **94** L14–6
Fereidouni F *et al* 2011 A modified phasor approach for analyzing time-gated fluorescence lifetime images *J. Microsc.* **244** 248–58
Förster T 1948 Zwischenmolekulare energiewanderung und fluoreszenz *Ann. Phys., Lpz.* **437** 55–75
Gaviola E 1926 Die abklingungszeiten der fluoreszenz von Farbstofflösungen *Ann. Phys., Lpz.* **386** 681–710
Gaviola E and Pringsheim P 1927 Zur frage nach dem Übergang von Fluoreszenz in Phosphoreszenz *Zeitschrift für Physik* **43** 384–93
Grecco H E and Bastiaens P I H 2008 Imaging protein states in cells *Energy* (New York: Cold Springer Harbor Laboratory Press)
Grecco H E *et al* 2009 Global analysis of time correlated single photon counting FRET-FLIM data *Opt. Express* **17** 6493

Grecco H E *et al* 2010 *In situ* analysis of tyrosine phosphorylation networks by flim on cell arrays *Nat. Meth.* **7** 467–72
Grecco H E *et al* 2011 Signaling from the living plasma membrane *Cell* **144** 897–909
Jares-Erijman E A and Jovin T M 2003 Fret imaging *Nat. Biotech.* **21** 1387–95
Knutson J R *et al* 1983 Simultaneous analysis of multiple fluorescence decay curves: a global approach *Chem. Phys. Lett.* **102** 501–7
Lakowicz J R 2006 *Principles of Fluorescence Spectroscopy* (New York: Springer)
Lin Z 2012 Fluorescence lifetime imaging camera: image analysis, optimization and enhancement *PhD Thesis* Ruperto-Carola Heidelberg University
Mallat S 1999 *A Wavelet Tour of Signal Processing* (Cambridge, MA: Elsevier)
Margineanu A *et al* 2016 Screening for protein-protein interactions using förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM) *Sci. Rep.* **6** 28186
Minsky M 1961 Microscopy apparatus *US Patent* 3,013,467
Paweletz N 2001 Walther Flemming: pioneer of mitosis research *Nat. Rev. Mol. Cell Biol.* **2** 72–5
Pelet S *et al* 2004 A fast global fitting algorithm for fluorescence lifetime imaging microscopy based on image segmentation *Biophys. J.* **87** 2807–17
Roudot P *et al* 2013 Noise modeling for intensified camera in fluorescence imaging: application to image denoising *2013 IEEE 10th International Symposium on Biomedical Imaging* IEEE
Shimomura O *et al* 1962 Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea *J. Cell. Comp. Physiol.* **59** 223–39
Spring B Q and Clegg R M 2009 Image analysis for denoising full-field frequency-domain fluorescence lifetime images *J. Microsc.* **235** 221–37
Tsien R Y 1998 The green fluorescent protein *Ann. Rev. Biochem.* **67** 509–44
Verveer P J and Bastiaens P I H 2003 Evaluation of global analysis algorithms for single frequency fluorescence lifetime imaging microscopy data *J. Microsc.* **209** 1–7
Verveer P J *et al* 2000 Global analysis of fluorescence lifetime imaging microscopy data *Biophys. J.* **78** 2127–37
Wouters F S *et al* 2001 Imaging biochemistry inside cells *Trends Cell Biol.* **11** 203–11