Edinburgh Research Explorer

Fast and robust single-exponential decay recovery from noisy fluorescence lifetime imaging

Fast and robust single-exponential decay recovery from noisy fluorescence lifetime imaging

Ali Taimori, Duncan Humphries, Gareth Williams, Kevin Dhaliwal, Neil Finlayson, James Hopgood, *Member, IEEE*

Abstract—Fluorescence lifetime imaging is a valuable technique for probing characteristics of wide ranging samples and sensing of the molecular environment. However, the desire to measure faster and reduce effects such as photo bleaching in optical photon-count measurements for lifetime estimation lead to inevitable effects of convolution with the instrument response functions and noise, causing a degradation of the lifetime accuracy and precision. To tackle the problem, this paper presents a robust and computationally efficient framework for recovering fluorophore sample decay from the histogram of photon-count arrivals modelled as a decaying single-exponential function. In the proposed approach, the temporal histogram data is first decomposed into multiple bins via an adaptive multi-bin signal representation. Then, at each level of the multi-resolution temporal space, decay information including both the amplitude and the lifetime of a single-exponential function is rapidly decoded based on a novel statistical estimator. Ultimately, a game-theoretic model consisting of two players in an "amplitude-lifetime" game is constructed to be able to robustly recover optimal fluorescence decay signal from a set of fused multi-bin estimates. In addition to theoretical demonstrations, the efficiency of the proposed framework is experimentally shown on both synthesised and real data in different imaging circumstances. On a challenging low photon-count regime, our approach achieves about 28% improvement in bias than the best competing method. On real images, the proposed method processes data on average around 63 times faster than the gold standard least squares fit. Implementation codes are available to researchers.

Index Terms—Adaptive filtering, binning, fluorescence lifetime imaging microscopy, game theory, lifetime estimation, multi-resolution representation, rapid lifetime determination, signal recovery, single-exponential decay.

I. Introduction

TIME-DOMAIN fluorescence lifetime imaging microscopy (FLIM) is a powerful signal acquisition technique to characterise biological and chemical samples such as cells, viruses and molecules. FLIM has diverse applications in fields such as biology, chemistry, physics, materials science, medicine, pharmacology, and cancer research [1–4]. Unlike in steady state fluorescence imaging which simply measures the intensity of fluorescence produced by a sample, lifetime imaging allows for probing of the molecular environment. The fluorescence lifetime of a molecule, the time between photon absorption and fluorescence emission, is often dependent on the environment in which the fluorophore is present. Such environmental properties include viscosity, pH, temperature and quenching interactions with surrounding molecular species. This can allow for the design of specific probes to detect and monitor these environmental effects, or the observation of such changes through autofluorescence. Furthermore, fluorescence lifetime is relatively insensitive to concentration effects and other effects on intensity that cause large fluctuations in steady state imaging. In FLIM, a location in a specimen is first excited by a pulsed laser. Absorption of excitation photons can lead to fluorescence, where electrons moved to an excited state return to the ground state resulting in photon emission. The average elapsed time from excitation pulse to the emission of a photon is known as the fluorescence lifetime. The detection of the emitted photons associated with the electronic transition from the excited state to the ground state is recorded as a time-stamped event via a sensitive sensor relative to the laser pulse. For a predefined exposure time extending over multiple laser pulses, the measurement cycle is repeated many times. Photon arrival times in a number of quantised time intervals are then counted. Finally, the whole process generates a raw data sequence consisting of a histogram of photon-count arrivals representing a decay function for a pixel of the imaged sample. The data sequence contains noise originating from various sources such as dark counts and shot noise. On-chip histogramming has been reported in a single photon avalanche diode line sensor which permits very high throughput of fluorescence lifetime signals [5]. Where the complexity of decay depends on the number of emitting species or environment within the image pixel, the histogram decay can be modelled by a single-, double- or generally multi-exponential function. The model of single-exponential decay is widely used across FLIM applications due to its simplicity and utilisability for high-speed imaging. In the literature, different methods have been proposed for estimating decay parameters including the fluorescence lifetime [6–15]. The lifetime is a physical characteristic that describes the decay rate of fluorescence. As a biomarker or chemomarker, it provides contrast for distinguishing substances in biology and chemistry sciences. In this paper, we focus on both speed and robustness issues for lifetime sensing in a wide range of photon-count regimes.

A. Review of related work

We categorise state-of-the-art available lifetime estimation methods into three general families consisting of fitting-based, non-fitting-based, and fit-free approaches. The fitting-based family tries to fit a curve to the histogram of photon counts. Here, the least squares (LS) method is generally employed to estimate parameters of a fluorescence decay model [6]. Maximum likelihood estimation (MLE) is another approach in this category that can account for a more realistic Poisson distribution for the histogram of photon counts, instead of the Gaussian assumption behind the LS fit [7].

In contrast to the direct fitting formulation and solution, non-fitting-based algorithms use specific techniques for finding fluorescence lifetime imaging characteristics. These include rapid lifetime determination (RLD) [8], RLD with overlapping windows (RLD-OW) [9], center of mass method (CMM) [10], and fluorescence lifetime estimation via rotational invariance techniques (FLERIT) [11] are all paradigms of non-fitting-based approaches. The methods available in the RLD class present closed-form solutions for estimating the lifetime parameter [8, 9, 16]. CMM uses the center of mass of the single-exponential distribution to derive a simplified lifetime calculation formula suitable for on-board clinical firmware applications. FLERIT models the problem of lifetime estimation into a general class of direction of arrival estimation in array signal processing based on singular value decomposition [17].

Fit-free mechanisms are the third group of lifetime estimation methods. As two well-known approaches in this family, the phasor approach represents the fluorescence lifetime information in a diagram-wise manner [12], and learning-based methods exploit function approximation capability of shallow and deep neural networks for lifetime parameter estimation based on raw decay training samples [13, 14]. In practical FLIM system, lifetime estimation performance and its computation time are two key components. Each of the lifetime calculation families have their own benefits and shortcomings for a compromise between performance and complexity [3].

B. Current challenges

Fitting-based methods are accurate to some extent but are computationally complex. Non-fitting-based algorithms, as addressed in this paper, are fast and provide an apt option for hardware realisation [18]; however, dealing with a range of uncertainties available in the acquired signal simultaneously remains a difficult task. Fit-free phasor and learning-based approaches suffer respectively from a need for intervention of experts for lifetime analysis, and a requirement for a large number of training time series instances. Moreover, non-automatic implementation and the lack of data present two main challenges for medicine-related researchers.

We can fundamentally summarise that the performance of a lifetime estimation approach is affected by three influential dynamics which originate from: 1) the nature of the sample being imaged; 2) the quality of the imaging instrument; and 3) the user experience on tuning acquisition parameters. All the dynamics appear in the captured histogram of photon counts in different forms, making lifetime determination a challenging task [19]. To name a few, low, mid, and high photon counting regimes lead to low-, mid- and high-level signal-to-noise ratio (SNR) acquisitions to be managed [20, 21]. Due to limiting the power of light source for preventing damage to sensitive specimens, low photon-count imaging is often observed which leads to amplification of noise components. Blurring effects due to convolving the original decay signal with the instrument response function (IRF) of the imaging device deform the head and tail of a decaying function, which may results in overestimation of the lifetime [15]. Different fluorophores expose varying lifetimes from short (less than 1ns) to long (more than 10ns) ranges. User tunable parameters during imaging such as bin width or size of the histogram of photon arrivals for exposure time control also influence the decay curve shape and ultimately the system's accuracy and precision [11]. These factors demand a robust lifetime estimation approach coping with the various regimes, and this is the goal of the method proposed in this paper.

C. Research contributions

To take account of the variability in fluorescence lifetime acquisition, we suggest a perturbation-robust lifetime estimation framework. We first mathematically model the fluorescence decay function in both deterministic and stochastic modes. Afterwards, we propose a fast algorithm to be able to estimate amplitude and lifetime parameters of a single-exponential decay function in the presence of both inevitable convolution and noise perturbations. To control the uncertainties existing in a decay signal, we decompose the temporal histogram data as an adaptive multi-bin signal representation. The decomposition is done via a set of adjustable, successive Savitzky-Golay (S-G) low-pass filtering and binning. At each level, decay parameters are estimated from the transformed signals. We formulate the problem of reliable detection of the best representative signal from a set of candidates determined from multiple temporal resolutions using a 2-player game model [22]. Game theory based approaches such as generative adversarial networks [23] have recently found beneficial applications in diverse areas including FLIM [24]. Here, our main motivation of using the game model is to handle intrinsic dynamics of the histogram of photon counts in different circumstances. Amplitude and lifetime as players act on estimated parameters based on their own payoff functions to be able to recover optimal decay. Our approach falls into the category of non-fitting-based approaches, where the decaying signal can be rapidly and robustly recovered for different regimes. Figure 1 illustrates the suggested method on an exemplar histogram of photon count arrivals. The robustness of our single-exponential decay recovery method at high level of noise is also shown in Fig. 2. Implementation codes are available online to researchers[1].

In summary, our main scientific contributions are:

- a fast algorithm to decode lifetime information from the histogram of photon counts modelled by a single-exponential decaying function;
- an adaptive multi-bin representation of histograms of photon counts to deal with uncertainties of decay variability;

[1]https://datashare.ed.ac.uk/handle/10283/4399

Fig. 1: The proposed multi-resolution fluorescence decay recovery method unfolded and visualised for an exemplar histogram of photon counts. S-G stands for Savitzky-Golay; L_{f_i} denotes window length of the S-G filter at the ith tunable branch; and the parameters A, τ, and M_i represents amplitude and lifetime of the exponential decay, and the number of time bins after binning at the ith level, respectively. The functions $g_i[n]$ and $y_i[m]$ are output signals of S-G filtering and binning at the ith level, respectively. The blue star in the game space denotes the point of Nash equilibrium. The plots in the output of the flow diagram include: the original decay $x[n]$ in the blue colour, corrupted measurements of $\tilde{x}[n]$ with the red stems, and the recovered decay $f^*[n]$ in the green colour.

- a game-theoretic model to permit robust recovery of single-exponential decay from a set of signal selection.

D. Paper organisation

The remainder of the paper is organised as follows. In Section II, we present our theoretical analysis behind the lifetime estimation problem, and provide details of the proposed algorithms for single-exponential decay recovery in Section III. In Section IV, a set of coherent experiments for demonstrating the efficiency of our suggested method is described. Section V provides an analysis of the computational complexity of the algorithms. Finally, conclusions are presented in Section VI.

II. THEORY

In this section, we model the decay signal of fluorescence lifetime both deterministically and stochastically. A derivation of FLIM parameter estimation is presented for each case.

A. Deterministic decay model and estimation

In a FLIM system, the fluorescence decay function is ideally modelled by a first-order linear differential equation as:

$$\frac{dx(t)}{dt} = -(k_r + k_n)x(t), \tag{1}$$

where $x(t)$, k_r and k_n denote the quantity at time t, the rate constant of a radiative process, and the cumulative rate constant of environment-dependent non-radiative processes, respectively [7, 25]. Using the separation of variables method on (1) gives the solution $x(t) = Ae^{-\frac{t}{\tau}}$, where $A \triangleq x(0)$ is the initial amplitude and $\tau \triangleq \frac{1}{(k_r+k_n)}$ is the fluorescence lifetime.

In practice, the fluorescence decay is measured at equally-spaced time intervals rather than a given time. In this case, a discrete-time representation of the exponentially decaying continuous-time function with unquantised intensities as $x[n] \in \mathbb{R}^+$ follows from:

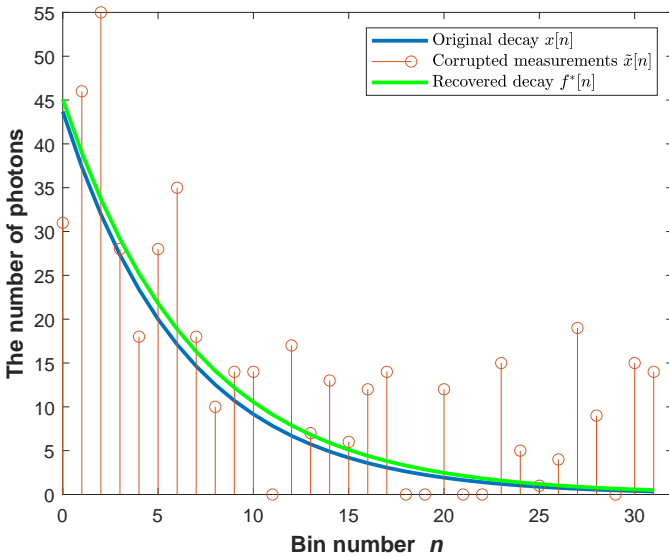$$x[n] = Ae^{-\frac{\Delta \cdot n}{\tau}}, \forall n = 0, 1, \ldots, N-1, \tag{2}$$

Fig. 2: An illustration of robustness of the proposed single-exponential decay recovery under severe perturbation. The total number of photons of the histogram is 460.

where Δ, N and $\mathbb{R}^+$ represent the measurement time interval (in order ranging from picosecond to nanosecond for FLIM devices), the number of measurements, and the set of all positive real numbers, respectively. It is obvious that $A = x[0]$. The function $x[n]$ is a strictly monotonically decreasing function with the descending sort property: $x[0] > x[1] > \cdots > x[N-1]$. For a detailed discrete-time signal analysis, we provide standard definitions below to make this paper more accessible to a wider audience.

Definition 1 (Linear difference operator): The linear difference operator $\mathcal{D}\{\cdot\}$ is defined as:

$$d[n] \triangleq \mathcal{D}\{x[n]\} = x[n] - x[n+1], \forall n = 0, 1, \ldots, N-1. \quad (3)$$

The descending sort property is held for $d[n]$, where $d[0] > d[1] > \cdots > d[N-1]$.

Definition 2 (Maximum fall of signal): We define the point with maximum difference, or maximum fall, of the signal by:

$$n^* = \operatorname{argmax}_{0 \leqslant n < N}(d[n]). \quad (4)$$

It is obvious that always $n^* = 0$ in the ideal perturbation-free case. In this special case, this means the greatest difference is between the first sample and the second one.

Theorem 1 (Lifetime decoding): Let $x[n] = Ae^{-\frac{\Delta \cdot n}{\tau}}, \forall n = 0, 1, \ldots, N-1$ be a histogram of photon-count arrivals. The slope tangent to the point with maximum fall of $x[n]$, i.e. n^* in (4), conveys lifetime information of decay. Geometrically interpreting, for the point n^*, its initial slope just tangent to the exponentially decaying curve intercepts the n-axis at a point near $n \approx \tau$, where the estimated lifetime is $\hat{\tau} = \frac{\Delta}{\ln(\frac{x[0]}{x[1]})}$.

Proof: By using a first-order approximation of the Taylor expansion around the point with maximum difference, the tangent slope can be captured by an appropriate precision, which finally decodes the fluorescence lifetime value. The Taylor expansion of $x[n]$ around the point n^* gives:

$$\hat{x}[n^* + T] = x[n^*] + Tx'[n^*] + \varepsilon, \quad (5)$$

where $x'[n^*] = \frac{x[n^*+T] - x[n^*]}{T}$ represents the first-order difference. Neglecting the higher order terms by setting $\varepsilon \approx 0$, the first-order approximation of $x[n]$ is reached. By substituting $x'[n^*]$ into (5), taking the discrete interval $T = 1$ and reducing to the Maclaurin series for $n^* = 0$ in the ideal model, we have:

$$\hat{x}[1] = x[0]e^{-\frac{\Delta}{\tau}}. \quad (6)$$

If the approximation $\hat{x}[1]$ is the same as the true value $x[1]$, taking the natural logarithm from two sides of (6) will yield:

$$\hat{\tau} = \frac{\Delta}{\ln(\frac{x[0]}{x[1]})}. \quad (7)$$

∎

For the special case $N = 2$, (7) is equal to that of RLD.

B. Stochastic noisy-blurred decay model and estimation

1) Corrupted measurements modelling: In real FLIM systems, the signal (2) is a discrete stochastic process, where both A and τ are random variables. Any FLIM device consists of two separate optical and electrical parts. Hence, two main sources of noise are contributable including: 1) signal-dependent electronic shot noise arising from photon sensing equipment; and 2) signal-independent background noise due to instrument ambient disturbances. Contrary to the former with an electronic and discrete nature, the latter has mainly a non-electronic, optical source with continuous nature. The diverse noise components are assumed to combine to each other additively making a whole noise of the FLIM device. Let us model the histogram of photon-count arrivals in the presence of both blur and noise perturbation in a pixel as the following quantised measurements (See also Fig. 3.):

$$\tilde{x}[n] = \lfloor x[n] * h[n] + \eta[n] \rfloor \, u[\lfloor x[n] * h[n] + \eta[n] \rfloor], \quad (8)$$

where $\tilde{x}[n] \in \mathbb{Z}^+, \forall n = 0, 1, \ldots, N-1$. Operators $\lfloor \cdot \rfloor$, $u[\cdot]$ and $*$ denote the round (quantiser) and step functions, and convolution, respectively. The step function in (8) acts as a signal clipper and ensures that the photon-count $\tilde{x}[n]$ takes on physical positive counts only. The functions $h[n]$ and $\eta[n] \triangleq e_{\mathrm{P}}[n] + e_{\mathrm{G}}[n]$ respectively represent the blurring kernel as the IRF of the system, and the noise term with errors of $e_{\mathrm{P}}[n] \in \mathbb{Z}^+$ and $e_{\mathrm{G}}[n] \in \mathbb{R}$. We assign a discrete Poissonian probability mass function to the signal-dependent noise and a continuous Gaussian probability density function to the signal-independent one [26]. The characteristic of shot noise distribution is defined as $e_{\mathrm{P}}[n] \sim \mathcal{P}(\lambda)$, where $\mathcal{P}(\cdot)$ denotes Poissonian distribution. As a property of Poisson process, one straightforward way, usually used in the literature, is to consider the expectation of shot noise as a constant value across bins. In this case, the standard deviation of shot noise equals to $\sqrt{I}$ and consequently total SNR $\triangleq \frac{I}{\sqrt{I}} = \sqrt{I}$, where $I \triangleq \frac{N_{\mathrm{p}}}{N}$ denotes the average number of arrived photons [27]. The parameter N_{p} represents the number of photons per histogram, i.e., $N_{\mathrm{p}} = \sum_{n=0}^{N-1} x[n]$. The average rate of the shot noise can be written as $\lambda = I$, which means its total dependency on the histogram signal.

For the signal-independent noise, we assume $e_{\mathrm{G}}[n] \sim \mathcal{N}(\mu, \sigma^2)$, namely a Gaussian distribution with the mean
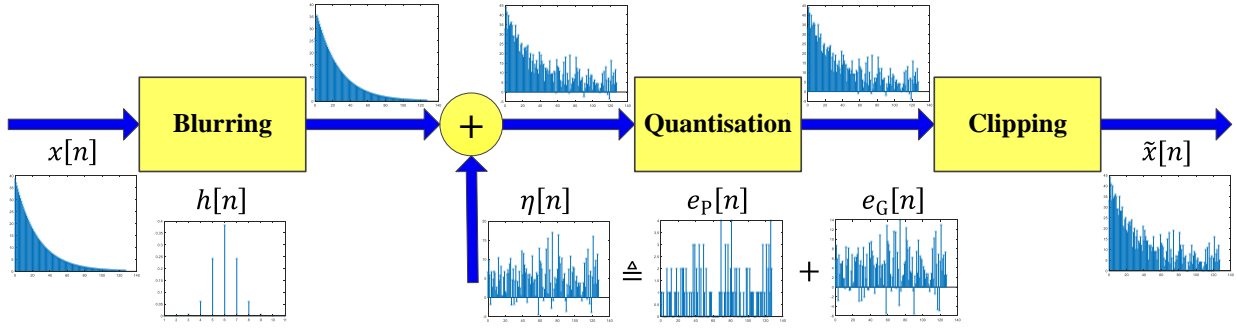
Fig. 3: The proposed process for generating synthesised histogram of photon counts. Signal of each point has visualised for a spacial experimental setting. The x-axis of plots denotes the bin index n.

$\mu = c\sigma$ and variance σ^2. The parameter $c \geqslant 0$ is a real constant, which, in case of $c = 1$, about 68.3% of random samples from the distribution cast over the zero level. To emphasise error resources in the system, we rewrite (8) as $\tilde{x}[n] = x[n] + e[n]$ with the total error of:

$$e[n] = e_{\mathrm{q}}[n] + e_{\mathrm{b}}[n] + \underbrace{e_{\mathrm{P}}[n] + e_{\mathrm{G}}[n]}_{\triangleq \eta[n]} + e_{\mathrm{u}}[n], \qquad (9)$$

where terms $e_{\mathrm{q}}[n]$, $e_{\mathrm{b}}[n]$ and $e_{\mathrm{u}}[n]$ denote errors introduced by quantisation, blurring and other uncategorised noise sources, respectively. It is notable that the quantisation noise originates from rounding amplitudes, due to the quantised nature of light in photon counting process.

2) Modelling of time-resolved shot noise: Effects of noise would appear more in points with high variations, especially in bins with low number of photons from low-photon count regimes. As evidence, taking the logarithm of histogram of photon counts reveals the locations of noise, where signal values in bins with low SNR are seen to have higher than proportional noise components relative to regions with higher SNR. Consequently, the histogram dependency in shot noise can be retrospectively assigned to binned photons based on different point-wise behaviours of SNR. To take this reality into account, one can change the constant noise valuation and mathematically customise the signal-dependent noise model for a special device by using weighting mechanisms. A general form for the average rate of shot noise can be defined as a combination of increasing ($\nearrow$) and decreasing ($\searrow$) terms as:

$$\lambda[n] \triangleq \zeta \left\lceil a \underbrace{\frac{x[0]}{x[n]}}_{\nearrow} + \beta \underbrace{\frac{x[n]}{x[0]}}_{\searrow} \right\rceil \in (0, \infty), \qquad (10)$$

in which the symbol $\lceil \cdot \rceil$ is the ceil function, and, parameters $a, \beta \in \mathbb{R}^+$ and $\zeta \in \mathbb{Z}^+$ are arbitrary scale factors.

Here, based on our device measurement observations, we have particularly modelled the variance of noise inversely proportional to the original signal, where we set $\beta = 0$ and $\zeta = 1$ in (10). In this case, the equation $\lambda[n]$ increases by increasing the bin index n. The coefficient a should be a small positive constant to control the growth of shot noise fluctuations with a power increase almost linear during falling

photons in the tail of distribution. The simulation in Fig. 2 tries to mimic the noise behaviour for the tail of the fluorescence decay. Figure 3 visualises the proposed process for generating synthesised histogram of photon counts with an example.

3) Estimation under perturbation: For signals with a sufficiently large average of photons, e.g., $I > 10$, Poissonian noise distribution can be approximated well by a Gaussian noise function [28], where $\mathcal{P}(\lambda) \approx \mathcal{N}(\lambda, \lambda)$. Therefore, based on the central limit theorem [29], the distribution of the error term in (9) consisting of the superposition of discrete and continuous noises can be approximated by i.i.d. $e[n] \sim \mathcal{N}(0, \sigma_e^2)$ components [30]. A substitution gives corrupted measurements as:

$$\tilde{x}[n] = Ae^{-\frac{\Delta \cdot n}{\tau}} + e[n], \forall n = 0, 1, \cdots, N - 1. \qquad (11)$$

A statistical analysis on the model in (11) reveals that an efficient minimum variance unbiased estimator (MVUE) [31] does not exist for estimating τ. In this case, based on the Cramér-Rao lower bound (CRLB) analysis [32], the variance of the estimator is given by:

$$\mathrm{var}(\hat{\tau}) \geqslant \frac{\sigma_e^2 \tau^4}{A^2 \Delta^2 \sum_{n=0}^{N-1} n^2 e^{-\frac{2\Delta \cdot n}{\tau}}}, \qquad (12)$$

where the derivation details are provided in the Appendix. But, a MVUE does exist for estimating the amplitude given the true τ, and is given by:

$$\hat{A} = \frac{1 - e^{-\frac{2\Delta}{\tau}}}{1 - e^{-\frac{2\Delta \cdot N}{\tau}}} \sum_{n=0}^{N-1} e^{-\frac{\Delta \cdot n}{\tau}} \tilde{x}[n], \qquad (13)$$

with CRLB of:

$$\mathrm{var}(\hat{A}) \geqslant \frac{\left(1 - e^{-\frac{2\Delta}{\tau}}\right) \sigma_e^2}{1 - e^{-\frac{2\Delta \cdot N}{\tau}}}. \qquad (14)$$

We refer the reader to a similar derivation of this in the Appendix. Despite the lack of existence of a MVUE for τ, the theory still inspires the design of a novel estimator for the lifetime and the histogram amplitude, as described below.

It is important to note that the point corresponding to maximum difference does not always occur at zero and its location may generally be a function of a number of parameters such as the IRF, the pulse amplitude and duration of the lighting source, response speed of fluorescence to the

excitation pulse, the time bin width, photon detector speed, and any perturbation. Therefore, a little displacement to the right in the histogram of photon counts may occur. Nevertheless, the descending sort property still holds to some extent after a point $m*$, and the lifetime estimation problem can be reformulated by a Taylor expansion around this arbitrary point, $m* \geqslant 0$. Consider the M-point signal of $y[m], \forall m, M \leqslant N$ as a transformed version, including low-pass filtering and dimensionality reduction, of corrupted measurements $\tilde{x}[n]$; See (21) for a complete definition of the transformation. By defining the parameters $r_0 \triangleq y[0]$, $r_1 \triangleq y[m*+1]$, the ratio $r \triangleq \frac{r_0}{r_1}$, and taking natural logarithms from the first-order Taylor expansion equation, the lifetime is estimated as:

$$\hat{\tau} = \frac{\Delta \cdot (m*+1)}{\ln(r)}. \tag{15}$$

Even in noisy measurements, points in the vicinity of $m*$ are somewhat robust to shot noise, because starting points from the falling signal have inherently stronger intensities than other points in the tail of the photon-counts histogram.

However, instead of selecting a specific single point from $y[m]$ for obtaining each of the parameters r_0 and r_1, then inspired by the CRLB analysis in the Appendix, we propose an estimator to suppress these noise effects. To do this, we employed two left and right weighted average mechanisms of $y[m]$ around the central point $m*$, so that:

$$r_0 = \sum_{i=1}^{m*} \omega_{\mathrm{L}}[i]\, y[i], \tag{16}$$

in which $\omega_{\mathrm{L}}[i] = \frac{f_{\mathrm{L}}[i]}{\sum_{i=1}^{m*} f_{\mathrm{L}}[i]}, \forall i = 1, \dots, m*$ is a weighting function. To obey the trend of data which at first exhibits rise and then fall, we utilise the function $f_{\mathrm{L}}[i] \triangleq \frac{1}{2b} e^{-\frac{(m*-i)}{b}}, \forall i$ for the left side of $m*$, called the left exponentially growth weighting function. Similarly, for the right side of $m*$, let:

$$r_1 = \sum_{k=\mathcal{T}(m*+1)}^{\mathcal{T}(2m*)} \omega_{\mathrm{R}}[k] y[k], \tag{17}$$

where $\omega_{\mathrm{R}}[k] = \frac{f_{\mathrm{R}}[k]}{\sum_{k=\mathcal{T}(m*+1)}^{\mathcal{T}(2m*)} f_{\mathrm{R}}[k]}, \forall k = \mathcal{T}(m* + 1), \dots, \mathcal{T}(2m*)$ and $f_{\mathrm{R}}[k] \triangleq \frac{1}{2b} e^{-\frac{(k-m*)}{b}}, \forall k$ is called the right exponentially decaying weighting function. The parameter b denotes scale of the exponential functions, which is in relation to τ. In experiments, we set the optimal $b = 3$ by trial and error. On real FLIM measurements, effects such as outlier affect the position of $m*$ and hence demand a control mechanism on lower and upper bounds of (17). The operator $\mathcal{T}(\cdot)$ performs as a signal-length truncation function such that:

$$\mathcal{T}(m) = \min(m, M - 1). \tag{18}$$

The weights in (16) and (17) are normalized, so that $\sum_{i=1}^{m*} \omega_{\mathrm{L}}[i] = \sum_{k=\mathcal{T}(m*+1)}^{\mathcal{T}(2m*)} \omega_{\mathrm{R}}[k] = 1$. Figure 4 illustrates
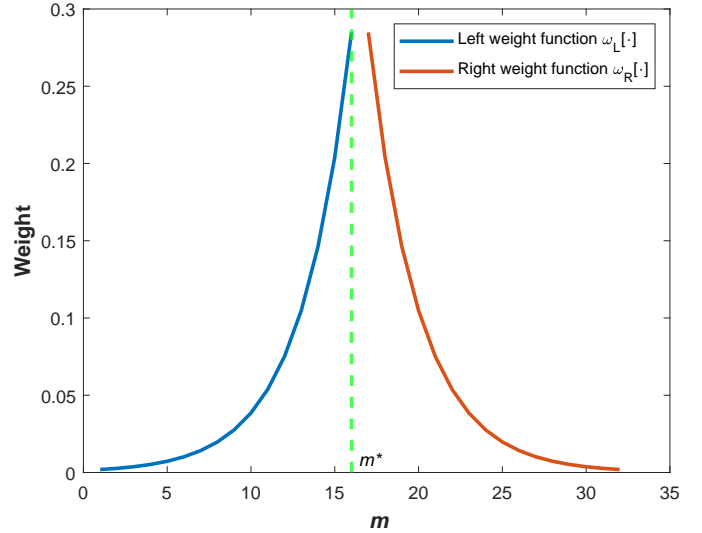

Fig. 4: An illustration of left and right weighting functions.

left and right weighting functions. Substituting (16) and (17) into (15) and simplifying by geometric series yields:

$$\hat{\tau} = \frac{\Delta \cdot (m*+1)}{\ln\left(\frac{\sum_k e^{-\frac{(k-m*)}{b}} \sum_i e^{-\frac{(m*-i)}{b}} y[i]}{\sum_i e^{-\frac{(m*-i)}{b}} \sum_k e^{-\frac{(k-m*)}{b}} y[k]}\right)} = \frac{\Delta \cdot (m*+1)}{\ln\left(\frac{c_\tau \sum_i e^{\frac{i}{b}} y[i]}{\sum_k e^{\frac{-k}{b}} y[k]}\right)}, \tag{19}$$

in which the constant $c_\tau \triangleq \frac{e^{-\frac{\mathcal{T}(m*+1)}{b}} - e^{-\frac{\mathcal{T}(2m*)+1}{b}}}{e^{\frac{m*}{b}} - 1}$. Also, the approximate amplitude is equal to $\hat{A} = r_0$, i.e.:

$$\hat{A} = \frac{\sum_i e^{-\frac{(m*-i)}{b}} y[i]}{\sum_i e^{-\frac{(m*-i)}{b}}} = c_A \sum_i e^{\frac{i}{b}} y[i], \tag{20}$$

where the constant $c_A \triangleq \frac{1 - e^{-\frac{1}{b}}}{e^{\frac{m*}{b}} - 1}$.

III. ALGORITHMS

The proposed FLIM parameter estimator is summarised in Algorithm 1. Due to applying the difference operator defined in (3), noise components are naturally amplified. Although decay parameters are usually estimated from the initial samples of the histogram curve, where the SNR is higher due to higher intensities in those time bins, pre-smoothing of the signal $\tilde{x}[n]$ by means of low-pass filtering mechanisms is required in practice to be able to alleviate perturbation effects. In this paper, to robustly recover the original fluorescence decay signal, two-step smoothing is used including a sequential S-G filter and temporal binning with an adaptive approach in a multi-resolution representation. Then, based on a game rule, the final decay signal is recovered from decisions made in the multi-resolution temporal space.

A. Multi-bin decay representation

Temporal binning of a histogram as a technique for information representation also concurrently performs dimensionality reduction and intrinsic smoothing tasks [11, 33]. However, finding an optimum bin size in histogram binning is generally

---

Algorithm 1 The proposed FLIM parameters estimator

1: **Inputs**: The M-point signal $\mathbf{y} = [y_0, y_1, \cdots, y_{M-1}]^{\mathrm{T}}$ and the bin width Δ.
2: Apply the difference operator in (3) on $y[m]$ to find $d[m]$.
3: Obtain the point with maximum fall m^* from (4).
4: Estimate lifetime from (19).
5: Estimate amplitude from (20).
6: **Outputs**: Decoded amplitude $\hat{A}$ and lifetime $\hat{\tau}$.

---

a challenging issue in statistics. Here, the bin size is simply defined as the number of bins in a histogram. On the one hand, if the bin size is very small, the appearance of a decay curve cannot be sketched well and the shape of the function may be too smooth. On the other hand, if the number of bins is too large, broken combs occur in the histogram of photo counts and a noisy representation of data is made. To address these problems, in this paper a multi-bin representation of the FLIM signal is proposed to be able to gain both smoothness and crispness of the signal without any temporal resolution loss. Naturally, higher-levels of temporal resolution represent crisper signals and lower ones expose smoother signals. To control levels of smoothness/crispness of a signal, we apply a S-G smoothing prior to histogram binning in an adaptive manner according to the degree of the inherent smoothing property in the binning mechanism at different resolution levels of signal.

1) Adaptive Savitzky-Golay filtering: S-G filter is a prominent type I finite impulse response signal smoother which uses a local LS polynomial approximation [34, 35]. Characteristics of interest of the S-G filter are functions of shape and peak preservation. Also, the cut-off frequency of the filter is proportional to the polynomial order, $f_c \propto O_f$, and inversely proportional to the window length of the filter, $f_c \propto \frac{1}{L_f}$. In practice, the polynomial order should be limited to a small number such as $O_f = 2$ to prevent ill-conditioned responses; and, for achieving a filtering process, the window length of the filter should satisfy the condition $L_f > O_f$. For adaptive filtering, at higher levels of temporal resolution, we use larger L_f to only pass very low frequencies. By moving to lower levels, we exponentially reduce the filter length to gradually widen the pass-band, so that mid- and finally high-frequency components are also passed. At lower levels of resolution, due to strong smoothing by binning itself, initial S-G pre-smoothing is not needed; thus, S-G filter will only mimic the signal and pass all frequencies. The behaviour of the adaptive filtering process is illustrated in Fig. 1, where the S-G filter is ineffective in the right branch, and, the same behaviour is seen for the binning stage in the left branch.

2) Adaptive temporal binning: Consider the functions of $\tilde{x}[n], \forall n$ as a N-sample histogram of photon counts and $g[n], \forall n$ as a S-G filtered version of the histogram, where $N = 2^l$ and l is the number of levels in the multi-bin representation of the decay signal. The data points at the ith level of binned space can be determined from the summing

formula of:

$$y_i[m] = \sum_{j=1}^{B_i} g[m \cdot B_i + j], \forall m = 0, 1, \ldots, M_i - 1, \quad (21)$$

where the function $y_i[m], \forall i = 1, \cdots, l$ represents a smoothed, dimensionality reduced version of the signal $g[n]$ and $M_i \triangleq 2^{(l-i+1)} \leqslant N$ is the number of time bins in the ith reduced dimension from the multi-bin representation. The parameter $B_i \triangleq \frac{N}{M_i}$ refers the number of consecutive time bins that are binned into a new one at the ith level.

3) Decision fusion: For each of the temporal resolutions, we can estimate amplitude and lifetime parameters based on Algorithm 1. However, estimated parameters at each level need to be rescaled to the original signal space. To do this for the ith level, we can multiply the estimated amplitude and lifetime values by B_i^{-1} and B_i, respectively. Specifically, rescaling the amplitude with B_i^{-1} assumes a uniform distribution which is not in agreement with a decaying trend. Instead, we use the scale s_i^{-1} for the amplitude, where:

$$s_i = \begin{cases} 1, & i = 1 \\ 2^{i-1-\gamma}, & i > 1 \end{cases}. \quad (22)$$

The bias γ accounts for an exponentially decaying distribution. In experiments, we set $\gamma = 0.25$ as an optimised value.

Initial estimates from diverse levels may encompass optimal parameters, where each level exposes its own properties. For instance, in large binning, uncertainty in estimating amplitude increases, due to combining more intensity values. This also means amplitudes are estimated well at higher levels of temporal binning. Conversely, lifetime estimation at top levels of binning will be generally more sensitive to perturbation, where lower levels of temporal resolutions are preferable. This brings a conflict of interest for a final decision making. Hence, a consensus strategy is suggested to synergically decode the original decay signal based on the game modelled below.

B. Game-theoretic modelling for decay recovery

Suppose the goal is to decide the best parameters for an optimal decay recovery based on the multi-bin histogram representation and corresponding initial estimates. Both of optimal parameters $A^\star$ and $\tau^\star$ are influential on the optimum selection. However, a single objective function is not available to consider both roles. Therefore, we suggest a game in which the problem is modelled by the "game" $\mathcal{G}(\upsilon, \mathcal{A}, d)$, where υ, $\mathcal{A}$, and d denote the number of players, the set of actions, and payoff function of the game, respectively. We consider the initial estimation behaviour of each of the amplitude and lifetime random variables as the players, i.e. there exists $\upsilon = 2$ players. The strategy of each player is to choose a set of known actions to be successful. Actions themselves are drawn from multiple temporal binning representations of histogram of photon-count arrivals, where we define the actions set $\mathcal{A} = \{\alpha_1, \alpha_2, \cdots, \alpha_k\}$, in which $k = l$ is the number of actions. The game space is shown in the decision fusion stage of Fig. 1 as a grid. The action α_i corresponds to $2^{(k-i+1)}$-dimensional temporal resolution from the multi-bin decay representation.

For the action i^{th}, a single-exponential decay can be recovered from FLIM data of a given pixel, i.e. the pair $(\hat{A}_i, \hat{\tau}_i)$. A Representation of actions taken by each player in the 2-D space $\Omega \in \mathbb{Z}^2$ constitutes a matrix called a reward matrix as $\mathbf{R} = [r_{ij}]$. For one repetition of the game, the matrix's entries takes 1 or 0 values, referring to prize or penalty, respectively.

Theorem 2 (Optimal decay recovery): Let the matrix $\mathbf{R} = [r_{ij}]$ be a reward matrix for the amplitude-lifetime game $\mathcal{G}(v, \mathcal{A}, d)$, where for a repetition of the game, the entry corresponding to an optimal parameter is 1 and other locations are 0. A place in the space Ω of the reward matrix exists by which the underlying optimal decay signal is recoverable.

Proof: Generally, Nash equilibrium proves that an equilibrium point exists for a game, where all rational players reach their maximum payoff [36]. To show that for the game $\mathcal{G}(v, \mathcal{A}, d)$, consider the equilibrium position as location of optimal estimates. Then, the optimal decay signal recovery is possible by players' payoff function optimisation. Here, optimality is constrained to quality of initial estimates of amplitude and lifetime parameters in the multi-bin decay representation. To detect optimal $A^\star$ and $\tau^\star$, we first measure distance between each recovered signal from the pair $(\hat{A}_i, \hat{\tau}_j), \forall i, j$ in the game space Ω and received corrupted measurements, and then, find the minimised one: see the blue star in the game space shown in Fig. 1 for illustration.

In terms of curve-like shape of an exponentially decaying function, we define two critical fixation points for the distance optimisation. The first one is connected to amplitude, which controls the initial position of the decay signal on the y-axis; and, another is related to lifetime, in which the steady-state location of the curve on the x-axis is controlled. It is difficult to seek a single objective function for finding the optimal pair $(A^\star, \tau^\star)$, in a manner that the metric concurrently copes with the two degrees of freedom well. Instead, we consider a specific distance measure for each player. Now, the problem is to search apt metrics for satisfying the two fixation points.

The scale of squared error at transition times of a decay curve is usually more than settling times. Hence, for fixing the amplitude, mean squared error (MES) criterion is apt as:

$$d^{\text{MSE}}(\tilde{\mathbf{x}}, \hat{\mathbf{f}}) \triangleq \frac{1}{N} \sum_{i=0}^{N-1} (\tilde{x}_i - \hat{f}_i)^2, \qquad (23)$$

where $\hat{\mathbf{f}} = [\hat{f}_0, \hat{f}_1, \cdots, \hat{f}_{N-1}]$ represents a single-exponential recovered signal in the space Ω. Conversely, for the tail of the decay curve, the scale is usually higher in terms of Neyman's chi-square test (CHI) metric χ^2 [37]. Therefore, it is appropriate for probing optimum lifetime. The distance χ^2 for the lifetime player is determined by:

$$d^{\text{CHI}}(\tilde{\mathbf{x}}, \hat{\mathbf{f}}) \triangleq \sum_{i=0}^{N-1} \frac{(\tilde{x}_i - \hat{f}_i)^2}{\tilde{x}_i}. \qquad (24)$$

It is worth to noting in case of observing division by zero in (24), we offset the effect by the trick of adding an extremely small number (e.g., 10^{-10}) to its denominator [38]. Assume vectors $\mathbf{d}^{\text{MSE}}$ and $\mathbf{d}^{\text{CHI}}$ contain MES and χ^2 distances ob-

---

Algorithm 2 The proposed single-exponential decay recovery

1: **Inputs**: The N-point histogram of photo-count arrivals $\tilde{\mathbf{x}}$ and the bin width Δ.
2: Set $O_{\text{f}} = 2$, $N_l = \log_2 N$, and $\mathbf{n} = [0, 1, \cdots, N-1]^{\text{T}}$.
3: Initialise $\mathbf{g} = \tilde{\mathbf{x}}$, $M = N$, and $k = 1$.
4: **while** $M \geqslant 2$ **do**
5: $\quad L_{\text{f}} = \frac{N}{2^k} + 1$
6: $\quad$ **if** $L_{\text{f}} > O_{\text{f}}$ **then**
7: $\quad\quad$ Smooth $\mathbf{g}$ by S-G filter with specs of O_{f} and L_{f}.
8: $\quad$ **end if**
9: $\quad$ Bin smoothed $\mathbf{g}$ into M bins by (21) to make a $\mathbf{y}$.
10: $\quad$ Estimate $\hat{A}$ and $\hat{\tau}$ by passing $\mathbf{y}$ to Algorithm 1.
11: $\quad \hat{\mathbf{A}}_k = (\frac{1}{s_k})\hat{A}$
12: $\quad \hat{\boldsymbol{\tau}}_k = (\frac{N}{M})\hat{\tau}$
13: $\quad$ Update $\mathbf{g} = \tilde{\mathbf{x}}$, $M \leftarrow \frac{M}{2}$ and $k \leftarrow k+1$.
14: **end while**
15: Reset counter $k = 1$.
16: **for** $i \leftarrow 1, N_l$ **do**
17: $\quad$ **for** $j \leftarrow 1, N_l$ **do**
18: $\quad\quad \hat{\mathbf{f}} = \hat{\mathbf{A}}_i e^{-\frac{\Delta}{\hat{\tau}_j}\mathbf{n}}$
19: $\quad\quad$ Measure distance d_k^{MSE} between $\tilde{\mathbf{x}}$ and $\hat{\mathbf{f}}$ by (23).
20: $\quad\quad$ Measure distance d_k^{CHI} between $\tilde{\mathbf{x}}$ and $\hat{\mathbf{f}}$ by (24).
21: $\quad\quad k \leftarrow k+1$
22: $\quad$ **end for**
23: **end for**
24: $A^\star = \underset{\hat{\mathbf{A}}, \hat{\boldsymbol{\tau}}}{\arg\min}(\mathbf{d}^{\text{MSE}})$
25: $\tau^\star = \underset{\hat{\mathbf{A}}, \hat{\boldsymbol{\tau}}}{\arg\min}(\mathbf{d}^{\text{CHI}})$
26: $\mathbf{f}^\star = A^\star e^{-\frac{\Delta}{\tau^\star}\mathbf{n}}$
27: **Output**: The recovered fluorescence decay signal $\mathbf{f}^\star$.

---

tained from the space Ω. Optimal influential parameters are selected from the minimisation of:

$$A^\star = \underset{\Omega}{\arg\min}(\mathbf{d}^{\text{MSE}}), \qquad (25)$$

$$\tau^\star = \underset{\Omega}{\arg\min}(\mathbf{d}^{\text{CHI}}). \qquad (26)$$

Finally, the optimal decay signal is recovered by:

$$f^\star[n] = A^\star e^{-\frac{\Delta \cdot n}{\tau^\star}}, n = 0, 1, \ldots, N-1. \qquad (27)$$

$\blacksquare$

Our robust recovery approach for a pixel is detailed in Algorithm 2. We term our method "Robust RLD".

Example 1: Consider the original decay signal in (2) with $A = 100$, $\tau = 3$ns, $\Delta = 468.8$ps, and $N = 32$. The signal was corrupted with the Gaussian blurring kernel $\mathbf{h} = [0.25, 0.5, 0.25]^{\text{T}}$ and noise parameters of $a = 0.05$ and $\sigma^2 = 25$. Figure 1 illustrates the whole process for a random run. For a million random plays of the game, Table I (a) reports accumulated reward matrix normalised on summation of prizes for each player. The Nash equilibrium point is bold in the table. To be able to compare detection performance of our model to that of ground truth decay, we have also repeated the game for the ideal case. For each round of the game, the prize entry in the ground truth reward matrix was determined as the place

in which the absolute difference between a ground truth value and corresponding multi-bin estimates is minimum. Outcomes are listed in Table I (b). Again, Nash equilibrium is bold in the table. Comparing bold numbers in Tables I (a) and (b) reveals the congruence of equilibria. Dominant success almost always occurs in upper triangle part of the reward matrices.

IV. EXPERIMENTS

In this section, using synthesised data with known ground truth, we evaluated the robustness of our method under different settings and circumstances through Monte Carlo simulation. We compared our proposed method to CMM, standard LS fit of MATLAB software, Poisson MLE, RLD, RLD-OW and FLERIT approaches. The effectiveness of the proposed method is also shown on real FLIM data.

A. Tests on synthesised data

1) Evaluation on different bin sizes: In this experiment, we changed the bin size of histogram of photon counts from 2^1 to 2^{10} for various decay signals and determined the performance of our method. The amplitude of functions was set in a manner that the number of photon counts per histogram remains approximately the same as $1,000$ photons/histogram before perturbation, where $A = \frac{N_{\mathrm{p}}}{\sum_{n=0}^{N-1} e^{-\frac{\Delta \cdot n}{\tau}}}$ in the deterministic model of (2). We adjusted the bin width to the formula $\Delta = \frac{5\tau}{N}$, which guarantees settling a decay curve into total measurement cycle for a pixel. All signals were first blurred with a Gaussian kernel of the length 3 and then noise was added with noise characteristics of $a = 0.1$ and $\mathcal{N}(0, 25)$. Figure 5 plots median lifetime, as a measure of bias or accuracy, vs bin size for three fluorescence decay functions with different ground truth lifetimes (ns) of 1, 2.5 and 4 ns. The median lifetime reported for each point of the plot was obtained for $100,000$ random runs. As seen from the figure, median lifetimes are inaccurate at initial points for the bin size of 2, where we have only a line segment for representing a fluorescence decay but not a desired curve. However, the estimations sharply improve with increasing bin size. Estimations reach their corresponding ground truth lines for bin sizes between 64 and 128. Meanwhile, the proposed method delivers appropriate estimations for $32 \leqslant N \leqslant 256$, where a sufficiently continuous decay curve apt for lifetime estimation exists. Due to photon-starved-like behaviours, an overestimation is seen for the bin size $1,024$. Hence, the bound can be considered as a user guide for clinical settings of instruments.

2) Effects of blurring and noise: As mentioned before, fluorescence decay signals are affected by perturbation of blurring and noise in practice. However, both the shape and level of signal's perturbation may differ from instrument to instrument. This experiment provides two simulation modes for analysing those effects. In the first simulation, for a fixed Gaussian blurring kernel with length 5, we evaluated statistics of estimated lifetimes from a corrupted signal under various shot and background noises. Parameters of the original signal were $N_{\mathrm{p}} \approx 2,600$, $A = 100$, $\tau = 1.5$ns, $\Delta = 58.6$ps and $N = 128$. The performance of our method is shown in Fig. 6
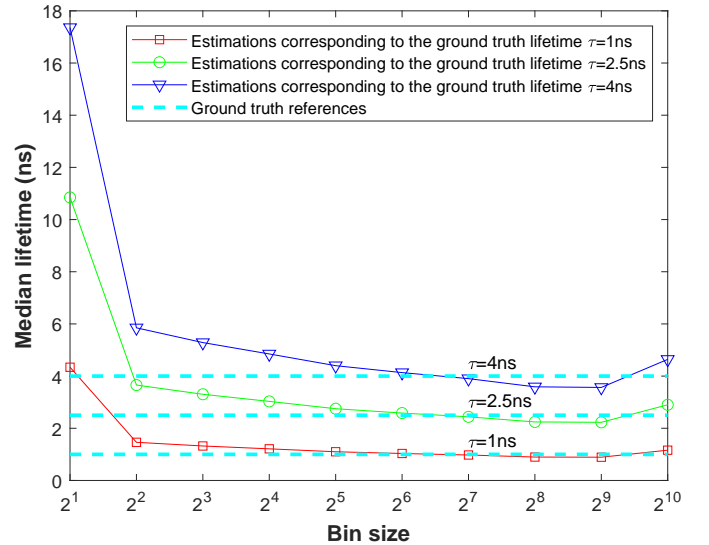

Fig. 5: The performance of our method for a diverse range of bin sizes from three lifetime-varying decay signals. For all the test points, the number of photon counts per histogram before perturbation is approximately the same as $1,000$ photons/histogram.

for $100,000$ random runs. The original signal encountered with a spectrum of perturbation including shot noises with both variable and constant rates ($\lambda = 10$) plus the background fluctuations. The curves related to median of lifetime show a stable behaviour of bias in our estimator even for severe noise. In this regard, the amount of averaged overestimation than the ground truth value is at most around 0.5ns in the worst-case scenario for the case $\mu = 0$ and $\lambda = 10$. As seen from Fig. 6, an increase at DC level of fluctuations leads to increasing level shift of the overestimation (bias). In terms of standard deviation, as a measure of precision, it remains under control with a tolerable rise during increasing noise power. The standard deviation of lifetime averaged over all noise powers is 0.2ns in the worst case. Results in $\sigma^2 = 0$ denote only the effect of shot noise in the absence of background fluctuations. Also, results of the single point $\sigma^2 = 0$ in the absence of shot noise (the case $\mu = 0$, $a = 0$) mean perturbation only by blurring.

In the second simulation, we fixed the parameter of shot noise on $a = 0.05$ and then evaluated mean and standard deviation of estimated lifetimes from a corrupted signal under different blurring kernel shapes and background noise powers. To model the IRF of a FLIM system, different blurring kernels may be utilised. Here, we employed impulse, box, approximate Gaussian and Airy smoothing functions with the same length of 7. The Airy IRF is the profile of a 2-D Airy pattern, which is approximately determined from the Bessel function of first kind of order [39]. It is notable that applying an impulse function means an ideal case without any blurring effect. Therefore, in that case, only the noise effect is present. Parameters of the original signal were the same as the first mode above. Figure 7 plots results for $100,000$ random runs. By increasing noise power, a decreasing trend is seen among

TABLE I: (a) Predicted and (b) ground truth normalised reward matrices for a million rounds of the game in Example 1

(a)

Player A \ Player τ	No binning	Binning with B = 2	Binning with B = 4	Binning with B = 8	Binning with B = 16
No binning	0.09 / 46.84	13.96 / 46.37	18.94 / 25.59	64.66 / 12.53	2.35 / 29.78
Binning with B = 2	5.64 / 41.01	26.07 / 14.39	33.26 / 13.05	34.78 / 10.96	0.25 / 15.45
Binning with B = 4	0.03 / 8.94	2.84 / 39.18	20.59 / 61.36	**76.51** / 76.51	0.03 / 54.33
Binning with B = 8	11.43 / 3.21	11.42 / 0.06	38.33 / 0	37.06 / 0	1.76 / 0.44
Binning with B = 16	0 / 0	0.36 / 0	22.53 / 0	74.51 / 0	2.6 / 0

(b)

Player A \ Player τ	No binning	Binning with B = 2	Binning with B = 4	Binning with B = 8	Binning with B = 16
No binning	0 / 0	4.16 / 15.94	19.38 / 4.94	75.62 / 6.15	0.84 / 1.97
Binning with B = 2	0 / 0	0.95 / 6.37	37.32 / 16.7	58.87 / 8.4	2.86 / 11.84
Binning with B = 4	0 / 0	1.43 / 77.59	21.73 / 78.29	74.26 / **85.33**	2.58 / 86.09
Binning with B = 8	0 / 0	1.38 / 0.1	16.22 / 0.07	80 / 0.12	2.4 / 0.1
Binning with B = 16	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0

Figure 6 (left plot): Median of lifetime (ns) / Standard deviation of lifetime (ns) vs Power of background Gaussian noise (σ^2). Legend:
- MED by $\mu=0$, a=0 (1.49) — SD by $\mu=0$, a=0 (.06)
- MED by $\mu=0$, a=.01 (1.54) — SD by $\mu=0$, a=.01 (.07)
- MED by $\mu=0$, a=.05 (1.55) — SD by $\mu=0$, a=.05 (.09)
- MED by $\mu=0$, a=.1 (1.62) — SD by $\mu=0$, a=.1 (.09)
- MED by $\mu=.5\sigma$, a=.01 (1.63) — SD by $\mu=.5\sigma$, a=.01 (.08)
- MED by $\mu=\sigma$, a=.01 (1.72) — SD by $\mu=\sigma$, a=.01 (.11)
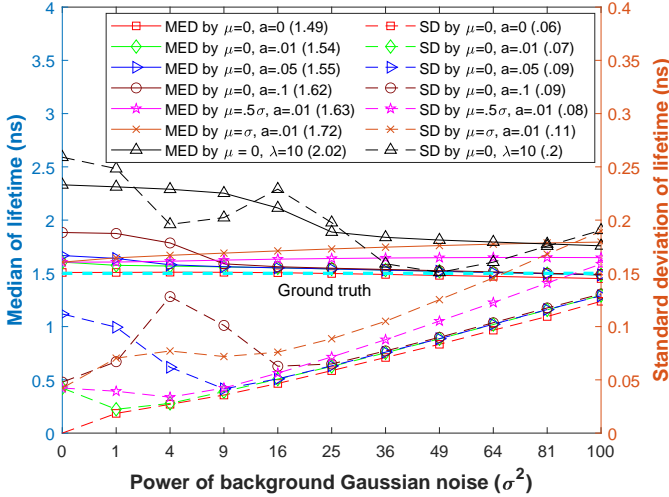- MED by $\mu=0$, $\lambda=10$ (2.02) — SD by $\mu=0$, $\lambda=10$ (.2)

Fig. 6: Evaluation of our method under noise variations. Numbers in parenthesis of the legend are average on all σ^2. The number of photon counts before perturbation is approximately 2,600.

Figure 7 (right plot): Estimated lifetime (ns) vs Blurring kernel (Impulse, Box, Gaussian, Airy). Legend:
- Background Gaussian noise with $\mu=0$ and $\sigma^2=4$
- Background Gaussian noise with $\mu=0$ and $\sigma^2=25$
- Background Gaussian noise with $\mu=0$ and $\sigma^2=64$
- Ground truth

Fig. 7: Evaluation of our method for various blurring kernels. The number of photon counts before perturbation is approximately 2,600.

estimations for all kernels. Responses show small deviations from the ground truth reference. In this regard, the lowest and highest absolute errors between the reference line and mean of lifetime averaged on different noise powers belong to Airy and box kernels, respectively. It is notable that throughout the experiments of this paper, we have not used any IRF cutoff, bin exclusion or deconvolution, where performing such procedures can potentially improve the quality of lifetime estimators.

3) Comparison to theoretical Cramér-Rao lower bound: This section investigates the variance of various estimators in comparison to the theoretical CRLB derived from Section II-B3 under the Gaussian noise assumption of $\mathcal{N}(0, \sigma_e^2)$, where bias analysis is provided in next sections as well. We used the same settings of experimental parameters applied for generating synthesised histograms of photon counts in the first simulation of Section IV-A2. Figures 8 (a) and (b) plot the variance of estimated lifetime and amplitude vs the variance of noise σ_e^2 for different compared approaches, respectively. The experiment repeated $10,000$ times for each σ_e^2. It is notable that in the implementation of LS fit, we utilised a nonlinear LS method of MATLAB that employs the "trust region" algorithm for optimisation [40, 41]. Also, for the Poisson MLE, we used an exhaustive 2-D search mechanism on possible amplitude and lifetime values to find those optimal parameters. It should be pointed out that the CMM method does not has any formulation for amplitude estimation; thus, it was ignored in Fig. 8 (b). As shown in the plots, variances of
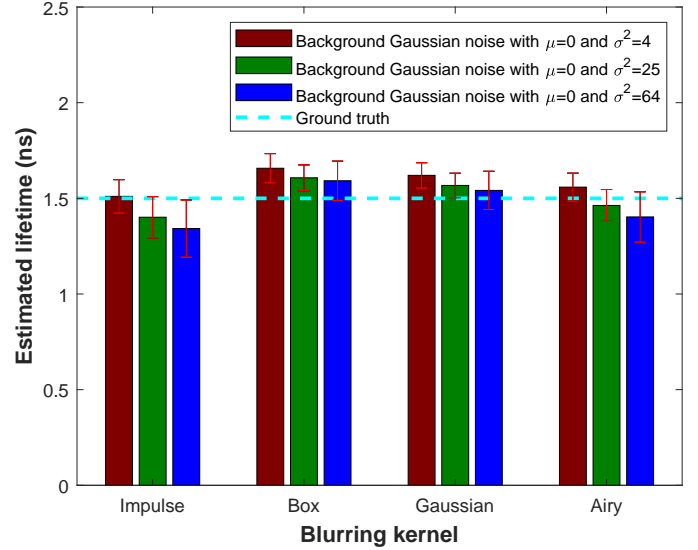
both lifetime and amplitude grow with increasing noise power. In terms of these variances, our proposed approach exposes a middle precision with under control variances, where the gold standard LS fit is ranked first; although, based on estimation bias, we will show in future experiments that the suggested method outperforms the LS fit on a vast range of variations.

4) Performance comparison under decreasing shot noise: In this experiment, we particularly evaluate the performance of a set of lifetime estimators under a decreasing shot noise. To do this, we set parameters $a = 0$, $\beta = 20$ and $\zeta = 1$ in (10). Other experimental settings are the same as those we used in Fig. 6. Table II reports (median, standard deviation) of different lifetime estimators for various background noises with the distribution $\mathcal{N}(\sigma, \sigma^2)$. Based on the ground truth lifetime of $\tau = 1.5$ns, our Robust RLD shows the best accuracy (the bold median) among all. CMM achieves the best rank in terms of precision (the bold standard deviation); but, it is biased. At the same time, the proposed approach remains robust and maintains the precision at an acceptable level. Also, a pair-wise comparison between the averages of (MED, STD) for the proposed method here, i.e., $(1.66, 0.09)$ in ns, and those obtained in Fig. 6 for the increasing shot noise with similar background noise, i.e., $(1.72, 0.11)$ in ns, reveals the increasing shot noise is the best worst case scenario of noise modelling though. The results indicate that the performance of Robust RLD is not heavily dependent on shot noise models used.
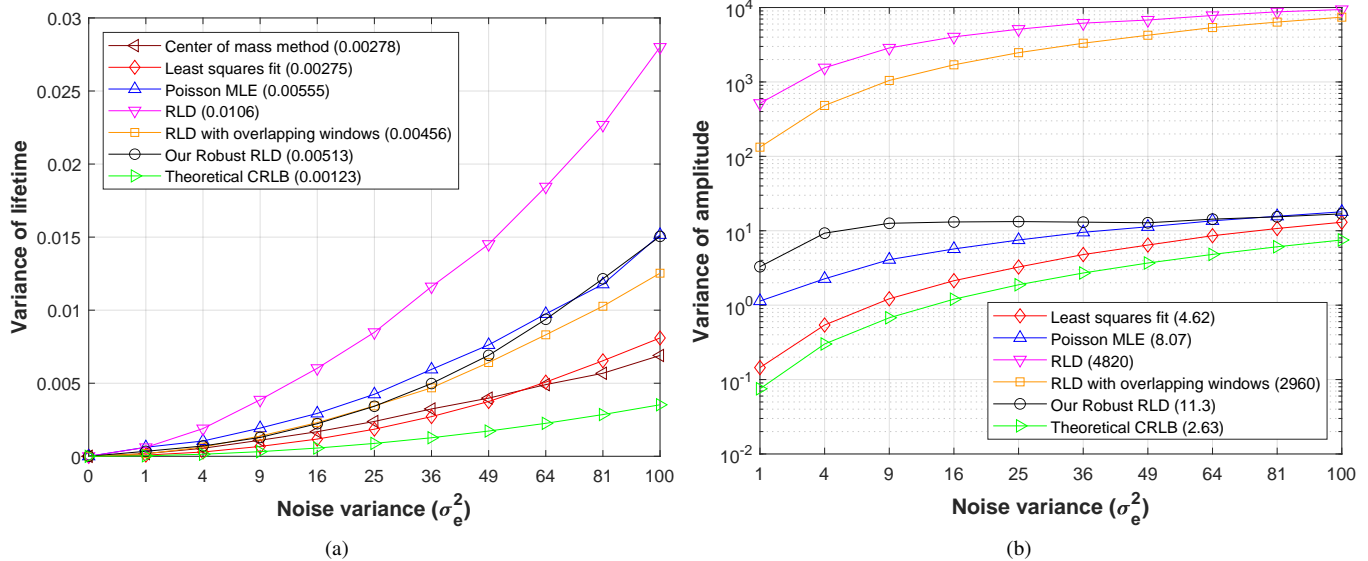
Fig. 8: Estimation variance vs noise variance of different approaches in comparison to theoretical Cramér-Rao lower bound for both parameters of (a) lifetime and (b) amplitude. Numbers in parenthesis of legends show mean on all noise variances. The number of photon counts before perturbation is approximately 2,600.

TABLE II: Comparisons of (median, standard deviation) for different approaches under a decreasing shot noise with the ground truth lifetime of $\tau = 1.5$ns

Method	(MED, STD) of lifetime in ns				
	$\sigma^2 = 1$	$\sigma^2 = 9$	$\sigma^2 = 25$	$\sigma^2 = 49$	$\sigma^2 = 81$
CMM	(1.61, **.02**)	(1.77, **.03**)	(1.9, **.04**)	(2.03, **.05**)	(2.14, **.06**)
LS fit	(1.64, .02)	(1.74, .03)	(1.86, .05)	(1.98, .07)	(2.12, .09)
RLD	(1.73, .03)	(1.99, .06)	(2.24, .09)	(2.48, .12)	(2.72, .16)
RLD-OW	(1.61, .02)	(1.74, .04)	(1.88, .06)	(2.01, .08)	(2.14, .11)
Robust RLD	(**1.59**, .06)	(**1.63**, .06)	(**1.67**, .08)	(**1.69**, .11)	(**1.72**, .14)

5) Performance comparison under lifetime variations: The goal of this experiment is to examine the performance of different approaches in short-, mid- and long-lifetime regimes. Figure 9 shows lifetime estimation statistics of various methods on three random signals with short, mid and long lifetimes of 0.5, 3 and 12 ns, respectively. Common parameters utilised for the signals were $N_p = 1,500$, $\Delta = 500$ps and $N = 64$. Perturbation characteristics were Gaussian blurring kernel of length 3 and noise parameters of $a = 0$, $\mathcal{N}(10, 100)$. Reported statistics were obtained for 10,000 random runs. As is shown in the box plots related to short- and mid-lifetime regimes in Fig. 9, our method outperforms other benchmark approaches. Our results are the nearest ones to the ground truth values, and for the short lifetime, the first quartile is aligned with the ground truth line. The optimised nonlinear LS fit ranks second for those regimes. For the long-lifetime regime, the lifetime value is too long, so that one can not see settling of a decay curve into the given time measurement window of 64×500ps $= 32$ns, which can be interpreted as an inappropriate experimental setting of the window due to the

small selection of the bin width. Contrary to the short- and mid-lifetime tests, CMM exposes the lowest bias than others in the special case of measurement window, where our estimator yet ranks second among all. It is important to note that all compared estimators were evaluated in a fair situation without compensating background components introduced by the bias due to the mean offset of the background noise, where such a mechanism can potentially alleviate the bias totally.

The FLERIT approach [11] mainly employs neighbouring pixels information of a central point in an image to construct the observation matrix for estimating the average lifetime. Hence, we separately compare the performance of lifetime estimation between FLERIT and our method on a synthesised image. To do this, we generated an image of dimensions 1024×1024 including three regions with various lifetimes as shown in the intensity map of Fig. 10. Ground truth lifetimes in ns for pixels inside regions of A, B and C are 10, 6 and 2, respectively. Amplitudes were set under a low photon count regime, so that they were $\frac{10}{3}$, 2 and $\frac{2}{3}$ for regions of A, B and C, respectively. Common characteristics of ground truth signals are $N = 32$ and $\Delta = 0.4$ns. The decay signals were corrupted by Gaussian noise of zero mean and 2.5×10^{-3} variance. In FLERIT, 3×3 neighbours and the number of consecutive merged bins equal to 4 were utilised. In Fig. 10, the left-side intensity map represents summation on all arrived photons, which is the same in FLERIT and our method for a fair comparison. The mid lifetime map visualises lifetime estimation performance. And, the right-side plot depicts lifetimes' histogram. Visual results demonstrate the superiority of our method over FLERIT. For regions A, B and C, Table III tabulates mean $\pm$ standard deviation from estimated lifetimes of FLERIT and our proposed method. The best results are bold in the table. Except for the mean of region C, the evaluation of numerical results demonstrates that the
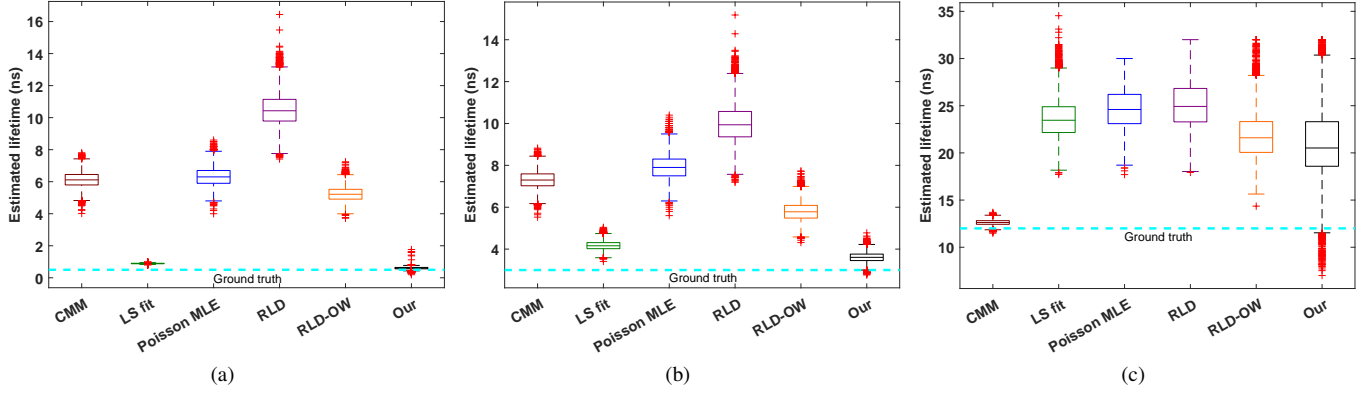
(a) (b) (c)

Fig. 9: Lifetime estimation statistics of various methods on (a) short-, (b) mid- and (c) long-lifetime regimes. For all plots, the number of photon counts before perturbation is approximately 1,500.
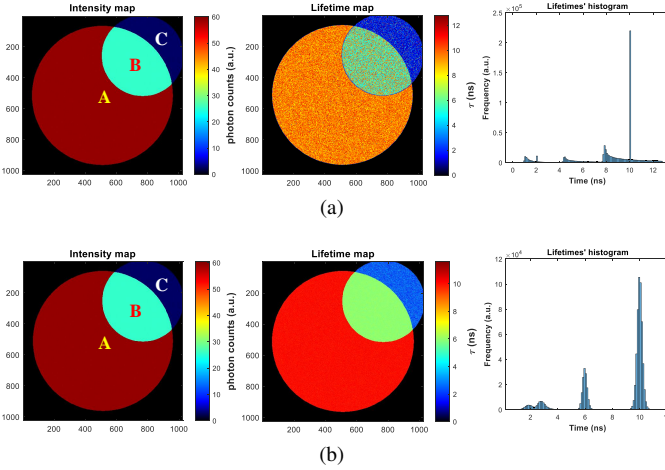

(a)

(b)

Fig. 10: Lifetime estimation performance comparison between (a) FLERIT and (b) our method on the same synthesised image. Ground truth lifetimes in ns for regions of A, B and C are 10, 6 and 2, respectively.

TABLE III: Numerical comparison of the methods in Fig. 10

Method	Estimated lifetime in ns ($\mu \pm \sigma$)		
	Region A	Region B	Region C
FLERIT [11]	9.51 ± 1.42	6.86 ± 2.42	$\mathbf{2.25 \pm 2.05}$
Our approach	$\mathbf{9.99 \pm 0.19}$	$\mathbf{5.97 \pm 0.17}$	2.52 ± 0.58
Ground truth lifetime (ns)	10	6	2

suggested approach brings desired lower bias and variance of lifetime estimates than those of FLERIT.

6) Comparison in various photon-count regimes: As mentioned earlier, SNR is the square root of the average number of arrived photons, I. This parameter is a function of amplitude. Hence, we change the amplitude to simulate low, mid and high photon count regimes, which are interpreted as different levels of SNR. To this intent, we considered three signals with amplitudes of 50, 250 and 1,000. We set other common parameters of signals as $\tau = 2$ns, $\Delta = 312.5$ps and $N = 32$. All signals were first blurred with a Gaussian kernel of

length 3 and then noise was added with characteristics of $a = 0.1$, $\mathcal{N}(0, 25)$. Those settings result in the values of I approximately equal to 10, 50 and 200 averaged on the measurement cycle for low, mid and high photon counts, respectively. For all of the regimes, Fig. 11 represents the mean and standard deviation of the estimated lifetimes from different approaches including our proposed methods in Algorithms 1 and 2. Due to the differentiation operator used in Algorithm 1, we have also evaluated the effect of pre-binning as a simple signal smoother before estimation. To do this, we reduced the number of bins from 32 to reasonably sized 8. For a fair comparison, the pre-binning was repeated for the standard LS fit. Statistics in the bar graph were obtained for $100,000$ times random run. For the challenging low photon-count situation, comparing results of different approaches to the ground truth reference reveals the lowest bias for the proposed Robust RLD with a significant difference from other competing approaches. In the regime, the percent mean bias of our approach is 7.55%, whereas, for the best competitor (RLD-WO), it is 35.13%, which shows 27.58% improvement. This specifically demonstrates the robustness of our method under severe perturbation, as seen in the example of Fig. 2. The accuracy of the estimator in Algorithm 1 is ranked second but in an uncontrolled variance, which is because of estimations beyond the measurement window caused by noise amplification. However, with an accuracy-precision trade-off, the pre-binning clearly improves the variance of the estimator in Algorithm 1. As shown in results of LS estimator with pre-binning, the smoothing also improves slightly the bias than the LS counterpart with no binning at the expense of increasing the standard deviation. Thanks to high levels of SNR for the high photon counts regime, most of the lifetime estimators expose close performance. Interestingly, the bias of Algorithm 1 is less than that of Algorithm 2 in this regime. This fact also highlights that for low noise levels, it may not be needed to employ more complex approaches for lifetime estimation. Meanwhile, results of Robust RLD show its appropriateness for utilising in a wide variety of photon-count regimes. Specifically, our approach makes better chance for robust estimation of the lifetime under photon starvation
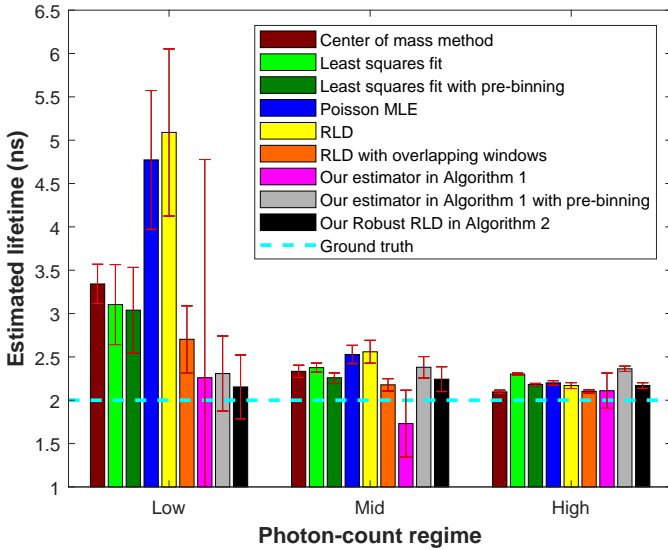
Fig. 11: Mean and standard deviation of estimated lifetimes of different approaches under various photon-count regimes. The number of photons per histogram is approximately 350, 1,800 and 6,900 for regimes of low, mid and high photon counts, respectively.

situations.

7) Comparison on decay recovery error: Although the lifetime is known as the most important parameter in FLIM, some fields like materials science deal with the amplitude information, too. However, as mentioned in Section IV-A3, all benchmark FLIM systems do not have the capability of amplitude estimation such as CMM. Or, if capable, they may work weakly under difficult situations like RLD and RLD-OW. This fact also means they are not effective for recovering a continuous or discrete version of a decay signal. Therefore, due to the best rank of LS fit for amplitude estimation, we have reported a recovery performance comparison between our method and that of LS fit in Table IV. For this experiment, we considered three signals with small, mid and large amplitudes of 20, 100 and 500, respectively. Other parameters of the three signals were $\tau = 5$ns, $\Delta = 781.3$ps and $N = 32$. Signals were first blurred with a Gaussian kernel of length 3 and then noise was added with characteristics of $a = 0.05$, $\mathcal{N}(0, 64)$. In Table IV, bold mean and standard deviation values represent the best performance. For the small amplitude, our method has lower bias than LS fit; and, in mid and large amplitudes, it follows the LS fit with little differences. The LS fit is inaccurate for lifetime estimation in case of the small amplitude, where signal perturbation is severe, whereas the proposed method is robust for diverse amplitudes. Regarding recovery error of χ^2, our method outperforms the LS fit for almost all cases.

B. Tests on real data

1) FLIM system: FLIM images were recorded using a confocal scanning imaging system based on previously reported work [42]. The system includes a 20MHz super-continuum laser filtered to produce excitation at 480nm. The system was used in conjunction with an optical imaging fibre bundle to allow for remote imaging. Light was guided down individual cores in the fibre and fluorescence collected from the same fibre core and directed onto a spectral, time resolved detector. The system can record between 2 and 512 spectral channels and between 2 and 32 temporal channels for each pixel in the image. For the data presented here, 2 spectral channels were used with an exposure time of 85μs per pixel and an image size of 128×128 pixels.

2) Experimental design: Neutrophil Activation Probe (NAP) [43] is an activatable fluorescent probe used for the detection of Human Neutrophil Elastase (HNE), an enzyme released by activated, pro-inflammatory neutrophils [44]. Consisting of three internally quenched fluorescein moieties each conjugated to a cleavable peptide sequence specific to HNE, NAP ($\lambda_{ex} = 488$nm and $\lambda_{em} = 525$nm) in its activated form amplifies its fluorescence signal due to the release of fluorescein. By incubating NAP with HNE, we sought to characterise its fluorescence lifetime properties and test the effects of Sivelestat, a specific inhibitor of HNE, as well as Nafamostat mesylate, an antiviral drug currently undergoing clinical trials as a potential treatment for coronavirus (COVID-19) (clinical trial identifier NCT04473053) as a potential inhibitor of HNE.

Another property of the experiment from the lifetime estimation viewpoint, is measuring fluorescence in a homogeneous solution as opposed to those of complex biological objects such as cells. This means that signals from each sample should be consistent, and we expect a uniform, flat lifetime map. Therefore, such partial information about samples' behaviour brings the ability to evaluate the performance of the proposed lifetime estimation method in real-world scenarios.

3) Samples preparation: 400nM HNE (HNE-L, Molecular Innovations, USA), diluted in HNE buffer (0.05M Na Acetate, pH 5.0 containing 0.1M NaCl), was incubated with 10μM Sivelestat (ONO-5046, Tocris, Bristol, UK) or 10μM Nafamostat mesylate (N0289, Sigma-Aldrich, UK) prior to the addition of 5μM NAP (diluted in phosphate buffered saline). Solutions were incubated for 15 minutes at $37°$C prior to fluorescence intensity/lifetime imaging. Experimental samples include 5μM NAP alone, 5μM NAP+400nM HNE, 5μM NAP+400nM HNE+10μM Sivelestat and 5μM NAP+400nM HNE+10μM Nafamostat mesylate, labelled respectively A, B, C and D for simplicity.

4) Sensing and outcomes: For each sample, microendoscopy was done using green and red spectral bands with wavelength ranges of (498nm $\sim$ 570nm) and (594nm $\sim$ 764nm), called bands 1 and 2, respectively. Each sample contains video sequences with N_f frames. Signal parameters are $N = 16$, $\Delta = 800$ps and the frames dimension of 128×128. Table V summarises statistics of samples for comparison of both our proposed and LS fit approaches side by side. Mean and standard deviation values were calculated on pixels inside the probe circle accumulated over all video frames of $f = 0, 1, \ldots, N_f - 1$; and, as a pattern of the plain specimens, Figs. 12 (a) and (b) show intensity and lifetime maps as well lifetimes' histogram from bands 1 and 2, 10th frame, sample C, respectively. The zigzag-wise patterns on intensity maps are mainly due to fibre-bundle artefacts.

TABLE IV: Decay recovery error of the proposed method than the best competitor among our compared methods

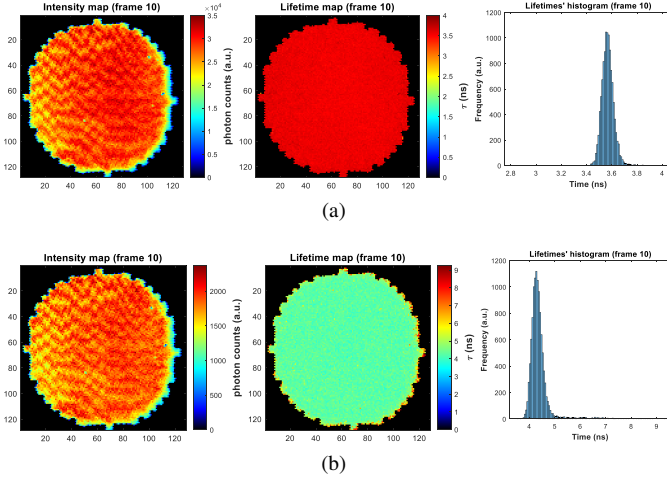| Method	Amplitude estimation $(\mu \pm \sigma)$			Lifetime estimation $(\mu \pm \sigma)$			Decay recovery error of χ^2 $(\mu \pm \sigma)$		
	$A = 20$	$A = 100$	$A = 500$	$A = 20$	$A = 100$	$A = 500$	$A = 20$	$A = 100$	$A = 500$
Least squares fit	$15.14 \pm \mathbf{4.66}$	$\mathbf{86.97} \pm \mathbf{5.08}$	$\mathbf{439.29} \pm \mathbf{5.05}$	15.35 ± 6.67	$6.24 \pm \mathbf{0.54}$	$5.77 \pm \mathbf{0.1}$	682.8 ± 625.9	44.52 ± 36.69	$62.17 \pm \mathbf{15.34}$
Our approach	$\mathbf{15.41} \pm 5.23$	86.55 ± 5.59	433.29 ± 7.3	$\mathbf{5.59} \pm \mathbf{3.48}$	$\mathbf{5.21} \pm 0.62$	$\mathbf{5.31} \pm 0.21$	$\mathbf{59.49} \pm 114.43$	$\mathbf{21.36} \pm \mathbf{19.49}$	$\mathbf{42.27} \pm 15.98$

Fig. 12: Results from left to right are intensity and lifetime maps, and lifetimes' histogram from (a) band 1 and (b) band 2, 10$^{\text{th}}$ frame, sample C in Table V, respectively.

Our obtained results are statistically significant for all samples. HNE activated NAP via the cleavage of peptide sequence to release fluorescein, resulting in increased fluorescence intensity and lifetime. 10μM Sivelestat reduced the fluorescence intensity; however, it had no effect on fluorescence lifetime due to the partial inhibition of HNE. This potentially highlights the limitation of single exponential fitting and the small number of time bins used as there was likely both cleaved and uncleaved probe in the signal, with the cleaved probe dominating in terms of amplitude as the uncleaved probe is quenched. The fitted lifetime is therefore dominated by the longer lifetime signal. No reduction in fluorescence intensity or lifetime was observed with 10μM Nafamostat mesylate, indicating that it is not an inhibitor of HNE. Due to the fluorescence emission spectrum properties of NAP, fluorescence was primarily detected in band 1. Thus, standard deviations of band 2 are greater than band 1 counterparts because of noise amplification, where the band 2 red spectrum has lower photon energy than green one and falls largely outside of the NAP emission band. There is, however, some evidence in band two to support the partial cleaving inhibition in sample B with an intermediate lifetime observed. The increased fluorescence lifetime signature of NAP makes it suitable for FLIM. In terms of LS fit results, the lifetime of different specimens exposes lower variance; however, they are affected again by the important issue of bias due to the underestimates.

V. COMPUTATIONAL COMPLEXITY ANALYSIS

Except the performance of a lifetime estimator, another important issue in FLIM is computational complexity/run-time of the estimator. In this regard, estimators with low complexity enable on-chip lifetime estimation capabilities, which lead to benefits from reducing data transfer rate between a sensor head and a distant computer to portability of a medical device. This section provides approximate computational complexity order of our algorithms in terms of the input size of bins number N, yielding an upper bound on their time complexity. We have done that task for methods presented in Section IV. Additionally, we have calculated run-time for different approaches as a complementary analysis, where all influential factors on the complexity are considered in a real scenario.

As shown in Fig. 1, the proposed Robust RLD algorithm consists of three processing stages at each level of the multi-temporal resolution space and a decision unit at the end. The processing steps include S-G filter, binning, and estimator. The input bin size is variable for each level; and, the complexity of the decider is negligible in comparison to the processing stages. Therefore, the computational complexity is practically bounded to processing applied at the level with the highest resolution. At this level, no binning exists and all photon time-stamps are individually recorded. The proposed estimator in Algorithm 1 has a liner complexity of the order $\mathcal{O}(N)$. Therefore, the dominant term is the complexity of S-G filter [45], i.e. the order of 1-D convolution as $\mathcal{O}(N \log_2 N)$. The compared algorithms expose two types of complexity, regardless of their implementation details. On the one hand, the complexity of analytic closed-form approaches of CMM, RLD, and RLD-OW grows linearly with increasing N, which is the same as the proposed estimator in Algorithm 1. On the other hand, LS fit, Poisson MLE, and FLERIT have all matrix decomposition-based solutions. Hence, they shows cubic order of complexity. Table VI summarises the complexity for different categories. Our proposed Algorithm 2 is ranked second among them with a linearithmic complexity between the fair order $\mathcal{O}(N)$ and the worst one $\mathcal{O}(N^3)$. Figure 13 reports the average run-time vs bin size for synthesised histograms having $\tau = 2.5$ns in experiments of Section IV-A1. For each bin size, we repeated the random experiment 1,000 times. In the exhaustive search-based Poisson MLE, which can be considered as the most complex optimiser, we set expected lower and upper bounds of 1 and 1,100 for the amplitude, respectively, with the resolution step of 4. Also, lower and upper bounds for the lifetime were respectively 1 and 10, with the step size of 0.2. All simulations were performed in MATLAB R2021b environment and run on an Intel Core i7 2.2GHz machine. The ranking of run-times justifies congruency of the two complexity analyses. On

TABLE V: Results of both our proposed and the least squares fit approaches side by side on real samples

Sample	Label	N_f	Spectral band 1 ($\mu \pm \sigma$)			Spectral band 2 ($\mu \pm \sigma$)		
			Intensity (a.u.)	Lifetime (ns)		Intensity (a.u.)	Lifetime (ns)	
				LS fit	Our method		LS fit	Our method
5μM NAP alone	A	19	13679 ± 2537	3.17 ± 0.03	3.23 ± 0.06	1083 ± 163	3.61 ± 0.22	5 ± 0.71
5μM NAP+400nM HNE	B	23	44922 ± 8475	3.39 ± 0.03	3.57 ± 0.09	2844 ± 547	3.65 ± 0.19	4.19 ± 0.27
5μM NAP+400nM HNE+10μM Sivelestat	C	31	27060 ± 5005	3.35 ± 0.02	3.57 ± 0.05	1753 ± 298	3.75 ± 0.16	4.39 ± 0.46
5μM NAP+400nM HNE+10μM Nafamostat mesylate	D	28	43389 ± 8822	3.36 ± 0.03	3.53 ± 0.07	2658 ± 566	3.64 ± 0.12	4.19 ± 0.3

TABLE VI: Computational complexity of different algorithms

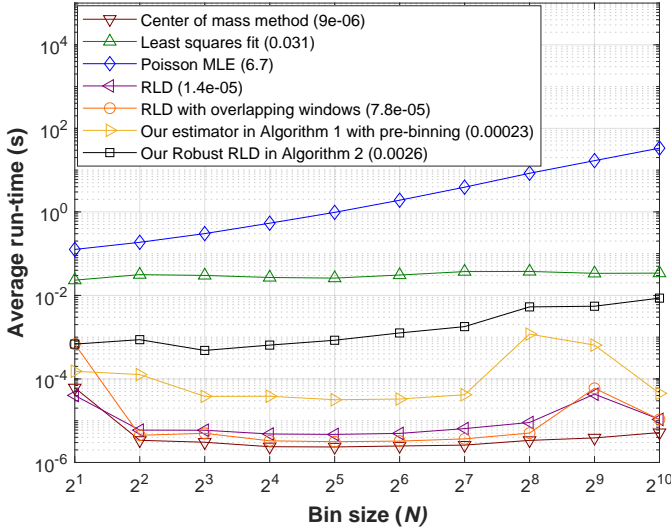| Algorithm(s)	Time complexity
CMM, RLD, RLD-OW, and our estimator in Algorithm 1	$\mathcal{O}(N)$
LS fit, Poisson MLE, and FLERIT [11]	$\mathcal{O}(N^3)$
Our Robust RLD in Algorithm 2	$\mathcal{O}(N \log_2 N)$

Fig. 13: Average run-time vs bin size for various methods. Numbers in parenthesis of the legend show averaging on all bin sizes.

real images in specimens of Section IV-B, our robust method processed data on average about 63 times faster than the gold standard LS fit. To sum up, the analyses demonstrate that both of our algorithms are suitable options for hardware-level implementations.

VI. CONCLUSION

This paper presented a robust and computationally efficient fluorescence decay signal recovery algorithm in the presence of inevitable blurring and noise occurring during time-resolved single photon sensing. The proposed framework first provides a multi-bin decay representation from the histogram of photon count arrivals using an adaptive signal smoothing approach. Subsequently, each representation is fed to our lifetime decoding algorithm. Finally, estimated parameters from different temporal resolutions are fused to each other based on game-theoretic modelling. We theoretically proved that our method

is capable of recovering optimal fluorescence decay robustly under a wide variety of imaging situations.

In addition to being robust, due to the non-fitting-based nature of our method, it can be considered as a rapid approach for hardware-level realisation. This capability is of great importance in real-time fluorescence lifetime sensing systems such as in vivo, in situ microendoscopy.

APPENDIX
DERIVATION OF CRAMÉR-RAO LOWER BOUND AND MINIMUM VARIANCE UNBIASED ESTIMATOR

Starting from the corrupted observations model represented in (11), likelihood function for the given τ is:

$$f_{\tilde{X}}(\tilde{x}|\tau) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma_e^2}} e^{-\frac{1}{2\sigma_e^2}\left[\tilde{x}[n] - Ae^{-\frac{\Delta \cdot n}{\tau}}\right]^2}. \quad (28)$$

The first-order partial derivative of the log-likelihood gives:

$$\frac{\partial \ln f_{\tilde{X}}(\tilde{x}|\tau)}{\partial \tau} = \frac{A\Delta}{\tau^2\sigma_e^2} \sum_{n=0}^{N-1} \left[ne^{-\frac{\Delta \cdot n}{\tau}} \tilde{x}[n] - Ane^{-\frac{2\Delta \cdot n}{\tau}} \right]. \quad (29)$$

Repeating the differentiation yields:

$$\frac{\partial^2 \ln f_{\tilde{X}}(\tilde{x}|\tau)}{\partial \tau^2} = \frac{A\Delta^2}{\sigma_e^2\tau^4} \sum_{n=0}^{N-1} \left[n^2 e^{-\frac{\Delta \cdot n}{\tau}} \tilde{x}[n] - 2An^2 e^{-\frac{2\Delta \cdot n}{\tau}} \right]. \quad (30)$$

The estimation variance satisfies $\mathrm{var}(\hat{\tau}) \geqslant \mathrm{CRLB}$, where:

$$\mathrm{CRLB} \triangleq \frac{1}{-\mathbb{E}\left[\frac{\partial^2 \ln f_{\tilde{X}}(\tilde{x}|\tau)}{\partial \tau^2}\right]} = \frac{\sigma_e^2\tau^4}{A^2\Delta^2 \sum_{n=0}^{N-1} n^2 e^{-\frac{2\Delta \cdot n}{\tau}}}. \quad (31)$$

Now, the condition of existence of a MVUE is to check $\psi(\tau)(\hat{\tau} - \tau) = \frac{\partial \ln f_{\tilde{X}}(\tilde{x}|\tau)}{\partial \tau}$ for a function like $\psi(\tau)$ as:

$$\psi(\tau)(\hat{\tau} - \tau) = \frac{A^2\Delta}{\sigma_e^2\tau^3} \sum_{N=0}^{N-1} ne^{-\frac{2\Delta \cdot n}{\tau}} \left\{ \frac{\tau \sum_{n=0}^{N-1} ne^{-\frac{\Delta \cdot n}{\tau}} \tilde{x}[n]}{A \sum_{n=0}^{N-1} ne^{-\frac{2\Delta \cdot n}{\tau}}} - \tau \right\}, \quad (32)$$

which $\hat{\tau} \triangleq \frac{\tau \sum_{n=0}^{N-1} ne^{-\frac{\Delta \cdot n}{\tau}} \tilde{x}[n]}{A \sum_{n=0}^{N-1} ne^{-\frac{2\Delta \cdot n}{\tau}}}$ shows the MVUE does not exist due to being a function of the true parameter τ. ∎

REFERENCES

[1] A. Perperidis, K. Dhaliwal, S. McLaughlin, and T. Vercauteren, "Image computing for fibre-bundle endomicroscopy: A review," *Med Image Anal*, vol. 62, p. 101620, 2020.

[2] C. Tardif, G. Nadeau, S. Labrecque, D. Côté, F. Lavoie-Cardinal, and P. De Koninck, "Fluorescence lifetime imaging nanoscopy for measuring Förster resonance energy transfer in cellular nanodomains," *Neurophotonics*, vol. 6, no. 1, p. 015002, 2019.

[3] R. Datta, T. M. Heaster, J. T. Sharick, A. A. Gillette, and M. C. Skala, "Fluorescence lifetime imaging microscopy: Fundamentals and advances in instrumentation, analysis, and applications," *J. Biomed. Opt.*, vol. 25, no. 7, p. 071203, 2020.

[4] C. Poudel, I. Mela, and C. F. Kaminski, "High-throughput, multiparametric, and correlative fluorescence lifetime imaging," *Methods Appl. Fluoresc.*, vol. 8, no. 2, p. 024005, 2020.

[5] A. T. Erdogan, R. Walker, N. Finlayson, N. Krstajić, G. Williams, J. Girkin, and R. Henderson, "A CMOS SPAD line sensor with per-pixel histogramming TDC for time-resolved multispectral imaging," *IEEE J. Solid-State Circuits*, vol. 54, no. 6, pp. 1705–1719, 2019.

[6] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, vol. 6, no. 2, pp. 418–445, 1996.

[7] A. Chessel, F. Waharte, J. Salamero, and C. Kervrann, "A maximum likelihood method for lifetime estimation in photon counting-based fluorescence lifetime imaging microscopy," in *IEEE 21st Eur Signal Process Conf*, 2013, pp. 1–5.

[8] R. M. Ballew and J. Demas, "An error analysis of the rapid lifetime determination method for the evaluation of single exponential decays," *Anal. Chem.*, vol. 61, no. 1, pp. 30–33, 1989.

[9] K. K. Sharman, A. Periasamy, H. Ashworth, and J. Demas, "Error analysis of the rapid lifetime determination method for double-exponential decays and new windowing schemes," *Anal. Chem.*, vol. 71, no. 5, pp. 947–952, 1999.

[10] D. U. Li, D. Tyndall, R. Walker, J. A. Richardson, R. K. Henderson, J. Arlt, D. Stoppa, and E. Charbon, "Video-rate fluorescence lifetime imaging camera with CMOS single-photon avalanche diode arrays and high-speed imaging algorithm," *J. Biomed. Opt.*, vol. 16, no. 9, p. 096012, 2011.

[11] H. Yu, R. Saleeb, P. Dalgarno, and D. D.-U. Li, "Estimation of fluorescence lifetimes via rotational invariance techniques," *IEEE. Trans. Biomed. Eng.*, vol. 63, no. 6, pp. 1292–1300, 2015.

[12] M. A. Digman, V. R. Caiolfa, M. Zamai, and E. Gratton, "The phasor approach to fluorescence lifetime imaging analysis," *Biophys. J.*, vol. 94, no. 2, pp. L14–L16, 2008.

[13] G. Wu, T. Nowotny, Y. Zhang, H.-Q. Yu, and D. D.-U. Li, "Artificial neural network approaches for fluorescence lifetime imaging techniques," *Opt. Lett.*, vol. 41, no. 11, pp. 2561–2564, 2016.

[14] J. T. Smith, R. Yao, N. Sinsuebphon, A. Rudkouskaya, N. Un, J. Mazurkiewicz, M. Barroso, P. Yan, and X. Intes, "Fast fit-free analysis of fluorescence lifetime imaging via deep learning," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 116, no. 48, pp. 24 019–24 030, 2019.

[15] P. Pande and J. A. Jo, "Automated analysis of Fluorescence Lifetime Imaging Microscopy (FLIM) data based on the Laguerre deconvolution method," *IEEE. Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 172–181, 2010.

[16] S. P. Chan, Z. Fuller, J. Demas, and B. DeGraff, "Optimized gating scheme for rapid lifetime determinations of single-exponential luminescence lifetimes," *Anal. Chem.*, vol. 73, no. 18, pp. 4486–4490, 2001.

[17] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.

[18] G. Wu, T. Nowotny, Y. Chen, and D. D.-U. Li, "GPU acceleration of time-domain fluorescence lifetime imaging," *J. Biomed. Opt.*, vol. 21, no. 1, p. 017001, 2016.

[19] M. I. Rowley, A. C. Coolen, B. Vojnovic, and P. R. Barber, "Robust Bayesian fluorescence lifetime estimation, decay model selection and instrument response determination for low-intensity FLIM imaging," *PLoS One*, vol. 11, no. 6, p. e0158404, 2016.

[20] O. DeGuchy, L. Adhikari, A. Kim, and R. F. Marcia, "Photon-limited fluorescence lifetime imaging microscopy signal recovery with known bounds," in *IEEE 7th Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process.*, 2017, pp. 1–5.

[21] A. Le Marois, S. Labouesse, K. Suhling, and R. Heintzmann, "Noise-corrected principal component analysis of fluorescence lifetime imaging data," *J. Biophotonics*, vol. 10, no. 9, pp. 1124–1133, 2017.

[22] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT Press, 1994.

[23] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.

[24] Y.-I. Chen, Y.-J. Chang, S.-C. Liao, T. D. Nguyen, J. Yang, Y.-A. Kuo, S. Hong, Y.-L. Liu, H. G. Rylander, S. R. Santacruz *et al.*, "Generative adversarial network enables rapid and robust fluorescence lifetime image analysis in live cells," *Commun. Biol.*, vol. 5, no. 1, pp. 1–11, 2022.

[25] M. Y. Berezin and S. Achilefu, "Fluorescence lifetime measurements and biological imaging," *Chem. Rev.*, vol. 110, no. 5, pp. 2641–2684, 2010.

[26] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data," *IEEE Trans Image Process*, vol. 17, no. 10, pp. 1737–1754, 2008.

[27] R. L. McCreery, "Chapter 4- Signal-to-noise in Raman spectroscopy," in *Raman spectroscopy for chemical analysis*. John Wiley Sons, Inc., 2000, pp. 49–71.

[28] S. W. Hasinoff, "Photon, Poisson noise," in *Computer vision: A reference guide*, K. Ikeuchi, Ed. Boston, MA: Springer US, 2014, pp. 608–610.

[29] P. Billingsley, *Probability and measure*, 3rd ed. John Wiley & Sons, 1995.

[30] J. Liu, Y. Sun, J. Qi, and L. Marcu, "A novel method for fast and robust estimation of fluorescence decay dynamics using constrained least-squares deconvolution with Laguerre expansion," *Phys. Med. Biol.*, vol. 57, no. 4, p. 843, 2012.

[31] S. M. Kay, *Fundamentals of statistical signal processing, Volume I: Estimation theory*. Prentice-Hall, Inc., 1993.

[32] D. Bouchet, V. Krachmalnicoff, and I. Izeddin, "Cramér-Rao analysis of lifetime estimations in time-resolved fluorescence microscopy," *Opt. Express*, vol. 27, no. 15, pp. 21 239–21 252, 2019.

[33] A. J. Walsh, J. T. Sharick, M. C. Skala, and H. T. Beier, "Temporal binning of time-correlated single photon counting data improves exponential decay fits and imaging speed," *Biomed. Opt. Express*, vol. 7, no. 4, pp. 1385–1399, 2016.

[34] A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures." *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[35] R. W. Schafer, "What is a Savitzky-Golay filter?[lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, 2011.

[36] J. F. Nash, "Equilibrium points in n-person games," *PNAS*, vol. 36, no. 1, pp. 48–49, 1950.

[37] X. Ji, W. Gu, X. Qian, H. Wei, and C. Zhang, "Combined Neyman–Pearson chi-square: An improved approximation to the Poisson-likelihood chi-square," *Nucl. Instrum. Methods. Phys. Res. A*, vol. 961, p. 163677, 2020.

[38] "IEEE standard for floating-point arithmetic," *IEEE Std 754-2019 (Revision of IEEE 754-2008)*, pp. 1–84, 2019.

[39] "Chapter 8- Linewidth and writing resolution," in *Three-dimensional microfabrication using two-photon polymerization*, ser. MNT, T. Baldacchini, Ed. Oxford: William Andrew Publishing, 2016, pp. 190–220.

[40] D. C. Sorensen, "Newton's method with a model trust region modification," *SIAM J Numer Anal*, vol. 19, no. 2, pp. 409–426, 1982.

[41] A. R. Conn, N. I. Gould, and P. L. Toint, *Trust region methods*. SIAM, 2000.

[42] G. O. Williams, E. Williams, N. Finlayson, A. T. Erdogan, Q. Wang, S. Fernandes, A. R. Akram, K. Dhaliwal, R. K. Henderson, J. M. Girkin *et al.*, "Full spectrum fluorescence lifetime imaging with 0.5 nm spectral and 50 ps temporal resolution," *Nat. Commun*, vol. 12, no. 1, pp. 1–9, 2021.

[43] N. Avlonitis, M. Debunne, T. Aslam, N. McDonald, C. Haslett, K. Dhaliwal, and M. Bradley, "Highly specific, multi-branched fluorescent reporters for analysis of human neutrophil elastase," *Org. Biomol. Chem.*, vol. 11, no. 26, pp. 4414–4418, 2013.

[44] T. H. Craven, T. Walton, A. R. Akram, E. Scholefield, N. McDonald, A. D. Marshall, D. C. Humphries, B. Mills, T. A. Campbell, A. Bruce *et al.*, "Activated neutrophil fluorescent imaging technique for human lungs," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, 2021.

[45] R. Miralles and R. Molina, "A comparative study of the impulsive noise reduction algorithms in ultrasonic B-scans," in *IEEE 17th Eur SignalProcess Conf*, 2009, pp. 1577–1581.