Research Article | Vol. 33, No. 11 / 2 Jun 2025 / Optics Express 22270
Optics EXPRESS

Speed-optimization strategies for time-correlated single photon counting experiments

ALESSANDRO COMINELLI,[1] GIULIA ACCONCIA,[2,*] AND IVAN RECH[2]

[1]Tecnosens S.p.A., Via Vergnano 16, Brescia 25125, Italy
[2]Department of Electronics, Information, and Bioengineering of the Politecnico di Milano, Piazza Leonardo da Vinci 32, Milan 20133, Italy
*giulia.acconcia@polimi.it

Abstract: Time-correlated single photon counting (TCSPC) is recognized as a gold-standard technique in many fields, including life science and remote sensing applications. In particular, it is the enabling technology to record extremely fast luminous signals with accuracy down to a few picoseconds. However, setup limits like the presence of pile-up and dead-time-related phenomena impose a strong limitation to the speed of TCSPC experiments. In particular, the average number of recorded events is typically limited below 0.05 photons per excitation period, and also in those cases where the sample can emit a significantly higher photon rate. Many efforts have been done in the past few years to maximize speed, including the design of fast detectors and associated electronics, with nanosecond and sub-nanosecond dead time and multi-hit capabilities. Moreover, new advanced techniques have been proposed to get rid of distortion phenomena, thus enabling unprecedented speed. The goal of this paper is to provide a deep understanding of the ultimate limits to measurement speed, considering the best solutions presented so far. A general step-by-step methodology is proposed to calculate maximum speed in different conditions and to select the best suited technology to push speed to the edge depending on specific operation cases.

1. Introduction

Time-correlated single photon counting (TCSPC) is a key measurement technique every time an application requires the intensity time profile of an extremely fast luminous signal with high accuracy and resolution, down to the picosecond time scale [1,2]. In this regard, TCSPC has been the enabling technology for many powerful analyses in life science, like fluorescence lifetime imaging microscopy (FLIM), Förster resonance energy transfer (FRET) or fluorescence lifetime correlation spectroscopy (FLCS) [3–6]. Moreover, in the industrial and military sphere, it is becoming a gold standard in remote sensing applications, like light detection and ranging (LIDAR) [7,8]. Unfortunately, the metrological advantages of TCSPC come along with a relatively high acquisition time, which limits its exploitation every time fast response is required, for instance when measured phenomena change quickly over time.

In particular, setup limitations such as dead time or classic pile-up (discussed in more detail below) limit the maximum rate of events that the system can record and thus impose a speed limitation, which is particularly evident in experiments involving the generation of high photon rates. For example, in clinical applications such as in vivo brain surgery, classical TCSPC setups are not suitable due to their inability to handle high photon rates [9]. Another example of an experiment where the photon rate can be relatively high for classical TCSPC is biological imaging using a single-pixel camera [10–12]. In this case, a whole image is reconstructed using a single pixel, so the whole light coming from the analyzed scene is measured by means of a single TCSPC channel and incoming light can be relatively high [12].

#550769 https://doi.org/10.1364/OE.550769
Journal © 2025 Received 7 Dec 2024; revised 16 Feb 2025; accepted 22 Feb 2025; published 20 May 2025

Optics EXPRESS

In a classic TCSPC experiment, a pulsed laser source illuminates a target under test, which reemits photons in response to each pulse. This can happen, for example, due to reflection (e.g. in LIDAR) or due to light-matter interaction phenomena (e.g. in fluorescence experiments). At this point, a single-photon detector records photons re-emitted by the target, while a dedicated time-measurement electronics measures the arrival time of the photon within the laser period, which is the time interval between the laser pulse and the instant when the photon is detected. Finally, each time a photon is recorded the system stores the event into a histogram, depending on the measured arrival time. After many periods, the result is a histogram with the same profile of the average light signal emitted by the target. In other words, the histogram is proportional to the probability distribution of photon arrivals over time. Finally, a fitting algorithm permits to extract physical quantities of interest from the histogram (e.g. fluorescence decay time-constants in FLIM experiments) with picosecond accuracy.

Considering the statistical nature of TCSPC, collecting many photons is required in order to get a histogram with a good signal-to-noise ratio. In many cases, e.g. in fluorescence experiments involving a mono-exponential decay, a few thousand photons is the minimum quantity to obtain satisfactory results and in some critical cases some hundred thousand photons are required to get a reliable result, for example when resolving a double exponential decay in which the lifetime components are not well apart from each other [13].

Unfortunately, no real system is able to detect every photon coming from the sample, so the time needed to collect such high numbers of photons in the histogram can be relatively high.

In particular, not all the photons generated by the sample move to the active area of the detector. Indeed photon generation can have a large emission angle (even the whole solid angle if the process is isotropic) and in many cases only a small portion of the emission angle is recorded. In addition, detectors can have a limited fill-factor, which is a typical case of detector arrays. In some cases, the design of a dedicated optical system can help reducing photon losses; anyway, it is always impossible to catch every photon. Additionally, any detector has a photon detection probability (PDP) lower than one, meaning that even when a photon impinges on the active area there is a probability that it does not trigger any signal.

Anyway, even considering ideal collection of photons due to aforementioned issues, two additional effects concur in limiting the maximum operating frequency: classic pile-up and counting loss due to the dead time of detector and conversion electronics. Due to these effects, every time the systems records a photon it goes blind for a certain time. It follows that the higher the photon rate the longer the time the system is blind. This effect leads to an efficiency drop at high rates, meaning that the rate of recorded counts, so the measurement speed, does not increase proportionally to the number of photons impinging on the detector. Even worse, operating the system at high rates can lead to a significant distortion of the recorded histogram (thus impairing the effectiveness of fitting algorithms) [2,14]. It follows that, in most cases, the need to collect a certain number of photons, combined with all the non-idealities of the setup, gives rise to a difficult balance of parameters to get the best achievable result. On the one hand, increasing the photon rate permits to observe short-lived phenomena, on the other hand it gives rise to higher distortion.

The scope of this paper is to analyze how much the speed of a TCSPC experiment can be pushed to the edge, with a proper sizing and design of the measurement setup. Before entering into the details of setup design, the major limits of photon counting technology will be briefly reviewed in the following. In addition, some useful expressions will be derived to quantify both efficiency and distortion in the major cases of interest, including classic TCSPC and some advanced techniques to minimize distortion, which we proposed in the last few years. Finally, the paper will provide a step-by-step methodology to select the best measurement technique and to find the correct balance between parameters, to get the maximum achievable speed.

Optics EXPRESS

2. Main limits of TCSPC

A simple example to quantify the main limits of a TCSPC setup is the case of fluorescence experiments, where light intensity has the profile of a decaying mono-exponential pulse. In general, the mono-exponential decay is an effective model to describe many cases where the optical signal is a pulse with a relatively short duration, that is the typical case in many TCSPC applications. The main advantage of this model, is that it is possible to draw quantitative conclusions about the impact of system non-idealities relying on a limited set of parameters, i.e. pulse amplitude and a single lifetime. Starting from this scenario, we will derive some general conclusions, valid also for other kinds of experiments. In any case, the paper will provide the tools to mathematically analyze any case of interest.

Mono-exponential decay is defined by the following equation (valid for a single laser period):

$$P_{gen}(t) = P_{gen,0} \cdot e^{-\frac{t}{\tau}} \cdot [u(t) - u(t - T_{laser})], \tag{1}$$

where $P_{gen}(t)$ is the average number of photons generated by the excited sample in an infinitesimal interval dt centered at time t, $P_{gen,0}$ is the amplitude of the exponential pulse, τ is the decay time-constant, T_{laser} is the duration of the laser period and $u(t)$ is the Heaviside step function. A waveform of this kind is an effective model for the case of fluorescence emission induced by a laser pulse. In this case, τ is directly linked to the fluorescence lifetime. Anyway, in a real scenario, the shape of $P_{gen}(t)$ comes from the convolution of the instrument response function (IRF) of the laser source and the fluorescence emission, therefore τ immediately corresponds to the fluorescence lifetime only if the laser IRF has a much shorter duration.

During a TCSPC experiment only a portion of generated photons is collected in a histogram, resulting in a curve $P_{rec}(t)$, while the other photons are lost.

It is possible to subdivide losses in three groups. First, some photons coming from the sample do not even reach the active area of the detector due to setup geometry. Second, some photons reaching the active area are not effectively absorbed by the device due to limited photon detection probability. Finally, at the instant when a photon reaches the detector a TCSPC setup can be either active or blind (due to dead time or classic pile-up, as detailed later). Photons reaching the detector when the system is blind are lost.

In this scenario, it is possible to model the system with a simple block diagram, as depicted in Fig. 1.

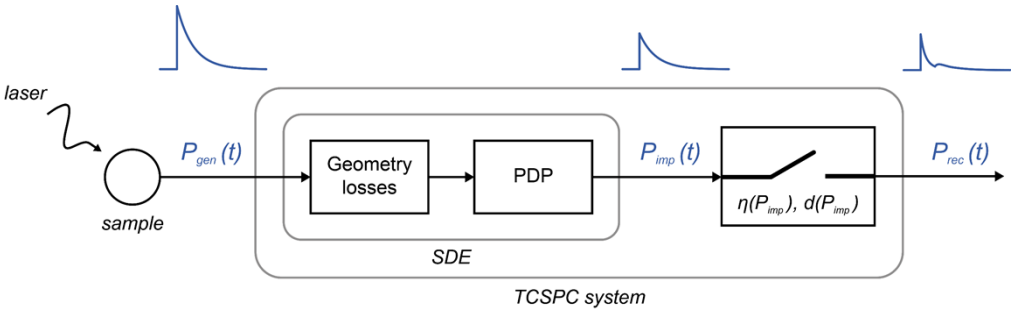

Fig. 1. Block diagram of a TCSPC system highlighting the main sources of photon losses: the laser excites a sample, which emits photons with a trend $P_{gen}(t)$. Then, some photons are lost due to geometry and photon detection efficiency (PDP), resulting in a curve $P_{imp}(t)$. Finally, other events are lost due to dead time or pile-up issues. The result is the recorded curve $P_{rec}(t)$. An example of curves is reported in blue above the block diagram.

The block *SDE* (which stands for system detection efficiency) condensates any loss due to geometry or PDP, while the second block of the TCSPC system symbolizes the bi-stable nature

Optics EXPRESS

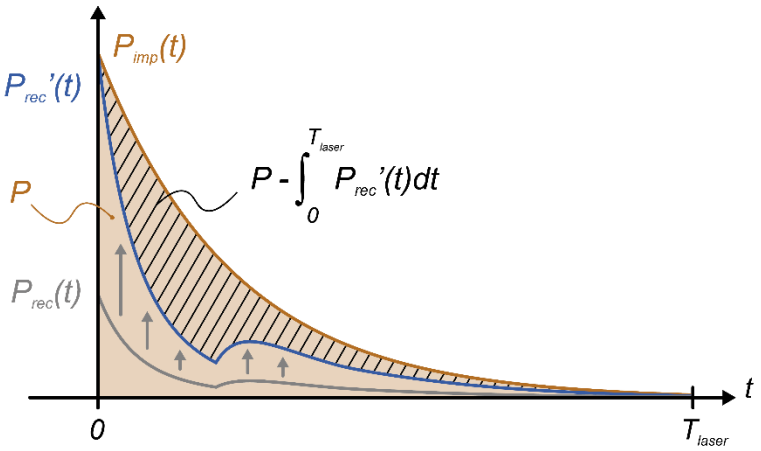

Fig. 2. Example of distortion calculation using the definition expressed in Eq. (4) [15]. The ideal curve $P_{imp}(t)$, featuring a mono-exponential trend, is highlighted in orange, while the recorded curve $P_{rec}(t)$ is depicted in grey. $P_{rec}'(t)$, i.e. the blue curve, results from a vertical stretch of $P_{rec}(t)$, expressed by Eq. (5). The stretch operation is highlighted by vertical grey arrows. Distortion is calculated dividing the area between $P_{imp}(t)$ and $P_{rec}'(t)$ (highlighted by oblique black lines) by P, i.e. the area of $P_{imp}(t)$ (highlighted in orange).

of the system (blind or active, hence the symbol of a switch). In other words, the second block represents a system where dead time or classic pile-up are the only sources of counting loss. $P_{imp}(t)$, equal to the product $P_{gen}(t) \cdot SDE$, is the average number of photons effectively impinging on this system. Here photons are lost only if they impinge when the system is blind, otherwise they are recorded. On average, this loss of photons is represented by a counting efficiency, η. The result is $P_{rec}(t)$, that is the recorded curve.

It is worth noting that losses due to geometry and PDP are uniform over time, meaning that every photon emitted by the sample has the same probability to be lost. In other words, *SDE* is a constant value. It follows that it does not introduce any distortion of the curve, meaning that $P_{imp}(t)$ is proportional to $P_{gen}(t)$. Instead, the transition from $P_{imp}(t)$ to $P_{rec}(t)$ is characterized by a distortion parameter d, meaning that $P_{rec}(t)$ does not feature the same shape of $P_{imp}(t)$. In general, both η and d depend on the impinging rate. In particular, η decreases when increasing the photon rate, while d increases, as will be clarified hereafter.

Measurement speed, expressed in photons per unit time, is proportional to the average number of recorded photons per period and to the laser frequency, as described by the following equation:

$$speed = \frac{1}{T_{laser}} \cdot \int_o^{T_{laser}} P_{rec}(t)dt, \tag{2}$$

where $1/T_{laser}$ is the laser frequency and the integral of $P_{rec}(t)$ is the average number of photons recorded in a period. The integral of $P_{rec}(t)$ can be obtained multiplying the integral of $P_{imp}(t)$ by η or, equivalently, multiplying the integral of $P_{gen}(t)$ by $SDE \cdot \eta$.

In this scenario, maximization of speed passes through three steps: i) maximizing *SDE*, ii) minimizing the excitation period, iii) maximizing the photon rate, but taking into account the unavoidably decrease of counting efficiency η when increasing the rate. In addition, photon rate cannot exceed a maximum value, not to introduce an unacceptable distortion of the curve.

Concerning *SDE*, the choice of the detector technology is fundamental to achieve the best PDP. In the same way, a proper design of the optical system to guide generated light to the detector is important to collect as many photons as possible. Nevertheless, the selection of the

Optics EXPRESS

detector with the best photon detection probability is not always the best choice in any experiment. Indeed, PDP is only one of many factors that have to be considered (other factors include timing performance, power dissipation, physical dimensions . . .). In any case, in real case scenarios, PDP and optical systems are usually far to be ideal, so *SDE* is always much lower than one.

Concerning the laser period, the minimum possible value is limited by the duration of the light waveform emitted by the sample in response to an excitation pulse. Indeed, the goal of a TCSPC measure is to observe the trend of intensity over time in its entirety, so it is fundamental that the luminous signal extinguishes before the end of the laser period. Considering the case of mono-exponential intensity, the minimum laser period needed to accommodate the whole pulse is conventionally defined equal to five times the decay time-constant. Indeed, considering an exponential pulse, $exp(-t/\ \tau)$, the area of the pulse in the interval $t \in [0, 5\tau]$ is slightly higher than 99% of the overall area in the interval $[0, +\infty)$, meaning that almost all photons are contained in the interval $[0, 5\tau]$. It follows that in this case 5τ can be considered the lower limit to T_{laser}.

Unfortunately, it is not always feasible to fix the laser period arbitrarily. Indeed, considering pulsed lasers suitable for TCSPC experiments, only a limited set of possible frequencies are currently available. Most typical values are 20, 40, 80 and 100 MHz. It follows that many times the laser period is much longer than the duration of the luminous pulse, thus limiting the maximum speed of the experiment. This is particularly true when analyzing extremely short time-constants.

Finally, concerning efficiency and distortion, they strongly depend on the TCSPC setup used to perform the measure. It follows that the calculation of both η and d is the starting point to analyze maximum speed in different setup conditions. Definitions of both counting efficiency and distortion are given in the following; then their application are shown for the cases of major interest.

The counting efficiency η is defined as the ratio between the total number of recorded photons and the total number of impinging photons, that is:

$$\eta = \frac{\int_0^{T_{laser}} P_{rec}(t)dt}{\int_0^{T_{laser}} P_{imp}(t)dt} = \frac{\int_0^{T_{laser}} P_{rec}(t)dt}{P}, \tag{3}$$

where P is the area of $P_{imp}(t)$, i.e. the average number of photons impinging on the detector during a period, ignoring losses due to geometry or PDP.

Instead, distortion is defined as the shape difference between $P_{rec}(t)$ and $P_{imp}(t)$: every time the experiment undergoes distortion, $P_{rec}(t)$ does not feature the same shape of $P_{imp}(t)$; conversely, when distortion is equal to zero $P_{rec}(t)$ is proportional to $P_{imp}(t)$.

Different expressions have been proposed for distortion calculation for mono-exponential pulses, based on a comparison of decay time-constants extracted from $P_{imp}(t)$ and $P_{rec}(t)$. In particular, barycenter estimation leads to an easy way for distortion calculation [2,14]. Unfortunately, the barycenter of the curve can be strictly used to estimate lifetime distortion only for mono-exponential pulses. Even worse, it can lead to a zero-distortion estimation even in cases where $P_{rec}(t)$ and $P_{imp}(t)$ feature a completely different shape [14].

In 2024, we proposed a novel definition of distortion [15], aiming at avoiding the aforementioned issues. In this case, distortion d is proportional to the area between $P_{imp}(t)$ and a normalized version of $P_{rec}(t)$:

$$d = \frac{\int_0^{T_{laser}} [P_{imp}(t) - P_{rec}{}'(t)] \cdot dt}{\int_0^{T_{laser}} P_{imp}(t)dt} = 1 - \frac{\int_0^{T_{laser}} P_{rec}{}'(t)dt}{P}, \tag{4}$$

where $P_{rec}{}'(t)$ is defined as:

$$P'_{rec}(t) = \frac{P_{rec}(t)}{\max\limits_{\{t\in[0,T_{laser}]\}} \left\{\frac{P_{rec}(t)}{P_{imp}(t)}\right\}}. \tag{5}$$

Optics EXPRESS

An example of distortion calculation is given in Fig. 2. In cases with zero distortion the curves $P_{imp}(t)$ and $P_{rec}{}'(t)$ overlap, so the area within them is zero; conversely, when the curves feature a different shape distortion results a positive non-zero value. Throughout the rest of the paper, this definition of distortion will be used.

3. Classic pile-up

Historically, in a classic TCSPC setup the major limit to efficiency and distortion has been related to the time-measurement electronics. Indeed, a typical acquisition channel is able to account only for a single photon during an excitation cycle [2,14]. In this scenario, every time more than a single photon impinge on the detector during the same period, only the first photon is recorded. This phenomenon, called "classic pile-up", gives rise to a loss of efficiency, together with a distortion of the histogram. Indeed, the probability to record photons in the first bins of the histogram is higher than the recording probability of later bins.

Considering a mono-exponential intensity, i.e. $P_{imp}(t) = P_{imp}(0)\cdot\exp(-t/\tau)\cdot[u(t) - u(t - T_{laser})]$, under classic pile-up it is possible to derive a closed-form expression for $P_{rec}(t)$ [2]:

$$P_{rec}(t) = P_{imp}(0) \cdot e^{-\frac{t}{\tau}} \cdot e^{-P} \cdot e^{P\cdot e^{-\frac{t}{\tau}}} \cdot [u(t) - u(t - T_{laser})]. \tag{6}$$

The result is that $P_{rec}(t)$ decays much faster than $P_{imp}(t)$, meaning that the equivalent time-constant of $P_{rec}(t)$ is lower than τ. An example is depicted in Fig. 3, considering P equal to 1 photon per period, τ equal to 1 ns and T_{laser} equal to 12.5 ns, corresponding to an 80-MHz excitation frequency.

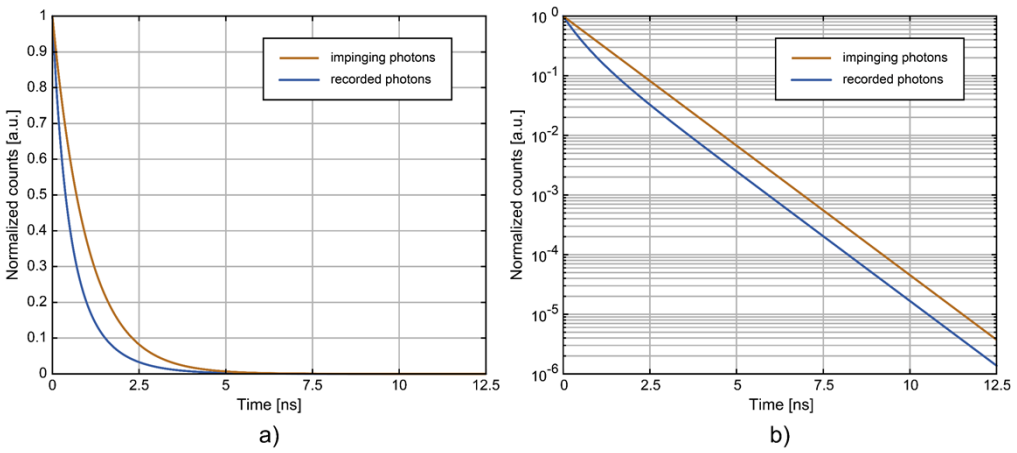

Fig. 3. Example of classic pile-up distortion for a mono-exponential $P_{imp}(t)$ with $P = 1$, $\tau = 1$ ns, $T_{laser} = 12.5$ ns. The ideal curve of impinging photons, which is proportional to $P_{imp}(t)$, is depicted in orange, while the recorded curve is depicted in blue. The same example is reported in linear scale (a) and in logarithmic scale (b)

Combining Eq. (6) with Eq. (3) it is possible to find a simple expression for efficiency:

$$\eta = \frac{1 - \mathrm{e}^{-P}}{P}. \tag{7}$$

Similarly, distortion can be calculated starting from Eqs. (4) and (6):

$$d = 1 - \frac{1 - \mathrm{e}^{-P}}{P}. \tag{8}$$

It is worth noting that distortion due to classic pile-up is equal to $1 - \eta$, meaning that zero distortion is feasible only when efficiency is 100%, i.e. when P approaches zero. The same

Optics EXPRESS

holds also for luminous shapes different from a simple mono-exponential pulse. This is the main reason why historically TCSPC experiments have been performed keeping the average number of photons per period near to zero (0.01-0.05 are typical values) and this has represented the main limitation to measurement speed so far.

Nevertheless, in the last years the landscape has changed. In particular, recent time-measurement electronics feature a very low dead time, along with a multi-hit capability, meaning they are able to measure more than a photon per period [16–19]. This evolution of time converters has opened the door to surpass the classic pile-up limit.

4. Detector dead time

Given a state-of-art timing electronics with low dead time and multi-hit capabilities, the ultimate limit to speed is related to the dead time of the detector. In particular, every time a photon is observed, the detector remains blind for a fixed amount of time, T_{dead}. For example, considering a Single-Photon Avalanche Diode (SPAD), just after a photon is detected, the diode is kept off by a dedicated quenching electronics, in order to turn off the avalanche current and bring the system back to the original condition [20,21].

In this kind of system, both efficiency and distortion strongly depend on the duration of dead time with respect to the laser period. Unfortunately, in this case there is no closed-form expression for $P_{rec}(t)$. Anyway, it is possible to describe the link between $P_{rec}(t)$ and $P_{imp}(t)$ with a recursive expression [14]:

$$P_{rec}(t) = P_{imp}(t) \cdot \left[1 - \int_{t-T_{dead}}^{t} \sum_{i=-\infty}^{+\infty} P_{rec}(t' + i \cdot T_{laser}) \cdot dt' \right], \tag{9}$$

where the summation takes into account the periodic nature of TCSPC.

Once the expression of $P_{imp}(t)$ is known, it is possible to find the trend of $P_{rec}(t)$ using an iterative method for non-linear equations to solve Eq. (9). At this point, it is possible to calculate both efficiency and distortion by means of Eqs. (3) and (4) respectively.

4.1. Efficiency

Even if there is no closed-form available for efficiency calculation, some general considerations can be derived. First of all, it can be proved that, independently of the shape of $P_{imp}(t)$, efficiency always decreases when increasing the photon rate. In particular, it follows a general trend expressed by the following equation [14]:

$$\eta_{trend} = \frac{1}{1 + P \cdot \frac{T_{dead}}{T_{laser}}}. \tag{10}$$

Figure 4 shows a comparison between η_{trend} and an example of efficiency curve η as a function of T_{dead}, calculated combining a numerical resolution of Eq. (9) with Eq. (3). In this example $P_{imp}(t)$ in Eq. (9) is a mono-exponential with $P = 1$, $\tau = 1$ ns and $T_{laser} = 12.5$ ns.

It can be shown that Eq. (10) is strictly equal to η only at T_{dead} values equal to any integer multiple of the excitation period (see red dots in Fig. 4). Instead, for other values of T_{dead} the efficiency curve shows a step-like behavior, formed by some plateaus, connected one to each other by means of smooth transitions centered at integer multiples of T_{laser}. This behavior is valid for any value of P and τ.

Concerning the first plateau, it corresponds to dead time values sufficiently high to cover the entire luminous pulse. For instance, in the example of Fig. 4 the plateau starts roughly at T_{dead} equal to 5 ns, since 5 ns is long enough to cover almost entirely the 1-ns-timeconstant pulse. In the plateau region only one photon per period (the first one) can be recorded, since the dead time

Optics EXPRESS

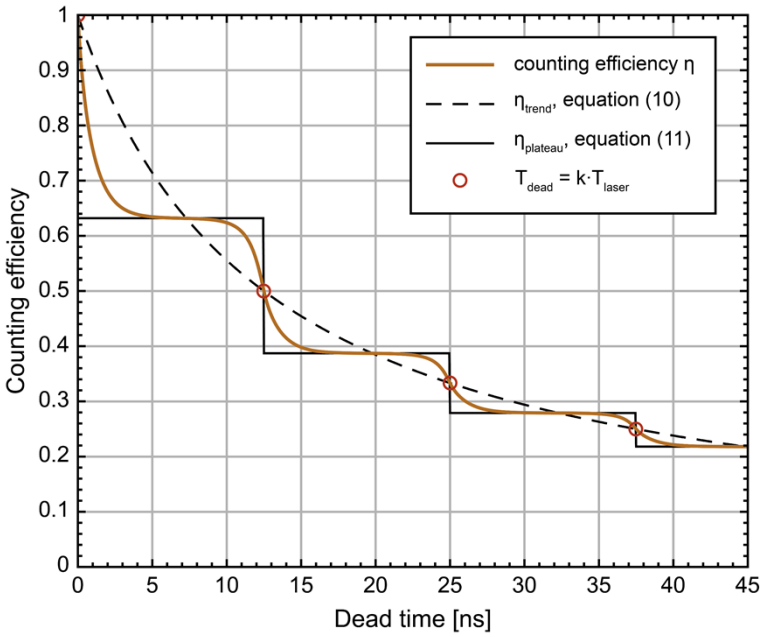

Fig. 4. Example of counting efficiency as a function of the detector dead time T_{dead}, calculated for a mono-exponential $P_{imp}(t)$ with $P = 1$, $\tau = 1$ ns, $T_{laser} = 12.5$ ns. The counting efficiency η is depicted along with η_{trend}, expressed in Eq. (10), and $\eta_{plateau}$, expressed in Eq. (11). Red dots indicate points where the dead time is an integer multiple of the laser period. All curves go through these points.

is sufficiently long to cover all the subsequent photons belonging to the same pulse. This effect is akin classic pile-up, so the value of efficiency in the first plateau can be calculated using Eq. (7). The plateau ends when the dead time becomes so long that it starts covering photons belonging to the subsequent period. At that point efficiency drops. In the example of Fig. 4 this happens at T_{dead} surpassing 10 ns. When dead time is long enough to cover entirely two subsequent pulses efficiency enters a second plateau and this behavior repeats for any integer multiple of the laser period.

A more general formula, valid for all plateaus, is given in Eq. (11):

$$\eta_{plateau} = \frac{\left(\frac{1-e^{-P}}{P}\right)}{1 + (1 - e^{-P}) \cdot \left\lfloor \frac{T_{dead}}{T_{laser}} \right\rfloor}, \tag{11}$$

where $\lfloor \cdot \rfloor$ is the floor function, which extracts the integer part of the argument. A complete mathematical demonstration of Eq. (11) is given in Appendix A.

The function $\eta_{plateau}$ is depicted in Fig. 4 (black continuous line) together with the efficiency curve.

Concerning transitions, they are located at integer multiples of the laser period and they recall the same exponential trend of $P_{imp}(t)$. For instance, in Fig. 4, τ is equal to 1 ns, so transitions feature a trend similar to an exponential decay with 1 ns time-constant.

Figure 5(a) shows efficiency curves for different values of τ, varying T_{dead} or P, together with Eq. (10) and Eq. (11). It is worth noting that efficiency curves stay always between Eqs. (10) and (11). In particular, η tends to $\eta_{plateau}$ for time-constant values near to zero. Conversely, η approaches η_{trend} for large values of τ. At the limit, η coincides with $\eta_{plateau}$ when τ is zero, i.e.

Optics EXPRESS

when $P_{imp}(t)$ is a Dirac delta, and it coincides with η_{trend} when τ tends to infinite, i.e. when the light signal is constant over time.

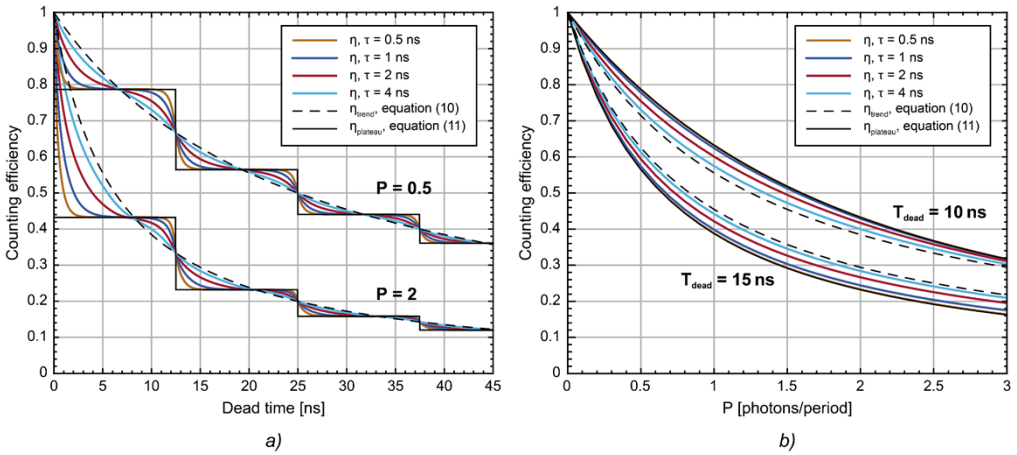

Fig. 5. Counting efficiency for mono-exponential $P_{imp}(t)$ with different values of τ ranging from 0.5 ns to 4 ns, considering a laser period of 12.5 ns. a) Efficiency as a function of the dead time T_{dead}; the two families of curves correspond to two different values of P (0.5 and 2 photons per period). b) Efficiency as a function of P; the two families of curves belong to two different values of T_{dead} (10 ns and 15 ns).

In general, efficiency always decreases as a function of both P and T_{dead}, meaning that an increase of the dead time or of the impinging photon rate leads to an efficiency drop. This is a general rule, valid also when the intensity profile is not mono-exponential. Nevertheless, even if η decreases when increasing P, the average number of recorded photons, given by the product $P \cdot \eta$, increases when increasing P. This means that measurement speed can be always increased by increasing luminous power, at the expense of a poor collection efficiency. This can be demonstrated observing that both $P \cdot \eta_{trend}$ and $P \cdot \eta_{plateau}$ steadily increase when increasing P. Figure 6 shows the product between P and η for the same cases of Fig. 5.

4.2. Distortion

Unlike efficiency, distortion features a periodic behavior when changing T_{dead}, with period equal to T_{laser} [14]. Figure 7 shows an example of distortion curves, calculated combining the numerical solution of Eq. (9) with distortion definition given in Eq. (4), considering a mono-exponential decay with τ equal to 1 ns, T_{laser} equal to 12.5 ns and different values of P.

In general, distortion increases when increasing P and strongly depends on dead time duration. In particular, it approaches zero for low values of T_{dead} (near zero), then by increasing the dead time distortion increases up to a plateau. Finally, it decreases back to zero at any dead time equal to an integer multiple of the excitation period, as demonstrated for the first time in 2017 [14] and shown experimentally using dedicated electronics [17,22,23].

It is worth recalling that plateaus correspond to a classic-pile-up region, so the distortion at plateaus can be calculated using Eq. (8), as highlighted by the dashed lines in Fig. 7.

Similarly to efficiency, transitions from 0 to plateau depend on the time-constant. For example Fig. 8 shows distortion considering a single value of P equal to 1 and different values of τ. It is evident that high values of time-constants have a slower transition from zero to the plateau value and in some cases the curve does not even reach the plateau, meaning that, in general, large values of τ lead to lower distortion.

Optics EXPRESS

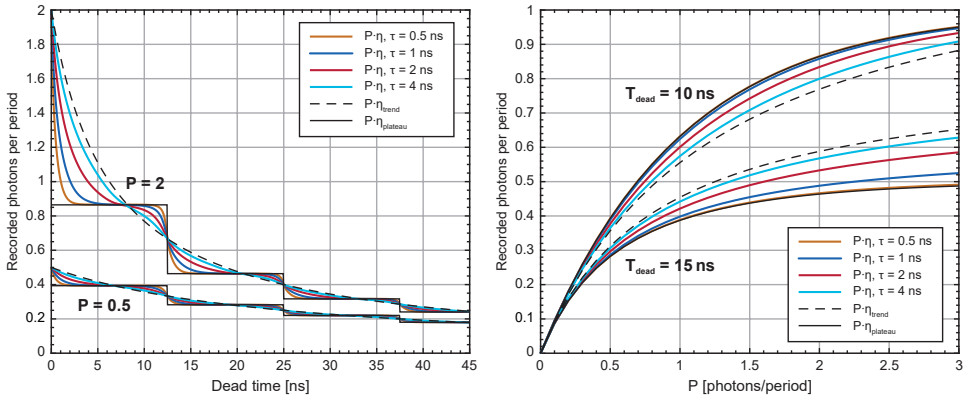

Fig. 6. Average number of recorded photons per period ($= P \cdot \eta$) for mono-exponential $P_{imp}(t)$ with different values of τ ranging from 0.5 ns to 4 ns, considering a laser period of 12.5 ns. a) $P \cdot \eta$ as a function of the dead time T_{dead}; the two families of curves correspond to two different values of P (0.5 and 2 photons per period). b) $P \cdot \eta$ as a function of P; the two families of curves belong to two different values of T_{dead} (10 ns and 15 ns).

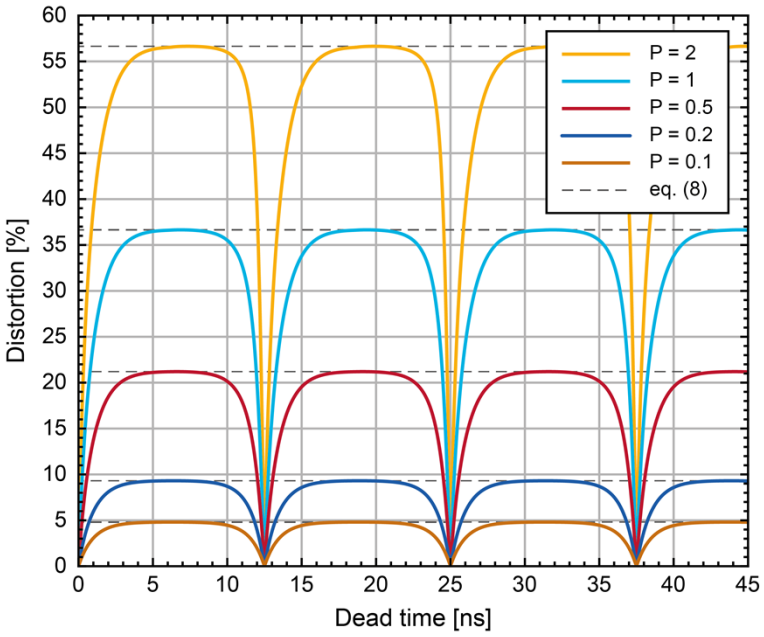

Fig. 7. Distortion calculated for mono-exponential $P_{imp}(t)$ with τ equal to 1 ns, considering a laser period of 12.5 ns and different values of P ranging from 0.1 to 2 photons per period. Equation (8) is depicted in dashed lines for all the considered values of P.

Optics EXPRESS

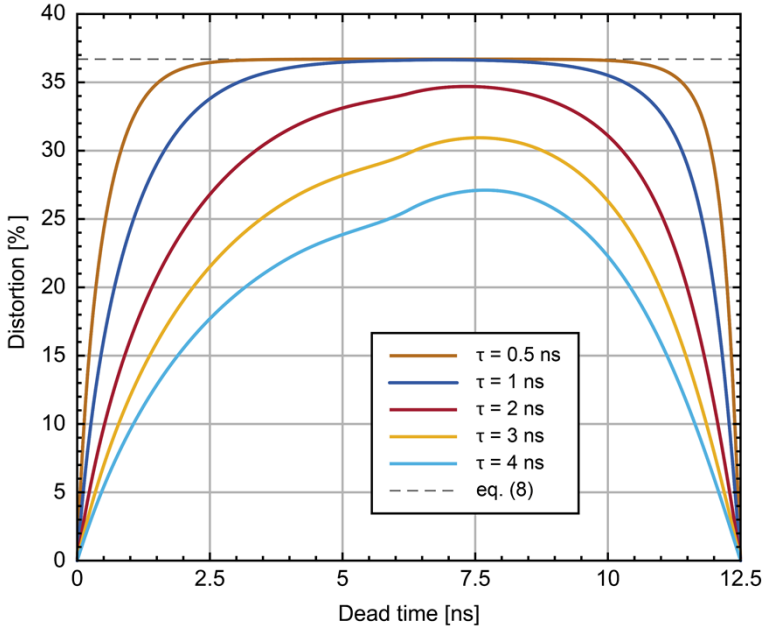

Fig. 8. Distortion calculated for mono-exponential $P_{imp}(t)$ with P equal to 1, considering a laser period of 12.5 ns and different values of τ ranging from 0.5 ns to 4 ns. Equation (8) is depicted in dashed line for $P = 1$.

It is worth noting that distortion curves are slightly asymmetrical. For example, each curve in Fig. 8 shows two different slopes for transitions at low values of T_{dead} and at T_{dead} approaching T_{laser}. This is due to the different nature of the distortion at different values of T_{dead}: when T_{dead} is much lower than T_{laser}, the dead time only masks photons belonging to the same pulse, thus mainly distorting the tail of the exponential pulse and not the peak; instead, when T_{dead} approaches T_{laser}, the dead time also masks early photons of the next laser period. In any case, the slopes decrease as the time-constant increases and the plateau tends to disappear.

5. Setup design for maximum speed

It has been shown that speed steadily increases when increasing the photon rate. Therefore, a solution to get maximum speed consists in boosting the laser power. Nevertheless, this solution can lead to high distortion and makes efficiency lower and lower. Both these effects can be an issue. Concerning distortion, in most cases 1-2% is considered an insurmountable limit, which translates in an upper limit to speed. Concerning efficiency, the minimum acceptable value strongly depends on the experiment nature.

On the one hand, in some experiments where efficiency is not important, the luminous power can be pushed to increase speed (e.g. in LIDAR applications, as long as the laser power is below eye-safety limits). However, it is worth saying that in any case luminous power cannot be increased indefinitely; indeed, at high rates any measurement setup experiences non-idealities that lead to a speed saturation or even decrement. For instance, at a certain point the time-measurement electronics reaches its maximum count rate (even most recent multi-hit electronics have a limit due to electronics dead time or data-bandwidth limitation). Moreover, the presence of a reset phase at the end of the dead time can lead to a decrease of speed at high counts [14]. Anyway, these limits are usually experienced at quite high rates.

Optics EXPRESS

On the other hand, in other kinds of experiments it is fundamental to keep efficiency high, especially when the number of generated photons is a limited quantity. For example, in many fluorescence experiments after the sample has absorbed and re-emitted a certain amount of luminous energy, it undergoes a photochemical alteration, which makes it unable to fluoresce anymore (photobleaching) [24]. In this case, it is important to keep the overall efficiency ($SDE \cdot \eta$) higher than a certain limit. In particular, given a maximum number of photons that the sample can generate, $N_{gen,max}$, and considering that a minimum number of photons, N, is required to have a good statistic of the histogram [13], the minimum acceptable efficiency is equal to $N/N_{gen,max}$. This means that the overall efficiency of the system, that is $SDE \cdot \eta$ (see Fig. 1), must be kept above this limit, as expressed in the following equation:

$$SDE \cdot \eta \geq \frac{N}{N_{gen,max}}. \tag{12}$$

If SDE is known, this equation translates in a limit to η. Considering that η is a function of the photon rate, Eq. (12) translates in turn to a limit to speed.

In the following both efficiency- and distortion-related limits to speed are analyzed in details.

5.1. Efficiency limit

To quantify how much the limit expressed by Eq. (12) impacts on speed, it is possible to map η as a function of both P and T_{dead} combining Eq. (3) with a numerical solution of Eq. (9).

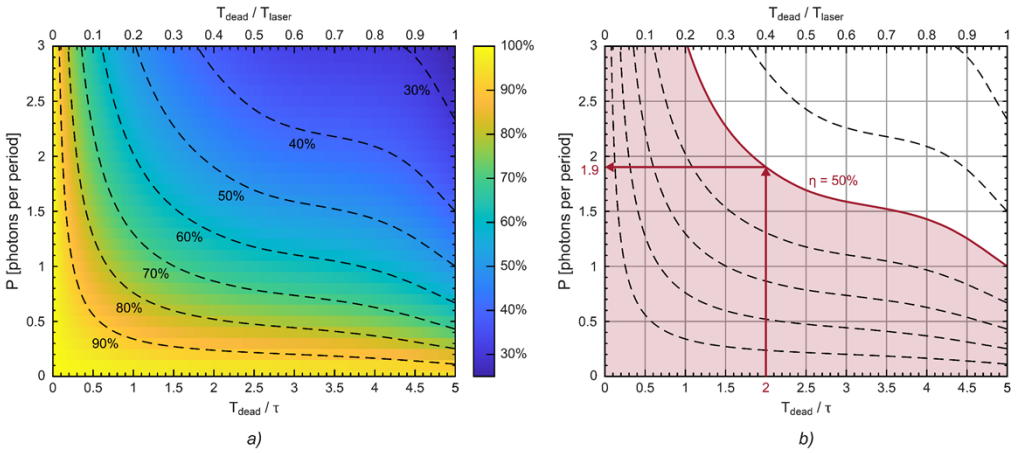

Fig. 9. Map of counting efficiency as a function of dead time and photon rate P for mono-exponential $P_{imp}(t)$ and $T_{laser} = 5\tau$. a) Counting efficiency is depicted in color scale, while contour lines are highlighted by black dashed lines. b) Example of maximum rate calculation at dead time equal to 2τ, which corresponds to dead time equal to $0.4 \cdot T_{laser}$. The curves are contour lines extracted from the color map in a). Given that $T_{laser} = 5\tau$, the maps show two x-axes: one expressing T_{dead} as a fraction of T_{laser} and another one expressing T_{dead} as a fraction of τ.

Considering the case of a mono-exponential decay with T_{laser} equal to 5τ (that is the optimal condition), the result is shown in Fig. 9(a). Here the value of η is shown in color scale, while dashed curves highlight contour lines, which connect points having the same values of η. Graphically the efficiency threshold expressed in Eq. (12) corresponds to a specific contour line and the region under that line represents the area where P and T_{dead} values satisfy the inequality. Starting from this consideration, it is possible to calculate maximum achievable speed for practical cases. For example, considering N equal to 500,000 photons, $N/N_{gen,max}$ equal to

Optics EXPRESS

0.4 (e.g. due to photobleaching) and *SDE* equal to 80%, the upper limit to η is 50%, as can be inferred from Eq. (12). The contour line corresponding to $\eta = 50\%$ is highlighted in red in Fig. 9(b), together with the region under the curve, that is the region where the values of P and T_{dead} are valid. For instance, if τ is equal to 2.5 ns and T_{dead} is 5 ns (that is $T_{dead}/\tau = 2$), the maximum value of P is 1.9 photons per period, which corresponds to maximum achievable speed. Recalling the model of Fig. 1 and the definition of speed expressed in Eq. (2), maximum speed is equal to $P \cdot \eta / T_{laser} = 76\text{photons}/\mu\text{s}$. Since N is equal to 500,000 photons, the minimum duration of the experiment is 6.58 ms. This procedure can be used for any given value of τ, T_{laser}, T_{dead} and η (assuming $T_{laser}/\tau = 5$) meaning that the diagram in Fig. 9 is a general way to compare maximum speed of different TCSPC setups. In addition, it is worth noting that maximum P at a given efficiency does not depend on absolute values of T_{dead}, T_{laser} and τ, but only on the ratios T_{dead}/τ and T_{laser}/τ, meaning that the same results are valid for any multiple of T_{dead}, T_{laser} and τ that preserves the ratios. For example, optimal P is 1.9 also for $T_{dead} = 40$ ns, $T_{laser} = 100$ ns and $\tau = 20$ns.

Unfortunately, there is no closed-form equation one can apply to extract P from η, so the graphical solution is the easiest way. Anyway, as explained in section 4.1, η stays always between η_{trend} and $\eta_{plateau}$, expressed in Eqs. (10) and (11) respectively. It follows that both Eqs. (10) and (11) can be used to approximate η. In the previous example, if η_{trend} is used instead of η, the resulting maximum value of P is 2.5 photons per period. Conversely, when using $\eta_{plateau}$, the result is 1.59 photons per period. The correct value (i.e. $P = 1.9$) stays in between.

It is worth recalling that η_{trend} is a better approximation in the case of luminous signals with high τ. Even better, η_{trend} expresses a general trend, so it is a valid approximation even if illumination is not an exponential decay. To be precise, η follows exactly η_{trend} when illumination is constant for the whole period [14], so this approximation fits better every time there is a strong background light or the signal is almost flat. Figure 10(a) shows a color map of η_{trend} as a function of P and T_{dead}.

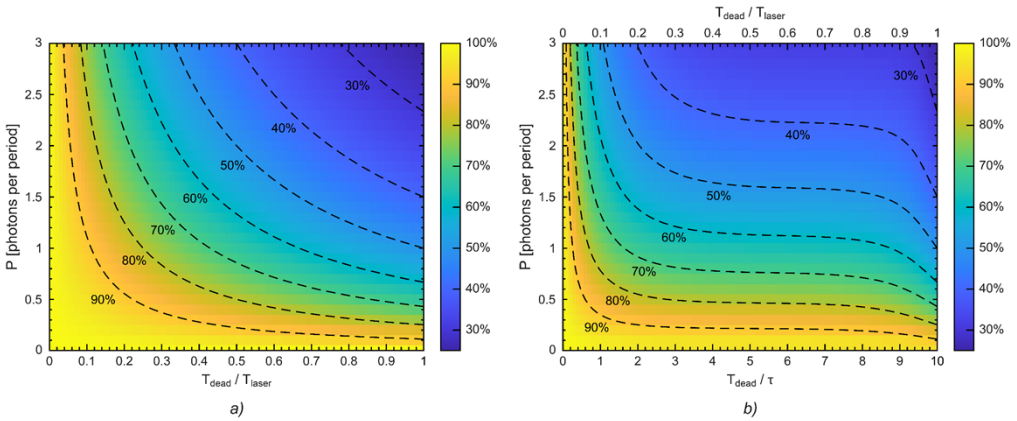

Fig. 10. a) Map of η_{trend}, expressed by Eq. (10), as a function of dead time and photon rate P. b) Map of counting efficiency η for a mono-exponential signal with $T_{laser} = 10\tau$ as a function of dead time and photon rate P. It is worth noting that in a) the x-axis expresses T_{dead} only as a fraction of T_{laser} and not as a function of τ, since η_{trend} does not depend on τ. Conversely, in b) two x-axes are shown: one expressing T_{dead} as a fraction of T_{laser} and another one expressing T_{dead} as a fraction of τ.

On the other hand, $\eta_{plateau}$ is a good approximation for η if the luminous signal has a short duration (e.g. short τ in case of mono-exponential signal) and the dead time is sufficiently high

Optics EXPRESS

to cover the entire signal. It is worth highlighting that this is valid for any waveform being much shorter than T_{dead}, not only for mono-exponential signals, as better clarified in Appendix A.

Figure 10.b shows a color map of η calculated for a mono-exponential signal with a relatively short τ equal to 10% of the laser period ($T_{laser} = 10\tau$). Contour lines are depicted in dashed lines for different values of η. In Fig. 109(b), it is evident that η shows plateaus for intermediate values of T_{dead}, i.e. when the dead time is much larger than τ and shorter than T_{laser}. Plateau values can be calculated using Eq. (11). In principle plateaus are present also in the case of T_{laser} equal to 5τ, but the effect is less pronounced (see Fig. 9).

It is worth noting that curves in Fig. 9(a) and Fig. 10(b), calculated for different T_{laser}/τ ratios, feature similar behaviors. To better understand how the behavior changes for different values of laser period, a comparison of contour lines for T_{laser} equal to 5τ, 10τ and 15τ is depicted in Fig. 11. For each family of curves, it is possible to distinguish three regions: i) a first transition before plateaus (for low values of T_{dead}), ii) the plateau region, iii) a last transition after plateaus (for T_{dead} approaching T_{laser}). During the first transition the dead time is much lower than T_{laser}, so the curves feature the same trend regardless of the laser period. Then, curves enter the plateau region. Here the values are the same for all families, but the duration of the plateau depends on the duration of the laser period. Indeed, the plateau region lasts until the dead time approaches T_{laser} (e.g. for $T_{laser} = 5\tau$ the plateau region is barely visible, while for a longer T_{laser}, e.g. 15τ, the plateau lasts for a wide range of T_{dead} values). Finally, just after plateau, the curves feature another transition (they bend downwards), which has the same shape for all values of T_{laser}. In short, curves belonging to different families feature the same trend during transitions and the only difference is the duration of plateaus.

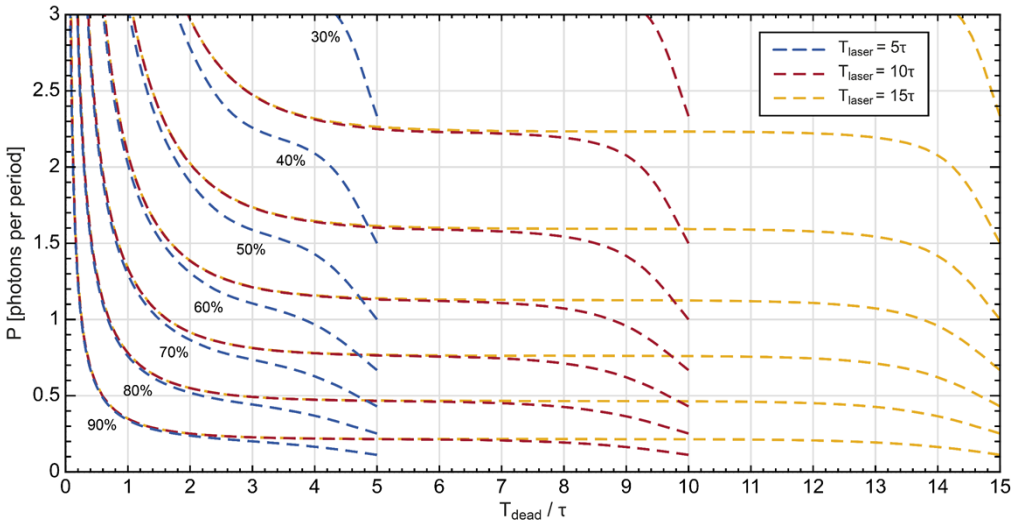

Fig. 11. Contour lines of counting efficiency in the P-T_{dead} map, considering P from 0 to 3 and T_{dead} going from 0 to T_{laser}. Here T_{dead} is represented as a fraction of τ (on x axis), while curves are subdivided in three families highlighted with different colors, each one corresponding to a different value of T_{laser}. In each family the curves represent points at constant efficiency ranging from 90% down to 30%.

Starting from this consideration, it is possible to condensate the curves shown in Fig. 11 in a single map, which is depicted in Fig. 12. Here the three different regions are well separated: efficiency decays in the first region, which is depicted for T_{dead}/τ values going from 0 to 5, then it is constant for intermediate values and it decays again in the third region, depicted for T_{dead}/τ values between $T_{laser}/\tau - 5$ and T_{laser}/τ.

Optics EXPRESS

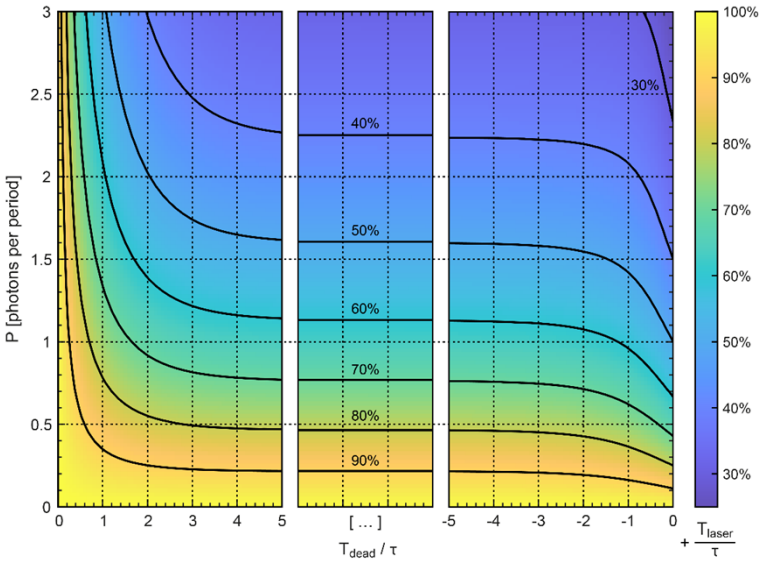

Fig. 12. Efficiency color map for a mono-exponential $P_{imp}(t)$, valid for generic values of T_{laser}, T_{dead} and τ. Contour lines are highlighted in black for efficiency values ranging from 90% down to 30%. The diagram is divided in three different regions: i) contour lines bending downwards for low values of T_{dead} (depicted from 0 to 5τ), ii) plateau region (constant efficiency), iii) contour lines bending downwards for the second time (for values of T_{dead} approaching T_{laser}). In the third region the x axis ranges from $-5 + T_{laser}/\tau$ to $0 + T_{laser}/\tau$.

In principle, starting from this map it is possible to estimate efficiency for any combination of T_{laser}, T_{dead} and τ. To be precise, the diagram is strictly valid for any laser period equal to or higher than 10τ, but it can be used to obtain an acceptable approximation of efficiency also for lower values. For example, considering $T_{laser} = 5\tau$ and $T_{dead} = 2\tau$ efficiency depicted in Fig. 12 is equal to 50% at P equal to 2, that is an approximation of the correct value ($P = 1.9$), shown in the example of Fig. 9(b).

All the diagrams presented so far refer to the case of mono-exponential light waveform. Anyway, if different cases have to be analyzed, it is always possible to calculate the correct map starting from a numerical solution of Eq. (9), considering the actual shape of $P_{imp}(t)$. Alternatively, in some cases it is possible to approximate $P_{imp}(t)$ to a mono-exponential decay to derive a simple estimation of efficiency from the maps proposed above. For example, when considering a multi-exponential decay, it is possible to estimate an effective time-constant, using a weighted average of lifetime components or a barycenter estimation [2].

In any case, regardless of the shape of the luminous signal, it is possible to derive some general considerations. First of all, *SDE* plays a key role in defining the speed of the system. In particular, a reduction of *SDE* leads to a decrease of maximum speed. For example, if we consider the same example depicted in Fig. 9(b), but with different *SDE* equal to 66.7% (reduction of 16.7% compared to the previous value), the minimum acceptable η becomes 60% instead of 50%, thus leading to a reduction of the working area (compare the areas under the curves labelled with $\eta = 50\%$ and $\eta = 60\%$ in Fig. 9(a)). In this case, maximum P at $T_{dead} = 2\tau$ is 1.3, so maximum speed becomes 62.4 photons/μs, corresponding to a speed reduction of 17.9%.

The other key parameter to get high speed is dead time associated to the detector. Indeed, efficiency shows a high sensitivity to dead time, especially when dead time is a low value compared to the decay time-constant. For instance, considering both Fig. 9 and Fig. 10, with

Optics EXPRESS

T_{dead} equal to few percent of τ counting efficiency can be higher than 90% even if the photon rate is much higher than 1 photon per period. In other words, very low dead times can lead to the exploitation of extremely high photon rates.

A mathematical approximation for maximum speed can be given, considering η equal to η_{trend}. In this case, the limit expressed by Eq. (12) translates in the following inequality:

$$speed \leq \frac{1}{T_{dead}} \cdot \left(1 - \frac{N/N_{gen,max}}{SDE}\right). \tag{13}$$

This expression reveals that maximum speed is inversely proportional to the dead time, while it is not exactly proportional to SDE, meaning that T_{dead} has a greater impact on speed.

Nevertheless, it is worth recalling that this is only an approximation, which is valid on average. Depending on the shape of the luminous signal and on dead time itself, the impact of an increase of dead time can be higher or lower than what expressed by Eq. (13). For instance, when considering a mono-exponential signal, efficiency can be almost independent of T_{dead} in certain ranges (see plateaus in Fig. 11) or can change very steeply for values of T_{dead} approaching zero.

5.2. Distortion limit

Similarly to efficiency, also distortion can be mapped as a function of dead time and photon rate. Figure 13(a) shows a color-map representing distortion as a function of P and T_{dead} for the case of a mono-exponential signal with $T_{laser} = 5\tau$. Figure 13(b) is a zoom for T_{dead} from 0 to 0.5τ, i.e. 10% of T_{laser}. The map has been calculated starting from Eqs. (9) and (4).

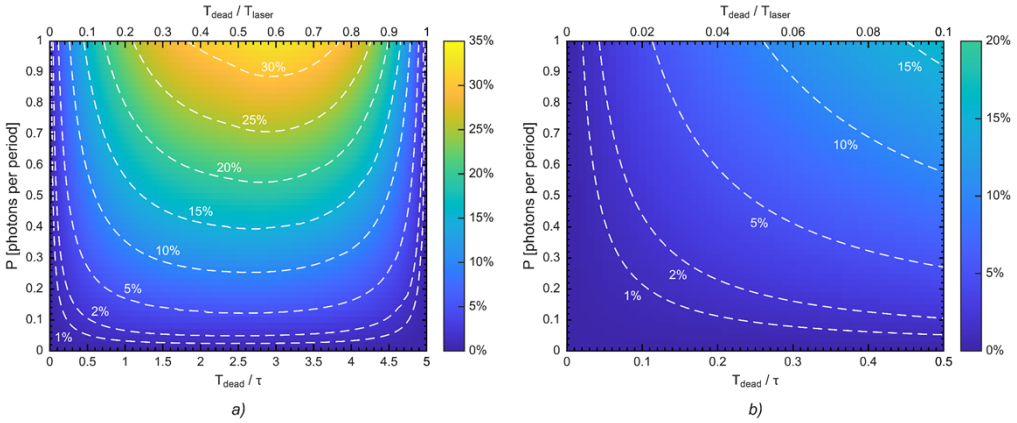

Fig. 13. Map of distortion as a function of dead time and photon rate P for mono-exponential $P_{imp}(t)$ and $T_{laser} = 5\tau$. Distortion is depicted in color scale, while contour lines are highlighted by white dashed lines, connecting points where distortion is constant an equal to 1%, 2%, 5%, etc. a) Map for T_{dead} ranging from 0 to 5τ (equal to T_{laser}). b) Zoom for T_{dead} ranging from 0 to 0.5τ, that is $0.1 \cdot T_{laser}$. Two x-axes are shown: one expressing T_{dead} as a fraction of T_{laser} and another one expressing T_{dead} as a fraction of τ.

Like in efficiency maps, also in the diagram of Fig. 13 it is possible to distinguish three zones: i) a first region where contour lines bend downwards, ii) a plateau region, iii) a latest region where contour lines bend upwards to come back to their initial values at $T_{dead} = T_{laser}$. This behavior is generally valid not only for mono-exponential $P_{imp}(t)$, but also for different shapes. Indeed, distortion is always zero at T_{dead} equal to 0 and to any integer multiple of T_{laser} (as discussed in section 4.2) and it is higher than zero elsewhere, so contour lines always bend downwards for low values of T_{dead} and bend upwards when T_{dead} approaches T_{laser}. In any case, it is possible to

Optics EXPRESS

derive the actual behavior of distortion for any shape of $P_{imp(t)}$, combining a numerical solution of Eq. (9) with Eq. (4).

Once distortion is calculated, the limits of P and T_{dead} to keep distortion below a defined threshold can be inferred from the distortion map (see Fig. 13), starting from contour lines: maximum acceptable distortion corresponds to a specific contour line and the area under it corresponds to the region where the system can operate. A typical value for distortion limit is 1%. Considering Fig. 13, it is evident that this requirement confines the possible values of P in a very narrow area. In particular, for most T_{dead} values P has to be selected below 0.05 photons per period, which is the typical case of TCSPC experiments. Only two regions exist where P can be increased above this limit: where T_{dead} approaches 0 (see Fig. 13(b)) and were it approaches T_{laser}. Those possibilities have been already discussed in section 4.2.

It is possible to combine the limits imposed by efficiency and distortion, to get the overall limit to speed. In particular, depending on which limit is more stringent, efficiency or distortion set the ultimate limit to speed. Figure 14(a) shows the contour lines of efficiency (in black) superposed on the color map of distortion (distortion contour lines highlighted in white). A red line delimits the region where efficiency is higher than 50% and distortion is lower than 1%. It is clear that distortion sets the limit in most cases. The only region where efficiency sets the major limit is where T_{dead} approaches T_{laser} (Fig. 14(b) shows a zoom in this region).

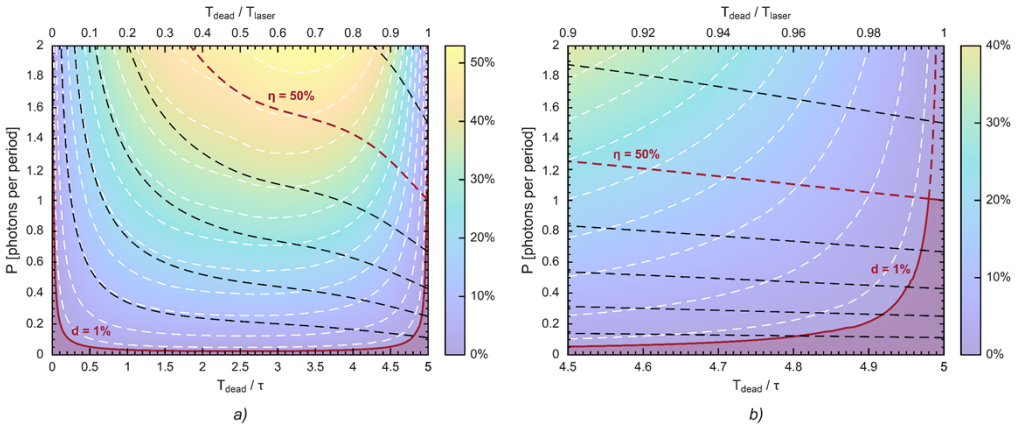

Fig. 14. Map of distortion with superposed efficiency contour lines (black dashed lines). The map has been calculated considering a mono-exponential $P_{imp}(t)$ and $T_{laser} = 5\tau$. The working region with $\eta \geq 50\%$ and $d \leq 1\%$ is highlighted in red. a) T_{dead} ranging from 0 to T_{laser}. b) Zoom for T_{dead} ranging from $0.9 \cdot T_{laser}$ to $1 \cdot T_{laser}$.

6. Recent advances in TCSPC measurements

In the last years many efforts have been done to design detectors and associated conditioning electronics with extremely short dead times, in order to push experiments to maximum efficiency and speed. For example, nowadays SPADs integrated in CMOS technology exhibit dead times in the order of few nanoseconds [25–29] and in some cases even less than 1 ns [30]. In addition, alternative technologies are arising, with sub-nanosecond dead time, e.g hybrid photodetectors [31].

Nevertheless, even with state-of-art dead times distortion can be a serious problem, especially when considering fast light signals (i.e. short τ if the signal is exponential). For example, considering τ equal to 2.5 ns and T_{laser} equal to 12.5 ns (that is five times τ) distortion fixes a strong limit to maximum speed even considering a state-of the art dead time of 1 ns (that is equal

Optics EXPRESS

to 8% of T_{laser} or, equivalently, to 0.4τ). In particular, it can be inferred from Fig. 11(b) that P cannot surpass 0.06 photons per second to keep distortion below 1%.

In addition, the choice of a very fast detector can result in a tradeoff with other characteristics, e.g. timing performance or PDP, so, depending on the application, it may not always be feasible.

On the other hand, the solution proposed in [14] completely gets rid of distortion, but at the expense of a relatively high dead time value, which corresponds to a relatively low efficiency. For example, considering a fast laser rate of 80 MHz, the dead time has to be set to 12.5 ns, which is much higher than dead times of state-of-the-art detectors.

To surpass this tradeoff, in 2023 we proposed a novel solution, which permits to avoid distortion regardless of the dead time duration [32], and in 2024 we experimentally proved its effectiveness [33]. The solution adds a second measurement channel to the TCSPC setup, whose aim is to store in an auxiliary histogram a curve tracking the status of the system. In particular, this histogram is incremented every time the system is active, that is when it is able to detect a photon. Conversely, if the system is blind due to a dead time the histogram gets no update. In this way, at the end of the experiment, the histogram contains, for each bin, how many times the system was able to observe photons.

Finally, the status histogram is divided by the total amount of laser periods composing the experiment to obtain a curve $\alpha(t)$, representing for each bin the fraction of time the system was active. In other words, $\alpha(t)$ represents the probability that a photon impinging at a certain time is recorded.

In can be shown that, regardless of the shape of luminous power over time, the impinging rate $P_{imp}(t)$ can be recovered starting from $P_{rec}(t)$ and $\alpha(t)$, with a simple division [32]:

$$P_{imp}(t) = \frac{P_{rec}(t)}{\alpha(t)}. \tag{14}$$

If the shape of $P_{rec}(t)$ is known, it is possible to derive a closed-form expression of $\alpha(t)$ from a direct comparison between Eq. (14) and Eq. (9):

$$\alpha(t) = 1 - \int_{t-T_{dead}}^{t} \sum_{i=-\infty}^{+\infty} P_{rec}(t' + i \cdot T_{laser}) \cdot dt'. \tag{15}$$

An example of $P_{imp}(t)$, $P_{rec}(t)$ and $\alpha(t)$ is reported in Fig. 15 for a mono-exponential $P_{imp}(t)$ with τ equal to 2.5 ns and P equal to 5 photons per period, considering a dead time of 1 ns and a laser period equal to 12.5 ns (= 5τ).

As expected, $\alpha(t)$ is always equal to or less than one. In particular, in this example, $\alpha(t)$ is almost 1 at t equal to zero, then it drops to a minimum, which is located at t equal to T_{dead} (= 1 ns). This immediate drop of $\alpha(t)$ shows that early photons have a better chance to be recorded, while later events are typically masked by dead time generated by early photons themselves. More in general, $\alpha(t)$ features a series of troughs located roughly at integer multiples of the dead time [32] (for example, the second trough in Fig. 15(a) is located at about 2 ns). It is worth highlighting that this behavior is valid for any value of τ and P. In addition, the higher is P, the deeper are troughs. It follows that at extremely high photon rates the minimum of $\alpha(t)$ can even approach zero, thus making the correction expressed in Eq. (14) impractical.

Even without reaching zero, low values of $\alpha(t)$ are not desirable, since they mean a low collection efficiency in some bins, which translates in a poor statistics. In general, the proposed correction permits to reconstruct perfectly the shape of $P_{imp}(t)$, but unfortunately it cannot enhance the statistics of a poor collection of photons, so bins with poor statistics keep the poor statistics also after correction. As a result, the reconstructed curve will show higher statistical noise at bins corresponding to low values of $\alpha(t)$.

Optics EXPRESS

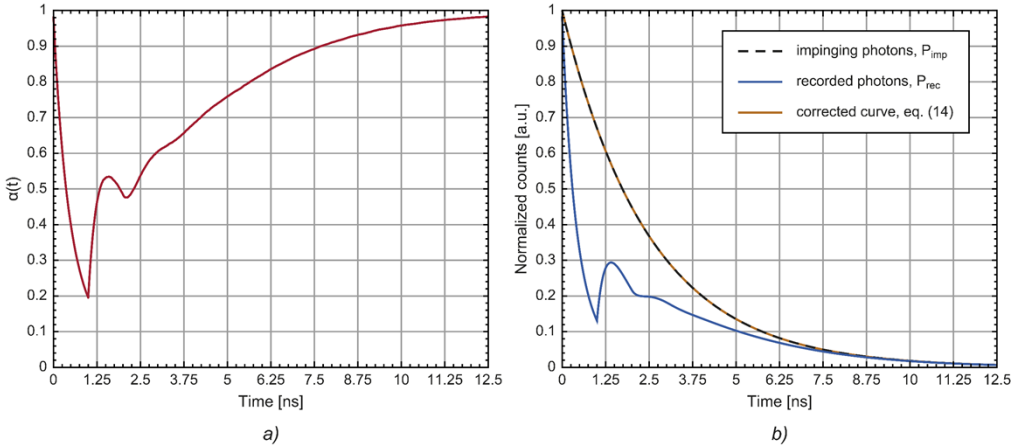

Fig. 15. a) $\alpha(t)$ calculated using Eq. (15) after solving Eq. (9) numerically, b) curves obtained from Eq. (9). The corrected curve is recovered dividing $P_{rec}(t)$ by $\alpha(t)$, as expressed by Eq. (14). All curves have been calculated considering $P_{imp}(t)$ equal to a mono-exponential decay with τ equal to 2.5 ns and P equal to 5 photons per period, a dead time of 1 ns and T_{laser} equal to 12.5 ns.

The effect of noise due to troughs of $\alpha(t)$ can be shown by means of a numerical simulation, which emulates the arrival of photons starting from a Poisson model [32]. An example of simulation result is reported in Fig. 16, considering the same case of Fig. 15.

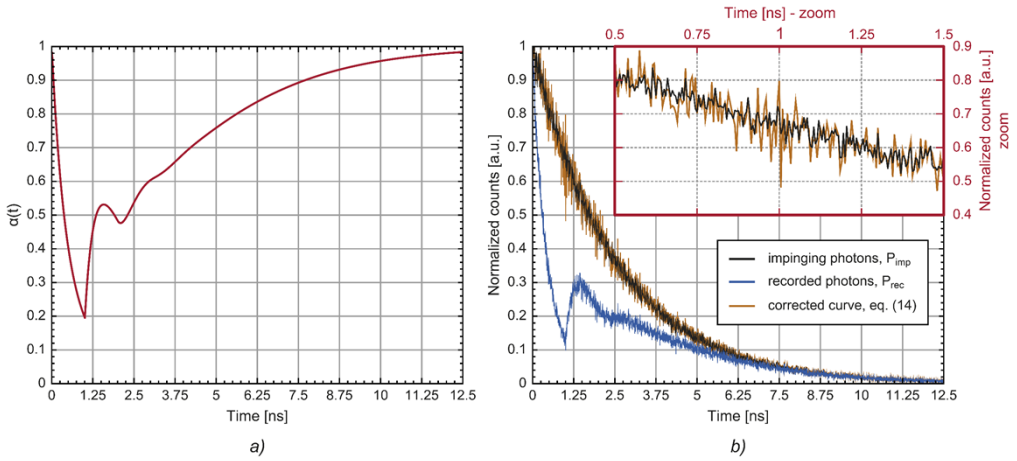

Fig. 16. Histograms obtained by numerical simulation: a) $\alpha(t)$ obtained as described in [32]; b) Histograms of recorded photons in case of ideal system (P_{imp}) and in case of dead time equal to 1 ns (P_{rec}). The corrected curve is recovered dividing the histogram $P_{rec}(t)$ by $\alpha(t)$. The plot outlined in red in the top right corner is a zoom of the main plot, aimed at better highlight the noise difference between impinging and corrected curves. The simulation considered $P_{imp}(t)$ equal to a mono-exponential decay with τ equal to 2.5 ns and P equal to 5 photons per period, a system dead time of 1 ns and T_{laser} equal to 12.5 ns. Histogram bins are 5-ps wide.

The trends are the same depicted in Fig. 15, but here the curves show a superposed noise due to the poissonian nature of light. It is worth noting that even if Poisson noise is present in the

Optics EXPRESS

experiment the noise superposed on $\alpha(t)$ is negligible, as discussed in [32]. As a result, $\alpha(t)$ in Fig. 16(a) is practically equal to the curve in Fig. 15(a), which was calculated mathematically. Finally, the reconstructed curve, obtained by the division $P_{rec}(t)/\alpha(t)$, shows the same behavior of $P_{imp}(t)$, but with enhanced noise, especially around the minimum of $\alpha(t)$ (see the zoom in Fig. 16(b)).

In short, the solution proposed in [32] completely avoids any tradeoff between speed and distortion, but a tradeoff still exists between speed and noise: the higher the photon rate, the higher the speed, but at the same time the deeper the troughs of $\alpha(t)$ and so the higher the noise.

In most cases having higher noise concentrated mainly in small portions of the histogram does not represent a significant issue. For instance, in fluorescence experiments involving a mono-exponential curve, fitting algorithms for the extraction of a single time-constant are typically very robust to noise, so a high localized noise enhancement can be accepted (e.g. barycenter estimation is based on area calculation, so noise is filtered out). Moreover, very high rates are needed to get a well visible noise enhancement, which is typically not the case in real experiments.

Nevertheless, depending on the experiment nature, a non-uniform noise enhancement can be an issue for fitting algorithms in some cases, so it is recommended to limit P to avoid having troughs of $\alpha(t)$ that are too deep. To this purpose, it is useful to quantify mathematically the impact of noise on the final histogram to find the ultimate limit to speed. It is well known that $P_{imp}(t)$ follows a Poisson distribution [13], so noise superposed on the curve has a standard deviation $\sigma_{imp}(t)$ equal to the square root of the curve itself:

$$\sigma_{imp}(t) = \sqrt{P_{imp}(t)}. \tag{16}$$

$P_{rec}(t)$ collects a sub-set of the events contained in $P_{imp}(t)$. In particular, if a photon hits the detector during a dead time, it is discarded, otherwise it is accumulated in $P_{rec}(t)$. In this scenario, the standard deviation of $P_{rec}(t)$ can be still estimated with the square root of $P_{rec}(t)$ itself. Recalling Eq. (14), the result is the following:

$$\sigma_{rec}(t) = \sqrt{P_{rec}(t)} = \sqrt{\alpha(t)} \cdot \sqrt{P_{imp}(t)}. \tag{17}$$

Finally, the shape of $P_{imp}(t)$ is recovered dividing $P_{rec}(t)$ by $\alpha(t)$, so the resulting standard deviation is equal to $\sigma_{rec}(t)/\ \alpha(t)$, that is $\sigma_{imp}(t)/\ \sqrt{\alpha(t)}$. It is worth highlighting that this is valid since standard deviation of measured $\alpha(t)$ is negligible [32], so it does not add its own contribution to the overall noise.

In this scenario, the inverse of the square root of $\alpha(t)$ represents a sort of noise amplification with respect to Poisson noise superposed on $P_{imp}(t)$. Since this amplification is not constant over time, the original noise profile gets distorted. In particular, maximum noise enhancement corresponds to the deepest trough of $\alpha(t)$. In this scenario, it is possible to quantify noise distortion starting from the depth of the deepest trough, which can be quantified as the ratio between the minimum and the maximum of $\alpha(t)$.

It follows that maximum noise amplification, $A_{noise,max}$ is equal to:

$$A_{noise,max} = \sqrt{\frac{\max\{\alpha(t)\}}{\min\{\alpha(t)\}}}\ . \tag{18}$$

It is worth noting that this parameter is always higher than or equal to 1, where 1 indicates that the noise profile of $P_{imp}(t)$ is not distorted. For example, considering $\alpha(t)$ in Fig. 15(a) and Fig. 16(a), the maximum of $\alpha(t)$ is located in the first bin (t = 0) and it is equal to 1, while the minimum is equal to 0.2, so $A_{noise,max}$ is roughly 2.2, meaning that noise around the minimum of

Optics EXPRESS

$\alpha(t)$ has been incremented 2.2 times more than noise in the first bin, as shown in the zoom of Fig. 16(b).

In cases where noise is an issue, it is important to fix a higher limit to $A_{noise,max}$, which translates into a limit to the maximum achievable speed.

In order to quantify the limit imposed by noise amplification to speed, it is possible to quantify $A_{noise,max}$ for different values of P and T_{dead}. Figure 17(a) shows how $A_{noise,max}$ varies as a function of both P and T_{dead} for a mono-exponential $P_{imp}(t)$ with $T_{laser} = 5\tau$. The result has been obtained combining Eq. (18) with Eq. (15).

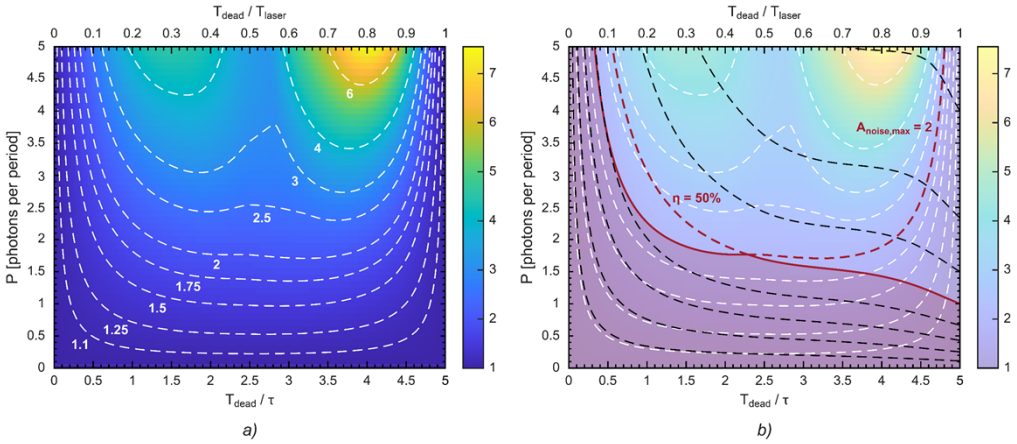

Fig. 17. Map of $A_{noise,max}$ as a function of P and T_{dead} for a mono-exponential $P_{imp}(t)$ with $T_{laser} = 5\tau$. a) Color map of $A_{noise,max}$. Contour lines are highlighted in white dashed lines for different values of $A_{noise,max}$. b) Map of $A_{noise,max}$ with superposed efficiency contour lines (black dashed lines). The working region with $\eta \geq 50\%$ and $A_{noise,max} \leq 2$ is highlighted in red.

Similarly to efficiency and distortion, a limit to noise amplification translates in a region delimited by a contour line. For example, considering a maximum noise enhancement equal to 2, the region where P and T_{dead} values are valid stays under the contour line corresponding to $A_{noise,max} = 2$.

Considering that distortion is always zero using the technique proposed in [32], noise amplification and efficiency represent the only limitations to speed. Figure 16(b) shows contour lines of η and $A_{noise,max}$ on the same plot. Here the area where η is higher than 50% and $A_{noise,max}$ is lower than 2 is highlighted in red. Comparing this area with the area in Fig. 14 (mainly limited by distortion), it is clear that the area in Fig. 16(b) is much wider, meaning that the solution proposed in [32] really allows the exploitation of much higher photon rates. Even better, every time noise enhancement is not an issue, the only limit is efficiency, meaning that the usable area is even larger.

7. Case study

The theory proposed in the previous sections permits to calculate the maximum operating rate for any case of interest. An example will be proposed in the following to show a comparison between different TCSPC setups.

Let us consider an exponential light signal with a relatively fast τ equal to 2 ns. In this case, the best choice for the laser period is 10 ns, corresponding to a 100-MHz excitation frequency. Once the *SDE* has been optimized, three solutions are available to maximize speed: i) selecting a

Optics EXPRESS

detector with state-of-art short dead time, ii) matching the dead time with the laser period, iii) using the solution proposed in [32].

Best state-of-art detectors have dead time values slightly lower than 1 ns. Let us consider the first solution and a very fast detector with $T_{dead} = 500ps$. As explained in section 5.2, in this case the ultimate limit to speed is imposed by distortion. Since $T_{laser} = 5\tau$, Fig. 13 and Fig. 14 are valid map of distortion as a function of both T_{dead} and P. In this case T_{dead} is equal to 0.25τ, so P has to be kept below 0.09 photons per period to keep distortion below 1% (it can be inferred from Fig. 14(b)). Referring to Fig. 9, $T_{dead} = 0.25\tau$ and $P = 0.09$ correspond to almost unitary efficiency (to be precise $\eta = 0.99$). This result is confirmed calculating η_{trend} and $\eta_{plateau}$, expressed by Eqs. (10) and (11). Indeed $\eta_{trend} = 0.996$ and $\eta_{plateau} = 0.96$. The correct value of η stays in between.

In short, the exploitation of a low dead time of 500 ps leads to a maximum rate of about 0.09 recorded photons per period, corresponding to a speed equal to 9 photons/µs. Figure 18 shows the result of a numerical simulation, which reproduces this situation. Here the histogram of recorded photons is depicted along the ideal histogram containing all the impinging photons, considering an experiment duration of 55.5 ms, which corresponds to 500,000 recorded photons. It is worth highlighting that the two curves practically overlap, since distortion is only 1% and efficiency approaches unity.

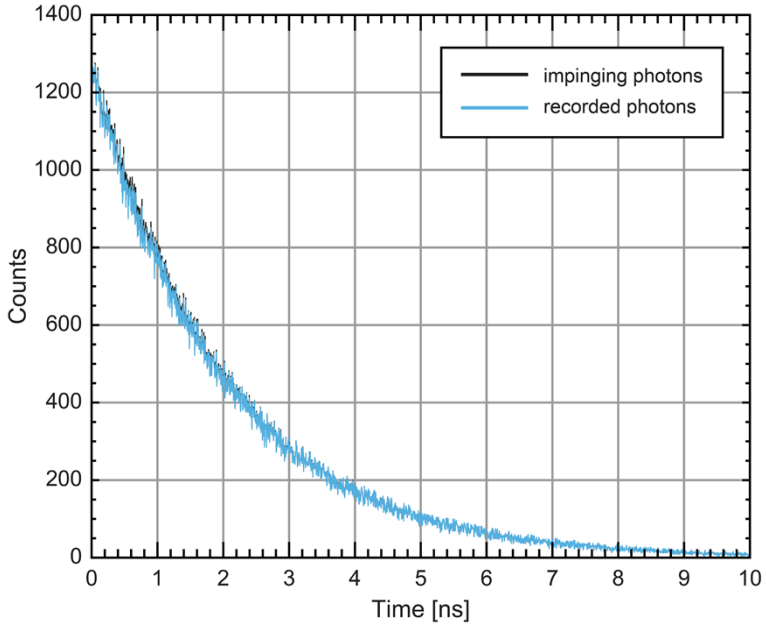

Fig. 18. Result of a numerical simulation lasting for 55.5 ms considering $P_{imp}(t)$ equal to a mono-exponential decay with τ equal to 2 ns, P equal to 0.09 photons per period, T_{laser} equal to 10 ns and T_{dead} equal to 500 ps. Histogram bins are 5-ps wide.

The second way to maximize speed is to tune the dead time to be equal to the laser period, i.e. 10 ns. In this case, distortion is not a limit. Nevertheless, the exploitation of a relatively long dead time leads to low efficiency. In this case, maximum speed is limited by the minimum efficiency one wants to guarantee. Let us consider this limit is equal to 50%. As explained in section 4.1, in case of dead time matched to the laser period η_{trend} coincides with η, so Eq. (10) can be used to calculate the photon rate P corresponding to $\eta = 50\%$. The result is a maximum P equal to 1 photon per period. This result is confirmed by Fig. 14(b).

Optics EXPRESS

Since efficiency is 50%, the average number of recorded photons per period is 0.5, corresponding to a speed equal to 50 photons/μs, that is much higher than speed reached with the first solution.

Figure 19 reports a numerical simulation reproducing this situation. In this case the experiment duration has been set to 10 ms to get 500,000 recorded photons. It is worth highlighting that in this case the two curves feature exactly the same trend, since distortion is 0. However, the recorded histogram contains half the counts compared to the curve of impinging photons.

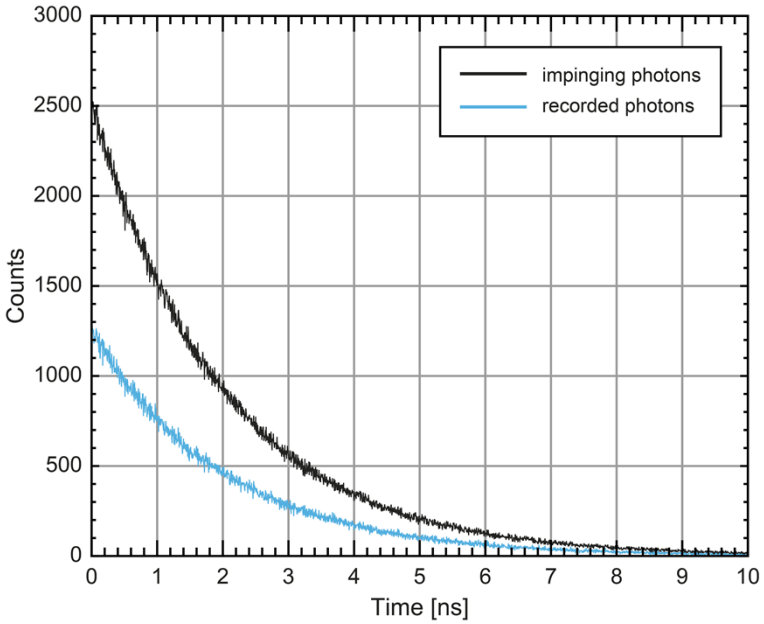

Fig. 19. Result of a numerical simulation lasting for 10 ms considering $P_{imp}(t)$ equal to a mono-exponential decay with τ equal to 2 ns, P equal to 1 photons per period, T_{laser} equal to 10 ns and T_{dead} equal to T_{laser}. Histogram bins are 5-ps wide.

Finally, the third solution [32] permits to combine short dead time (e.g. 500 ps) with zero distortion. In this case, the ultimate limit to speed is given either by efficiency or by noise enhancement, as explained in section 6. Considering a minimum efficiency of 50% and a noise amplification limit equal to 2, the useful P-T_{dead} region is depicted in Fig. 16(b). Referring to the figure and considering T_{dead} equal to 500 ps (5% of T_{laser} or 0.25τ) P can be raised above 5 photons per period (the exact value is $P = 6.3$ photons per period, limited by noise amplification). At that value of P efficiency is 61%, meaning that maximum speed is equal to 384 photons/μs.

Figure 20 shows the result of a numerical simulation of this situation, considering an experiment duration of 1.3 ms, which corresponds to 500,000 recorded photons. In this case the recorded histogram (light blue curve) does not feature the same shape of the ideal histogram of the impinging photons. However, the system also records $\alpha(t)$ and calculates a corrected histogram using Eq. (14). This histogram features the same shape of the ideal curve, but with higher superposed noise. In particular, the peak-to-peak noise is doubled in correspondence of the minimum of $\alpha(t)$. It is worth noting that this effect is barely visible observing the histogram as a whole and a zoom is required to note it.

The proposed case study is only an example to show how the mathematical background and the graphical methods provided in this paper can be used for maximum speed estimation with different TCSPC setups. In any case, it shows that, in general, novel techniques have paved the way to overcome classical TCSPC limitations, thus opening the way for an effective exploitation

Optics EXPRESS

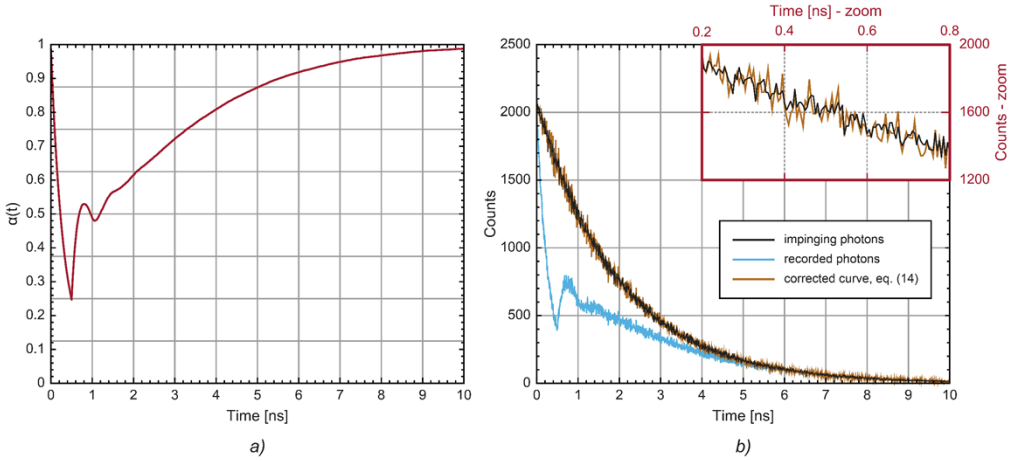

Fig. 20. Result of a numerical simulation lasting for 1.3 ms considering $P_{imp}(t)$ equal to a mono-exponential decay with τ equal to 2 ns, P equal to 6.3 photons per period, T_{laser} equal to 10 ns and T_{dead} equal to 500 ps. Histogram bins are 5-ps wide. a) Simulated $\alpha(t)$. b) Histograms of recorded photons in case of ideal system (P_{imp}) and in case of dead time equal to 500 ps (P_{rec}). The corrected curve is recovered dividing the histogram $P_{rec}(t)$ by $\alpha(t)$. The plot outlined in red in the top right corner is a zoom of the main plot, aimed at better highlight the noise difference between impinging and corrected curves.

of TCSPC in those cases where the photon rate coming from the sample is not a constraint, e.g. LiDAR, in vivo surgery, single-pixel camera applications and many others [7–12].

8. Conclusions

Historically, the main limit of TCSPC has been low speed due to distortion phenomena. In particular, in a classic TCSPC experiment the average number of recorded photons is typically limited to few percent of the excitation period.

Nowadays, the recent advances in single photon detectors and associated timing electronics permit to surpass this limit, but a proper sizing of the measurement setup is fundamental to reach the best result.

In this paper we analyzed the main causes that limit speed in TCSPC experiments and we proposed a detailed method to calculate maximum speed in different operating conditions.

The proposed method, based both on graphical analysis and analytical calculations, permits to have a better insight in speed-related issues in TCSPC experiments and to select the best suited solution to maximize speed depending on the specific case.

A. Appendix

Equation (11) describes a situation where the dead time is sufficiently long to cover an integer number N of cycles, meaning that a single photon is recorded during a laser period, then no other photon is recorded for the *N-1* consecutive periods. An example of this situation is depicted in Fig. 21 for $N = 3$. In this scenario, for each recorded photon there are *N-1* blind periods, which do not contribute to recording events (for example, in Fig. 21 two periods are completely blind due to dead time: periods #2 and #3). If P_{rec} is the average number of recorded photons per period, the fraction of blind periods during a TCSPC experiment is equal to $(N-1)\cdot P_{rec}$, meaning that the fraction of non-blind periods is equal to $1-(N-1)\cdot P_{rec}$.

Optics EXPRESS

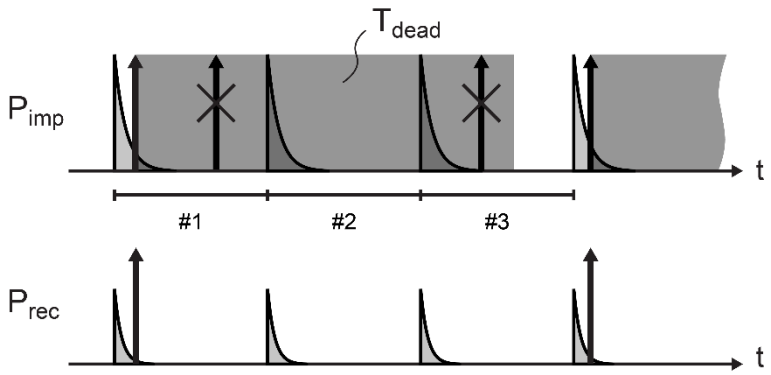

Fig. 21. Example of situation described by Eq. (11), which corresponds to third plateau in Fig. 4. Here photon events are highlighted with arrows, which are crossed when the photon is not recorded. On top exponential pulses represent the average number of impinging photons as a function of time, while the average number of recorded photons is depicted below. In this example every time a photon is recorded a dead time starts (dark grey rectangle), that is sufficiently long to cover all the events for three consecutive periods (#1, #2 and #3).

It is worth noting that P_{rec} is always lower than 1, since maximum one impinging photon can be recorded in a period (see Fig. 21). It follows that P_{rec} represents the probability to record a photon during a period, so it can be calculated as the product between the probability that at least a photon impinges on the detector during a period and the probability that that period is not blind. Considering a Poisson distribution of light, the probability to observe at least an impinging photon during a period is equal to $1 - exp(-P)$, while the probability that the photon does not impinge during a blind period is equal to the fraction of non-blind periods, i.e. $1 - (N-1) \cdot P_{rec}$. As a result:

$$P_{rec} = (1 - e^{-P}) \cdot [1 - (N-1) \cdot P_{rec}] \,. \tag{19}$$

It follows that

$$P_{rec} = \frac{(1 - e^{-P})}{1 + (1 - e^{-P}) \cdot (N-1)}. \tag{20}$$

Counting efficiency η is equal to the ratio between P_{rec} and P. Considering a duration of the dead time T_{dead}, the term $(N-1)$ can be calculated applying the floor function on the ratio T_{dead}/T_{laser}. The result is Eq. (11).

It is worth noting that Eq. (20) has been derived without relying on a specific shape of the luminous signal. It follows that Eq. (11) is valid for any shape, not only for mono-exponential signals.

Funding. Ministero dell'Università e della Ricerca (PRIN2022, Proactive); European Research Council (HÈRMES, 101116943).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. D. V. O'Connor and D. Phillips, *Time-correlated single photon counting* (Academic Press, 1984).
2. B. Wolfgang, "Advanced time-correlated single photon counting techniques," in *Springer Series in Chemical Physics*, (Springer, 2005).
3. K. Suhling, L. M. Hirvonen, J. A. Levitt, *et al.*, "Fluorescence lifetime imaging (FLIM): Basic concepts and some recent developments," Medical photonics. **27**, 3–40 (2015).
4. X. Liu, D. Lin, W. Becker, *et al.*, "Fast fluorescence lifetime imaging techniques: A review on challenge and development," J. Innov. Opt. Health Sci. **12**(05), 1930003 (2019).

Optics EXPRESS

5. H. Wallrabe and A. Periasamy, "Imaging protein molecules using FRET and FLIM microscopy," Curr. Opin. Biotechnol. **16**(1), 19–27 (2005).
6. P. Kapusta, M. Wahl, Aleš Benda, *et al.*, "Fluorescence lifetime correlation spectroscopy," J. Fluoresc. **17**(1), 43–48 (2006).
7. A. M. Pawlikowska, A. Halimi, R. A. Lamb, *et al.*, "Single-photon three-dimensional imaging at up to 10 kilometers range," Opt. Express **25**(10), 11919–11931 (2017).
8. Z. Li, E Wu, C. Pang, *et al.*, "Multi-beam single-photon-counting three-dimensional imaging lidar," Opt. Express **25**(9), 10189–10195 (2017).
9. X. Zhou, J. Bec, K. Ehrlich, *et al.*, "Pulse-sampling fluorescence lifetime imaging: evaluation of photon economy," Opt. Lett. **48**(17), 4578–4581 (2023).
10. F. Rousset, N. Ducros, Françoise Peyrin, *et al.*, "Time-resolved multispectral imaging based on an adaptive single-pixel camera," Opt. Express **26**(8), 10550–10558 (2018).
11. A. Perri, J. H. Gaida, A. Farina, *et al.*, "Time-and frequency-resolved fluorescence with a single TCSPC detector via a Fourier-transform approach," Opt. Express **26**(3), 2270–2279 (2018).
12. A. Ghezzi, E. Avanzi, A. García Fleitas, *et al.*, "High throughput compressive fluorescence lifetime imaging with a silicon photomultiplier detector," Opt. Express **32**(14), 24553–24562 (2024).
13. M. Köllner and W. Jürgen, "How many photons are necessary for fluorescence-lifetime measurements?" Chem. Phys. Lett. **200**(1-2), 199–204 (1992).
14. A Cominelli, G Acconcia, P Peronio, *et al.*, "High-speed and low-distortion solution for time-correlated single photon counting measurements: A theoretical analysis," Rev. Sci. Instrum. **88**(12), 123701 (2017).
15. A. Bovolenta, A. Cominelli, G. Acconcia, *et al.*, "Quantifying distortion in time-correlated single photon counting: a universal parameter," Opt. Lett. **49**(6), 1563–1566 (2024).
16. P Peronio, G Acconcia, I Rech, *et al.*, "Improving the counting efficiency in time-correlated single photon counting experiments by dead-time optimization," Rev. Sci. Instrum. **86**(11), 113101 (2015).
17. S. Farina, G. Acconcia, I. Labanca, *et al.*, "Toward ultra-fast time-correlated single-photon counting: A compact module to surpass the pile-up limit," Rev. Sci. Instrum. **92**(6), 063702 (2021).
18. F. Malanga, G. Acconcia, S. Farina, *et al.*, "Versatile multichannel time-to-amplitude converter for high-speed and high-precision timing applications," Advanced Photon Counting Techniques XVI. SPIE **22**, 22 (2022).
19. Time–Correlated, High Resolution, and High-Speed Time Tagger. "PicoHarp 330."
20. S. Cova, M. Ghioni, A. Lacaita, *et al.*, "Avalanche photodiodes and quenching circuits for single-photon detection," Appl. Opt. **35**(12), 1956–1976 (1996).
21. A. Gallivanoni, I. Rech, and M. Ghioni, "Progress in quenching circuits for single photon avalanche diodes," IEEE Transactions on nuclear science. **57**(6), 3815–3826 (2010).
22. S. Farina, I. Labanca, G. Acconcia, *et al.*, "Above pile-up fluorescence microscopy with a 32 Mc/s single-channel time-resolved SPAD system," Opt. Lett. **47**(1), 82–85 (2022).
23. S. Farina, A. Ghezzi, I. Labanca, *et al.*, "Pile-up free microscopy imaging with a single-pixel camera," Single Molecule Spectroscopy and Superresolution Imaging XVII. SPIE **PC12849**, 24 (2024).
24. A. Diaspro, G. Chirico, C. Usai, *et al.*, "Photobleaching," *Handbook of biological confocal microscopy* 690–702 (2006).
25. N. Cristiano and M. Soga, "A miniature actively recharged single-photon detector free of afterpulsing effects with 6 ns dead time in a 0.18 μm CMOS technology," *International Electron Devices Meeting* (IEEE, 2010).
26. F. Ceccarelli, G. Acconcia, A. Gulinatti, *et al.*, "Fully integrated active quenching circuit driving custom-technology SPADs with 6.2-ns dead time," IEEE Photon. Technol. Lett. **31**(1), 102–105 (2019).
27. F. Gramuglia, M. L. Wu, C. Bruschini, *et al.*, "A low-noise CMOS SPAD pixel with 12.1 ps SPTR and 3 ns dead time," IEEE Journal of Selected Topics in Quantum Electronics. **28**(2: Optical Detectors), 1–9 (2022).
28. S. Farina, I. Labanca, G. Acconcia, *et al.*, "10-nanosecond dead time and low afterpulsing with a free-running reach-through single-photon avalanche diode," Rev. Sci. Instrum. **93**(5), 053102 (2022).
29. A. Giudici, G. Acconcia, I. Labanca, *et al.*, "4 ns dead time with a fully integrated active quenching circuit driving a custom single photon avalanche diode," Rev. Sci. Instrum. **93**(4), 043103 (2022).
30. F. Severini, I. Cusini, D. Berretta, *et al.*, "SPAD pixel with sub-ns dead-time for high-count rate applications," IEEE Journal of Selected Topics in Quantum Electronics. **28**(2: Optical Detectors), 1–8 (2022).
31. W. Hwang, D. Kim, S. Moon, *et al.*, "Achieving a high photon count rate in digital time-correlated single photon counting using a hybrid photodetector," Opt. Express **29**(7), 9797–9804 (2021).
32. I. Rech, A. Bovolenta, A. Cominelli, *et al.*, "Toward Constraintless Time-Correlated Single-Photon Counting Measurements: A New Method to Remove Pile-Up Distortion," IEEE Journal of Selected Topics in Quantum Electronics (2023).
33. G. Fratta, P. Daniele, I. Labanca, *et al.*, "Near-zero distortion in TCSPC at more than one photon per excitation period: experimental validation," Opt. Lett. **49**(17), 4958–4961 (2024).