Silicon Photon-Counting Avalanche Diodes for Single-Molecule Fluorescence Spectroscopy

Xavier Michalet, Antonino Ingargiola, Ryan A. Colyer, Giuseppe Scalia, Shimon Weiss, Piera Maccagnani, Angelo Gulinatti, *Member, IEEE*, Ivan Rech, *Member, IEEE*, and Massimo Ghioni, *Senior Member, IEEE*

Manuscript received February 24, 2014; revised June 26, 2014; accepted July 14, 2014. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Awards 5R01 GM095904 (UCLA and Politecnico di Milano) and 5R01 GM069709 (UCLA), and by UCLA-DOE Institute for Genomics and Proteomics, under grant DE-FC02-02ER63421 (UCLA).

X. Michalet, A. Ingargiola, and S. Weiss are with the Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90046 USA (e-mail: michalet@chem.ucla.edu; tritemio@gmail.com; sweiss@chem.ucla.edu).

R. A. Colyer was with the Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90046 USA. He is now with the Department of Science, Cabrini College, Radnor, PA 19087 USA (e-mail: ryan.colyer@cabrini.edu).

G. Scalia was with the Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90046 USA. He is now with the Département de Physique, Université de Fribourg, 1700 Fribourg, Switzerland (e-mail: giuseppe.scalia@unifr.ch).

P. Maccagnani is with the Istituto per la Microelettronica e Microsistemi, Sezione di Bologna, 40129 Bologna, Italy (e-mail: maccagnani@bo.imm.cnr.it).

A. Gulinatti, I. Rech, and M. Ghioni are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milano, Italy (e-mail: angelo.gulinatti@polimi.it; ivan.rech@polimi.it; massimo.ghioni@polimi.it).

I. Introduction

THE need to detect single-molecules arises in multiple fields [1]–[3]. In basic science, elucidating the steady-state characteristics of heterogeneous mixtures or studying intrinsically stochastic molecular mechanisms is best done observing one molecule at a time and reproducing this observation on a statistically significant number of molecules. This allows sampling the complete distribution of configurations characterizing the static or dynamic distribution of the system's properties. Single-molecule detection is also the only way to detect rare events or rare species in a mixture. This capability is of direct relevance to the field of ultrasensitive detection of pathogens, biomarkers or chemicals, due to its ultimate sensitivity.

Single-molecule detection can be achieved using different physical effects, but optical means in particular have the advantage of allowing non-destructive detection and being usable with different sample states (gaseous, liquid, solid) and experimental arrangements (flow, diffusion, surface immobilization).

Single-molecule fluorescence spectroscopy (SMFS) refers to a now wide-spread set of techniques having in common dealing with the usually very weak fluorescence signal emitted by individual fluorophores [4]–[6]. Fluorophores can be intrinsic molecular moieties or external labels chemically attached to the target species. Several types of spectroscopic signatures can be used to distinguish different molecular species. Most can be measured at the single-molecule level: fluorescence emission intensity, spectrum, polarization (or polarization anisotropy), lifetime, energy transfer, etc., can all be extracted from individual molecule bursts, although data analysis can also be performed at the level of many such molecule bursts (or subpopulation level) rather than at the true single-molecule level.

A number of reviews have been published over the past few years describing the various fields of applications of this general approach and the typical setups employed in these experiments [7]–[9]. We will therefore limit ourselves to a very basic presentation of SMFS. In particular, we will emphasize the constraints imposed by current detector technologies, a topic that has been recently addressed in a broader context in [10]. We will then present our work on new silicon single-photon avalanche diodes (SPADs) and SPAD arrays, which meet many of the detector requirements for SMFS.

Detecting a single-molecule by optical means requires that the number of photons emitted by that molecule is large enough so that the signal-to-noise ratio (SNR) is significantly larger than 1. Additionally, it is important to be able to separate successive or nearby molecules.

This second *separability* criterion depends on the details of each experiment [4]. Detector characteristics play an important role in this respect, as we will discuss in Section VII.

The first *detectability* condition has several facets, which have been discussed in detail in [10]. From a detector point of view, it requires high sensitivity and low readout noise, and in the case of interest in this article, reviewed in the next section, high temporal resolution. These goals are best achieved with a photon-counting detector.

Photon-counting detectors exploit an internal amplification mechanism, which, in response to single-photons, generates macroscopic electrical signals that are much larger than the circuit noise. In addition to this beneficial characteristic, they also provide the possibility to not only count the number of photons, but measure their individual arrival time. This allows studying intensity fluctuations at very short as well as longer time scales. In the case of pulsed excitation with femtosecond (fs) to picosecond (ps) laser pulses, the best photon-counting detectors and associated time-correlated single-photon counting (TCSPC) techniques give access to the time elapsed since excitation of the fluorophore with ps resolution. This latter information is of particular interest when studying the environment of a fluorophore, as the presence of additional non-radiative relaxation channels reduces the measured fluorescence lifetime. To keep this discussion short, we will not review this important aspect of fluorescence spectroscopy and refer the interested reader to an extensive literature on fluorescence lifetime measurements [11], [12]. However, from a detector point of view, these measurements require instrument response function (IRF) in the sub-ns range.

Photon-counting and TCSPC techniques were originally developed using photomultiplier tubes (PMT), which are vacuum tube detectors with high internal gain. PMTs have been produced industrially with sophisticated technologies since the 1960. Commercially available devices can provide remarkable performance, with maximum count rates of millions of counts per second (MHz). Amongst their advantages, the most significant and distinct is their wide sensitive area ($\sim$cm^2), which in some cases greatly simplifies the design of the optical system. Micro-channel plate (MCP) PMTs also offer very narrow IRF with width in the ps range. However, PMTs suffer from low photon detection efficiencies (PDE), especially in the red and near-infrared wavelength region: conventional multialkali photocathodes have quantum efficiencies that typically fall from $\sim$20% at 400 nm to less than 1% at 800 nm. GaAsP photocathodes reach a peak quantum efficiency of 45% at 500 nm, but they are almost insensitive to wavelengths longer than 720 nm. GaAs photocathodes provide a relatively flat quantum efficiency of about 25% between 500 and 850 nm [11], [13].

Semiconductor-based detectors represent a valuable alternative to PMTs. Besides the well-known advantages of solid state versus vacuum tube devices (small size, ruggedness, low power dissipation, low supply voltage, high reliability, better ability to handle large photon flux without irreversible damage, etc.), semiconductor detectors provide inherently higher quantum efficiency, particularly in the red and near-infrared spectral regions [11].

Avalanche multiplication of carriers in reverse-biased p-n junctions is used in ordinary avalanche photodiodes (APDs) to obtain internal amplification in the detector similar to that in PMTs. However, in an ordinary APD the multiplication of both holes and electrons causes an inherent positive feedback, which produces strong fluctuations in the avalanche gain. In the best case, the achievable gain is limited to a few hundreds, as opposed to a gain of $\sim 10^6$ easily reached by PMTs. Such a gain is not sufficient to detect single-photons because the resulting single-photon signal is of the same order of magnitude as the electronic circuit noise.

Yet the positive feedback can be exploited to detect single-photons by giving up the linear relationship between input (photons) and output signals in a different kind of avalanche diodes. In these single-photon avalanche diodes (SPADs), the device is biased above the breakdown level and the photodiode is no longer a 'detector with amplifier inside', but is instead a 'detector with flip-flop inside'. The 'flip-flop' is set when a single-photon is absorbed and the photo-generated carrier succeeds in triggering the avalanche process. An electronic circuit associated to the detector (quenching circuit) then senses the avalanche current pulse, produces a standard high level output pulse, drives down the bias voltage to quench the avalanche and finally resets it to the operating level. Compared to PMTs, SPADs provide higher photon detection efficiencies (see Section IV) but smaller sensitive areas ($\sim$5 to 500 μm diameter), which might make optics alignment more challenging in some cases (see Section II). Indeed, in SMFS applications, the image of the emission spot needs to be aligned onto the SPAD active area, which becomes increasingly difficult as the size of the active area is reduced, due to the finite precision of standard micrometer translation stages used for alignment purpose.

Before discussing detector requirements in detail, Section II will briefly examine the technique and data analysis used in standard solution-based SMFS. Section III then discusses the sensitivity limit in single-molecule burst analysis. Section IV analyzes key detector parameters such as detection efficiency, dark count rate, afterpulsing probability and timing jitter. Sectio V then reviews available SPAD technologies and Section VI compares the performance of two state-of-the-art SPAD detectors in single-spot SMFS experiments. Section VII will then report new SPAD array detectors for diffusing molecules measurements in multispot geometries, which overcome throughput limitations of single-spot geometry. We conclude this article with a brief overview of future prospects for the field.

II. Solution-Based Single-Molecule Fluorescence Spectroscopy

SMFS can be performed on immobilized or freely diffusing molecules. The detectors discussed in this article were targeted to the latter approach. Since detailed presentation of this kind of experiments can be found in reviews [1]–[3], [7]–[10], [14], we limit ourselves to a few basic concepts needed for understanding detector performance targets.

A simple way to facilitate the detection of freely diffusing fluorescent molecules is by ensuring that the only region in which molecules are excited is small enough and the concentration of molecules sufficiently low that there is at most one molecule in it at any given time. One way of achieving this is to use

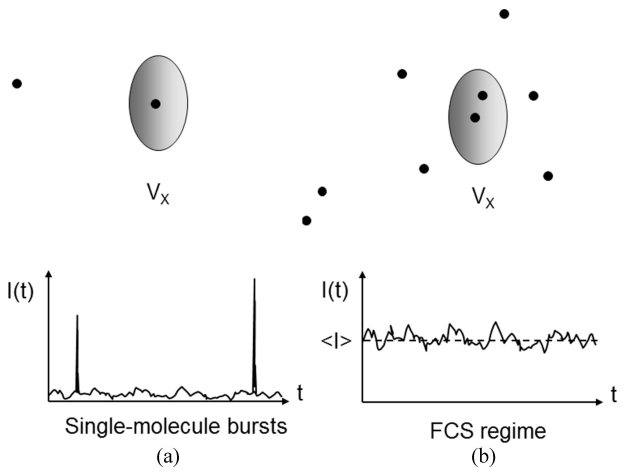


Fig. 1. (a) Single-molecule burst detection requires at most one molecule at any given time in the excitation volume V_x. The resulting signal consists of brief bursts corresponding to the transit of a single-molecule through the excitation volume. (b) The FCS regime is characterized by a larger concentration of $\sim$1 or more molecules in the excitation volume V_x at any time. This larger concentration prevents identifying single-molecule bursts in the signal time trace, which looks like a noisy, fluctuating signal.

a fixed point-like excitation geometry, where single-molecules stochastically diffuse in and out of the excitation region. The transit duration through the excitation volume, τ_D, depends on the lateral dimension of the region, *R*, and the molecule's diffusion coefficient *D*:

$$\tau_D \sim \frac{R^2}{4D}. \tag{1}$$

Typical values for single molecule in aqueous solutions are of the order of 100 μs to 1 ms. The mean separation between molecular transits, τ_R, is given by the inverse of the transit frequency, which depends linearly on *R* and *D* as well as on sample concentration *C*. The ratio τ_D/τ_R is simply equal to the average number *N* of molecules within the excitation region, which thus clearly needs to be small compared to 1 in order to allow separating successive molecules [15].

A variety of microscopy approaches can been used to achieve this condition [5], [6], [16]–[19]. The simplest consists in tightly focusing an expanded and collimated laser beam into the sample using a high numerical aperture lens. For instance, a diffraction-limited excitation volume of a fraction of femtoliter (1 fl $=$ 10^{-15} L $= 1\ \mu\text{m}^3$) can easily be achieved in a standard confocal microscope. The corresponding single-molecule concentration regime is of the order of 100 pM (picomolar) or less.

A. Single Burst Detection and Analysis

With the appropriate setup, distinct photon bursts can be detected as individual molecules transit across the excitation/detection volume [Fig. 1(a)]. For diffusing molecules, burst duration depends on molecule size, solvent viscosity and excitation volume. Typical values range from a few tens of microseconds to a few milliseconds, with an approximately exponential distribution of burst durations [20]. Burst intensity (or burst size), i.e., the total number of detected photons during the transit of a single-molecule through the excitation volume, depends on excitation power, absorption cross section and fluorescence quantum yield of the molecule as well as the total detection efficiency [20]. Burst sizes can reach up to a few hundred photons, but as for burst durations, their typical distribution is quasi-exponential, resulting in a large fraction of bursts having few dozen photons or less [20]. Since small bursts have small SNR, they increase the variance of quantities computed by averaging data from many single-molecule bursts [21], [22] and are generally rejected.

Single-molecule burst detection can be used to perform more than mere photon-counting when combined with various spectroscopic techniques. For instance, the conformation of a molecule (protein, nucleic acid, etc.) or the respective location of two interacting molecules can be studied by monitoring the distance between two dyes attached to specific sites of the molecule(s) using the phenomenon of fluorescence resonant energy transfer (FRET) [8], [23], [24]: when the two dyes have overlapping absorption and emission spectra and are in close proximity, intermolecular non-radiative energy transfer can occur by dipole-dipole interaction, with an efficiency varying with the inverse of the 6th power of the distance between the two dyes [12], [25]. This phenomenon has been extensively used in bulk in the past, but its use at the single-molecule level has exploded since its first demonstration on surface-immobilized molecules [26] and on freely-diffusing molecules in solution [27].

In standard single-molecule FRET (smFRET) experiments, doubly-labeled molecules or molecular complexes are excited by a single laser exciting the donor dye. The fluorescence signals from the donor and acceptor dyes are collected simultaneously in their respective emission spectral band by two separate detectors [27]. After identification of individual bursts, the donor and acceptor signals (I_D and I_A) resulting from donor excitation, corrected for background and other contaminating signals, are used to compute the FRET efficiency *E*:

$$E = \frac{I_A}{I_A + \gamma I_D}, \tag{2}$$

where γ is a correction factor, typically close to 1, accounting for the different quantum yield and detection efficiency of both fluorophores [28]. Although a rich and fascinating topic, its discussion would distract us from our main topic. We refer the interested reader to a recent introductory discussion in the context of detector development [14] as well as specialized and review articles [29]–[41]. We will briefly return to smFRET in Section VII, when discussing experimental results with SPAD arrays.

B. Fluorescence Correlation Spectroscopy (FCS)

A related experimental regime, illustrated in Fig. 1(b), is encountered when the excitation volume contains one or more molecules on average at any time. In a confocal microscope, this situation corresponds to sample concentration larger than 1 nanomolar (1 nM). In this regime, individual molecule bursts cannot be distinguished anymore and are replaced by a highly fluctuating signal *I(t)* centered on an average value $<I>$.

Analysis of the autocorrelation function (ACF) of these fluctuations ($\delta I(t) = I(t) - \langle I \rangle$) [42], [43] can yield physical as well as photochemical information (molecule size and concentration, blinking or binding/unbinding rates) about the diffusing

molecules. They can also give access to the microsecond time scale, intermediate between the nanosecond time scale accessible by TCSPC measurements and the typical single-molecule burst duration, of the order of 1 ms.

FCS techniques have the advantage of simplicity due to the existence of plug-and-play hardware correlators and well-established data analysis methods. They are also capable of detecting populations of molecules with sufficiently different sizes, the ACF of a mixture being a weighted sum of the ACFs of each population [42], [43]. We will not address this topic here.

III. Signal-to-Background and Signal-to-Noise Ratios in Single-Molecule Burst Analysis

Single-molecule fluorescence experiments represent a very particular case of low-light level situation: the single-molecule signal is low (a few hundred photons at most), brief (a few millisecond or less), and sparse (a burst every few tens of millisecond). In this regime, it is important to carefully examine all possible sources of background and noise, as well as their respective contributions to the SNR and SBR. There are two main sources of background counts: the sample and the detector. Whereas the user can try to reduce sample background, it is never totally absent. On the detector side, there are two sources of background signal: dark count and afterpulsing, whose origin and characteristics will be discussed in the next section. As both SNR and SBR play an important role in signal detection and data analysis, it is useful to examine all components of the recorded detector counts and their influence on the SNR and SBR, before discussing how the latter affect single-molecule experiments.

A. Signals

Counts recorded by a photon-counting detector during a single-molecule burst can have different sources:

1) The detector spontaneously generates dark counts, which can be characterized by a temperature- and excess bias voltage-dependent count rate n_D (see Section IV). Given a duration τ, the average number of dark counts, N_D, is:

$$N_D = n_D \tau. \tag{3}$$

2) A single-molecule may have a variable emission rate depending on a number of optical and photophysical parameters, which will mainly depend on its location with respect to the center of the excitation spot. Far away from the spot, the collected single-molecule photon rate $n_P^S(t)$ is negligible, while it can reach several 100 kHz at the center of a focused laser beam in a confocal microscope [see Fig. 2(a) for examples of typical single-molecule bursts]. The rate varies stochastically due to diffusion and other possible phenomena such as conformational fluctuations, photophysics, etc., resulting in a total number N_S of detected single-molecule photons during a burst:

$$N_S = \eta \int_{\text{start}}^{\text{stop}} n_P^S(t)\, dt = \eta\, \bar{n}_P^S\, \tau, \tag{4}$$

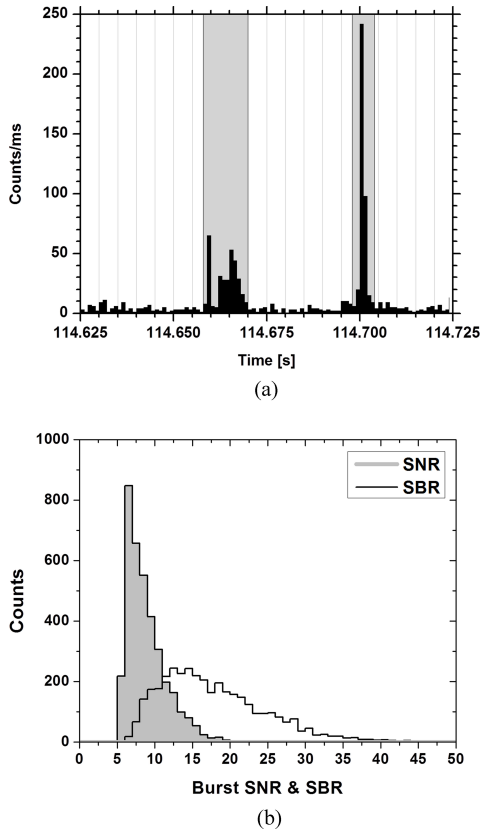
where η is the average detector PDE in the wavelength range selected by the optics in front of the detector, $\bar{n}_P^S$ the average photon rate impinging on the detector during the burst, and τ the burst duration, both depending on the burst under consideration.

Note that the definition of the beginning and end of a burst is not a trivial task and is to some extent arbitrary. We will get back to this issue in a later paragraph and simply assume for the time being that bursts have already been identified, each one being characterized by its own specific duration τ_i and total detected counts.

3) In addition to signal coming from single-molecules diffusing through the excitation volume, photons emitted by out-of-focus molecules, impurities or scattering processes contribute an average background photon rate, n_P^B, incident on the detector. The resulting average number of background counts during a single-molecule burst of duration τ is therefore:

$$N_B = \eta\, n_P^B\, \tau. \tag{5}$$

4) Finally, detector afterpulsing (discussed in Section IV) will add an average fraction ε ($\ll 1$) to the total number



(a)

(b)

Fig. 2. (a) 100 ms segment of a single-molecule time trace (black, bin: 1 ms) corresponding to a 30 pM solution of ATTO550-labeled dsDNA molecules. The detector used was one of the SPAD of the 8-SPAD linear array discussed in the text. Incident 532 nm laser power on the LCOS spatial light modulator: 320 mW. The gray boxes identify bursts detected according to the algorithm discussed in the text. Notice the very low background rate visible in between the bursts and the large count rate of single-molecules diffusing through the excitation/detection volume. (b) Histogram of SNR (gray, filled histogram) and SBR (black) corresponding to the time trace shown in part in (a).

of counts, irrespective of their origin:

$$N_A = \varepsilon \left(N_S + N_D + N_B \right). \tag{6}$$

Formally, we can clump together the dark count, sample background and afterpulsing rates into a single total background rate n_T^B:

$$n_T^B = n_D + \eta n_P^B + n_A, \tag{7}$$

where $n_A = N_A/\tau$. This rate turns out to be the only one of practical interest in experiments (see next section).

The total number of counts during a burst, *N*, is the sum of these four quantities N_S, N_D, N_B and N_A:

$$N = N_S + N_D + N_B + N_A = (1+\varepsilon)\left(N_S + N_D + N_B \right). \tag{8}$$

There is obviously no possibility to experimentally identify the source of a particular count (signal, sample background, dark count or afterpulsing). We can nonetheless formally define the SBR and SNR for each burst as follows.

B. Signal-to-Background Ratio

In single-molecule measurements, the only relevant SBR is that measured for individual bursts. Due to the stochastic nature of diffusion (as well as other possible dynamics), each burst has a different duration and photon count, therefore a potentially wide range of SBR is expected, as illustrated in Fig. 2(b). For a given burst, the theoretical SBR is simply given by:

$$\text{SBR} = \frac{N_S}{N_D + N_B + N_A}, \tag{9}$$

which can be rewritten as:

$$\text{SBR} = \frac{\text{SBR}_0}{(1+\varepsilon)(1+\alpha) + \varepsilon\ \text{SBR}_0}, \tag{10}$$

where we have introduced the burst signal-to-background ratio before detection (a characteristic of the sample and optical setup only), SBR_0:

$$\text{SBR}_0 = \frac{\bar{n}_P^S}{n_P^B}, \tag{11}$$

and dark count rate to detected background rate ratio:

$$\alpha = \frac{n_D}{\eta n_P^B}. \tag{12}$$

SBR_0 is the best signal-to-background that can be expected from the burst under consideration, given the sample background and average burst emission rate. A typical experimental distribution of SBR for single-molecule bursts is shown in Fig. 2(b), with a peak at SBR$\sim$15 and some bursts exhibiting SBR > 30.

From Eq. (10), it is clear that for typical values of ε of the order of 1% or less and $\text{SBR}_0 < 10$, the SBR depends very little on the afterpulsing probability, but is potentially sensitive to values of α of the order of 1. Since α (Eq. (12)) depends on detector parameters (dark count rate, n_D, and PDE, η) as well as sample parameters, we will discuss their respective influence on SBR when reviewing detector parameters in Section IV.

These theoretical definitions cannot be used directly in experimental measurements, due to the unknown source of individual counts, therefore a practical definition needs to be provided. The total experimental background rate $n_{T,\text{exp}}^B$ can be estimated by considering interburst regions. The experimental burst SBR can then be computed from the total number of recorded counts in the burst, *N*, its duration τ and the afterpulsing probability, ε, as:

$$\text{SBR}_{\text{exp}} = \left(1 + \varepsilon \frac{N}{n_{T,\text{exp}}^B \tau} \right)^{-1} \left(\frac{N}{n_{T,\text{exp}}^B \tau} - 1 \right) \sim \frac{N}{n_{T,\text{exp}}^B \tau} - 1. \tag{13}$$

C. Signal-to-Noise Ratio

In single-molecule burst measurements, like for the SBR, the relevant SNR is computed at the single-burst level and thus only involves counts collected during the burst duration. Since all the underlying processes are Poisson processes (the single-molecule signal itself being an inhomogeneous Poisson process), the noise, defined as the root-mean-square of the total counts, is given by:

$$\text{noise} = N^{1/2} = (1+\varepsilon)^{1/2} \left(N_S + N_D + N_B \right)^{1/2}, \tag{14}$$

and the SNR by:

$$\text{SNR} = N_S \left(1+\varepsilon\right)^{-1/2} \left(N_D + N_S + N_B \right)^{-1/2}. \tag{15}$$

We obtain the standard relationship between SNR and SBR:

$$\text{SNR} = \frac{\sqrt{\eta n_P^S \tau}}{\sqrt{1 + \text{ SBR}^{-1}}}. \tag{16}$$

Eliminating τ from this expression, we can rewrite it as:

$$\text{SNR} = \text{SNR}_0 \left(\frac{\eta}{1+\varepsilon} \times \frac{1 + \text{SBR}_0}{1 + \alpha + \text{SBR}_0} \right)^{1/2}, \tag{17}$$

where we have introduced the signal-to-noise before detection, SNR_0:

$$\text{SNR}_0 = \left(\frac{\bar{n}_P^S \tau}{1 + \text{SBR}_0^{-1}} \right)^{1/2}. \tag{18}$$

These apparently unwieldy expressions have the advantage of separating sample plus experimental setup properties (SBR_0 and SNR_0) from detector characteristics (η, ε), with the addition of α (Eq. (12)), the ratio of detector dark counts to detected sample background counts. We will use these expressions when discussing detector characteristics, in order to study their influence on SBR and SNR at the end of Section IV.

D. Single-Molecule Burst Detection

The importance of SBR and SNR in single-molecule experiments is twofold. On one hand, a good SBR facilitates identification of bursts. On the other hand, the relative uncertainty (standard deviation to mean ratio) of single-burst observables such as the FRET efficiency ratio (Eq. (2)) in a two-color experiment, decreases for large SNR. Since SNR increases with increasing SBR (Eq. (16)), large SBR and SNR are highly desirable.

It is beyond the scope of this article to discuss in detail the influence of SNR and SBR on burst detection and analysis.

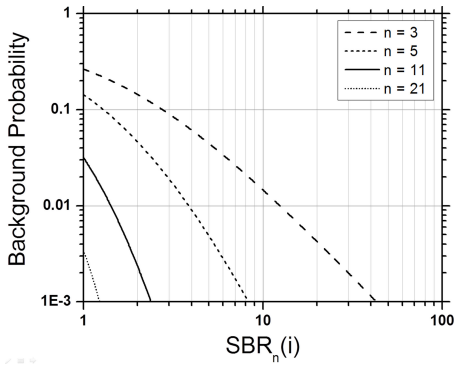


Fig. 3. Probability that the observed SBR calculated over n photons corresponds to background. To guarantee that the observed SBR is not due to background (i.e., that it is due to a single-molecule burst), a smaller SBR needs to be observed when a larger number of photons are used for the estimation.

The following is a brief summary of the main results of the literature, which the interested reader is invited to consult for details.

The definition of a single-molecule burst is to some extent arbitrary [44]. A molecule entering and exiting the excitation volume and reentering it again after a few ms could indeed be considered as two bursts or as a single one, depending on the intended purpose of the analysis [45]. If the molecule is expected to change state over a time scale longer than the average burst duration, fusing two consecutive bursts emanating from the same molecule in different states could provide a confusing observable. On the other hand, if no such change is expected, it might be advantageous to sum the signal of the two consecutive bursts in order to increase the signal N_S [as done for the first burst in Fig. 2(a)]. In doing so, the SBR would decrease compared to the largest SBR of the original bursts, but the overall SNR would in general increase.

A common burst definition criterion consists of looking for regions of the stream of detector counts exhibiting a local count rate significantly larger than the average total background rate n_T^B, where the local count rate $r_n(t_i)$ around count i (with time coordinate t_i) is estimated using a number $n = 2p + 1$ of consecutive counts [20], [14]:

$$r_n\left(t_i\right) = n/\left(t_{i+p} - t_{i-p}\right). \tag{19}$$

The criterion for the observed rate to not be due to background photons only (and therefore to be most likely due to a burst), can be expressed as an implicit condition on the minimum local signal-to-background ratio (see supporting information of [14] for details):

$$\mathrm{SBR}_n = r_n \;\; n_T^B - 1. \tag{20}$$

This relation is represented on Fig. 3 for different values of n as a function of the local SBR, SBR_n. For instance, if $n = 11$ photons are used (plain curve), a SBR larger than 2 is sufficient to ensure that the counts are due to a single-molecule burst (background probability $<3.10^{-3}$). With $n = 5$ (small dashed curve), this threshold would be increased to 4 to ensure this with a confidence level of 99%. In other words, relatively modest local SBR are needed to be able to unambiguously identify single-molecule bursts.

E. Single-Molecule Burst Analysis

The precise influence of SNR and SBR on single-burst data analysis will depend on the type of observable studied and can become rapidly complex. In the case of FRET efficiency, which involves counts from (at least) two different spectral channels (Eq. (2)), the problem is compounded by the existence of two distinct sets of count rates, and sample as well as detector parameters. However, the exact dependence of the observable on single-molecule burst SNR and SBR is not critical, because single-molecule data analysis is always performed on large numbers of single-molecule signals. In other words, single-molecule bursts are accumulated over time and their individual observable values are histogrammed. Consequently, we are interested in the distribution of SNR and SBR values [Fig. 2(b)] rather than in a particular molecule's SNR and SBR.

Since the overall variance of the observable distribution will obviously be improved if the SNR (and/or SBR) of all bursts is improved, we will later examine which detector parameters can be optimized to *globally* improve the SNR or SBR of *most* molecules.

IV. Key Detector Parameters

According to Eq. (17), given a SNR and a SBR before detection equal to SNR_0 and SBR_0, the ideal photon detector for SMFS applications would feature 100% photon detection efficiency ($\eta = 1$), zero dark count rate ($\alpha = 0$, see Eq. (12)) and no afterpulsing ($\varepsilon = 0$). Though such requirements cannot be strictly fulfilled, they can be fairly well approached by good SPAD devices coupled to suitably designed quenching circuits [46]. PDE, dark count rate and afterpulsing are however not strictly independent and any "tweak" aimed at improving one will generally affect the others. Understanding their origin and which parameters control their behavior is therefore key to optimizing them simultaneously for single-molecule applications. A complete overview of the key performance parameters is given in the following.

A. Photon Detection Efficiency

The PDE is the probability with which an incident photon triggers an avalanche (true detection event). Besides the physical phenomena that determine the performance of semiconductor photodiodes in general (reflection, absorption, etc.), other physical effects have key importance in SPADs. For a photon to be detected, not only it must be absorbed in the detector's active volume and generate a primary electron-hole pair. It is also necessary that the primary electron-hole pair succeeds in triggering an avalanche. The avalanche triggering probability, η_{at}, depends on both the device structure and the excess bias voltage (or overvoltage) V_E, i.e., the difference between the reverse bias voltage and the breakdown voltage [46], [47]. Theoretical and experimental studies [48], [49] have shown that this probability first increases linearly with low V_E and then tends to saturate at high V_E.

Fig. 4 shows the PDE as a function of wavelength for several SPAD detectors manufactured with different technologies (see Section V). Fig. 5 (dots) shows the typical dependence on excess bias voltage of the PDE of one of the SPAD devices developed

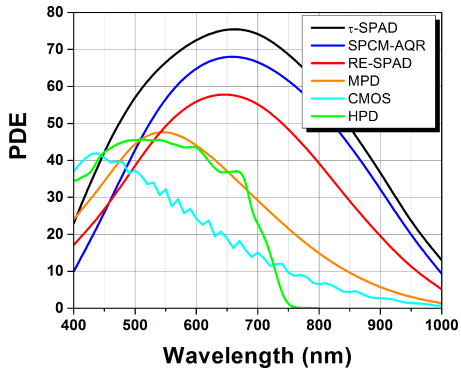


Fig. 4. Photon detection efficiencies (PDE) of SPAD detectors manufactured with different technologies. The PDE of a state-of-the art Hybrid Photodetector [145] is shown for comparison. Curves were provided by the manufacturers. τ-SPAD: SPAD module sold by PicoQuant GmbH, based on a thick reach-through SPAD manufactured by Laser Components. SPCM-AQR: thick reach-through SPAD manufactured by Excelitas Technologies. RE-SPAD: red-enhanced epitaxial SPAD developed by Politecnico di Milano. MPD: Former version of the epitaxial thin SPAD manufactured by MPD. CMOS: 32×32 CMOS SPAD array prototype developed by Politecnico di Milano (now available from MPD). HPD: Hybrid Photo Detector prototype manufactured by Hamamatsu Photonics based on a GaAsP photocathode. Commercial devices including this type of modules are now available from different vendors.

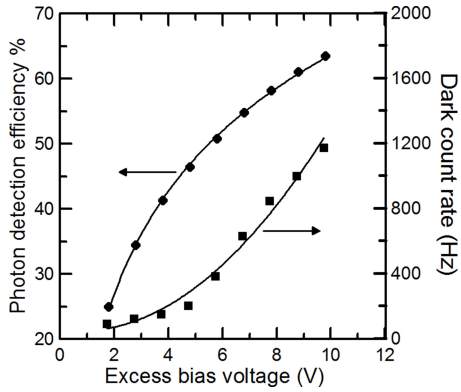


Fig. 5. PDE at 550 nm wavelength (●) and DCR (■) as a function of the excess bias voltage for a 50 μm diameter SPAD fabricated in a custom epitaxial technology [50] and operated at room temperature.

by our group [50], for which the photon detection efficiency, η, at 550 nm can be modeled as:

$$\eta = \eta_\infty \left(1 - \mathrm{e}^{-\frac{V_E}{V_C}}\right), \tag{21}$$

with $\eta_\infty = 0.694$ and $V_C = 4.23$ V.

B. Dark Count Rate

The dark count rate (DCR) is defined as the number of avalanche pulses per unit time that occur in the absence of incident photons. In SPAD devices operating at room temperature or below, the dark count rate is dominated by thermal generation of carriers and by trap-assisted and band-to-band tunneling processes in the depletion layer [51].

The total generation rate G per unit volume is given by:

$$G = G_{\mathrm{trap}} + G_{\mathrm{bbt}}, \tag{22}$$

where G_{trap} is the contribution of transitions via localized deep energy levels (also called traps), including the conventional Shockley-Read-Hall thermal generation. G_{bbt} is the band-to-band tunneling contribution [51]–[53]. It has been shown theoretically and demonstrated experimentally that G_{bbt} becomes the dominant contribution only for electric field amplitudes higher than $7{\cdot}10^5$ V/cm [52], [53]. This contribution can be made negligible by properly reducing the peak electric field within the active area of the SPAD device [50]. In this condition, the total generation rate per unit volume is given by:

$$G \approx G_{\mathrm{trap}} = N_T \frac{e_n\, e_p}{e_n + e_p}, \tag{23}$$

where N_T is the volume density of deep levels, e_n and e_p are the electron and hole emission rates, respectively. Both the quality of the starting material and the technological processes used in the device fabrication have a strong impact on N_T and therefore on the generation rate G. Transition metal impurities are the most common source of deep levels. Metal contamination may occur during silicon handling, high-temperature heat treatments or ion implantations. Unintentional contaminants as Fe, Cu, Ti or Ni are usually found in silicon in concentrations of $\sim 10^{11}$–10^{12} cm^{-3} [54]. In addition, the generation rate is strongly enhanced by the electric field acting on deep levels [52], [53], [55], [56]. Three different mechanisms cause the increase of emission rates e_n and e_p under the application of an electric field: the Poole–Frenkel effect (this effect is present only when the deep impurity level behaves like a Coulombic well, i.e., the emission of electrons or holes ionizes the impurity), the phonon-assisted tunneling, and the direct tunneling from deep levels into conduction or valence bands. As a consequence, even trace levels of metal impurities that would be normally undetectable with state-of-the-art analytical techniques might impair the DCR performance of SPAD detectors, due to the inherently high electric field within the depletion layer ($\approx 10^5$ V/cm). Under the simplifying hypothesis of one type of trap, the DCR, n_D expected from a SPAD can be calculated as follows [57]:

$$n_D = \frac{n_i}{\tau_{g,0}} \Gamma_{\mathrm{eff}}\, Aw = \frac{n_i}{\tau_{g,\mathrm{eff}}}\, Aw, \tag{24}$$

where Γ_{eff} is an effective field enhancement factor that may range from tens to thousands depending on the electric field strength and on the trap type and spatial distribution [57]; $n_i = 1.45\ 10^{10}$ cm^{-3} is the intrinsic carrier concentration at room temperature; A is the junction area, w is the depletion region width; and $\tau_{g,0}$ is the low-field generation lifetime (tens of millisecond in high quality silicon wafers [58]). Equation (24) shows that manufacturing SPADs with low DCR and fairly wide area requires a high quality fabrication technology, leading to a consistent reduction of impurity concentration (i.e., a substantial increase of $\tau_{g,0}$) with respect to standard high-quality silicon. In addition, the electric field profile within the depletion region must be properly designed for ensuring a low value of Γ_{eff} and negligible band-to-band tunneling effects. This reduces the detector noise at room temperature and makes detector cooling more effective in reducing the DCR [50].

Fig. 5 (squares) shows the DCR of a 50 μm diameter SPAD fabricated in a custom epitaxial technology [50] as a function of the excess bias voltage, V_E. The DCR curve can be qualitatively

modeled as:

$$n_D = n_D^0 \left(\mathrm{e}^{\frac{V_E}{V_D}} - 1 \right. \quad , \tag{25}$$

with $n_D^0 = 1.75$ Hz, $V_D = 4.64$ V.

However, it is worth noting that not only the DCR, but also the DCR dependence on temperature and excess bias voltage may vary considerably in SPAD devices manufactured with the same process. This is due to a combination of different factors, namely: i) the spatially-variable electric field profile, ii) the different types of deep levels and their random spatial distribution within the depleted region; iii) the dependence of the emission rate on the electric field strength and on the temperature. Therefore the validity of Eq. (25) is not general, but limited to a specific SPAD sample.

C. Afterpulsing

Deep levels located at intermediate energies between mid-gap and band edge may act as carrier traps. During each avalanche pulse carriers can be trapped in these levels and subsequently released with a statistically fluctuating delay, whose mean value depends on the deep levels actually involved. The released carrier can re-trigger the avalanche, thereby generating correlated afterpulses [46], [59].

The probability of detecting an afterpulse in the time interval between t and $t + \Delta t$ after an avalanche current pulse is $P_a(t)\Delta t$, where $P_a(t)$ is the afterpulsing probability density. SPADs typically exhibit a multi-exponential afterpulsing probability density with longest decay lifetimes in the microsecond range [59]. The total probability, ε, of observing an afterpulse after the triggering of an avalanche can be calculated as follows:

$$\varepsilon = \int_0^{+\infty} P_a(t)dt. \tag{26}$$

The afterpulsing effect introduces a positive feedback loop that can significantly increase the total dark count rate, unless the overall probability of generating an afterpulse is small. Since the population of traps is far from saturation [59], [60], the afterpulsing linearly increases with the charge that flows during the avalanche pulse. Afterpulsing can be effectively reduced by limiting the avalanche charge (hence the trapped charge) through the use of mixed passive-active quenching circuits with minimal parasitic capacitance added to the SPAD terminals [46].

If the trapped charge cannot be reduced to a sufficiently low level, a feature of the quenching circuit can be exploited to reduce the afterpulsing rate to negligible, or at least acceptable level [46]. After an avalanche, by deliberately maintaining the voltage at the quenching level during a so-called hold-off time, T_{ho} the carriers that are released during this period are prevented from retriggering the device. In this case, the total afterpulsing probability is given by:

$$\varepsilon = \int_{T_{ho}}^{+\infty} P_a(t)dt. \tag{27}$$

At room temperature, a hold-off time ranging from few tens to few hundreds ns is usually sufficient to reduce the total afterpulsing probability to 1% or lower. However, this approach is less effective for SPADs operating at cryogenic temperatures, since the trap-release process becomes slower [59], making the required hold-off time much longer and seriously limiting the dynamic range in photon-counting measurements.

Although the exact dependence of the total afterpulsing probability on the excess bias voltage is quite complex, to a first approximation it can be qualitatively modeled by a linear relationship:

$$\varepsilon = \beta \, V_E, \tag{28}$$

with $\beta \sim 2{\times}10^{-3}$ (at $T_{\mathrm{ho}} = 50$ ns) for the SPAD device whose DCR and PDE characteristics are represented in Fig. 5.

D. Timing Jitter

Picosecond resolution may be needed for time-correlated single photon counting measurements, in which each photon is timed with respect to the exciting laser pulse or other detected photons [11]. In SPAD detectors, the onset of the avalanche pulse is correlated with the arrival time of the photon that generates the primary electron-hole pair. Due to various physical effects, the delay of the instant at which the onset is sensed with respect to the true arrival time of the photon is not constant, but subject to statistical fluctuations whose amplitude depends on several factors, the main ones being the excess bias voltage, the device structure and the device area [61]. The timing jitter is usually quoted as the full-width at half maximum (FWHM) of the photon detection time distribution or Instrumental Response Function [11], [62].

SPAD devices with thick depleted region (20–30 μm) typically have significant timing jitter of 200–600 ps FWHM, which is however adequate for most time-resolved single-molecule experiments. SPADs with thin depleted region (1–2 μm) have much narrower timing jitter, which can reach a few tens of ps FWHM. This performance is comparable with that of the best micro-channel plate (MCP) PMTs [11], and it may be useful in demanding applications such as photon antibunching measurements [63] or very short lifetime measurements as encountered in electron transfer processes [64].

E. SPAD Parameter Optimization for SMFS

As a final remark, it is worth noting that in SPADs, PDE, DCR and afterpulsing probability all increase with the excess bias voltage V_E, albeit at vastly different rates. The question therefore arises whether V_E can be optimized to improve the SNR or SBR of single-molecule measurements. This question can be addressed by combining the results of this section (Eqs. (21), (25), and (28) with those of Section III ((10) and (17)).

The dependence of the ratio $\mathrm{SNR}/\mathrm{SNR}_0$ on V_E, using the numerical parameters introduced in the previous paragraphs, is shown in Fig. 6(a) for different sample signal-to-background ratio before detection SBR_0 and different incident sample background rate n_P^B. The SBR_0 values are taken from the distribution shown in Fig. 2, while the sample background rates are typical of single-molecule fluorescence experiments.

For each set of sample parameter (SBR_0, n_P^B), there is clearly an optimum excess bias voltage V_E maximizing $\mathrm{SNR}/\mathrm{SNR}_0$. This optimal voltage is represented in Fig. 6(b) as a function of n_P^B for different values of SBR_0. Note that because of dark counts and afterpulsing, the optimal SNR is lower than the sample SNR_0. For low incident sample background rate n_P^B (typically when $n_P^B < n_D^0$), the SNR is significantly degraded

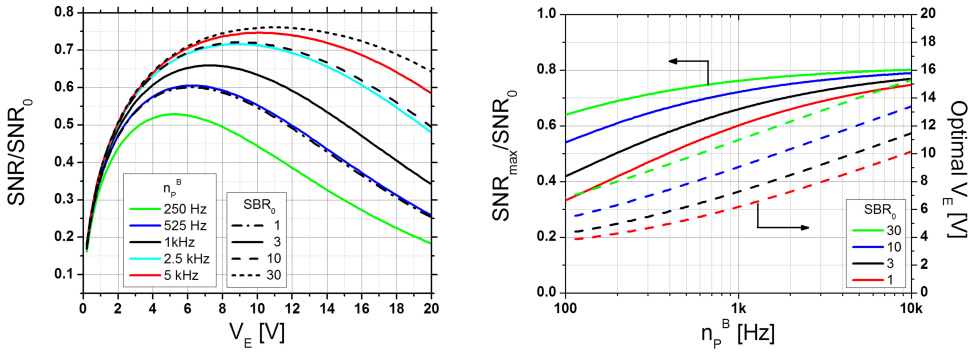


Fig. 6. (a) Dependence of the detected SNR to sample SNR_0 ratio as a function of excess bias voltage V_E for different incident sample background rates n_P^B, when the incident sample $\mathrm{SBR}_0 = 3$ (plain curves) or for different SBR_0 when $n_P^B = 1$ kHz (black curves, different styles). (b) Dependence of the relative V_E-optimized SNR on incident sample background rates (plain curves, left axis) and corresponding optimal excess bias voltage (dashed curves, right axis) for the typical detected SBR to sample SBR_0 ratios.

compared to SNR_0. For larger n_P^B, the optimal SNR approaches the sample SNR_0 because the DCR at the optimal V_E becomes negligible. As expected, the optimal $\mathrm{SNR}/\mathrm{SNR}_0$ ratio increases with SBR_0.

In order to avoid any confusion, it is worth stating that these trends do not imply that a large sample background rate is beneficial! For a given incident signal, a larger sample background rate will always decrease SBR_0 (Eq. (11)) and SNR_0, everything else being equal. However, a large sample background rate renders the detector dark count rate negligible in comparison, therefore allowing an increase of V_E to be used in order to get as close as possible to the sample SNR_0.

Importantly, Fig. 6 shows that, for a given sample n_P^B, the optimal excess voltage depends on the incident burst signal-to-background ratio, SBR_0. Since experimentally, a distribution of such values is expected [Eq. (11), Fig. 2(b)], no excess bias voltage can in principle optimize the SNR of all bursts. Luckily, SNR is not too sensitive to the exact value of the excess bias voltage, making such an optimization a realistic goal.

For instance, if a typical incident background rate of $n_P^B = 1$ kHz is observed, an excess bias voltage $V_E \sim 9$ V optimizes the SNR for bursts such that $\mathrm{SBR}_0 = 10$ [Fig. 6(b)]. Fig. 6(a) shows that the resulting $\mathrm{SNR}/\mathrm{SNR}_0$ is in fact close to optimal for a large range of SBR_0 ($1 < \mathrm{SBR}_0 < 30$) representative of most of the observed bursts. However, since SBR is a decreasing function of V_E, any adjustment aimed at increasing SNR may also reduce SBR and thus slightly affect the bursts detected using the criterion of Eq. (20). A trade-off between these two effects will obviously be needed, whose discussion goes beyond the scope of this article.

The previous analysis shows that, in the case of single-molecule burst fluorescence, adjusting the excess bias voltage of the detector after estimation of the sample background rate n_P^B and typical burst SBR_0 might be advantageous in some cases, and can be straightforwardly adapted to different experimental criteria.

V. Available SPAD Technologies

Silicon SPADs have been extensively investigated and are now well developed. Considerable progress has been achieved in SPAD design and fabrication techniques, and devices with good characteristics are commercially available for applications across the visible spectral range, up to 1 μm. The SPAD devices reported to date can be grouped in three categories, depending on their fabrication technology:

1) non-planar SPADs fabricated in a custom technology;
2) planar SPADs fabricated in a custom technology;
3) SPADs fabricated in standard complementary metal-oxide-semiconductor (CMOS) technologies.

Their features and performance will be reviewed in the following sections.

A. Non-Planar SPAD Devices Fabricated in a Custom Technology

The Slik™ device devised by R. J. McIntyre and P. Webb at the former RCA Optoelectronics (now Excelitas Technologies Corporation), are employed to produce highly successful single-photon counting modules (SPCM-AQR) [47], [65]. The device represents a remarkable evolution of the previous reach-through avalanche diode structure pioneered by the same team [66], [67]. It is built in special ultra-pure high-resistivity silicon wafers with a dedicated technological process; various device features and technological process steps are proprietary and covered by patents [68], [69]. The active area of the detector is fairly wide (diameter about 180 μm). The breakdown voltage is high and strongly varies from sample to sample over a wide range from 250 to 500 V. Thanks to the thick depletion layer (20–30 μm), the PDE is very high in the visible region and fairly good in the NIR up to about 1 μm. The typical value is significantly higher than 50% over the entire range from 540 to 800 nm (see Fig. 4) [65]. Notwithstanding the remarkable volume of the depletion layer, the DCR is very low, ranging from a few ten to a few thousand Hz at –15 °C silicon chip temperature. The afterpulsing probability is also strongly reduced, typically well below 1%. The timing performance is modest: the IRF has a relatively broad peak with about 450 ps FWHM with count rate-dependent walk [63]. However, a significant improvement in the timing performance can be gained by focusing the light at the center of the active area and by using the current pick-up circuit discussed in [70].

Devices similar to the Slik™ have been recently developed by Laser Components Inc., USA. First generation devices have been exploited in commercially available τ-SPAD photon-counting modules (PicoQuant GmbH, Germany) [71], [72]. A second generation of these devices, having similar PDE but larger active area (500 μm diameter) and reasonably low timing jitter of about 200 ps FWHM [73], [74], are available in the id-120 photon counting modules (ID Quantique SA, Geneva, Switzerland) [75].

The Slik™ devices have very good performance, but also a number of practical drawbacks. Due to the high breakdown voltage, the power dissipation during the avalanche is high, from 5 to 10 W, and very effective cooling of the detector under normal operating conditions is mandatory (with Peltier stages, or other means) [47]. The special fabrication technology is inherently complex, the production yield of good devices, low, and the cost, high. The devices are delicate and degradable, and integrating multiple detectors, or associated circuitry, is out of the question.

B. Planar SPADs Fabricated in a Custom Technology

During the past three decades, a number of custom planar technologies have been developed for fabricating SPAD devices with optimized performance [76]–[82]. A definite advantage of these technologies is their inherent suitability for the fabrication of SPAD arrays.

Continued improvement of the planar epitaxial technology described in [77] made it possible to reliably fabricate SPAD devices with large active-area diameters (up to 500 μm), providing an excellent compromise between breakdown voltage (typically around 30 V), PDE (50% peak at 550 nm, declining to 25% at 730 nm, and 12% at 850 nm wavelength), DCR (from a few Hz to a few kHz at -15 °C for SPAD diameters ranging from 50 μm to 500 μm), total afterpulsing probability (about 1% at -15 °C), and timing jitter (better than 40 ps FWHM). These custom SPAD devices are now exploited in commercially available photon detection modules from Micro Photon Devices [83].

New developments in planar-epitaxial SPAD technology are mainly concerned with the improvement of the PDE in the red wavelength range (600 to 900 nm), either by incorporating a resonant cavity in the device structure [84] or by increasing the thickness of the absorption region [85]. The latter approach resulted in a red-enhanced (RE) SPAD device having a separate absorption and multiplication structure. This structure provides a thick (about 10 μm) absorption region with low electric field (hence no multiplication and negligible field-enhanced carrier generation) and a shallow multiplication region with a peaked electric field profile (devised to enhance the avalanche triggering probability and to reduce the photon timing jitter). The electric field profile in the two regions was designed to achieve the optimal trade-off between operating voltage, avalanche triggering probability (thus PDE), DCR, and timing jitter. Experimental measurements on 50 μm RE-SPAD devices showed a significantly improved PDE in the red region, reaching 40% at 800 nm wavelength (i.e., a factor 2.5 higher than the PDE of standard planar SPADs) and 60% at 550 nm wavelength. The devices exhibit a remarkably low DCR less than 1 kHz at room temperature, decreasing to a few ten Hz at 5 °C. Although the active volume of RE-SPADs is considerably larger than that of standard SPADs due to the increased thickness of the absorption layer, the DCR of the two detectors is comparable at room temperature. The thicker absorption region of the RE-SPAD does not significantly contribute to the DCR, since the low electric field practically rules out field-enhanced generation of carriers. The dominant contribution to the DCR comes from the high-field multiplication region, whose design remained substantially unchanged. A timing jitter of 93 ps FWHM was obtained at room temperature, which is larger than the typical figure of standard SPADs (30–50 ps). This is due to the increased thickness of the absorption and drift region, resulting in increased transit times for photo-generated electrons of about 10 ps/μm at the saturation speed of 10^7 cm/s. Since photons are absorbed randomly in the drift region, a timing jitter of about 100 ps FWHM can be attributed to the 10 μm thick absorption region. Total afterpulsing probability lower than 1.5% was measured over the entire temperature range of operation.

C. SPADs Fabricated in Standard CMOS Technologies

Planar Si-SPADs with remarkable performance can nowadays also be developed by fab-less laboratories, relying on industrial microelectronic fabrication services of silicon foundries offering advanced CMOS technologies. Photon-counting modules based on SPAD devices fabricated in standard CMOS technologies are commercially available from SensL [86] and IdQuantique [87]. Monolithic integration of SPAD devices and CMOS circuits offers manifest advantages, from the availability of a fully supported, mature and reliable technology at reasonably low cost, to the possibility of developing complete systems on chip with a high degree of complexity. The mandatory requirement for SPAD integration is that a suitable subset of fabrication steps can be specified within a complex CMOS process flow and used to build a planar p-n junction free from edge effects.

A number of small-area SPADs were obtained by independent research groups first using a 0.8 μm high-voltage CMOS technology (HV-CMOS) [88]–[90] and then a 0.35 μm HV-CMOS technology [91]–[94]. These 0.35 μm HV-CMOS SPADs exhibit a moderate DCR ($\sim$1 kHz for a 20 μm detector) and a maximum PDE of 35% at 450 nm wavelength when biased 4 V above breakdown (Fig. 4). The PDE however drops to $\sim$20% at 600 nm and it is small ($<$5%) at 800 nm [92]. An intrinsic limitation of HV-CMOS SPAD devices comes from their "inverted" structure. Namely, the active p+n junction is formed between an n-well (e.g., a standard high-voltage PMOS well), and a shallow p+ implantation (e.g., a PMOS source/drain region). In this device geometry, visible and near-infrared photons are mainly absorbed in the n-well region, thus generating holes that act as avalanche triggering carrier. In silicon, holes have a lower avalanche triggering probability compared to electrons [49], leading to a reduced PDE.

CMOS technology with deep submicron (DSM) resolution is mandatory for the fabrication of dense SPAD arrays with large numbers of pixels, adequate fill-factor (i.e., the active-to-total area ratio of a single pixel), and smart pixels which include integrated electronics. However, a challenging basic issue must be faced: the inherent features of DSM CMOS technologies,

namely the relentless trend toward higher doping levels, lower thermal budget, and thinner p- and n-well layers, conflict with SPAD detector performance. The smaller depth of carrier-collection layers limits the PDE, and the high electric fields arising from higher doping result in strongly enhanced DCRs due to band-to-band and trap-assisted tunneling effects. In addition, the reduced thermal budget and the lack of external gettering processes [95] also have adverse effects on afterpulsing.

Several SPAD structures in DSM CMOS technologies have been demonstrated in recent years [96]–[100]. Richardson *et al.* [101] recently introduced three SPAD structures based on a novel retrograde buried n-well guard ring, capable of scaling from 32 to 2 μm in diameter. A remarkable DCR of $\sim$100 Hz for an 8 μm diameter SPAD devices was achieved at room temperature with 0.8 V excess bias, a maximum PDE of 25% at 560 nm, and negligible afterpulsing. SPAD devices have been demonstrated in 90 nm CMOS technologies, but with significantly lower performance [102]. A notable exception is the 90 nm SPAD device reported by Webster *et al.* [103], where the deep n-well/p-epi junction is used as the active junction, achieving a peak PDE of 44% at 690 nm and better than 20% at 850 nm. Timing jitter as low as 50 ps FWHM was demonstrated, although the timing distribution was affected by a relatively long diffusion tail.

The integration of SPAD devices and associated electronics in submicron and deep-submicron CMOS technologies paved the way for the fabrication of large SPAD arrays. Several array detectors have been reported so far [89], [93], [94], [104]–[109] with steadily increasing pixel numbers, and more and more complex "in-pixel" integrated electronics and efficient readout architectures. Guerrieri *et al.* [109] designed and fabricated a high-speed single-photon camera based on a monolithic array of 32 $\times$ 32 smart pixels fabricated in a 0.35 μm HV-CMOS technology. Each pixel (100 μm $\times$ 100 μm) is a completely independent photon-counting channel that includes a 20 μm diameter SPAD, analog sensing and avalanche quenching electronics, digital processing for counting the incoming photons, and memory and buffer stages for global shutter readout with no dead-time. Better than 35% peak PDE is attained at 450 nm, decreasing to 8% at 800 nm, with DCR in the range of 1 kHz for more than 75% of the SPADs. The 32 $\times$ 32 SPAD array can operate up to 100 000 frames/s with a dynamic range of 8 bits for counting. Noteworthy results have been obtained with this detector in multi-channel FCS experiments [14]. However, due in part to its limited sensitivity, detection of single-molecule bursts turned out to be impossible [14].

A remarkable, fully-parallel SPAD array of 32 $\times$ 32 pixels was recently developed in a 130 nm CMOS technology, capable of simultaneously evaluating the photon time-of-arrival with a time-bin width of 119 ps and a 10 bit range [110]. The array exploits 8 μm low-noise SPAD devices, having a median DCR of $\sim$100 Hz and a peak PDE of 25% at 460 nm. Each channel operates independently, and contributes to an overall data rate from the chip of up to 10 Gb/s in time-correlated operation mode. The fill-factor does not exceed 2%. To mitigate this limitation, an array of microlenses based on a design described in [111] was used. This SPAD array was successfully used in wide-field fluorescence-lifetime imaging (FLIM) [12] experiments in the blue/green wavelength range [110]. A 32 $\times$ 32 SPAD array [108] was used in combination with a custom selective plane illumination microscope (SPIM) [112] for performing FCS measurements [113], [114]. However, no true single-molecule experiments have been reported so far with these advanced SPAD array detectors, likely due in part to their limited sensitivity.

VI. SPAD Detectors for Single-Point Geometry Single-Molecule Experiments

Although early single-molecule fluorescence experiments were performed using PMTs [115], [116], the introduction of SPADs for single-point geometry single-molecule detection [117]–[119] and FCS [120], [121] significantly improved the detection efficiency in the visible range and replaced high-voltage, bulky detectors by more compact, robust and simpler devices, permitting a rapid development of the field. Nowadays, $\mathrm{Slik}^{\mathrm{TM}}$ SPAD (available in the SPCM-AQR modules by Excelitas) and RE-SPAD (to be commercialized by MPD) offer the best sensitivity. We compared these two devices in smFRET experiments on freely-diffusing molecules using a single-spot confocal geometry with either continuous-wave lasers (for intensity-based FRET measurements using a μs-ALEX approach [28], [31]) or pulsed lasers (for lifetime-based measurements using a ns-ALEX approach [32]). The samples consisted of a common 40 base-pair (bp) long double-stranded DNA (ds-DNA) backbone doubly-labeled with the donor (ATTO 550) on one strand and the acceptor (ATTO 647N) on the other strand at different bp distances in order to cover a wide range of FRET efficiencies [122]. For intensity-based measurements (μs-ALEX), we have verified that, provided the size of the emission point-spread function (PSF) is matched to the sensitive area of the detectors, the detected count rates scale according to the measured PDE. This requirement is not very demanding, considering the relatively large size of the RE-SPAD (50 μm diameter, compared to $\sim$200 μm diameter for the $\mathrm{Slik}^{\mathrm{TM}}$ SPAD). The burst size distributions reflected the $\sim$15% difference in PDE at 670 nm in favor of the SPCM-AQR. However, this difference turned out to have a negligible impact on the measured FRET efficiency E and on the shot-noise broadening of FRET histograms [122].

For time-resolved measurements, the much narrower IRF of the RE-SPADs ($\sim$90 ps FWHM compared to 330 ps FWHM for the SPCM-AQR) did not provide any significant advantage in the range of FRET values studied [122]. However, we were able to verify that for very short fluorescence lifetimes such as that of Erythrosin B (95 ps in aqueous buffer), the narrow IRF of RE-SPADs is essential for accurate extraction of the correct decay constant. This characteristic could be critical for lifetime-based FRET measurements at large E-values (i.e., short lifetimes) or photoinduced electron-transfer measurements [64], [123].

The single-point excitation/single-point detection design results in highly sensitive detection but requires in general long acquisition times (typically tens of minutes per data point) in order for sufficient statistics to be accumulated. This confines SMFS to niche fundamental research and in general limits its application to equilibrium state studies such as conformational fluctuations occurring at time scales much shorter or longer than the diffusion time and, with some additional efforts, at

intermediate time scales [22], [44]. However, fast irreversible reactions cannot be studied by this approach, as they are over before enough single-molecule bursts have been acquired. In other words, only reactions with time scale significantly longer than the minimum duration of a measurement (a few minutes or more) can be conveniently studied in single-point detection geometry (see for instance [124], [125]). Furthermore, long acquisition time poses significant throughput constraints for clinical or biopharmaceutical applications. For instance, rapid response to bacterial infection is critical for vital prognosis; therefore, reducing the time needed for pathogen nucleic acid detection by one order of magnitude or more would have immediate clinical impact. Similarly, being able to quantify large numbers of rare disease biomarkers in multiplexed single-molecule detection assays using minute sample amount would help simplify early diagnosis and treatment of diseases.

A solution to this problem is to acquire the same kind of data from several distinct locations in the same sample and pool the data together in order to obtain the same statistics in a shorter amount of time. In principle wide field observation of immobilized molecules using total internal reflection (TIR) microscopy could represent a solution [126]. However, (i) additional chemistry is required to bind molecules to a substrate, (ii) there is potential interference of the substrate with the attached molecules and (iii) the temporal resolution of cameras is limited. In practice, the best approach consists of multiplexing data acquisition in solution, by simultaneously observing the same sample at different locations (or observing several samples simultaneously) using an array of SPAD detectors. However, such an approach requires fundamental efforts in several fields including excitation and collection optics, detector technology and data analysis, which can only be achieved through a collaborative effort between research groups with different expertise.

The next section describes our approach for high-throughput SMFS resulting from the collaboration between our research groups. Preliminary results illustrating the tremendous potential of this technology will be presented as well.

VII. New Detectors for Single Diffusing Molecules Measurements in Multispot Geometries

Parallelization is the simplest way to address the throughput limitations of point-like detection. This problem was in fact initially encountered in fast confocal imaging and has received a number of technical solutions in the past. Single-molecule detection brings some peculiarities that make some of these solutions inadequate. Parallelization is challenging because it needs appropriate technology to: (i) create high-quality multispot excitation patterns; (ii) detect the corresponding signals in parallel; (iii) acquire this high-throughput data stream and process the data efficiently. Parallel data acquisition and processing can be suitably addressed by exploiting *field programmable gate arrays* (FPGA) to implement real-time data reduction algorithms and strategies. Multispot excitation and detection raise several technical challenges. We will therefore briefly discuss these points, before presenting some results recently obtained with multipixel (i.e., array) detectors.

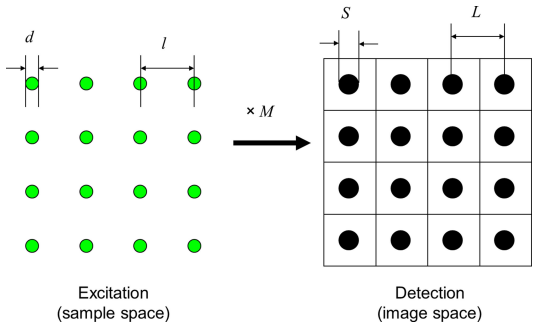


Fig. 7. Multispot excitation and detection geometry requirements. On the excitation side, spots need to be separated by at least the PSF dimension *d*. On the detection side, the detector pitch *L* needs to match that of the spot times the magnification, $L = Ml$, while the relation between detector and spot size depends on the specific application.

A. Multispot Excitation

Multiple diffraction-limited excitation spots can be obtained using different approaches, among which cascaded beamsplitters [127], microlens arrays [128], digital micromirror devices [129], diffractive optics element [130], spatial light modulators [131] or zero-mode waveguides (ZMWG) [18] are a few proven solutions. A first practical requirement is that each spot is sufficiently far away from its neighbors to avoid cross-excitation and detection of the same molecule by adjacent spots. A simple rule of thumb to avoid this problem is to ensure that the inter-spot distance, *l*, is at least a few time its diameter, *d* (Fig. 7) [132].

The spot size itself is defined based on the excitation PSF, which to first approximation can be modeled as a 3-D Gaussian slightly elongated along the optical axis. Noting $\sigma_{XY} \sim 0.21$ *λ/NA* the PSF's standard deviation in the focal plane ($d \sim 2 \times$ FWHM $\sim$ *λ/NA*), the separation criterion can be expressed as $l > d \sim 5\sigma_{XY}$. In practice larger separation may be needed to reduce cross-talk.

B. Signal Detection Parallelization and Detector Alignment

Detectors used in multispot experiments need to be able to collect light from each single spot, with minimum contamination from other spots emission. This can be achieved by using detectors with distinct sensitive elements (or pixels) associated with each individual spot. In this case, the detector's geometry should reflect that of the excitation spots, scaled up by the optical magnification, *M*. In particular, the detector pitch, *L* (distance between the center of sensitive areas collecting signal from nearest-neighbor spots), divided by the optical magnification, *M*, needs to match the spot separation, *l*. Since the minimum spot size *d* is fixed by the focusing optics (*NA*) and excitation/emission wavelength (λ), this previous constraint also fixes the ratio between the diameter of the sensitive area of the detector, *S* and the size of the spot image, *Md*. This ratio, $\sigma = S/Md$, can be interpreted as the detector size in image PSF unit or as a *reduced detector size*. This is an important parameter in both single-molecule burst detection and FCS applications, for which the optimal values are different, but also of the order of one. In other words, *d* being fixed by the excitation optics and sample characteristics (wavelength), choosing the spot separation fixes

the pitch to size ratio (or *aspect ratio*, $\alpha = L/S$) of the detector. The magnification M can then be adjusted to ensure that the optimal reduced detector size δ is used.

Obviously, the condition $l > d$ results in $L > S$, which means that the *fill-factor* of an ideal detector for single-molecule spectroscopy, $\pi/4\alpha^2 < 1$, a requirement that distinguishes these applications from traditional imaging applications, for which a fill-factor as close to 1 as possible is generally sought.

Once the optimal parameters needed to match excitation spot size and pitch to detector geometry have been selected, a few more challenging steps are needed to ensure that single-molecule signals will be efficiently collected. In particular, aligning all excitation volumes with their respective detector pixels happens to be a non-trivial task, as anyone having aligned a single-spot single-molecule setup will easily understand. With detector pixel sizes in the range of 10 to 100 μm, a task that is relatively easy to accomplish at the single pixel level with standard micrometer-resolution translation stages is rendered more challenging by the introduction of one or more additional degrees of freedom for multipixel detectors, corresponding to the orientations of the detector and its pitch. This additional alignment complexity can be easily solved using a programmable pattern generator allowing complete control on the position, scale and orientation of the excitation pattern. A liquid crystal on silicon spatial light modulator (LCOS SLM) used in a direct space pattern generation mode (rather than a Fourier space or holographic approach as commonly used) makes it straightforward to either interactively or automatically orient and shift, as well as adjust the pitch of simple patterns, as we and others have recently demonstrated with 1D and 2D patterns [131], [133], [134]. However, the alignment of more than one multipixel detector to a common excitation pattern, as needed for multicolor detection used in smFRET experiments, remains challenging.

C. Multipixel Detectors

A cornerstone for building a high-throughput and high-performance SMFS setup is provided by a fully-parallel monolithic array of SPAD detectors. We exploited the custom planar epitaxial technology described in [50] to fabricate both linear and 2-D arrays. The key advantage of this technology is the availability of a highly doped isolation region that completely surrounds the detector. As a result, each SPAD pixel is electrically isolated from other pixels fabricated on the same chip and, even more important, the direct component of the optical crosstalk between adjacent pixels is reduced to less than 2% [135], [136]. The adopted technology provides a rather thin depletion region ($\sim$1 μm), resulting in a good PDE in the green region of the visible spectrum. Work is ongoing to incorporate the red-enhanced technology in SPAD arrays [137], in order to improve the sensitivity at longer wavelengths.

1) 8-Pixel SPAD Array: The detector is comprised of eight individual SPADs (diameter $S = 50$ μm, pitch $L = 250$ μm) [138]. Their PDE is identical to that of single-SPAD detectors manufactured by Micro Photon Devices (MPD, Fig. 4) and peaks at 550 nm (PDE $\sim$50%), reaching $\sim$45% at 580 nm, the emission peak of Rhodamine 6G (R6G) and ATTO 550 (Atto-Tec Gmbh, Germany), and only $\sim$30% at 700 nm, the emission peak of ATTO 647N, dyes used in the experiments

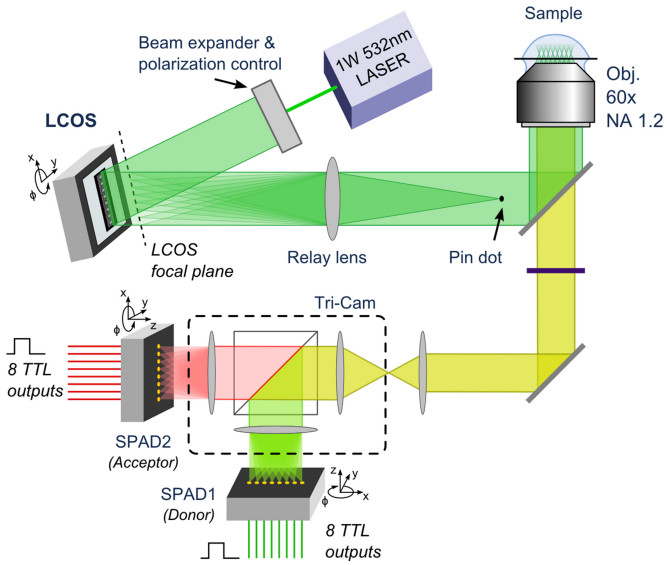


Fig. 8. Schematic representation of the multispot setup. After laser beam expansion, 8 separate spots are formed by a programmable LCOS-SLM, recollimated and directed into an inverted microscope, where a water-immersion objective lens refocuses the spots into the sample. A 300 μm diameter pindot was used to block the fraction of unmodulated light reflected off the LCOS. The 8–spot emission signal is focused by the microscope tube lens and relayed by a multi-camera port system to two 8-pixel SPAD modules, after spectral separation.

described below. TTL signals generated by each SPAD upon photon detection were routed to a reconfigurable multichannel counting board (PXI-7813R, National Instruments), which was programmed to provide a time-stamp with 12.5 ns resolution for each pulse. Data were asynchronously streamed to a host PC using a communication bridge (PXI-PCI 8330, NI), saved to disk and processed in real-time (time trace binning) or off-line (ACF calculation, FRET histogram).

We first successfully demonstrated parallel FCS measurements from 8 quasi-diffraction-limited spots separated by $\sim$5 μm using a sample of R6G molecules diluted in various concentration of sucrose in aqueous buffer [131], [139]. During the course of FCS experiments, study of a much less concentrated sample of R6G (100 pM) yielded intensity time traces exhibiting clear single-molecule bursts [131].

We therefore moved on to single-molecule FRET experiments using first a single 8-SPAD array (one half of which was dedicated to a different color) for 4-spot measurements [133] and then two 8-SPAD arrays to collect donor and acceptor photons emitted from 8 spots [140]. In these experiments, a 532 nm pulsed laser (IC-532–1000 ps, High Q Laser GmbH, Rankweil, Austria, P $>$ 1.5 W) and a programmable LCOS-SLM (model X10468-01, Hamamatsu, Bridgewater, NJ, USA) with relay optics, were used to generate a linear array of diffraction-limited spots with software-adjustable pitch and orientation (see Fig. 8) [140]. The fluorescence emission generated by molecules diffusing through the 8 excitation spots was relayed by a multi-camera port system (Tricam, Cairn Research, Kent, U.K.) to two 8-pixel SPAD modules, after spectral separation. Each module was individually aligned so that each pixel's active-area received the emission signal from one of the conjugated excitation spot. We

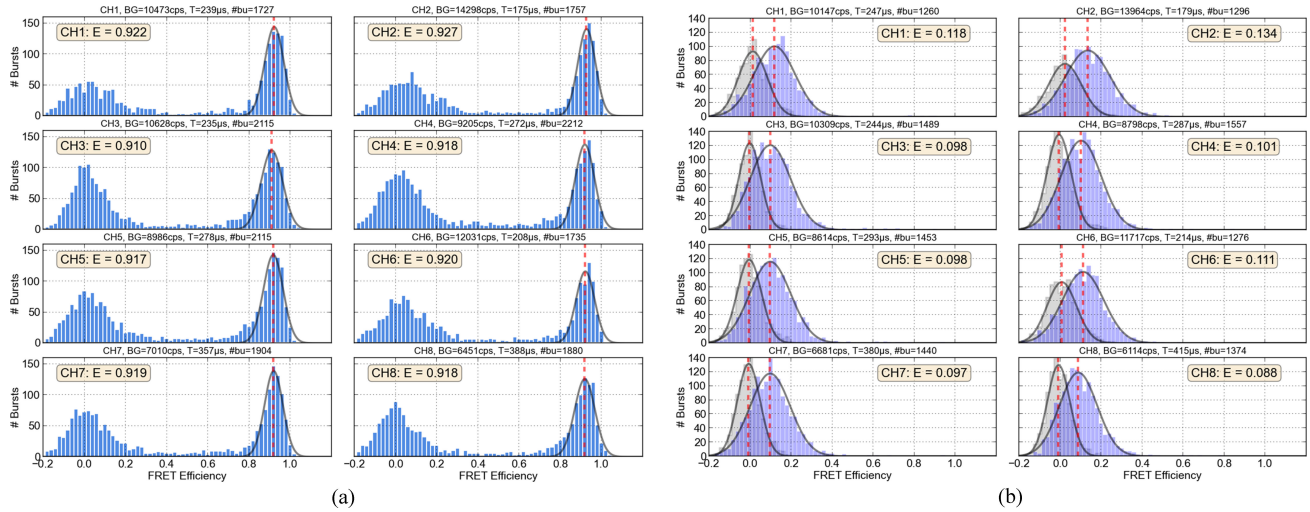


Fig. 9. FRET histograms for the 7 bp sample (a) and the 27 bp sample (b). Each plot corresponds to one channel of the multispot system. A Gaussian fit (grey line) of the FRET population and the fitted peak position (red dashed lines) are shown with the estimated FRET efficiency for each channel indicated in a text box. Histograms for the donor-only sample are shown in light gray for comparison with the FRET histograms for the 27 bp sample.

developed an iterative and semi-automated strategy for precise alignment of the linear arrays using a concentrated (100 nM) Cy3B solution with significant signal in both channels [140].

We demonstrated the capabilities of this system by studying the same FRET samples already investigated by using red-enhanced SPAD detectors in single-point geometry (see Section VI). In order to cover a wide range of FRET efficiencies [28], we prepared 40 bp long dsDNA samples having a common "upper" strand labeled with the acceptor dye (ATTO 647N) attached to the 5' end, whereas the donor dye (ATTO 550) was attached on the complementary strand at five different positions, resulting in a donor-acceptor separation of 7, 12, 17, 22 and 27 bp. Furthermore, for comparison with the low FRET samples, a donor-only sample using an unlabeled upper strand was prepared. It is worth noting that compared to the single-point experiments performed with RE-SPADs (see Section VI), multi-spot experiments are more challenging since the PDE of the 8×1 SPAD array is significantly lower in the far red region of the spectrum (Fig. 4). Therefore the low FRET efficiency molecules emitting less red (acceptor) photons than donor photons are expected to be difficult to distinguish from donor-only molecules. Fig. 9(a) reports FRET efficiency histograms (one for each channel) for samples having a donor-acceptor separation of 7 and 27 bp, respectively the highest and the lowest FRET samples of the series. In the 7 bp sample, we observed both a FRET population peak and a donor-only peak (centered on $E = 0$). The donor-only population was not present for the 27 bp sample and the FRET histograms only showed a low-FRET peak. To demonstrate this point, we also reported in Fig. 9(b) the measured histograms for the donor-only sample (light grey), which showed peaks exactly centered on the $E = 0$ value. The low-FRET peaks of the 27 bp sample were thus clearly separated from the donor-only peaks. Therefore, although we expected a lower sensitivity on the multi-spot system for low-FRET samples, these results demonstrated that FRET efficiencies as low as 0.1 could clearly be identified.

These experiments illustrate some potential applications of these (and future larger) arrays. For instance, higher throughput measurement on a single sample can be achieved by averaging the fitted parameters of measurements performed at n distinct spots in parallel. In this case, the standard deviations of the fitted parameters are reduced by a factor $1/\sqrt{n}$. Alternatively, a single measurement could be performed in n samples in parallel. Although this is an exciting prospect for high-throughput screening applications, one needs to keep in mind that it requires the samples to be located at a distance $l \sim \alpha d$ from one another in the object plane. For $d \sim 0.5$ μm and a typical detector aspect ratio $\alpha \sim 10$, this represents a 5 μm distance, within reach of current microfluidic technology, but challenging [141].

Ongoing improvements of the setup include increasing the number of excitation spots as new SPAD arrays with more pixels become available as well as implementing an ALEX excitation scheme in order to be able to perform measurements on lower FRET samples as well as distinguish between singly- and doubly-labeled species.

2) 48-Pixel SPAD Array: Fig. 10(a) illustrates the microphotograph of the SPAD array including 48 pixels arranged in 4 rows of 12 detectors. By using one or more of these arrays combined with commercially available camera multiport modules (e.g., OptoSplit III or TriCam, Cairn Research, UK), we can envision several detection geometries suitable for high-throughput 2- or 3-color smFRET/ALEX measurements [41], illustrated in Fig. 10(b)–(d): 2-color, 2×24 channels [1 detector, Fig. 10(b)]; 3-color, 3×16 channels [1 detector, Fig. 10(c)]; 3-color, 3×48 channels [3 detectors, Fig. 10(d)]. In the latter case, the total number of channels to be monitored simultaneously is 144, which is well within the capability of the PXI-7813R counting board (160 channels, 20 million signals per second or MS/s). Indeed, the average SMFS count rate per channel is rarely larger than 100 kHz, which amounts to a maximum required data (time-stamp) throughput of $\sim$14.4 MS/s.

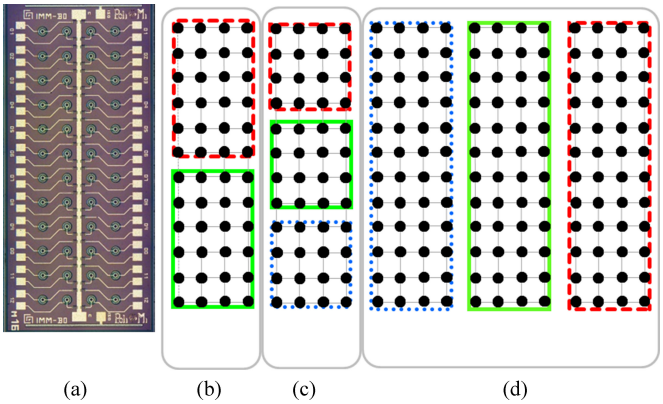


Fig. 10. (a) Microphotograph of the 48-pixel SPAD array including 50 μm-diameter SPAD pixels with a pitch of 500 μm. The chip size is 6.5 mm $\times$ 2.4 mm. (b)–(d) Example of possible use of 1 to 3 detectors in 2 to 3-color experiments. (b) 2-color, 2 $\times$ 24 channels (1 detector); (c) 3-color, 3 $\times$ 16 channels (1 detector); (d) 3-color, 3 $\times$ 48 channels (3 detectors).

Two array formats were fabricated at the IMM-CNR silicon facility: i) 50 μm-diameter SPAD pixel with 500 μm pitch and, ii) 25 μm-diameter SPAD pixel with 250 μm pitch [142]. While a smaller diameter is a definite advantage in terms of detector noise, a larger diameter allows for an easier mechanical alignment of the detector. However matching the pitch of the excitation spots with the larger pitch of such a detector requires a larger expansion of the light beam into the optical path and is thus more prone to optical aberrations. This can in turn result in a reduction of the collection efficiency [14]. The two different structures have been designed to further investigate the trade-off between these phenomena.

For each SPAD pixel, the anode is connected to a bonding pad and made available for the connection to an external integrated AQC (iAQC) [142], [143] by means of wire bonding. In contrast, each cathode is connected to a common supply voltage by means of an integrated resistor of about 2 kΩ, which is used to limit the current flowing through the detector during the avalanche pulse. This, in turn, reduces the afterpulsing probability to about 1% and the optical crosstalk probability between adjacent pixels to less than 2%.

We designed new iAQCs by using a highly-scaled, 0.18 μm high-voltage CMOS technology, in order to reduce area occupancy, power dissipation and dead time. Two iAQC arrays each one including 32 elements were placed at the opposite sides of the SPAD array to reduce the size and wiring complexity of the printed circuit board.

To easily integrate the detector into the multispot single-molecule setup, we developed a compact module for parallel single photon-counting (see Fig. 11). The SPAD array and the two iAQC chips (see Fig 11, zoom) are housed into a hermetically sealable chamber separated from the remaining part of the module [142]. Sealing in a dry atmosphere makes it possible to mount the SPAD array on a double-stage Peltier and to cool the detector down to temperatures of about -20 °C, thereby reducing the dark count rate of SPAD devices and increasing the sensitivity of the instrument. A C-mount thread at the top of the detector chamber allows for a simple and reliable connection to the single-molecule optical setup.

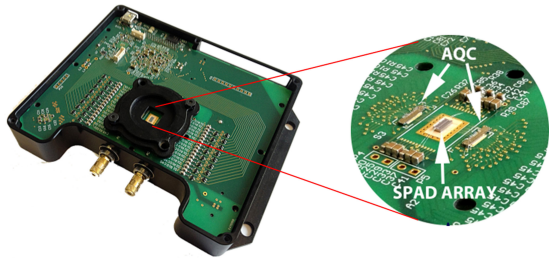


Fig. 11. Picture of the 48-pixel SPAD array module. The SPAD array wire bonded to two integrated AQCs is visible in the zoom. After [142] reproduced with permission from the Society of Photo-Optical Instrumentation Engineers (SPIE).

All the electronic circuitry needed for a proper operation of the detection system is housed outside the sealed chamber. An onboard FPGA (Xilinx Spartan 6, model SLX150) is used to synchronize the 48 channels, time-stamp each detected count with a time resolution of 10 ns and send this information to a remote PC at sufficiently high rate to avoid data loss. Since the nominal detection rate is around 50–100 kHz per channel and the time-stamping counter is 32-bit wide (where 8 bits are used to identify the triggered detector) the FPGA must be able to drive the USB connection to a rate of 20 MByte/s.

Preliminary measurements on the prototype module showed that the PDE was comparable to that of the 8-pixel SPAD array. The DCR was assessed by operating the 25 μm-diameter SPAD array at room temperature and the 50 μm-diameter SPAD array at -15 °C. For the 25 μm-diameter SPAD array, more than 90% of the devices have a DCR lower than 1 kHz, with just a pixel having a DCR slightly higher than 7 kHz. For the 50 μm diameter SPAD array, about 90% of the devices have a DCR lower than 1 kHz, whereas the worst pixel has a fairly low DCR of about 2.5 kHz.

Work is currently ongoing to incorporate the red-enhanced technology in a SPAD array, which is not straightforward. Indeed the increased thickness of the active region needed to improve the detection efficiency at the longer wavelengths, prevents from using the "junction isolation" process used in the standard epitaxial technology to electrically isolate SPAD pixels [137]. We addressed this limitation by replacing the classical junction isolation with a dielectric isolation obtained by etching deep trenches into the silicon substrate and refilling them with silicon dioxide and poly-silicon. Besides providing the desired isolation, this structure is also compact, easy to fabricate and suitable for further developments of the detectors.

VIII. Conclusions and Perspectives

The future progress of silicon SPADs will be mainly driven by user demands for detector performance, which naturally set requirements for the device design and fabrication technology.

Single-molecule detection is a demanding application requiring a combination of high PDE and low DCR. It is probable that such high-end applications will continue to rely on SPAD devices fabricated with dedicated technologies. Future developments in these technologies will likely focus on

improving the PDE in the red region of the spectrum. For example, a combination of red-enhanced [85] and resonant-cavity-enhanced [84] technologies might be used to develop frequency-selective SPAD devices with unprecedented PDE at a desired wavelength.

There is further room for improvement as far as SPAD arrays are concerned. Work is ongoing in multiple directions to increase the pixel number, extend the spectral sensitivity and improve the photon-performance [137]. However, if on one hand the custom planar technology has proven to be very promising in term of sensitivity, on the other hand the number of SPAD pixels which can be included in an array is limited to about one hundred. Above that limit, the array design, implementation and exploitation would become cost-prohibitive, overly complex and in the end, of little use to the everyday single-molecule biophysicist.

A substantial increase in the pixel number can be achieved by resorting to more complex and sophisticated technologies, such as advanced multi-wafers and three-dimensional technologies, which make the integration of custom SPAD arrays with high performance CMOS electronics for quenching/timing onto a single device possible. Aull *et al.* [144] reported a fully parallel laser radar imager based on a 64×64 custom SPAD array coupled to high-speed SOI CMOS circuits by using 3-D integration techniques.

On the other hand, CMOS integration has enabled progressively smaller feature sizes, to the point where it is now possible to envision extremely large imaging systems based on SPADs. Standard CMOS technologies will definitely outperform custom SPAD technology in all applications requiring a large number ($>$1000) of small pixels with good fill-factor ($>$10%) and integrated electronics. Future research activity in this area will be aimed at the development of denser arrays with larger formats ($>10^6$ pixels) by exploiting sub-100 nm CMOS technologies [102]. It appears likely that large CMOS SPAD arrays of the type discussed here but with improved sensitivity [103], coupled to microlens arrays [110] or designed to achieve fill factors close to 100% by alternative architectures, will play an important role in time-resolved photon-counting imaging.

Single-molecule spectroscopy raises a number of interesting challenges from a detector point of view. We have presented our current efforts to take advantage of custom silicon technologies to increase the throughput and information content achievable in these experiments. Clearly, a lot more needs to be done to reach optimal performance. What we have learned from our efforts is that a dialog between detector developers and single-molecule experimentalists is essential to optimize detector design and performance, and inversely, knowing the existing potential of detector technologies, suggests new types of experiments or data analysis approaches.

At the same time, it is clear that many of the requirements of single-molecule imaging and spectroscopy overlap with a large number of other low-light level applications and therefore ideal detectors for our purposes will certainly be used in many others. The potential benefits of these technologies for biomedical, biopharmaceutical and biodetection applications in general will hopefully motivate more research and development of these novel detectors, as well as funding for this type of activities.

ACKNOWLEDGMENT

Conflict of interest statement: S. Weiss discloses equity in Nesher Technologies and intellectual property used in the research reported here. The work at UCLA was conducted in Dr. Weiss's Laboratory. M. Ghioni discloses equity in Micro Photon Devices S.r.l. (MPD). No resources or personnel from MPD were involved in this work. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, National Science Foundation or Department of Energy.

REFERENCES

[1] X. S. Xie and J. K. Trautman, "Optical studies of single molecules at room temperature," *Annu. Rev. Phys. Chem.*, vol. 49, p. 441–480, Oct. 1998.

[2] W. E. Moerner and M. Orrit, "Illuminating single molecules in condensed matter," *Science*, vol. 283, no. 5408, pp. 1670–1675, Mar. 1999.

[3] S. Weiss, "Fluorescence spectroscopy of single biomolecules," *Science*, vol. 283, no. 5408, pp. 1676–1683, Mar. 1999.

[4] E. Betzig, "Proposed method for molecular optical imaging," *Opt. Lett.*, vol. 203, pp. 237–239, 1995.

[5] J. B. Pawley, Ed., *Handbook of Biological Confocal Microscopy*, 2nd ed. New York, NY, USA: Plenum, 1995.

[6] E. Betzig and J. K. Trautman, "Near-field optics: Microscopy, spectroscopy, and surface modification beyond the diffraction limit," *Science*, vol. 257, no. 5067, pp. 189–195, Jul. 1992.

[7] W. E. Moerner and D. P. Fromm, "Methods of single-molecule fluorescence spectroscopy and microscopy," *Rev. Sci. Instrum.*, vol. 74, no. 8, pp. 3597–3619, Aug. 2003.

[8] X. Michalet, S. Weiss, and M. Jäger, "Single-molecule fluorescence studies of protein folding and conformational dynamics," *Chem. Rev.*, vol. 106, no. 5, pp. 1785–1813, May 2006.

[9] R. Roy, S. Hohng, and T. Ha, "A practical guide to single-molecule FRET," *Nature Methods*, vol. 5, no. 6, pp. 507–516, Jun. 2008.

[10] X. Michalet, O. H. W. Siegmund, J. V. Vallerga, P. Jelinsky, J. E. Millaud, and S. Weiss "Detectors for single-molecule fluorescence imaging and spectroscopy," *J. Mod. Opt.*, vol. 54, no. 2/3, pp. 239–281, Jan. 2007.

[11] W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques*. Berlin, Germany: Springer, 2005.

[12] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 3rd ed. New York, NY, USA: Plenum, 2006.

[13] (2012). "Photocathode technology", Application note, Hamamatsu Photonics. [Online]. Available: https://www.hamamatsu.com/us/en/technology/innovation/photocathode/index.html

[14] X. Michalet, R. A. Colyer, G. Scalia, A. Ingargiola, R. Lin, J. E. Millaud, S. Weiss, O. H. W. Siegmund, A. S. Tremsin, J. V. Vallerga, A. Cheng, M. Levi, D. Aharoni, K. Arisaka, F. Villa, F. Guerrieri, F. Panzeri, I. Rech, A. Gulinatti, F. Zappa, M. Ghioni, and S. Cova, "Development of new photon-counting detectors for single-molecule fluorescence microscopy," *Phil. Trans. R. Soc., B*, vol. 368, no. 1611, pp. 20120035-1–20120035-22, Dec. 2012.

[15] M. Eigen and R. Rigler, "Sorting single molecules. Applications to diagnostics and evolutionary biotechnology," *Proc. Nat. Acad. Sci. USA*, vol. 91, pp. 5740–5747, Jun. 1994.

[16] R. Kasper, B. Harke, C. Forthmann, P. Tinnefeld, S. W. Hell, and M. Sauer, "Single-molecule STED microscopy with photostable organic fluorophores," *Small*, vol. 6, no. 13, pp. 1379–1384, Jul. 2010.

[17] C. Eggeling, C. Ringemann, R. Medda, G. Schwarzmann K. Sandhoff, S. Polyakova, V. N. Belov, B. Hein, C. von Middendorff, A. Schöle, and S. W. Hell, "Direct observation of the nanoscale dynamics of membrane lipids in a living cell," *Nature*, vol. 457, no. 7233, pp 1159–1162, Feb. 2009.

[18] M. J. Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, and W. W. Webb, "Zero-mode waveguides for single-molecule analysis at high concentrations," *Science*, vol. 299, no. 5607, pp. 682–686, Jan. 2003.

[19] P. R. Selvin and T. Ha Eds., *Single-Molecule Techniques: A Laboratory Manual*, 1st ed. Cold Spring Harbor, NY, USA:Laboratory Press, 2007.

[20] J. R. Fries, L. Brand, C. Eggeling, M. Kollner, and C. A. M. Seidel, "Quantitative identification of different single molecules by selective

time-resolved confocal fluorescence spectroscopy," *J. Phys. Chem. A*, vol. 102, no. 33, pp. 6601–6613, Jul. 1998.

[21] M. Antonik, S. Felekyan, A. Gaiduk, and C. A. M. Seidel, "Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis," *J. Phys. Chem. B*, vol 110, no. 13, pp. 6970–6978, Mar. 2006.

[22] E. Nir, X. Michalet, K. M. Hamadani, T. A. Laurence, D. Neuhauser, Y. Kovchegov, and S. Weiss, "Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments," *J. Phys. Chem. B*, vol. 110, no. 44, pp. 22103–22124, Oct. 2006.

[23] L. Stryer and R. P. Haugland, "Energy transfer: A spectroscopic ruler," *Proc. Nat. Acad. Sci. USA*, vol. 58, no. 2, pp. 719–726, May 1968.

[24] S. Weiss, "Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy," *Nat. Struct. Biol.*, vol. 7, no. 9, pp. 724–729.

[25] T. Förster, "Energy migration and fluorescence," *J. Biom. Opt.* vol. 17, no. 1, pp. 011002-1–011002-10, Feb. 2012.

[26] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, and S. Weiss, "Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor," *Proc. Nat. Acad. Sci. USA*, vol. 92, no. 13, pp. 6264–6268, Jun. 1996.

[27] A. A. Deniz, M. Dahan, J. R. Grunwell, T. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss, and P. G. Schultz, "Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations," *Proc. Nat. Acad. Sci. USA*, vol. 96, no. 7, pp. 3670–3675, Mar. 1999.

[28] N. K. Lee, A. N. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. H. Ebright, and S. Weiss, "Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation," *Biophys. J.*, vol. 88, no. 4, pp. 2939–2953, Apr. 2005.

[29] C. Eggeling, S. Berger, L. Brand, J. R. Fries, J. Schaffer, A. Volkmer, and C. A. Seidel, "Data registration and selective single-molecule analysis using multi-parameter fluorescence detection," *J. Biotech.*, vol. 86, no. 3, pp. 163–180, Apr. 2003.

[30] J. Widengren, V. Kudryavtsev, M. Antonik, S. Berger, M. Gerken, and C. A. M. Seidel, "Single-molecule detection and identification of multiple species by multiparameter fluorescence detection, " *Anal. Chem.*, vol. 78, no. 6, pp. 2039–2050, Mar. 2006.

[31] A. N. Kapanidis, N. K. Lee, T. A. Laurence, S. Doose, E. Margeat, and S. Weiss, "Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 24, pp. 8936–8941, Jul. 2004.

[32] T. A. Laurence, X. X. Kong, M. Jager, and S. Weiss, "Probing structural heterogeneities and fluctuations of nucleic acids and denatured proteins," *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 48, pp. 17348–17353, Nov. 2005.

[33] B. K. Muller, E. Zaychikov, C. Brauchle, and D. C. Lamb, "Pulsed interleaved excitation," *Biophys. J.*, vol. 89, no. 5, pp. 3508–3522, Nov. 2005.

[34] V. Kudryavtsev, M. Sikor, S. Kalinin, D. Mokranjac, C. A. M Seidel, and D. C. Lamb, "Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements," *Chem. Phys. Chem.* vol. 13, no 4, pp. 1060–1078, Mar. 2012.

[35] T. E. Tomov, R. Tsukanov, R. Masoud, M. Liber, N. Plavner, and E. Nir, "Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis," *Biophys. J.*, vol. 102, no. 5, pp. 1163–1173, Mar. 2012.

[36] H. Neubauer, N. Gaiko, S. Berger, J. Schaffer, C. Eggeling, J. Tuma, L. Verdier, C. A. Seidel, C. Griesinger, and A. Volkmer, "Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy," *J. Amer. Chem. Soc.*, vol. 129, no. 42, pp. 12746–12755, Oct. 2007.

[37] S. Kalinin, E. Sisamakis, S. W Magennis, S. Felekyan, and C. A. M. Seidel, "On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits," *J. Phys. Chem. B*, vol. 114, no. 18, pp. 6197–6206, Apr. 2010.

[38] S. Sindbert, S. Kalinin, N. Hien, A. Kienzler, L. Clima, W. Bannwarth, B. Appel, S. Müller, and C. A. Seidel, "Accurate distance determination of nucleic acids via Forster resonance energy transfer: Implications of dye linker length and rigidity," *J. Amer. Chem. Soc.*, vol. 133, no. 8, pp. 2463–2480, Mar. 2011.

[39] N. K. Lee, A. N. Kapanidis, H. R. Koh, Y. Korlann, S. O. Ho, Y. Kim, and S. Weiss, "Three-color alternating-laser excitation of single molecules: Monitoring multiple interactions and distances," *Biophys. J.*, vol. 92, no. 1, pp. 303–312, Jan. 2007.

[40] S. W. Yim, T. Kim, T. A. Laurence, S. Partono, D. S. Kim, Y. Kim *et al.*, "Four-color alternating-laser excitation single-molecule fluorescence spectroscopy for next-generation biodetection assays," *Clin. Chem.*, vol. 58, no. 4, pp. 707–716, Apr. 2012.

[41] A. N. Kapanidis, T. A. Laurence, N. K. Lee, E. Margeat, X. X. Kong, and S. Weiss, "Alternating-laser excitation of single molecules," *Acc. Chem. Res.*, vol. 38, no. 7, pp. 523–533, Apr. 2005.

[42] E. L. Elson and D. Magde, "Fluorescence correlation spectroscopy—Part I: conceptual basis and theory," *Biopolymers*, vol. 13, no. 1, pp. 1–27, Jan. 1974.

[43] O. Krichevsky and G. Bonnet, "Fluorescence correlation spectroscopy: The technique and its applications," *Rep. Prog. Phys.*, vol. 65, no. 2, pp. 251–297, 2002.

[44] I. V. Gopich and A. Szabo, "Single-molecule FRET with diffusion and conformational dynamics," *J. Phys. Chem. B*, vol. 111, no. 44, pp. 12925–12932, Oct. 2007.

[45] A. Hoffmann, D. Nettels, J. Clark, A. Borgia, S. E. Radford, J. Clarke, and B. Schuler, "Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: Recurrence analysis of single particles (RASP)," *Phys Chem. Chem. Phys.*, vol. 13, no. 5, pp. 1857–1871, Feb. 2011.

[46] S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single photon–detection," *Appl. Opt.*, vol. 35, no. 12, pp. 1956–1976, Apr. 1996.

[47] H. Dautet, P. Deschampes, B. Dion, A.D. MacGregor, D. MacSween, R. J. McIntyre, C. Trottier, and P. Webb, "Photon counting techniques with silicon avalanche photodiodes," *Appl. Opt.*, vol. 32, no. 21, pp. 3894–3900, Jul. 1993.

[48] R. J. McIntyre, "On the avalanche initiation probability of avalanche diodes above the breakdown voltage," *IEEE Trans. Electron Devices*, vol. 20, no. 7, pp. 637–641, Jul. 1973.

[49] W. O. Oldham, R. R. Samuelson, and P. Antognetti, "Triggering phenomena in avalanche diodes," *IEEE Trans. Electron Devices*, vol. 19, no. 6, pp. 1056–1060, Sep. 1972.

[50] M. Ghioni, A. Gulinatti, I. Rech, F. Zappa, and S. Cova, "Progress in silicon single-photon avalanche diodes," *IEEE J. Sel. Topics Quantum Electron.*, vol. 13, no. 4, pp. 852–862, Jul./Aug. 2007.

[51] S. M. Sze, *Physics of Semiconductor Devices*. New York, NY, USA: Wiley, 1981, pp. 520–527.

[52] G. A. M. Hurkx, H. C. de Graaff, W. J. Kloosterman, and M. P. G. Knuvers, "A new analytical diode model including tunneling and avalanche breakdown," *IEEE Trans. Electron Devices*, vol. 39, no. 9, pp. 2090–2098, Sep. 1992.

[53] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, Feb. 1992.

[54] K. Graff, *Metal Impurities in Silicon-Device Fabrication*, 2nd ed. Berlin, Germany: Springer-Verlag, 1995.

[55] G. Vincent, A. Chantre, and D. Bois, "Electric field effect on the thermal emission of traps in semiconductor junctions," *J. Appl. Phys.*, vol. 50, no. 8, pp. 5484–5487, Aug. 1979.

[56] P. A. Martin, B. G. Streetman, and K. Hess, "Electric field enhanced emission from non-Coulombic traps in semiconductors," *J. Appl. Phys.*, vol. 52, no. 12, pp. 7409–7415, Dec. 1981.

[57] M. Ghioni, A. Gulinatti, I. Rech, P. Maccagnani, and S. Cova, "Large-area low-jitter silicon single photon avalanche diodes," *Proc. SPIE*, vol. 6900, pp. 69001D-1–69001D-13, 2008.

[58] D. Schroder, *Semiconductor Material and Device Characterization*, 3rd ed. New York, NY, USA: Wiley, 2006.

[59] S. Cova, A. Lacaita, and G.Ripamonti, "Trapping phenomena in avalanche photodiodes on nanosecond scale," *IEEE Electron. Device Lett.*, vol. 12, no. 12, pp. 685–687, Dec. 1991.

[60] M. Ghioni, A. Giudice, S. Cova, and F. Zappa, "High-rate quantum key distribution at short wavelength: Performance analysis and evaluation of silicon single photon avalanche diodes," *J. Mod. Opt.*, vol. 50, no. 14, pp. 2251–2269, 2003.

[61] A. Spinelli and A. L. Lacaita, "Physics and numerical simulation of single photon avalanche diodes," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 1931–1943, Nov. 1997.

[62] S. Cova, A. Longoni, and A. Andreoni, "Towards picosecond resolution with single-photon avalanche diodes," *Rev. Sci. Instrum.*, vol. 52, no. 3, pp. 408–412, Mar. 1981.

[63] S. Felekyan, R. Kuhnemuth, V. Kudryavtsev, C. Sandhagen, W. Becker, and C. A. M. Seidel, "Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection," *Rev. Sci. Instrum.*, vol. 76, no. 8, pp. 083104-1–083104-14, Jul. 2005.

[64] H. Yang, G. B. Luo, P. Karnchanaphanurach, T. M. Louie, I. Rech, S. Cova, L. Xun, and X. S. Xie, "Protein conformational dynamics probed by single-molecule electron transfer," *Science*, vol. 302, no. 5643, pp. 262–266, Oct. 2003.

[65] (2014). SPCM, Single-photon counting module data sheet, Excelitas Technologies Corporation, [Online]. Available: http://www.excelitas.com/Downloads/DTS_SPCM-AQRH.pdf

[66] P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of avalanche photodiodes," *RCA Rev.*, vol. 35, pp. 234–278, Jun. 1974.

[67] R. J. McIntyre, "The distribution of gains in uniformly multiplying avalanche photodiodes: Theory," *IEEE Trans. Electron Devices*, vol. 19, no. 6, pp. 703–713, Jun. 1972.

[68] R. J. McIntyre, "Silicon avalanche photodiode with low multiplication noise," U.S. Patent 4,972,242, 1990.

[69] R. J. McIntyre and P. P. Webb, "Low-noise, reach-through, avalanche photodiodes," U.S. Patent n. 5,583,352, 1996.

[70] I. Rech, I. Labanca, M. Ghioni, and S. Cova, "Modified single photon counting modules for optimal timing performance," *Rev. Sci. Instrum.*, vol. 77, pp. 033104-1–033104-5, 2006.

[71] G. Kell, A. Bülter, M. Wahl, and R. Erdmann, "τ-SPAD: A new red sensitive single-photon counting module," *Proc. SPIE*, vol. 8033, pp. 803303-1–803303-8, May 2011.

[72] (2013). τ-SPAD, Single photon counting module data sheet, PicoQuant GmbH, [Online]. Available: http://www.picoquant.com/images/uploads/downloads/tau-spad.pdf

[73] (2014). SAP500, Silicon Geiger Mode Avalanche Photodiode data sheet, Laser Components Inc., [Online]. Available: http://www.lasercomponents.com/fileadmin/user_upload/home/Datasheets/lcd/sap-series.pdf

[74] M. Stipcević, H. Skenderović, and D. Gracin, "Characterization of a novel avalanche photodiode for single photon detection in VIS-NIR range," *Opt. Exp.*, vol. 18, no. 16, pp. 17448–17459, Aug. 2010.

[75] (2013). id120, Photon counting module datasheet, ID Quantique SA, [Online]. Available: http://www.idquantique.com/images/stories/PDF/id120/id120-specs.pdf

[76] M. Ghioni, S. Cova, A. Lacaita, and G. Ripamonti, "New silicon epitaxial avalanche diode for single-photon timing at room temperature," *Electron. Lett.*, vol. 24, no. 24, pp. 1476–1477, Nov. 1988.

[77] A. Lacaita, M. Ghioni, and S. Cova, "Double epitaxy improves single-photon avalanche diode performance," *Electron. Lett.*, vol. 25, no. 13, pp. 841–843, Jun. 1989.

[78] W. J. Kindt and H. W. van Zeijl, "Modelling and fabrication of geiger mode avalanche photodiodes," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 715–719, Jun. 1998.

[79] J. C. Jackson, A. P. Morrison, D. Phelan, and A. Mathewson, "A novel silicon Geiger-mode avalanche photodiode," in *Proc. Dig. Int. Electron Devices Meeting*, San Francisco, CA, USA, Dec. 2002, pp. 797–800.

[80] B. F. Aull, A. H. Loomis, J. A. Gregory, and D. J. Young, "Geiger-mode avalanche photodiode arrays integrated with CMOS timing circuits," in *Proc. Dig. 56th Ann. Device Res. Conf.*, Jun. 1998, pp. 58–59.

[81] E. Sciacca, A. C. Giudice, D. Sanfilippo, F. Zappa, S. Lombardo, R. Consentino, C. Di Franco, M. Ghioni, G. Fallica, G. Bonanno, S. Cova, and E. Rimini, "Silicon planar technology for single-photon optical detectors," *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 918–925, Apr. 2003.

[82] A. Spinelli, M. Ghioni, S. Cova, and L. M. Davis, "Avalanche detector with ultraclean response for time-resolved photon counting," *IEEE J. Quantum. Electron.*, vol. 34, no. 5, pp. 817–821, May 1998.

[83] (2013). PDM-series, Photon detection module data sheet, Micro-Photon-Devices S.r.l., [Online] Available: http://www.micro-photon-devices.com/Docs/Datasheet/PDM.pdf

[84] M. Ghioni, G. Armellini, P. Maccagnani, I. Rech, M. K. Emsley, and M. Selim Unlu, "Resonant-cavity-enhanced single-photon avalanche diodes on reflecting silicon substrates" *IEEE Photon. Technol. Lett.*, vol. 20, no. 6, pp. 413–415, Mar. 2008.

[85] A. Gulinatti, I. Rech, F. Panzeri, C. Cammi, P. Maccagnani, M. Ghioni, and S. Cova, "New silicon SPAD technology for enhanced red-sensitivity, high-resolution timing and system integration," *J. Mod. Opt.*, vol. 59, no. 17, pp. 1489–1499, Oct. 2012.

[86] PCD-Mini user manual. (2012). SensL, Ireland, [Online]. Available: http://www.sensl.com/downloads/ds/UM-PCDMini.pdf

[87] (2014). id100 series data sheet, ID Quantique SA, [Online] Available: http://www.idquantique.com/images/stories/PDF/id100-single-photon-detector/id100-specs.pdf

[88] A. Rochas, M. Gani, B. Furrer, P. A. Besse, R. S. Popovic, G. Ribordy, and N. Gisin "Single photon detector fabricated in a complementary metal–oxide–semiconductor high-voltage technology," *Rev. Sci. Instrum.*, vol. 74, no. 7, pp. 3263–3270, Jul. 2003.

[89] A. Rochas, M.Gösch, A. Serov, P. A. Besse, R. S. Popovic, T. Lasser, and R. Rigler, "First fully integrated 2-D array of single-photon detectors in standard CMOS technology," *IEEE Photon. Technol. Lett.*, vol. 15, no. 7, pp. 963–965, Jul. 2003.

[90] F. Zappa, S. Tisa, A. Gulinatti, A. Gallivanoni, and S. Cova, "Complete single-photon counting and timing module in a microchip," *Opt. Lett.*, vol. 30, no. 11, pp. 1327–1329, Jun. 2005.

[91] A. Rochas, A. Pauchard, L. Monat, A. Matteo, P. Trinkler, R. Thew, and R. Ribordy "Ultra-compact CMOS single photon detector," *Proc. SPIE*, vol. 6372, pp. 63720N-1–63720N-8, Oct. 2006.

[92] S. Tisa, F. Guerrieri, and F. Zappa, "Variable-load quenching circuit for single-photon avalanche diodes," *Opt. Exp.*, vol. 16, no. 3, pp. 2232–2244, Feb. 2008.

[93] C. Niclass, M. Sergio, and E. Charbon, "A single photon avalanche diode array fabricated in 0.35 μm CMOS and based on an event-driven readout for TCSPC experiments," *Proc. SPIE,*, vol. 63720, pp. 63720S-1–63720S-12, Oct. 2006.

[94] D. Mosconi, D. Stoppa, L. Pancheri, L. Gonzo, and A. Simoni, "CMOS single-photon avalanche diode array for time-resolved fluorescence detection," Proc. Eur. Solid-State Circuits Conf., Montreux, Switzerland, Sep. 2006, pp. 564–567.

[95] M. J. Binns, S. Bertolini, R. Wise, D.J. Myers, and T. A. McKenna, "Effective intrinsic gettering for 200 mm and 300 mm P/P-wafers in a low thermal budget 0.13 μm advanced CMOS logic process," in *Semiconductor Silicon* (The Electrochemical Society Proceedings Series), H. R. Huff, L. Fabry, and S. Kishino, Eds. 2002, Pennington, NJ, USA, pp. 647–657

[96] N. Faramarzpour, M. J. Deen, S. Shirani, and Q. Fang, "Fully integrated single photon avalanche diode detector in standard CMOS 0.18 μm technology," *IEEE Trans. Electron Devices*, vol. 55, no. 3, pp. 760–767, Mar. 2008.

[97] M. A. Marwick and A. G. Andreou, "Single photon avalanche photodetector with integrated quenching fabricated in TSMC 0.18 μm 1.8 V CMOS process," *Electron. Lett.*, vol. 44, no. 10, pp. 643–644, May 2008.

[98] H. Finkelstein, M. J. Hsu, and S. C. Esener, "STI-bounded single-photon avalanche diode in a deep submicron CMOS technology," *IEEE Electron Device Lett.*, vol. 27, no. 11, pp. 887–889, Nov. 2006.

[99] M. Gersbach, J. Richardson, E. Mazaleyrat, S. Hardillier, C. Niclass, R. Henderson, L. Grant, and E. Charbon, "A low-noise single-photon detector implemented in a 130 nm CMOS imaging process," *Solid State Electron.*, vol. 53, no. 7, pp. 803–808, Jul. 2009.

[100] C. Niclass, M. Gersbach, R. K. Henderson, L. A. Grant, and E. Charbon, "A single photon avalanche diode implemented in 130 nm CMOS technology," *IEEE J. Sel. Topics Quantum Electron.*, vol. 13, no. 4, pp. 863–869, Jul./Aug. 2007.

[101] J. Richardson, E. A. G. Webster, L. A. Grant, and R. Henderson, "Scaleable single-photon avalanche diode structures in nanometer CMOS technology," *IEEE Trans. Electron. Devices*, vol. 58, no. 7, pp. 2028–2035, Jul. 2011.

[102] M. A. Karami, M. Gersbach, H. J. Yoon, and E. Charbon, "A new single photon avalanche diode in 90nm standard CMOS technology," *Opt. Exp.*, vol. 18, no. 21, pp. 22158–22166, Aug. 2010.

[103] E. A. G. Webster, J. Richardson, L. Grant, D. Renshaw, and R. Henderson, "A single- photon avalanche diode in 90 nm CMOS imaging technology with 44% photon detection efficiency at 690 nm," *IEEE Electron Device Lett.*, vol. 33, no. 5, pp. 694–696, May 2012.

[104] C. Niclass, A. Rochas, P. A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes," *IEEE J. Solid-State Circuits*, vol. 40, no. 9, pp. 1847–1854, Sep. 2005.

[105] D. Stoppa, L. Pancheri, M. Scandiuzzio, L. Gonzo, G.-F. Dalla Betta, and A. Simoni, "A CMOS 3-D imager based on single photon avalanche diode," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 1, pp. 4–12, Jan. 2007.

[106] C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, "A 128$\times$128 single photon image sensor with column-level 10-bit time-to-digital converter array," *IEEE J. Sol.-State Circ.*, vol. 43, no. 12, pp. 2977–2989, Dec. 2008.

[107] C. Veerappan, J. A. Richardson, R. J. Walker, D. Li, M. W. Fishburn, Y. Maruyama, D. Stoppa, F. Borghetti, M. Gersbach, R. K. Henderson, and E. Charbon, "A 160$\times$128 single-photon image sensor with on-pixel

55ps 10b time-to-digital converter," in *Proc. IEEE Intl. Conf. Solid-State Circuits*, San Francisco, CA, USA, 2011, pp. 312–314.

[108] L. Carrara, C. Niclass, N. Scheidegger, H. Shea, and E. Charbon, "A gamma, X-ray and high energy proton radiation-tolerant CMOS image sensor for space applications," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2009, pp. 40–41.

[109] F. Guerrieri, S. Tisa, A. Tosi, and F. Zappa, "Two-dimensional SPAD imaging camera for photon counting," *IEEE Photon. J.*, vol. 2, no. 5, pp. 759–774, Oct. 2010.

[110] M. Gersbach, Y. Maruyama, R. Trimananda, M. W. Fishburn, D. Stoppa, J. A. Richardson, R. J. Walker, R. Henderson, and E. Charbon, "A time-resolved, low-noise single-photon image sensor fabricated in deep-submicron CMOS technology," *IEEE J. Sol.-State Circ.*, vol. 47, no. 6, pp. 1394–1407, Jun. 2012.

[111] S. Donati, G. Martini, and M. Norgia, "Microconcentrators to recover fill-factor in image photodetectors with pixel on-board processing circuits," *Opt. Exp.*, vol. 15, no. 26, pp. 18066–18075, Dec. 2007.

[112] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrod, and E. H. K. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination," *Microscopy. Science*, vol. 305, no. 5686, pp. 1007–1009, Aug. 2004.

[113] J. Buchholz, J. W. Krieger, G. Mocsár, B. Kreith, E. Charbon, G. Vámosi *et al.*, "FPGA implementation of a 32×32 autocorrelator array for analysis of fast image series," *Opt. Exp.*, vol. 20, no. 16, pp. 17767–17782, Jul. 2012.

[114] A. Singh, J. Krieger, J. Buchholz, E. Charbon, J. Langowski, and T. Wohland, "The performance of 2D array detectors for light sheet based fluorescence correlation spectroscopy," *Opt. Exp.*, vol. 21, no. 7, pp. 8652–8668, Apr. 2013.

[115] M. Orrit and J. Bernard, "Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal," *Phys. Rev. Lett.*, vol. 65, no. 21, pp. 2716–2719, Nov. 1990.

[116] E. B. Shera, N. K. Seitzinger, L. M. Davis, R. A. Keller, and S. A. Soper, "Detection of single fluorescent molecules," *Chem. Phys. Lett.*, vol. 174, no. 6, pp. 553–557, Nov. 1990.

[117] L. Q. Li and L. M. Davis, "Single-photon avalanche-diode for single-molecule detection," *Rev. Sci. Instrum.*, vol. 64, no. 6, pp. 1524–1529, Jun. 1993.

[118] E. Betzig and R. J. Chichester, "Single molecules observed by near-field scanning optical microscopy," *Science*, vol. 262, no. 5138, pp. 1422–1425, Nov. 1993.

[119] A. Spinelli, L. M. Davis, and H. Dautet, "Actively quenched single-photon avalanche diode for high repetition rate time-gated photon counting," *Rev. Sci. Instrum.*, vol. 67, no. 1, pp. 55–61, Jan. 1996.

[120] R. Rigler and U. Mets, "Diffusion of single molecules through a Gaussian laser beam," *Proc. SPIE*, vol. 1921, pp. 239–248, Jun. 1993.

[121] R. Rigler, U. Mets, J. Widengren, and P. Kask, "Fluorescence correlation spectroscopy with high count rate and low background—Analysis of translational diffusion," *Eur. Biophys. J.*, vol. 22, no. 3, pp. 169–175, Aug. 1993.

[122] F. Panzeri, A. Ingargiola, R. R. Lin, N. Sarkhosh, A. Gulinatti, I. Rech, M. Ghioni, S. Cova, S. Weiss, and X. Michalet, "Single-molecule FRET experiments with a red-enhanced custom technology SPAD," *Proc. SPIE*, vol. 8590, pp. 85900D-1–85900D-11, Feb. 2013.

[123] S. Doose, H. Neuweiler, and M. Sauer, "Fluorescence quenching by photoinduced electron transfer: A reporter for conformational dynamics of macromolecules," *Chem. Phys. Chem.*, vol. 10, no. 9/10, pp. 1389–1398, May 2009.

[124] A. N. Kapanidis, E. Margeat, T. A. Laurence, S. Doose, S. O. Ho, J. Mukhopadhyay, E. Kortkhonjia, V. Mekler, R. H. Ebright, and S. Weiss, "Retention of transcription initiation factor sigma(70) in transcription elongation: Single-molecule analysis," *Mol. Cell*, vol. 20, no. 3, pp. 347–356, Nov. 2005.

[125] W. J. A. Koopmans, R. Buning, T. Schmidt, J. van Noort, "spFRET using alternating excitation and FCS reveals progressive DNA unwrapping in nucleosomes," *Biophys. J.*, vol. 97, no. 1, pp. 195–204, Jul. 2009.

[126] E. Margeat, A. N. Kapanidis, P. Tinnefeld, Y. Wang, J. Mukhopadhyay, R. H. Ebright, and S. Weiss, "Direct observation of abortive initiation and promoter escape within single immobilized transcription complexes," *Biophys. J.*, vol. 904, pp. 1419–1431, Feb. 2006.

[127] D. N. Fittinghoff, P. W. Wiseman, and J. A.Squier, "Widefield multiphoton and temporally decorrelated multifocal multiphoton microscopy," *Opt. Exp.*, vol. 7, no. 8, pp. 273–279, Oct. 2000.

[128] J. Bewersdorf, R. Pick, and S. W. Hell, "Multifocal multiphoton microscopy," *Opt. Lett.*, vol. 23, no. 9, pp. 655–657, May 1998.

[129] R. Heintzmann, Q. S. Hanley, D. Arndt-Jovin, and T. M. Jovin, " A dual path programmable array microscope (PAM): Simultaneous acquisition of conjugate and non-conjugate images," *J. Microsc.*, vol. 204, pt. 2, pp. 119–135, Nov. 2001.

[130] M. Gösch, H. Blom, S. Anderegg, K. Korn, P. Thyberg, M. Wells, T. Lasser, R. Rigler, A. Magnusson, and A. Hård, "Parallel dual-color fluorescence cross-correlation spectroscopy using diffractive optical elements," *J. Biomed. Opt.*, vol. 10, no. 5, pp. 054008-1–054008-7, Oct. 2005.

[131] R. A. Colyer, G. Scalia, I. Rech, A. Gulinatti, M. Ghioni, S. Cova, S. Weiss, and X. Michalet, "High-throughput FCS using an LCOS spatial light modulator and an 8×1 SPAD array," *Biomed. Opt. Exp.*, vol. 1, no. 5, pp. 1408–1431, Dec. 2010.

[132] X. Michalet, R. A. Colyer, G. Scalia, T. Kim, M. Levi, D. Aharoni, A. Cheng, F. Guerrieri, K. Arisaka, J. Millaud, I. Rech, D. Resnati, S. Marangoni, A. Gulinatti, M. Ghioni, S. Tisa, F. Zappa, S. Cova, and S. Weiss, "High-throughput single-molecule fluorescence spectroscopy using parallel detection," *Proc. SPIE.*, vol. 7608, pp. 76082D-1–76082D-12, Jan. 2010.

[133] A. Ingargiola, R. A. Colyer, D. Kim, F. Panzeri, R. Lin, A. Gulinatti, I. Rech, M. Ghioni, S. Weiss, and X. Michalet, "Parallel multispot sm-FRET analysis using an 8-pixel SPAD array," *Proc SPIE*, vol. 8228, pp. 82280B-1–82280B-8, Feb. 2012.

[134] D. Tyndall, R. Walker, K. Nguyen, R. Galland, G. Jie, I. Wang, M. Kloster, A. Delon, and R. Henderson, "Automatic laser alignment for multifocal microscopy using a LCOS SLM and a 32 × 32 pixel CMOS SPAD array," *Proc. SPIE,* vol. 8086, pp. 80860S-1–80860S-6, Jun. 2011.

[135] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "A new approach to optical crosstalk modeling in single-photon avalanche diodes," *IEEE Photon. Technol. Lett.*, vol. 20, no. 5, pp. 330–332, Mar. 2008.

[136] I. Rech, A. Ingargiola, R. Spinelli, I. Labanca, S. Marangoni, M. Ghioni, and S. Cova, "Optical crosstalk in single photon avalanche diode arrays: A new complete model," *Opt. Exp.*, vol. 16, no. 12, pp. 8381–8394, Jun. 2008.

[137] A. Gulinatti, I. Rech, P. Maccagnani, S. Cova, and M. Ghioni, "New silicon technologies enable high-performance arrays of single photon avalanche diodes," *Proc. SPIE*, vol. 8727, pp. 87270M-1–87270M-10, May 2013.

[138] I. Rech, S. Marangoni, D. Resnati, M. Ghioni, and S. Cova, "Multi-pixel single-photon avalanche diode array for parallel photon counting applications," *J. Mod. Opt.*, vol. 56, no. 2–3, pp. 326–333, Jan. 2009.

[139] R. A. Colyer, G. Scalia, T. Kim, I. Rech, D. Resnati, S. Marangoni, M. Ghioni, S. Cova, S. Weiss, and X. Michalet, "High-throughput multispot single-molecule spectroscopy," *Proc. SPIE*, vol. 7571, pp. 75710G-1–75710G-10, Feb. 2010.

[140] A. Ingargiola, F. Panzeri, N. Sarkhosh, A. Gulinatti, I. Rech, M. Ghioni, S. Weiss, and X. Michalet, "8-spot smFRET analysis using two 8-pixel SPAD arrays," *Proc. SPIE*, vol. 8590, pp. 85900E-1–85900E-11, Feb. 2013.

[141] S. K. Das, M. D. Austin, M. C. Akana, P. Deshpande, H. Cao, and M. Xiao, " Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes," *Nucleic Acids Res.*, vol. 38, no. 18, pp. 1–8, Aug. 2010.

[142] A. Gulinatti, I. Rech, P. Maccagnani, and M. Ghioni, "A 48-pixel array of single photon avalanche diodes for multispot single molecule analysis," *Proc. SPIE*, vol. 8631, pp. 86311D-1–86311D-8, Feb. 2013.

[143] F. Zappa, A. Lotito, A. C. Giudice, S. Cova, and M. Ghioni, "Monolithic active-quenching and active-reset circuit for single-photon avalanche detectors," *IEEE J. Solid-State Circuits*, vol. 38, no. 7, pp. 1298–1301, Jul. 2003.

[144] B. Aull, J. Burns, C. Chen, B. Felton, H. Hanson, C. Keast, J. Knecht, A. Loomis, M. Renzi, A. Soares, V. Suntharalingam, K. Warner, D. Wolfson, D. Yost, and D. Young, "Laser radar imager based on three dimensional integration of Geiger-mode avalanche photodiodes with two SOI timing-circuit layers," in *Proc. IEEE Int. Conf. Solid-State Circuits*, San Francisco, CA, USA, Feb. 2006, pp 1179–1188.

[145] X. Michalet, A. Cheng, J. Antelman, M. Suyama, K. Arisaka, and S. Weiss, "Hybrid photodetector for single-molecule spectroscopy and microscopy," *Proc. SPIE*, vol. 6862, pp. 68620F-1–68620F-12, Feb. 2008.