COPYRIGHT NOTICE

L. Di Sieno, A. Behera, S. Rohilla, E. Ferocino, D. Contini, A. Torricelli, B. Krämer, F. Koberling, A. Pifferi, A. Dalla Mora

Probe-hosted large area silicon photomultiplier and high-throughput timing electronics for enhanced performance time-domain functional near-infrared spectroscopy

L. Di Sieno,[1,*] ◉ A. Behera,[1] S. Rohilla,[2,3] E. Ferocino,[1] ◉ D. Contini,[1] ◉ A. Torricelli,[1,4] B. Krämer,[5] F. Koberling,[5] A. Pifferi,[1,4] ◉ and A. Dalla Mora[1] ◉

[1]*Politecnico di Milano, Dipartimento di Fisica, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*
[2]*PicoQuant Innovation GmbH, Rudower Chaussee 29, 12489 Berlin, Germany*
[3]*Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu and Berlin Institute of Health, Department of Internal Medicine/Infectious Diseases and Respiratory Medicine, Charitéplatz 1, 10117 Berlin, Germany*
[4]*Consiglio Nazionale delle Ricerche, Istituto di Fotonica e Nanotecnologie, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*
[5]*PicoQuant GmbH, Rudower Chaussee 29, 12489 Berlin, Germany*
[*]laura.disieno@polimi.it

Abstract: Two main bottlenecks prevent time-domain diffuse optics instruments to reach their maximum performances, namely the limited light harvesting capability of the detection chain and the bounded data throughput of the timing electronics. In this work, for the first time to our knowledge, we overcome both those limitations using a probe-hosted large area silicon photomultiplier detector coupled to high-throughput timing electronics. The system performances were assessed based on international protocols for diffuse optical imagers showing better figures with respect to a state-of-the-art device. As a first step towards applications, proof-of-principle in-vivo brain activation measurements demonstrated superior signal-to-noise ratio as compared to current technologies.

1.   Introduction

In the last years, new low cost photonics instruments for clinical monitoring of vital parameters (such as tissue oxygenation) as well as for diagnostics (e.g. tumor detection) have been proposed [1–5]. In this context, diffuse optics aims to study the composition of biological tissues through the recovery of the absorption (μ_a) and reduced scattering (μ_s') coefficients [6,7], which are related to tissue chemical composition and microstructure, respectively.

Most of the proposed systems (even wearable ones) are based on the continuous wave (CW) approach due to the inherent lower cost and simplicity of the instrumentation [8]. Yet, CW instruments suffer from issues like vulnerability to superficial artifacts [9,10], lower depth sensitivity [11] and lack in precise quantitation [12]. However, some of these issues have been solved using suitable strategies (e.g. multi-distance or tomographic measurement schemes to uncouple superficial artifacts from deep signals [13] or hybrid approaches to improve quantitation [14]). In the last years, CW instruments have reached very high performances as demonstrated for example in Ref. [15]. However, the time-domain (TD) approach is considered to offer higher informative content and more promising performances [11,16].

The TD method relies on the injection of short laser pulses (duration of few tens/hundreds of ps) in the sample and on the reconstruction of the temporal distribution of the re-emitted photons (distribution of time-of-flight, DTOF). This approach has several advantages such as

disentangling the specific effects of absorption and scattering phenomena, and encoding the mean depth probed by photons in their arrival time. As demonstrated in Ref. [11], the TD technique has the potential to reach depths of about 6 cm in the ideal case – i.e. i) a dense coverage of pulsed sources and detectors onto the probe to maximize the amount of injected light; ii) a high-throughput timing electronics to process all photon counts without losses; iii) a time-gating capability to collect photons at null source-detector separation distance [17]. Conversely, using an ideal CW system under the same conditions, the penetration depth is limited to less than 3 cm. The improvement in depth sensitivity achievable with new generation TD instruments will allow one to explore completely new applications for *in-vivo* non-invasive diagnostics (e.g. monitoring of deep organs such as lungs or heart), which are not easily addressed by optical methods.

Compared to CW, TD systems are however more complicated, cumbersome and costly. Indeed, many of them are laboratory prototypes (see, for example, Ref. [18–22]) hosted inside rack cabinets. However, in the last years, several papers showing first attempts for source miniaturization and detector probe-hosting [23–25], as well as a first compact and cost-effective systems have been published [26], thus demonstrating the possibility to decrease both cost and size of TD systems, still maintaining excellent performances.

Two main bottlenecks prevent the full exploitation of the TD technique, that are: i) the limited light harvesting capability of the detection chain; ii) the limited data throughput of the timing electronics. These two limitations have to be overcome jointly for achieving improved performances. A large area detector can actually provide a large number of photon counts to a timing electronics, but this latter must sustain very large count-rates (CRs) to avoid losses of information.

For what concerns the detectors, the choice is between photo-cathode based devices (e.g. Photomultiplier Tubes –PMTs–, hybrid PMTs and Micro-Channel Plate PMTs) and solid-state detectors (Single-Photon Avalanche Diodes –SPADs– and Silicon Photomultipliers –SiPMs–). The former ones feature a large active area, which allows them to have a high light harvesting capability. However, they are bulky and not scalable in dimensions (preventing probe-hosting, thus requiring the use of optical fibers with limited numerical aperture, reducing the photon collection efficiency), sensitive to electromagnetic fields and mechanical shocks and require a high biasing voltage (in the order of kV) to be operated [16]. All these characteristics make them not suitable for next generation of TD-instruments which aim to be scalable, low-cost and wearable. In the last years, solid-state detectors have been considered the best solution for new TD-systems [16]. Up to few years ago, SPADs were the only solid-state detectors used in TD diffuse optics [27]. However, they were not the first choice for diffuse optical measurements due to their small active area (diameter of few hundreds of micrometers in the best cases [28,29]) to maintain performances suitable for diffuse optical applications). In 2015 this bottleneck was overcome thanks to the introduction of SiPMs in TD diffuse optics [30]. Indeed, SiPMs are arrays of several SPADs (hundreds or thousands) all connected in parallel and with a high fill factor (usually larger than 50%). For this reason, they have the advantages of solid-state detectors (scalability, ruggedness and immunity to electromagnetic field), while overcoming their main limitation, that is the small active area. The typical active area of the SiPMs used in diffuse optics are in the order of 1 mm^2 (see, for example, Ref. [31–34]). The improvement in the active area (from μm^2 to about one mm^2), coupled to the possibility to place the detector directly in contact with the sample, increase the light harvesting capability reaching figures even higher than those obtained with state-of-the-art fiber-based systems [30].

The maximum achievable throughput is another issue which sets a limit to further boost the performances of TD systems. Actually, in modern TD instruments the timing electronics adopts two different strategies: time-to-analog converters [35,36] and time-to-digital (TDC) converters [37]. The former usually features better timing resolution and linearity at the expenses of cost and size. On the other hand, TDC based electronics can be integrated in a single chip, thus making it

suitable for monolithic and/or multi-channel systems [16]. Despite the different characteristics of those two technologies, nowadays they show overall comparable performances in terms of maximum achievable count-rate, whose limits are usually around few millions of counts per second (Mcps).

Several works have been published showing the possibility to integrate detectors (usually a matrix of small area SPADs) with high speed TDCs [16]. Restricting the field to the devices specifically designed for photon timing application and considering the use of the entire SPAD array as a unique larger area detector (i.e. not for imaging purposes), Table 1 reports the main performances (in terms of overall sensitive area and throughput) of the high-throughput devices recently published. In some cases, single basic units (a small array of detectors and TDC) are parallelized to improve the overall performances of the system [38,39]. From Table 1, it is possible to notice that in all cases the overall detector area is always below 1 mm^2.

Table 1. Comparison between different integrated detection chains especially designed for photon timing applications and working at sustained count-rate. For the detector, the cell size and fill factor (FF) (i.e. the ratio between the effective active area and the nominal detector area) as well as the overall sensitive area are reported.

WORK	DETECTOR		TDC
	cell parameters	sensitive area	throughput
Niclass et al. 2008 [38]	25×25 μm FF = 6%	0.614 mm^2	320 Mcps
Niclass et al. 2011 [39]	25×25 μm FF = 70%	0.168 mm^2	400 Mcps
Tyndall et al. 2012 [43]	22.5×22.5 μm FF = 10%	0.050 mm^2	100 Mcps
Dutton et al. 2015 [44]	21×21 μm FF = 43%	0.194 mm^2	1700 Mcps
Krstajic et al. 2015 [45]	50×50 μm [46] FF = 1.5%	0.384 mm^2	500 Mcps

Focusing on discrete detectors, Ferocino and co-workers [40] demonstrated the possibility to increase the light harvesting capability using an arrangement of 8 separate SiPMs, thus covering an area of 13.5 mm^2 (with a fill factor of 74%, thus resulting in about 10 mm^2 sensitive area). However, in this case, each SiPM provides a single output which has to be summed up with the other seven in post processing. This solution allows one to reach an overall throughput of 40 Mcps, but at the expenses of system complexity and cost as eight separate timing channels are needed (in this case, 8 inputs of a commercial TDC).

Concerning developments in data throughput capabilities of TD systems, a high-throughput gated counter capable to sustain up to 100 Mcps was published and validated for diffuse optical measurements [41,42]. In this case, however, it is not possible to reconstruct the whole DTOF since only the overall number of photons within two well defined time-windows are saved. Such a solution can be profitably used in applications where the whole waveform is not needed (e.g. muscle or brain monitoring), but it cannot be used in applications where a spectroscopic approach is required since it does not permit to recover both μ$_a$ and μ$_s$'.

As shown above, recent works show improvements for both detectors and timing electronics but, at the best of our knowledge, the two bottlenecks have been addressed separately. In this paper, we demonstrate for the first time the possibility to couple a large area detector with a high-throughput electronics.

To this purpose, here we use a 3×3 mm^2 SiPM detector (i.e. the largest area device from which we were able up to now to obtain sub-nanosecond single-photon timing resolution, with 82% fill-factor, resulting in an overall active area of 7.38 mm^2 in a single device) to improve the light harvesting and an innovative time-correlated single-photon counting (TCSPC) device, which can sustain up to 180 Mcps with a very short dead time (650 ps). This system is the first prototype laying the foundation for the next generation of TD system with high performances.

In this paper, we present the system and characterize its performances using internationally agreed protocols for TD instruments such as BIP, MEDPHOT and nEUROPt [47–49]. Further, we challenge it for *in-vivo* motor cortex activation tests (finger tapping exercise) to show the improvement in signal as compared to a state-of-the-art system based on a 1 mm^2 fiber-coupled SiPM detector.

The paper is organized as follows: Section 2 deals with the description of the system, the phantoms and the protocols for *in-vivo* measurements, together with data analysis; Section 3 reports the results and discussion of phantom characterization, as well as of the *in-vivo* acquisitions. Finally, in Section 4 conclusions and future perspectives are presented.

2.　Material and methods

2.1.　Setup

Figure 1(a) reports the schematics of the experimental setup. A pulsed laser driver (PDL 828 Sepia II, Picoquant GmbH, Germany) provided pulses at 40 MHz to two high power laser heads operating at 670 and 830 nm (LDH-P-C-670M and LDH-P-C-830M, Picoquant GmbH, Germany). Each laser head output was directly coupled to an optical fiber (core diameter 600 μm, step index), which serves as the input of a fiber-to-fiber u-bench where a variable optical attenuator ("VOA" in Fig. 1(a)) was hosted to set the proper photon counting rate. The output of each u-bench is another 600 μm diameter core fiber, which was fixed to the 3D printed probe through an SMA connector (as reported in Fig. 1(b)). The maximum average power at the fiber tip was about 2.7 and 2 mW for 670 and 830 nm, respectively.
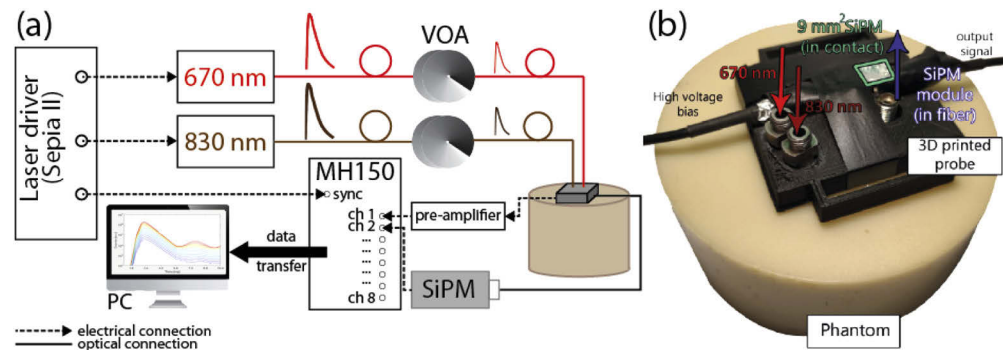

Fig. 1. (a) Schematic of the setup. (b) Picture showing the position of the main components on the 3D printed probe.

The 3D printed probe also hosted the detector, a 3×3 mm^2 SiPM (nominal area $= 9$ mm^2, S13360-3075CS by Hamamatsu Photonics K.K., Japan) with an overall effective active area of 7.38 mm^2 considering a FF of 82% and a dark count rate (DCR) of 550 kcps (at a temperature of the room of about 20°C).

The detector is mounted on a ceramic chip carrier and buried inside a suitable optical resin that provides both optical access and electrical isolation. Thus, the detector was directly put in contact with the sample to exploit at best its large numerical aperture to maximize the collection of the backscattered light. The source-detector separation (SDS, computed as the distance between the injection fiber and the center of the detector) was set to about 3 cm (more in detail, 28.8 and 30.1 mm for the 670 and 830 nm respectively). Under this condition, the lateral size of the detector (3 mm) is expected to have negligible effects on the spatial resolution of investigation [49].

As can be seen in Fig. 1(b), two radiofrequency electrical cables were arranged on the probe and connected to the SiPM, providing the biasing to the detector (56.3 V, i.e. 4 V above the SiPM breakdown level), and conveying the output signal to a fast low-noise broadband pulse amplifier (Ortec VT120A, AMETEK Inc, USA, "pre-amplifier" in Fig. 1(a)).

The timing electronics used for the proposed system was a high-throughput 8-channel TCSPC module (MultiHarp 150, Picoquant GmbH, Germany, "MH150" in Fig. 1(a)) featuring a sub-nanosecond dead-time (650 ps), thus allowing to sustain up to 180 Mcps in histogramming mode [50]. The temporal resolution of its time scale (i.e. the size of the time bin) was 80 ps. The sync input signal to the TCSPC module was provided by the laser synchronization output (as depicted in Fig. 1(a)).

The present system is bulky, because of laser and TCSPC electronics. Yet, suitable compact probe-hosted laser sources have been already demonstrated [23], as well as integrated TCSPC devices capable of extremely high throughput [38,39,44,45,51]. All these advances can permit in the near future the fabrication of extremely compact and wearable TD diffuse optics devices.

Since *in-vivo* measurements are affected by biological fluctuations that can confound the information, diffused photons were simultaneously recorded also by a 1 mm^2 fiber-based SiPM module [31]. The detection fiber (core diameter 1 mm, step index) was screwed in the same 3D-printed probe and the collected light was focused onto the active area of the SiPM module. Also for this control channel, the SDS was set to about 3 cm (28.8 and 30.1 mm for 670 and 830 nm, respectively) and the arrangement of the detection fiber was chosen so as to investigate the same portion of the sample probed by the 9 mm^2 SiPM. The DCR of the 1 mm^2 module is about 125 kcps. The output signal of the SiPM module was then sent to a second channel of the MH150 to record the DTOF. In this way, during *in-vivo* measurements it is possible to record the same biological fluctuation for a better comparison of differences in terms of signal-to-noise ratio from the two detectors.

2.2. Pile-up effect

2.2.1. Description

It is well known and widely discussed in the literature that classical TCSPC instruments are characterized by their dead-time (T_{dead}), which is usually in the order of tens to hundreds of nanoseconds [52,53]. During T_{dead}, the detection chain is blind and all imping photons are neglected. The pile-up effect is a phenomenon that rises from the capability of the classical TCSPC board to detect only one photon per laser cycle. Since only one photon can be detected, the early photons are those that have a higher probability to be recorded, while the late ones are more prone to be lost, thus leading to an artificial increase in the slope of the DTOF tail [52]. Obviously, the distortion in the curve is more important as the number of impinging photons increases, thus conventionally the recorded count-rate is kept within the 1-5% of the laser repetition rate [52] and, as recently discussed by Cominelli et al. [54], the laser period should preferably be an integer multiple of dead-time to avoid further distortion of the DTOF. If this condition is not matched, some portions of the DTOF are recorded with lower efficiency, resulting in regions of the DTOF where almost no photons are detected (e.g. a sort of undershoot of counts followed by a bump, that is a fingerprint of the recovery of the detection chain).

In this setup, the dead-time of the timing board is 650 ps [50], thus the pile-up effect is dominated by the T_{dead} of the 9 mm^2 SiPM detector and its front-end and amplification electronics, which was estimated to be ≈3.6 ns (assessed using an oscilloscope from the minimum distance between two consecutive distinguishable avalanche pulses). Such a value is much smaller than the standard dead-time of TD-systems, thus allowing a strong potential increase of the count-rate (the maximum saturated sustainable CR of the detector alone is around 270 Mcps, if it is retriggered as soon as it recovers from dead-time), provided that possible non-idealities (e.g. pile-up) are taken into account.

2.2.2. Correction algorithm

The Coates algorithm was applied to the acquired curves before any other processing (e.g. background noise subtraction) [55]. It is based on the calculation of the true probability (S_i) that a photon can be detected at the *i-th* time bin under the hypothesis that more than one photon can be detected in one cycle (i.e. no dead-time). This probability is computed from the number of photons (N) recorded both in the *ith* channel (N_i) and in the previous *i-1* channels, but considering also the overall number of cycles of the measurements (M). The Coates' formula states:

$$S_i = -\ln(1 - P_i) \tag{1}$$

where

$$P_i = \frac{N_i}{M - \sum_{j=1}^{i-1} N_j} \tag{2}$$

The corrected number of counts in the *i-th* channel can be computed as $N_i = M \, S_i$. Although the algorithm is simple and easy to be implemented, it does not need any assumption such as the apriori knowledge of the correct photon distribution or the uniformity in the time-bins [55].

Since the dead-time of the detector is about 3.6 ns, the correction is valid for a temporal range stretching no more than 3.6 ns from the beginning of the rising edge of the DTOF. In practice, we decided to consider the 3.6 ns validity range starting from the peak of the Instrument Response Function (IRF) since very few photons are detected before that, particularly with SDS=3 cm. After pile-up correction, all other post-processing analysis (e.g. background subtraction) were performed.

2.3. Measurements and data analysis

To rigorously characterize the system, we made use of internationally-agreed protocols (BIP, MEDPHOT and nEUROPt [47–49]) for performance assessment of diffuse optical instruments. In the following, measurements and data analysis procedures are presented.

Unless specified, all the tests for all protocols were repeated at different CRs (namely, 1, 15 and 30 Mcps nominal values) by acting on the VOA and observing the total photon counting rate displayed by the TCSPC module software. Only one wavelength (670 nm) was used for phantom characterization as the phantoms in use do not show any relevant spectral fingerprint and the characteristics of components were similar at both wavelengths.

2.3.1. BIP protocol

The BIP protocol was designed to assess the basic hardware performances of the system [47]. In this case, we adopted the following figures of merit: the detector responsivity, the Differential Non Linearity (DNL), the IRF, and the temporal stability.

The responsivity quantifies the overall harvesting capability of the system. It is computed as the ratio between the number of photons detected at the output surface of a calibrated reference phantom to mimic a nearly Lambertian angular profile and the power injected on the opposite side. Responsivity depends on many different factors, such as quantum efficiency, area and numerical aperture of the detector, collection optics, and dead-time of the timing electronics [47]. More details on the measurement procedure and the phantom used can be found in Ref. [47]. In our case, the responsivity was computed at the following nominal CRs: 1, 10, 15, 20, 30 and 40 Mcps (i.e. the maximum CR compatible with a laser operating at 40 MHz repetition rate). For all cases, the calculation was repeated both with and without pile-up correction to check the effect of the pile-up.

The DNL test assesses the non-uniformity in the time bins. The number of photons detected in one channel is proportional to its width, thus any non-uniformity creates distortions in the DTOF. To quantify the DNL with a good Signal-to-Noise Ratio (SNR), a CW illumination yielding at

least 10^5 counts per channel is needed. In our case, due to the different CR of the measurements, we modified the number of repetitions so to have a number of counts per channel of about 10^6, independently of the CR. In the ideal case, the number of detected photons in different time bins should be the same so, to assess any deviation, the BIP protocol suggests to compute the peak-to-peak variation in the number of counts recorded per bin with respect to the mean value of counts (ε_{DNL}), as reported in Eq. (3):

$$\varepsilon_{DNL} = \frac{N_{DNL,max} - N_{DNL,min}}{\overline{N_{DNL}}} \tag{3}$$

where $N_{DNL,max}$ ($N_{DNL,min}$) is the maximum (minimum) number of counts recorded across the different time bins, while $\overline{N_{DNL}}$ is the mean number of photons detected over the whole time scale.

Due to the not flat background of the 9 mm^2 detector (as highlighted in the results and discussion section), to assess the DNL we used as a detector the 1 mm^2 fiber-based SiPM (see Fig. 1(a)). This procedure is justified as the DNL value depends only on the timing electronics board.

To characterize the time resolution of the whole system, the BIP protocol dictates to record the IRF (i.e. the temporal response of the instrument when no diffusive sample is present). To correctly replicate the measurement conditions, a thin layer of Teflon was inserted to fill the whole acceptance area of the detector with all the light propagation modes of the injection fiber. As stated in the BIP protocol, 20 repetitions of 1 s each were acquired. The IRF was taken for the three CRs in order to assess any possible issue (e.g. pile-up effect, see Sect. "Correction algorithm") arising from the use of a count-rate above the single-photon statistics.

Another important feature that can have a large impact on the system performances is the temporal stability, which was monitored acquiring continuously the IRF for 1 hour starting from the system start-up. The changes in the temporal position of the IRF as well as its amplitude (overall number of counts in the pulse after background subtraction) and shape (assessed by its Full-Width at Half Maximum, FWHM) were considered.

2.3.2. MEDPHOT protocol

The MEDPHOT protocol [48] aims to assess the capability of a system to recover μ_a and μ_s' of a homogenous medium over a wide range of optical properties. The MEDPHOT phantom kit is composed of 32 phantoms divided in 4 series (each one having a fixed scattering value, nominally 5, 10, 15 and 20 cm^{-1} at 800 nm). For each series, there are 8 phantoms featuring different absorption coefficients (nominally from 0 to 0.35 cm^{-1} at steps of 0.05 cm^{-1} at 800 nm). The phantoms are made of epoxy-resin and calibrated quantities of titanium dioxide and black toner were added to provide scattering and absorption properties. For further details about the phantoms see Ref. [48]. Considering the typical optical properties of the human tissues (such as brain and muscle), we adopted a subset of the MEDPHOT phantom kit featuring a range of absorption from 0.07 to 0.28 cm^{-1} and reduced scattering from 7 to 16 cm^{-1} [56] at 670 nm.

The whole phantom subset was measured to assess the linearity in the recovery of μ_a and μ_s' as well as the possible coupling between retrieved absorption (scattering) vs. nominal scattering (absorption): 20 repetitions of 1 s were acquired for each phantom. To retrieve the optical properties, we first summed up all 20 repetitions to improve SNR and then we fitted the data using the analytical model of the photon transport in a diffusive semi-infinite medium under the diffusion approximation [57]. The fitting algorithm reaches the convergence minimizing the χ^2 using a Levenberg-Marquardt routine [58]. To take into account the non-idealities of the system, the IRF shape (recorded at the same CR as the measurement) was convolved to the analytical model. The curves were fitted in the range spanning from 20% of the peak on the rising edge down to the 5% of the tail of the curve. The fitting range was chosen as the one giving the best

results for phantoms with low signal (i.e. characterized by high scattering and absorption) at 1 Mcps and was kept constant for all measurements. To check the effect of the pile-up and the possibility to properly correct its effects, the fitting procedure was repeated for curves with and without pile-up correction.

2.3.3.   nEUROPt protocol

The nEUROPt protocol [49] was meant to objectively assess the capability of an optical brain imaging system to detect, localize and quantify the absorption changes occurring in the brain. In this paper, we adopted only the sensitivity tests (contrast and contrast-to-noise-ratio), which are related to the capability of the system to detect a localized inhomogeneity buried in depth within a homogeneous diffusive medium.

The contrast (C) is defined as the relative change in the number of photons when the perturbation is present (heterogeneous case) with respect to the homogeneous state (i.e. no perturbation), as reported in Eq. (4):

$$C = \frac{N_0 - N}{N_0} \tag{4}$$

where N_0 and N are the overall number of counts recorded in the homogeneous and heterogeneous case, respectively.

The contrast-to-noise ratio (CNR) is, on the other hand, an index of the robustness of the contrast with respect to the noise (e.g. Poisson noise and other possible sources of fluctuations):

$$CNR = \frac{<N_0 - N>}{\sigma(N_0)} \tag{5}$$

where $\langle N_0 - N \rangle$ is the mean value of the difference over the repetitions between the homogeneous and heterogeneous case while $\sigma(N_0)$ is the standard deviation of the number of counts over several acquisitions of the homogeneous curve.

The typical criteria for detectability of absorption perturbations buried inside scattering media are: i) $C \geq 1\%$ (to distinguish the change in N_0 in living tissues) and ii) $CNR \geq 1$ (to grant that the change in N_0 exceeds its fluctuation, due to e.g. Poisson noise) [11].

Both contrast and CNR can be computed inside different time gates (i.e. selecting given portions of the DTOF with respect to the IRF peak) to improve the depth selectivity since the arrival time of the photons encodes the mean visited depth. In this case, N_0 and N in Eq. (4) and (5), are no longer referred to the whole DTOF but must be computed for the specific gate. For the analysis, we used 5 time gates with a 480 ps width.

As for the MEDPHOT protocol, the calculation of C and CNR was repeated with and without pile-up correction to check its effect and the possibility to correct distortions.

For the measurements, we made use of a mechanically switchable phantom [59]. It is composed of a main block with homogeneous optical properties ($\mu_a = 0.1$ cm^{-1} and $\mu_s' = 10$ cm^{-1} at 670 nm) in which a cylindrical cavity hosts a movable cylinder. Such a rod has the same optical properties of the bulk phantom and it embeds a totally absorbing inclusion of 98 mm^3, featuring an absorption perturbation equivalent to an absorption increase of 0.17 cm^{-1} over a 1 cm^3 volume (see Ref. [60] for further details). The probe in Fig. 1(b) was placed on the phantom in order to have the perturbation always centered with the middle point of the source-detector couples, independently of its depth. The rod was moved so as to have the center of gravity of the inclusion ranging from 2 to 40 mm at steps of 2 mm below the phantom surface, thus testing the depth sensitivity of the system. For each position, 20 repetitions of 1 s each were recorded. The homogeneous measurement was acquired when the inclusion was completely out from the bulk phantom, thus its effect was completely negligible. The targeted CRs (i.e. 1, 15 or 30 MHz) were set acting on the VOA in this last condition and therefore they are referred to the CRs in the homogeneous cases. The CR in the heterogeneous case (i.e. when the effect of the perturbation is not negligible) is typically lower due to the presence of the absorption perturbation.

2.3.4.   In-vivo measurements

After validating the system on phantoms, we challenge it with *in-vivo* motor cortex activation adopting a finger tapping exercise [61]. In this case, both wavelengths were used to enable functional measurements by recovering both oxygenated and deoxygenate hemoglobin concentrations (HbO$_2$ and HHb, respectively). For comparison purposes, the signal was concurrently recorded using the external 1 mm^2 fiber-based SiPM module (see Setup section for details).

Three healthy subjects were enrolled in the experiment. Written informed consent was given by participants and measurements were approved by the Ethical Committee of Politecnico di Milano and conducted in compliance with the Declaration of Helsinki.

Each experiment consisted of 5 repetitions of 3 consecutive phases: 20 s of baseline (i.e. the subject was resting), 20 s of finger tapping task (at a self pace of about 2-3Hz) and finally 20 s of recovery (with the subject resting again). The overall duration was 5 minutes.

The probe was located on the motor cortex (C3 position according to the 10/20 EEG international system) and, to measure both ipsilateral and contralateral hemodynamics changes, in two different experiments the subject was asked to perform the full 5-minute experiment first with left hand and then with the right hand.

DTOFs were acquired every 1 s simultaneously at both wavelengths. A method based on the computation of the mean time pathlengths travelled by the photon in a 2-layer medium [62] was used to derive the absorption coefficient changes $\Delta\mu_{a,j}(\lambda, T)$ in each layer j (with $j = 1$, extra-cerebral tissue, and $j = 2$, intra-cerebral tissue) at each experimental time T. Absolute values for the baseline absorption coefficient $\mu_{a0}(\lambda)$ were then obtained by fitting the average DTOFs in the baseline to a homogeneous model for photon diffusion [57]. The time course of the absorption coefficient was then derived as $\mu_{a,j}(\lambda, T) = \mu_{a0}(\lambda) + \Delta\mu_{a,j}(\lambda, T)$. Finally, hemodynamic concentration time courses HbO$_2$(T) and HHb(T) were obtained by Beer's law.

3.   Results and discussion

Here the results and the discussion of the measurements described in the previous paragraph are presented and commented.

3.1.   BIP protocol

Table 2 reports the responsivity values obtained at the different target CRs, both using the not-corrected DTOFs and the corrected ones. In the "count rate (MHz)" column is reported both the "true value" (i.e. the actual TCSPC module CR obtained due to the VOA discretization) and the "corrected value" (i.e. the total amount of counts in the DTOF after pile-up correction). Looking at the values obtained using the not-corrected curves, it is clear that the increase in the CR is responsible for pile-up effects, thus reducing the ratio between the number of photons detected and those actually impinging on the detector. Such a reduction results in a decrease of the responsivity for the measurement at 40 Mcps, which is less than half of the supposed true value (i.e. at 1 Mcps, within single-photon statistics). This means that more than half of the actual photons are not detected due to pile-up.

To remedy this issue, we used Coates algorithms which corrects for the number of detected photons at high throughput, thus leading to a fairly constant value of the responsivity up to CR=30 Mcps (variation always lower than 1.5% with respect to 1 Mcps, in agreement with the expected experimental variability). When a count-rate of about 40 Mcps is reached, the recovered value is still a bit underestimated (about 7%) with respect to the supposed true value, meaning that the correction is effective even if not perfect. This may be due to other sources of photon losses that arise at such large CRs. For example, it is well-known and widely discussed in literature that SiPMs suffer from avalanche pulse pile-up at high detection rates [63]. Despite sharing a similar name, this phenomenon is completely different from the statistical distortion

Table 2. Responsivity values and percentage change with count-rates before and after applying pile-up correction.

Count rate (Mcps)			Percentage of the laser rep rate (%)	Responsivity (m²sr)		Responsivity change (%)	
nominal value	true value	corrected value		not-corrected	corrected	not-corrected	corrected
1	1.1	1.1	2.8	3.27E-06	3.29E-06	reference	reference
10	10.2	11.8	25.5	2.93E-06	3.29E-06	89.8	100.0
15	15.5	19.6	38.8	2.77E-06	3.33E-06	84.7	101.2
20	20.0	27.8	50.0	2.56E-06	3.30E-06	78.3	100.2
30	30.6	57.9	76.5	2.11E-06	3.30E-06	64.6	100.1
40	38.3	127.3	95.8	1.56E-06	3.04E-06	47.7	92.5

affecting TCSPC systems as it is related to the decay tail of the detector electrical response, which may generate severe baseline fluctuations affecting the detector response. Clearly, the avalanche pulse pile-up cannot be addressed by the Coates correction, which assumes the IRF to be constant. To avoid confusions, in the following we will always refer to the statistical distortion of the DTOF in TCSPC instruments simply as "pile-up effect", and to the electrical detector behavior with the term "avalanche pulse pile-up".

The DNL error was computed to be always lower than 3.5% and it is independent of the CR used in measurements. This independence was predictable since the DNL error depends only on the time-bin width of the timing electronics, thus there is no expected variation with the CR of the measurement.

As a general comment, these values of responsivity at 670 nm are, to the best of our knowledge, the highest values ever reported for a TD diffuse optics system, thus allowing us to easily reach the desired CR of 30 Mcps with many of the phantom tests described in the following and during in-vivo measurements despite the use of a large SDS (3 cm).

The IRF at different CRs is reported in Fig. 2. More in detail, Fig. 2(a) reports the raw curves acquired, while Fig. 2(b) shows the same curves after pile-up correction and background subtraction.
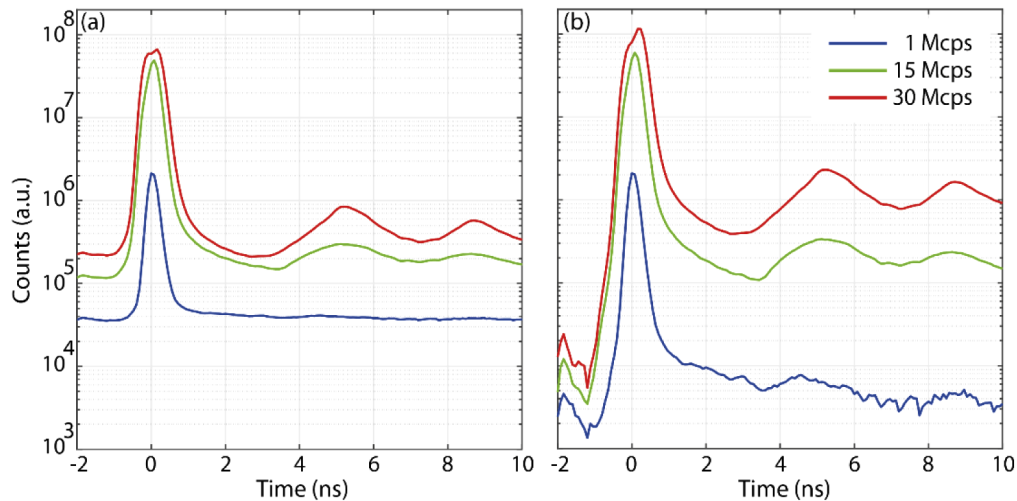

Fig. 2. (a) IRFs recorded at different CRs (colors) without any post-processing. (b) Same IRFs after pile-up correction and background subtraction.

Looking at the IRFs before pile-up correction and background subtraction (Fig. 2(a)), one could notice the expected peak increase with the CR. At the same time, also the background level raises, but with a lower rate with respect to the signal, which is potentially due to the afterpulsing of the detector [64]. Another feature to be underlined is the presence of the bumps mainly visible for 15 and 30 Mcps, which occur after about 4 ns and 8 ns from the peak. As discussed in pile-up section, they are due to the dead-time of the detector and the effect is in line with what reported by Cominelli et al. in Ref. [54].

As usually done in TD-measurements, to improve the dynamic range (DR) of the measurements, the average value of the background noise is removed. The IRF curves after pile-up correction and background subtraction are reported in Fig. 2(b). From comparison with the previous graph, it is evident the improvement in DR and also the change in the DTOF shape due to the pile-up correction (mostly evident in the 30 Mcps curve around the peak region). In this condition, the FWHM of the IRFs is 308, 356 and 556 ps at 1, 15 and 30 Mcps of nominal count rates, respectively. This change is again most probably due to the avalanche pulse pile-up effect, that is a well-known source of distortions in the IRF shape of SiPMs operated at high CRs [63]. As specified in Sect. 2.3, when the IRF shape is relevant for the analysis (as for MEDPHOT tests), the IRF is taken at the same CR (therefore needing pile-up correction) as that of the phantom measurements in order to have a similar detector response. This however cannot grant an identical behavior of the detector during IRF and phantom measurements since the avalanche pile-up effect is expected to depend on the probability that one photon can arrive soon after a previous photon detection and therefore on the shape of the optical waveform under measurement.

Regarding the temporal stability of the IRF, the results obtained are reported in Fig. 3. The stability of the IRF was evaluated in terms of counts, pulse width (FWHM) and temporal position of the center of gravity after 10 minutes of warm up. For counts and pulse width, data are displayed as a percentage variation with respect to the mean value (computed on the 50 minutes of measurements after warm-up), while for the temporal position we plot the variation (in ps) with respect to the average value.
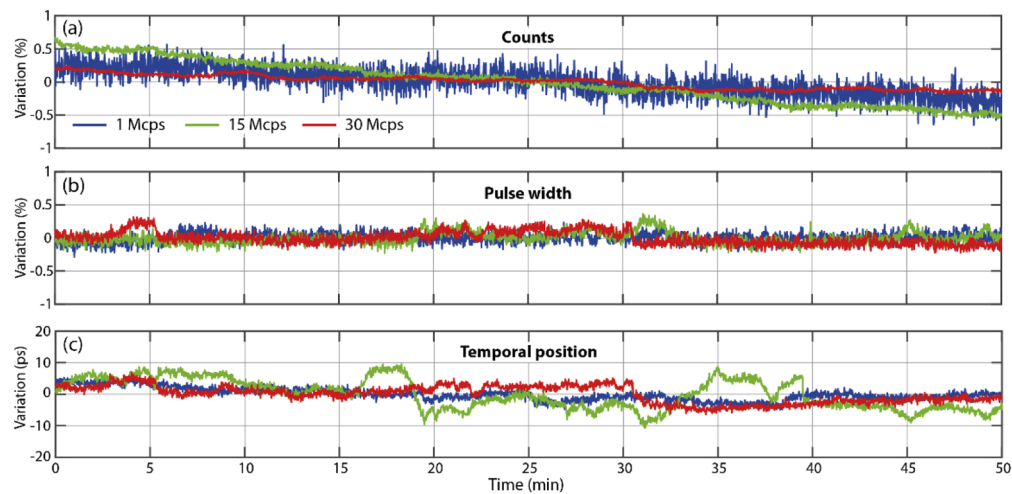

Fig. 3. IRF stability computed on 50 minutes (after warm-up) for the 3 considered count-rates (colors). The first two panels (a,b) represent the percentage variation with respect to their mean values of counts and pulse width (computed as FWHM), respectively. The third panel (c) displays the variation of the temporal position of the curve (center of gravity) in ps with respect to the mean value.

For all parameters considered, we can clearly notice a good system stability with a peak-to-peak variation within the ±0.5% for the counts and pulse width and ±10 ps for the temporal position of the curves. Such results are completely in line with other state-of-the-art instruments for brain imaging [21,56].

3.2. MEDPHOT protocol

Figure 4 reports the results of the MEDPHOT protocol tests. In the first and fourth rows are shown the linearity in the retrieval of absorption and reduced scattering properties respectively for the different count-rates (columns). Table 3 reports the slope and R^2 obtained from linear fitting for the linearity measurements for both not-corrected (NC) and corrected (C) DTOFs at all CRs. We should warn here that the nominal values are those obtained in a multi-laboratory exercise for performance assessment of diffuse optics instruments [48]. This is the best estimate of the true optical parameters of the MEDPHOT kit available so far. It surely permits to properly assess the linearity of the system, while precise estimate of the accuracy (i.e. discrepancy in absolute values) is still questionable since no clear consensus was reached, differently from the characterization of liquid phantoms where a discrepancy <2% was reached in a different multi-laboratory exercise [65].

Table 3. Value of slope and R^2 obtained from the linear fitting of the absorption/scattering linearity measurements for all considered CRs. Values obtained considered both not-corrected and corrected curves (NC and C columns, respectively) are reported.

| Phantom series (cm^{-1}) | Count-rate: 1 Mcps | | | | Count-rate: 15 Mcps | | | | Count-rate: 30 Mcps | | | |
| | Slope | | R^2 | | Slope | | R^2 | | Slope | | R^2 | |
	NC	C	NC	C	NC	C	NC	C	NC	C	NC	C
absorption linearity												
μ_s' = 7.3	0.93	0.93	0.9939	0.9940	0.79	0.74	0.9895	0.9886	0.91	0.76	0.9872	0.9979
μ_s' = 11.9	1.09	1.09	0.9956	0.9956	0.98	0.96	0.9973	0.9976	0.99	0.85	0.9916	0.9981
μ_s' = 16	0.85	0.85	0.9997	0.9997	0.87	0.83	0.9973	0.9964	1.09	0.94	0.9999	0.9959
scattering linearity												
μ_a = 0.07	1.04	1.04	0.9289	0.9290	1.06	1.05	0.9396	0.9412	1.11	1.06	0.9632	0.9505
μ_a = 0.14	0.96	0.96	0.8301	0.8311	1.02	1.03	0.8979	0.8949	1.12	1.06	0.9672	0.9397
μ_a = 0.21	0.92	0.92	0.7377	0.7387	1.00	1.00	0.8651	0.8649	1.07	0.98	0.9582	0.9417
μ_a = 0.28	0.81	0.81	0.6338	0.6351	0.92	0.91	0.7981	0.7787	1.10	0.94	0.9611	0.8963

For absorption linearity of not-corrected curves, there is in all cases (scattering series and CR) a good linearity (R^2 always larger than 0.987) while in most cases (for 1 and 15 Mcps) there is a general underestimation of the recovered μ_a (slope <1). For CR of 30 Mcps, an improvement in the linearity (both in terms of recovered value – i.e. slope value close to 1 – and R^2) is present for all scattering series. For what concerns the corrected curves, it is possible to notice that for the 1 Mcps there is no effect in the retrieved absorption value.

This was expected since at 1 Mcps the system is working within single-photon statistics, thus there is no relevant pile-up effect to be corrected. On the other hand, for measurements taken at 15 and 30 Mcps, the Coates correction seems to further underestimate the retrieved μ_a, without improving the linearity (R^2 is constant or even slightly lower). We cannot provide a definitive explanation for this effect. It might be due to the fit procedure which weighs more the first part of the tail, where the pile-up is affecting less the DTOF. Indeed, we adopted the reduced χ^2 as penalty term in the fitting procedure assuming shot-noise for data. Yet, whenever the model is not completely adequate to describe the data, the points with higher number of counts gain more weight with respect to the pure shot-noise case.
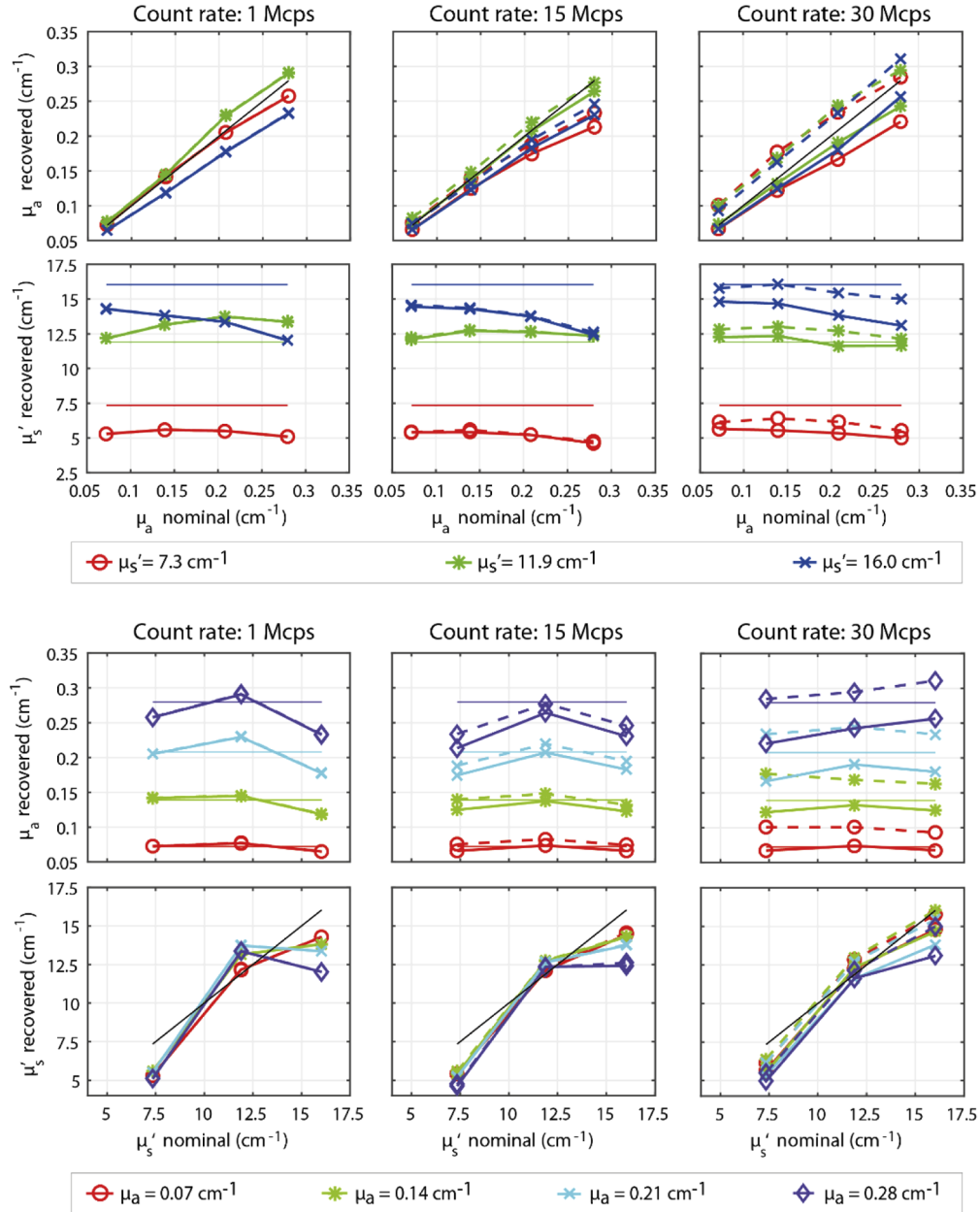
Fig. 4. Results of the linearity measurements for the three sets of count-rate (columns). In all cases, both results obtained with and without the pile-up correction (solid and dashed lines, respectively) are shown. In all plots, the thin solid lines represent the expected values.

At any rate, at 30 Mcps the underestimation of both the absorption and scattering coefficients obtained after Coates correction allows one to improve the accuracy in the μ_a retrieval and also to get more uniform values of μ_s' as compared to what obtained with the other CRs (see plot for absorption linearity at 30 Mcps and scattering to absorption coupling graphs in Fig. 4).

On the other hand, results obtained in scattering linearity show the weaknesses of the 9 mm^2 area system in the retrieval of μ_s', mostly for low CRs. Indeed, for 1 Mcps the recovered μ_s' are not much linear and this trend is more evident for phantom series with larger absorption (R^2 going from 0.9289 with $\mu_a = 0.07$ cm^{-1} to 0.6338 for $\mu_a = 0.28$ cm^{-1}). This might be due to the large amount of background noise which reduces the DR for fitting, thus impairing the retrieval of the optical properties, especially for the scattering coefficient which is more prone to artefacts related to non-idealities of the system. Yet, we have no proof this is the key reason. The increase in the CR improves the linearity of the μ_s' retrieval both in terms of slope (close to 1) and R^2. For 15 Mcps measurements, the effect of Coates correction is almost negligible, while for 30 Mcps ones the correction compensates the slight overestimation of the μ_s' but at the expenses of the R^2.

Regarding the absorption to scattering (scattering to absorption) coupling, conclusions can be drawn looking at the second (third) row of Fig. 4. It is evident that for the most scattering phantoms there is a strong coupling between the true μ_a and the retrieved μ_s' of the phantom for all CRs (second row). For phantom series featuring μ_s' = 7.3 and 11.9 cm^{-1}, this effect is less evident. On the other hand, the coupling between retrieved the true μ_s' and the retrieved μ_a (third row) is less pronounced.

In all cases, in the fitting procedure the detector was considered as point-like one. Indeed, the use of the correct value of the detector area showed no improvement in the retrieval of the optical properties (data not shown) while the computational complexity increased.

3.3. nEUROPt protocol

Figure 5 reports the computed contrast (top row) and CNR (bottom row) as a function of the perturbation depth obtained for different CRs (columns) for an early and late gate (from 0.72 to 1.2 and from 2.64 to 3.12 ns with respect to IRF peak, respectively).

Overall, for 15 and 30 Mcps the effect of the Coates correction is evident while, as expected, for 1 Mcps is negligible. Generally speaking, without correction, pile-up effect is expected to artificially reduce the number of detected photons in the DTOF. This decrease is typically larger for homogeneous-case DTOFs as heterogeneous-case ones are often characterized by a lower count rate due to the presence of the absorption perturbation. By analyzing the effect of this phenomenon in Eq. (4), one can notice that pile-up distortion will result into a reduction of the contrast. Indeed, this is clearly visible in Fig. 5 for the contrast computed at early gate since, upon increasing the CR, the amplitude of the contrast function decreases. After pile-up correction, it recovers the same value as expected. Actually, as also reported in [66], the contrast should not depend on the number of detected photons. On the other hand, without correction the CNR for the early gate (slightly) increases for 15 Mcps while, due to pile-up, it decreases for the 30 Mcps. The expected trend, that is an increase in the CNR proportional to the square root of the number of counts (see [66]), is recovered upon applying the Coates correction.

The situation is more complicated for the late gate (i.e. 2.64 to 3.12 ns from IRF peak*)*. Indeed, for 15 Mcps the contrast compression is clearly visible, while at 30 Mcps the not-corrected contrast shows an anomalous profile with an undershoot whose peak is below -0.1 (precisely -0.18 at 14 mm, not shown). To better understand this behavior, Fig. 6 shows the homogeneous and heterogeneous curves (orange and blue curves, respectively) acquired at 30 Mcps for three significant depths (14, 20 and 24 mm) both before and after applying the Coates correction (top and bottom row, respectively). The temporal position of the late gate is highlighted with the gray rectangle while the inset shows a zoom of curves in the gray regions. For not-corrected curves, due to pile-up effect, the DTOF of the homogeneous case presents a steeper decay with respect to
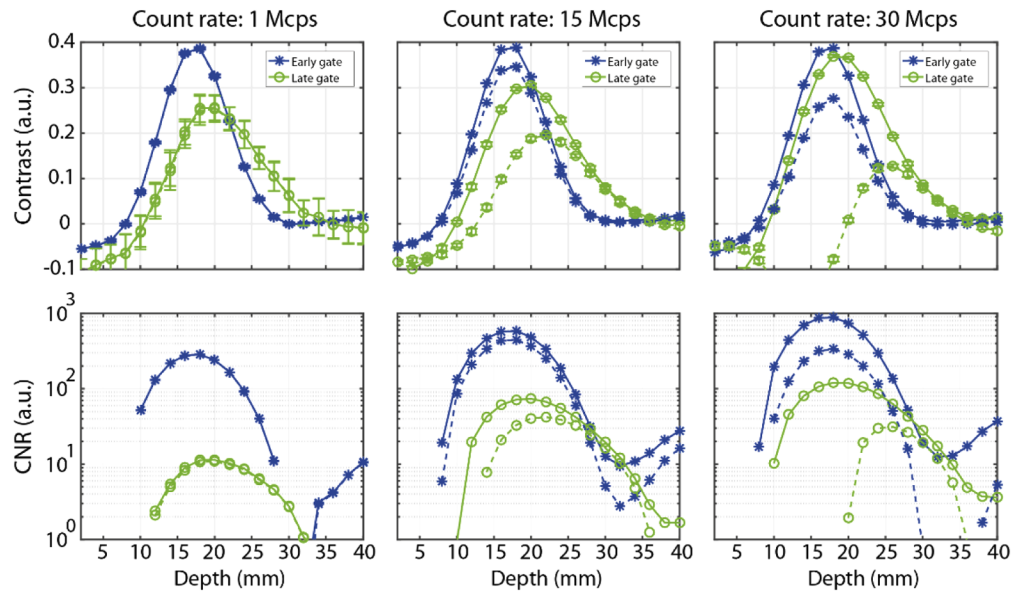
Fig. 5. Contrast (top row) and CNR (bottom row) computed for all count rates (columns) at an early (blue stars, 0.72 to 1.20 ns from IRF peak) and late (green circles, 2.64 to 3.12 ns from IRF peak) gate. Both results with (solid line) and without (dashed line) Coates correction are shown. Error bars are reported for all contrast measurements, even if in some cases are not noticeable.

the heterogeneous one. This is expected as in the heterogeneous case the number of detected photons decreases due to the presence of the absorbing perturbation, thus reducing the effect of the pile-up. For this reason, at 14 mm depth, the late gate for the homogeneous case is lower than the heterogeneous one, thus resulting in a negative contrast. This does not happen at early times as the crossing of the curves is at about 2 ns. When the perturbation is deeper, its effect on the DTOF is weaker, thus reducing the difference in terms of pile-up distortion between the homogeneous and heterogeneous cases and shifting the crossing point toward later times. For this reason, at a depth of 20 mm, the contrast computed in the late gate is almost null, while for the depth of 24 mm, it goes back to the positive (expected) value. When the Coates algorithm is applied (second row of Fig. 6), the corrected curves in the homogeneous case are always larger than the heterogeneous ones, thus leading to a positive contrast, which is more marked at 20 mm (i.e. separation between curves in the homogeneous and heterogeneous case is evidently larger).

Considering the contrast in the late gate for corrected DTOF, from Fig. 5 it is evident that the peak value increases along with the count-rate, which is not in agreement with theoretical expectations in ideal conditions (i.e. when the IRF shape does not change with the CR) [66]. Such an effect can be due to the avalanche pulse pile-up effect, which produces a change in the IRF shape when increasing the CR. Indeed, as it was simulated in Ref. [66], larger FWHM of the IRF does not improve the overall contrast, but leads to a shift of the contrast as a function of the photons arrival time towards later times [67]. Therefore, for certain delays of the late gate chosen for computing the contrast, it is possible to observe an improvement of the contrast when increasing the IRF FWHM.

For what concerns the CNR for late gate, the increase in the CR is beneficial, provided that the pile-up correction is applied. Considering the corrected curves, the improvement of the CNR with respect to the standard case (i.e. 1 Mcps) is of 546% (105%) and 940% (210%) for 15 and 30 Mcps respectively at the late (early) gate.
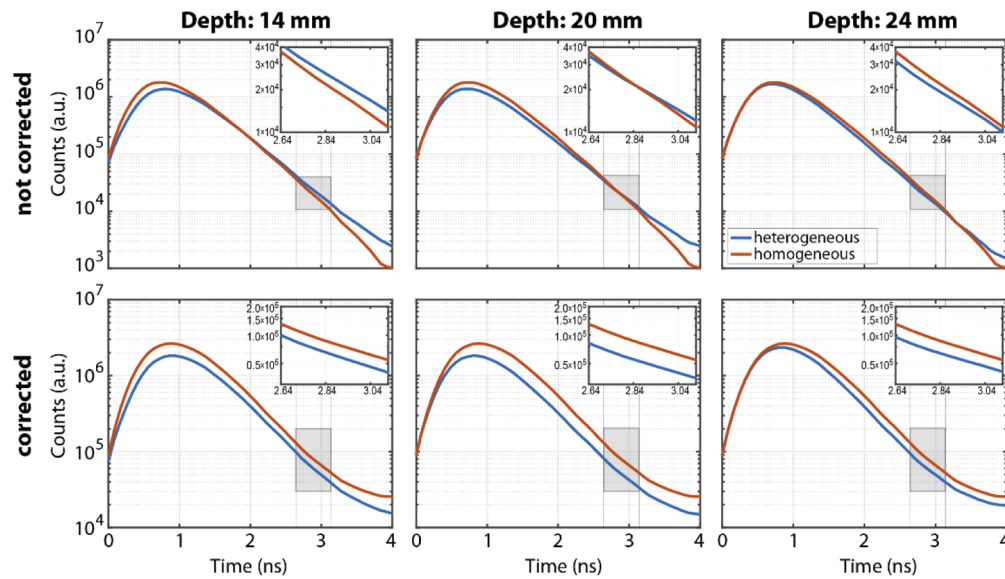
Fig. 6. DTOFs acquired at 30 Mcps at different depths of the inclusions (columns) both without and with Coates correction (first and second column, respectively). In both cases, the background was subtracted. In each graph, both the DTOFs acquired in the homogeneous and heterogeneous cases are displayed (orange and blue, respectively), while the temporal region corresponding to the late gate (2.64 to 3.12 ns from IRF peak) is enlighten in gray. The inset reports the portion of the DTOF in the late gate considered.

Considering typical detectability criteria (see Sect. 2.3), Fig. 5 shows that the perturbation can be detected down to a depth of about 30 mm at standard count rate and of about 35 mm at 15 and 30 Mcps. In these standard optical conditions, this penetration is similar to the maximum ever reported [27]. However, the investigation of the maximum achievable depth of investigation is beyond the scope of this paper as the phantom presents a weak heterogeneity in its structure [59], limiting the performance in this kind of investigation. This is for instance the reason for the negative contrast reported in Fig. 5 at shallow depths, a characteristic always detected when it has been previously used for the implementation of the nEUROPt protocol (see e.g. [24]).

In order to limit the additional distortion due to the avalanche pulse pile-up effect, for *in-vivo* measurements we decided to use a count-rate for each wavelength of 15 Mcps. For this reason, Fig. 7 reports a more comprehensive view of the contrast (top row) and CNR (bottom one) obtained for the depth scan at 15 Mcps after applying the Coates algorithm for pile-up correction. Here, contrast and CNR have been computed considering five different time windows. It is possible to better appreciate that upon increasing the delay of the gate used for analysis, the peak of the contrast is shifted toward deeper positions and the contrast has a larger value if considering later gates. This is due to the fact that the analyzed photons are progressively remitted later, thus they statistically explored a deeper region of the phantom.

For the same reason, the shift toward deeper positions can be seen also for the CNR. In this case, however, due to the reduced number of photons, the value decreases, but for the latest gate (2.64-3.12 ns) it is still about 12 (i.e. well above the value of 1 usually considered as the lower limit for detectability, together with contrast > 1% [11]) even for depth of 32 mm (where contrast is 5%).
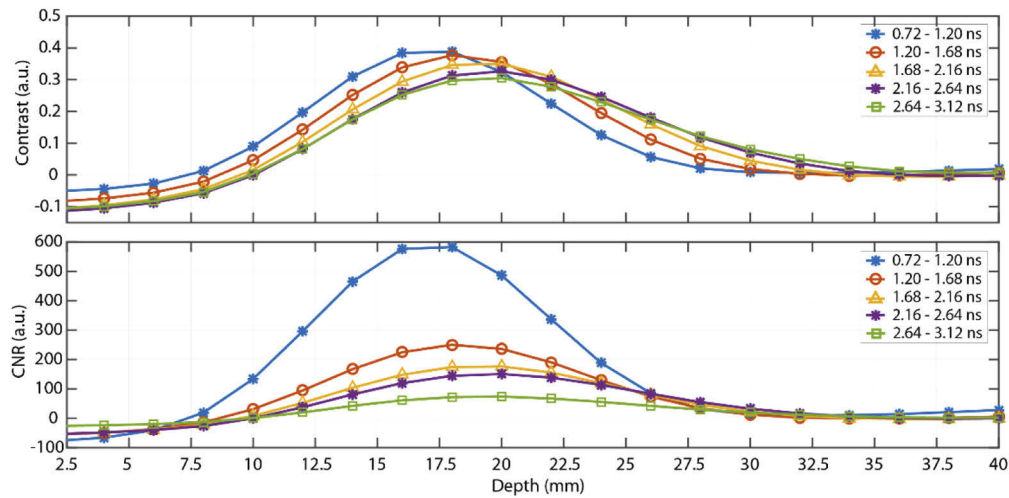
Fig. 7. Contrast (top) and CNR (bottom) as a function of the perturbation depth obtained after applying Coates correction for different time-gates (colors and symbols) for curves acquired at 15 Mcps. Errorbars in the contrast are not displayed since their value is almost negligible.

3.4. In-vivo measurements

Figure 8 reports the contrast at different time-gates for the two wavelengths (colors), for both the 1 mm^2 area fiber-based SiPM and for the 9 mm^2 area in-contact SiPM, during the finger tapping experiment performed by Subject 1. Each gate has a width of 0.5 ns and it is set at delays increasing from -0.5 to 3.5 ns with respect to the IRF peak (0 ns time).

As expected, for both detectors we observe an increased contrast for late time gates with an almost comparable amplitude. The visible difference between the two detectors is the noise level that is appreciably higher in the 1 mm^2 area fiber-based SiPM. Indeed, the noise level is dependent on the total number of collected photons. Table 4 shows the average number (N), standard deviation ($\sigma(N)$) over the five repetitions, and coefficient of variation (*real CV* $= \sigma(N)/N$, *Poisson CV* $= (N)^{-0.5}$) of photons in the baseline collected by the 1 mm^2 fiber-based SiPM and by the 9 mm^2 in-contact SiPM during the *in-vivo* motor task for the two wavelengths and three subjects. Data refers to the folding average of 5 repetitions. The ability of the 9 mm^2 in-contact SiPM to collect a larger number of photons is clear. Actually, the larger area is not sufficient to explain the difference between the two detectors. An additional feature to be taken into account is the fact that the 1 mm^2 SiPM is fiber based (1 mm core diameter step index silica optical fiber, NA = 0.39), while the 9 mm^2 SiPM is directly in contact with the tissue: indeed, the limited NA plays a key role in determining the collected signal. Regarding the CV, we notice that the real CV over the five repetitions is comparable for the two detectors (being influenced, as expected, by the physiological variations), while the ideal Poisson CV (being related only to the number of collected photons) is of course much less for the 9 mm^2 SiPM, therefore helping discriminability of activation.

Finally, to present the results of the in vivo test in a way that is directly understandable by researchers in the field of functional near infrared spectroscopy, Fig. 9 shows the time course of O$_2$Hb and HHb variation (ΔO$_2$Hb and ΔHHb) during the tapping experiment for the three subjects. For Subject 1 the canonical hemodynamic response (increase in O$_2$Hb and decrease of HHb) is clearly visible both with the 1 mm^2 fiber-based SiPM and the 9 mm^2 in-contact SiPM, the former being definitely much noisier. Conversely, for Subject 3 the activation is only
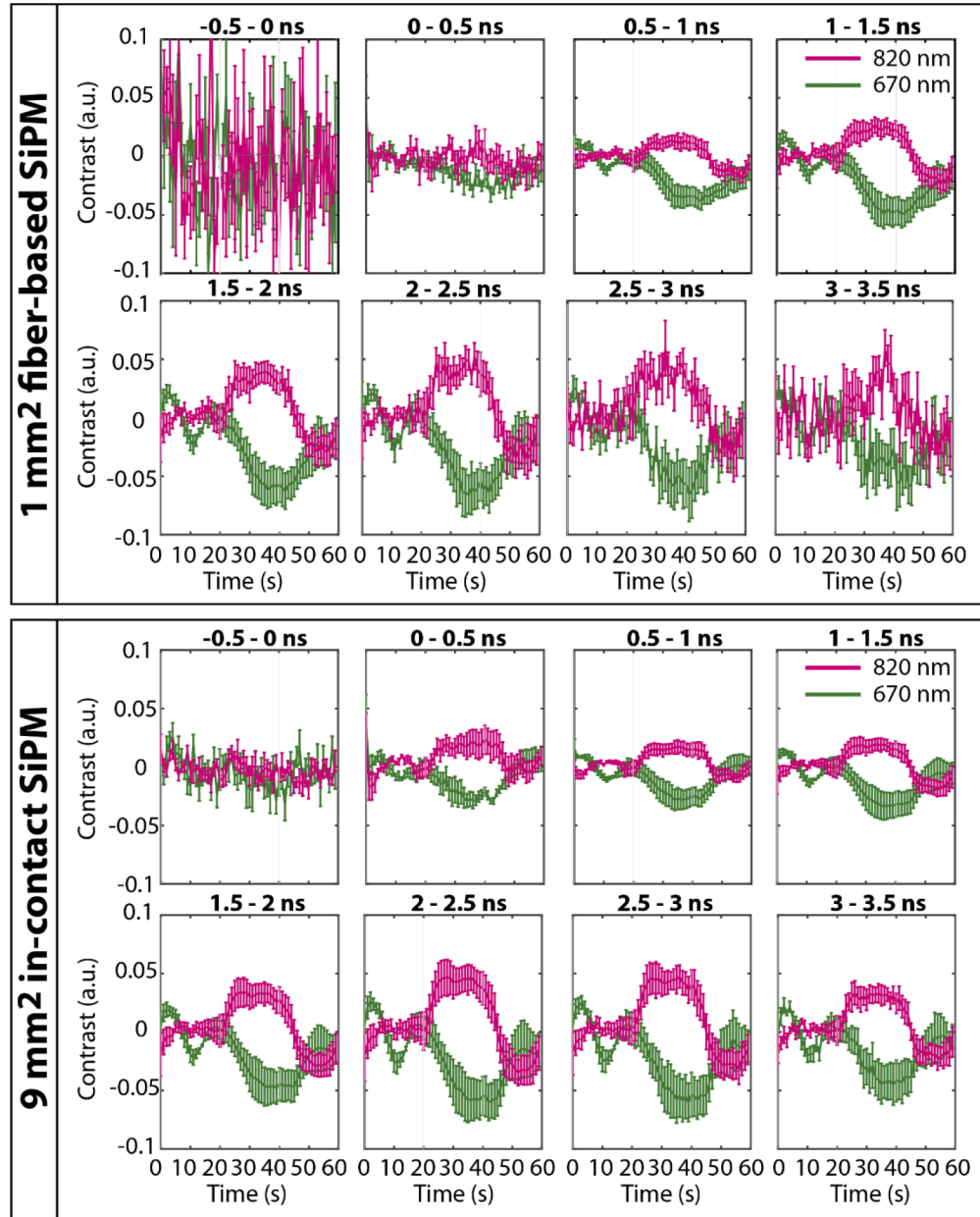
Fig. 8. Contrast obtained for the two wavelengths (colors) in the different gates (gate width: 500 ps, position referred to the IRF peak). Both contrast obtained for 1 mm² fiber-based SiPM and 9 mm² in-contact one are shown (top and bottom panel, respectively). The error bars are computed over the 5 repetitions of the task.

Table 4. Average number (N), standard deviation (σ(N)), and coefficient of variation (real CV = σ(N)/N, Poisson CV = (N)$^{-0.5}$) of photons in the baseline collected by the 1 mm^2 fiber-based SiPM and by the 9 mm^2 in-contact SiPM during the in vivo motor task for the two wavelengths and three subjects. Data refers to the folding average of 5 repetitions in the right hand tapping experiment. Similar results were obtained for the left hand tapping.

Wavelength (nm)	Subject #	SiPM area	N (ph)	$\sqrt{N}$ (ph)	*CV Poisson* (%)	$\sigma(N)$ (ph)	*CV real* (%)
670	1	1 mm^2	607,120	779	0.13%	8,052	1.3%
		9 mm^2	17,502,078	4,184	0.02%	235,977	1.3%
	2	1 mm^2	845,226	919	0.11%	3,097	0.4%
		9 mm^2	17,466,097	4,179	0.02%	79,341	0.5%
	3	1 mm^2	219,309	468	0.21%	6,773	3.1%
		9 mm^2	3,333,183	1,826	0.05%	27,845	0.8%
830	1	1 mm^2	241,023	491	0.20%	3,760	1.6%
		9 mm^2	14,839,034	3,852	0.03%	185,285	1.2%
	2	1 mm^2	264,018	514	0.19%	3,016	1.1%
		9 mm^2	14,132,310	3,759	0.03%	176,994	1.3%
	3	1 mm^2	185,522	431	0.23%	3,565	1.9%
		9 mm^2	13,616,515	3,690	0.03%	310,597	2.3%

distinguishable with the 9 mm^2 SiPM, while Subject 2 is an intermediate case with the activation visible for both systems but noisier than for Subject 1. For all subjects, considering the left hand, there is no evidence of task-related variation of O$_2$Hb and HHb, thus confirming the lateralization of the activation due to the exercise.

4.   Conclusions and future perspectives

In this paper we presented a new detection chain for TD diffuse optics, which features a large sensitive area (9 mm^2 with 82% fill-factor, resulting in an overall effective active area of 7.38 mm^2), coupled to a high-throughput timing electronics (180 Mcps), thus greatly improving the light harvesting capability (i.e. responsivity) with respect to the state-of-the-art. To the best of our knowledge, this is the first system which addresses together both technological bottlenecks of the current TD instruments (i.e. limited light harvesting and throughput). Making use of internationally shared protocols for performances assessment, we validated the use of this system for contrast measurements showing an improvement in the achievable CNR. On the other hand, the system showed limited performances for spectroscopy applications due to unsatisfactory retrieval of scattering properties. However, for many applications (e.g. detection of brain activation) the scattering coefficient, differently from the absorption coefficient, is not expected to vary during the tests and –if needed– a characterization of the static baseline optical properties can be easily performed before the test with a proper spectroscopic instruments or a dedicated instrument channel.

The improved performances of the system during a functional test were demonstrated with *in-vivo* motor cortex measurements, where the robustness of the signal recorded with the 9 mm^2 in-contact SiPM is greatly enhanced with respect to a system using a fiber-based 1 mm^2 SiPM.

The system can further evolve in the near future. The current version of the system is susceptible to ambient electromagnetic interferences (in particular from pulsed voltage signals in the laser heads, which can slightly modulate the background noise if the laser heads are too close to the detector, limiting the possibility of miniaturization in the present implementation)
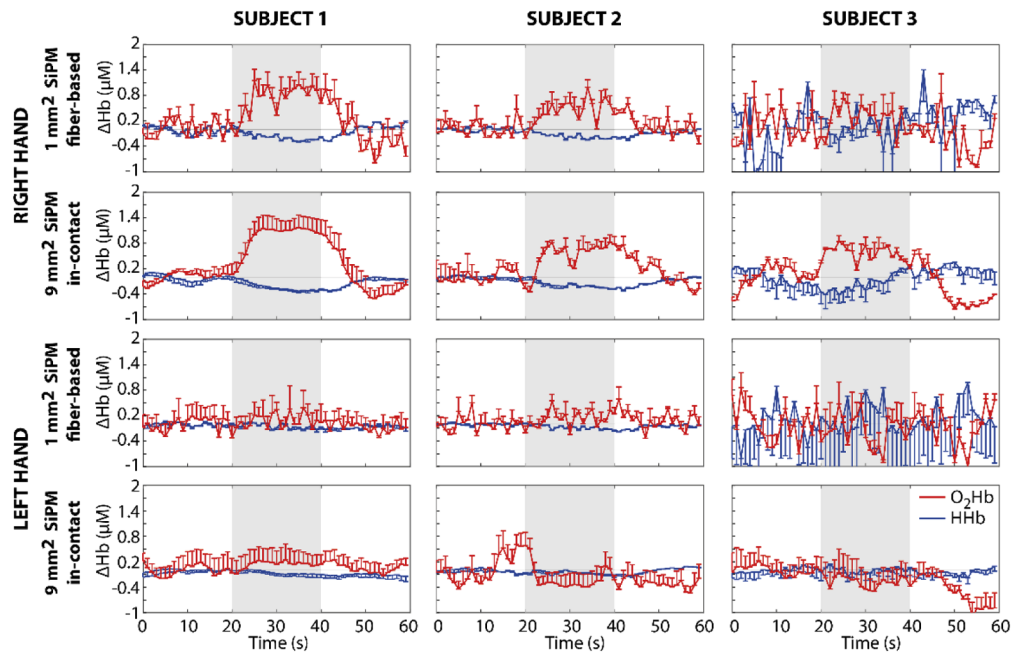
Fig. 9. Time courses of $\Delta O_2 Hb$ and ΔHHb measured over the left motor cortex for 3 subjects (columns) during finger tapping exercise done with right and left hand (first two and last two rows, respectively). Both results obtained with 1 mm^2 fiber-based SiPM and 9 mm^2 in-contact SiPM (odd and even rows, respectively) are displayed. The error bars are computed over the 5 repetitions of the task.

and its behavior is affected by the detector dead time as well as its DCR. For all those issues, an improvement in the detector readout electronics as well as a cooling down of the device will permit to push the performances even further. Regarding the post-processing analysis of data, more advanced algorithms for pile-up correction may be used, even if in some cases they require the acquisition of data in time-tagging mode [68].

As future perspectives, the proposed single channel device can be easily extended to 8 channels (the TCSPC board can in fact host up to 8 inputs), thus allowing fast acquisitions with higher count-rate with respect to state-of-the-art TD probe-hosted tomographic system as, for example, the one presented in [69].

We envisage that the adoption of larger area detectors – possibly with also time-gated capability – will dramatically improve the performances of the system, approaching the theoretical limits. This will result in the development of a new generation of devices which may open the way to several unexplored applications as for instance the exploration of the human head and chest in transmittance geometry.

Funding

Acknowledgments

Disclosures

Sumeet Rohilla conducted his PhD at PicoQuant Innovations GmbH, financed by a joint European funding project (BitMap). Benedikt Krämer and Felix Koberling are employees of PicoQuant GmbH. PicoQuant produces time-resolved fluorescence microscopes and spectrometers, picosecond pulsed diode lasers as well as time correlated single photon counting equipment.

References

1. A. A. P. Pinto Lima and D. De Backer, "Optical Monitoring," in *Monitoring Tissue Perfusion in Shock: From Physiology to the Bedside*, A. A. Pinto Lima and E. Silva, eds. (Springer International Publishing, 2018), pp. 153–171.
2. Y. Hoshi and Y. Yamada, "Overview of diffuse optical tomography and its clinical applications," J. Biomed. Opt. **21**(9), 091312–11 (2016).
3. T. Durduran, R. Choe, W. B. Baker, and A. G. Yodh, "Diffuse optics for tissue monitoring and tomography," Rep. Prog. Phys. **73**(7), 076701 (2010).
4. R. Srivastava, "Tissue Optics and Other Optical Image Technology," in *Atlas on Narrow Band Imaging in Upper Aerodigestive Tract Lesions* (Springer Singapore, 2019), pp. 3–5.
5. S. Mosca, P. Dey, T. A. Tabish, F. Palombo, N. Stone, and P. Matousek, "Determination of inclusion depth in ex vivo animal tissues using surface enhanced deep Raman spectroscopy," J. Biophotonics **13**(1), e201960092 (2020).
6. B. Chance, *Photon Migration in Tissues* (Plenum Press, 1989).
7. A. Gibson and H. Dehghani, "Diffuse optical imaging," Philos. Trans. R. Soc., A **367**(1900), 3055–3072 (2009).
8. F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. M. Pavia, U. Wolf, and M. Wolf, "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," NeuroImage **85**, 6–27 (2014).
9. D. Canova, S. Roatta, D. Bosone, and G. Micieli, "Inconsistent detection of changes in cerebral blood volume by near infrared spectroscopy in standard clinical tests," J. Appl. Physiol. **110**(6), 1646–1655 (2011).
10. A. I. R. Maas and G. Citerio, "Noninvasive monitoring of cerebral oxygenation in traumatic brain injury: A mix of doubts and hope," Intensive Care Med. **36**(8), 1283–1285 (2010).
11. A. Dalla Mora, D. Contini, S. Arridge, F. Martelli, A. Tosi, G. Boso, A. Farina, T. Durduran, E. Martinenghi, A. Torricelli, and A. Pifferi, "Towards next-generation time-domain diffuse optics for extreme depth penetration and sensitivity," Biomed. Opt. Express **6**(5), 1749–1760 (2015).
12. M. Firbank, C. E. Elwell, C. E. Cooper, and D. T. Delpy, "Experimental and theoretical comparison of NIR spectroscopy measurements of cerebral hemoglobin changes," J. Appl. Physiol. **85**(5), 1915–1921 (1998).
13. T. Funane, H. Atsumori, T. Katura, A. N. Obata, H. Sato, Y. Tanikawa, E. Okada, and M. Kiguchi, "Quantitative evaluation of deep and shallow tissue layers' contribution to fNIRS signal using multi-distance optodes and independent component analysis," NeuroImage **85**, 150–165 (2014).
14. S. Vasudevan, F. Forghani, C. Campbell, S. Bedford, and T. D. O'Sullivan, "Method for Quantitative Broadband Diffuse Optical Spectroscopy of Tumor-Like Inclusions," Appl. Sci. **10**(4), 1419 (2020).
15. A. T. Eggebrecht, S. L. Ferradal, A. Robichaux-Viehoever, M. S. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. P. Culver, "Mapping distributed brain function and networks with diffuse optical tomography," Nat. Photonics **8**(6), 448–454 (2014).
16. A. Pifferi, D. Contini, A. Dalla Mora, A. Farina, L. Spinelli, and A. Torricelli, "New frontiers in time-domain diffuse optics, a review," J. Biomed. Opt. **21**(9), 091310 (2016).
17. A. Pifferi, A. Torricelli, L. Spinelli, D. Contini, R. Cubeddu, F. Martelli, G. Zaccanti, A. Tosi, A. Dalla Mora, F. Zappa, and S. Cova, "Time-Resolved Diffuse Reflectance Using Small Source-Detector Separation and Fast Single-Photon Gating," Phys. Rev. Lett. **100**(13), 138101 (2008).

18. R. Re, D. Contini, M. Caffini, R. Cubeddu, L. Spinelli, and A. Torricelli, "A compact time-resolved system for near infrared spectroscopy based on wavelength space multiplexing," Rev. Sci. Instrum. **81**(11), 113101 (2010).

19. D. Milej, A. Gerega, M. Kacprzak, P. Sawosz, W. Weigl, R. Maniewski, and A. Liebert, "Time-resolved multi-channel optical system for assessment of brain oxygenation and perfusion by monitoring of diffuse reflectance and fluorescence," Opto-Electron. Rev. **22**(1), 55–67 (2014).

20. R. J. Cooper, E. Magee, N. Everdell, S. Magazov, M. Varela, D. Airantzis, A. P. Gibson, and J. C. Hebden, "MONSTIR II: A 32-channel, multispectral, time-resolved optical tomography system for neonatal brain imaging," Rev. Sci. Instrum. **85**(5), 053105 (2014).

21. F. Lange, L. Dunne, L. Hale, and I. Tachtsidis, "MAESTROS: A Multiwavelength Time-Domain NIRS System to Monitor Changes in Oxygenation and Oxidation State of Cytochrome-C-Oxidase," IEEE J. Sel. Top. Quantum Electron. **25**(1), 1–12 (2019).

22. S. Konugolu Venkata Sekar, A. Farina, E. Martinenghi, A. Dalla Mora, P. Taroni, A. Pifferi, E. Negredo, J. Puig, R. Escrig, Q. Rosales, C. Lindner, M. Pagliazzi, and T. Durduran, "Time-resolved diffused optical characterization of key tissue constituents of human bony prominence locations," in *Diffuse Optical Imaging V* (Optical Society of America, 2016), p. 95380X.

23. L. Di Sieno, J. Nissinen, L. Hallman, E. Martinenghi, D. Contini, A. Pifferi, J. Kostamovaara, and A. Dalla Mora, "Miniaturized pulsed laser source for time-domain diffuse optics routes to wearable devices," J. Biomed. Opt. **22**(08), 1 (2017).

24. R. Re, E. Martinenghi, A. Dalla Mora, D. Contini, A. Pifferi, and A. Torricelli, "Probe-hosted silicon photomultipliers for time-domain functional near-infrared spectroscopy: phantom and in vivo tests," Neurophotonics **3**(4), 045004 (2016).

25. S. Saha, Y. Lu, S. Weyers, M. Sawan, and F. Lesage, "Compact Fast Optode-Based Probe for Single-Photon Counting Applications," IEEE Photonics Technol. Lett. **30**(17), 1515–1518 (2018).

26. M. Renna, M. Buttafava, A. Behera, M. Zanoletti, L. Di Sieno, A. Dalla Mora, D. Contini, and A. Tosi, "Eight-wavelength, dual detection-channel instrument for near-infrared time-resolved diffuse optical spectroscopy," IEEE J. Sel. Top. Quantum Electron. **25**(1), 1–11 (2019).

27. A. Dalla Mora, D. Contini, S. R. Arridge, F. Martelli, A. Tosi, G. Boso, A. Farina, T. Durduran, E. Martinenghi, A. Torricelli, and A. Pifferi, "Towards dense and wearable time-domain diffuse optical devices with ultimate contrast and depth sensitivity," Biomed. Opt. Express (2015).

28. S. Cova, M. Ghioni, A. Lacaita, C. Samori, and F. Zappa, "Avalanche photodiodes and quenching circuits for single-photon detection," Appl. Opt. **35**(12), 1956–1976 (1996).

29. R. H. Hadfield, "Single-photon detectors for optical quantum information applications," Nat. Photonics **3**(12), 696–705 (2009).

30. A. Dalla Mora, E. Martinenghi, D. Contini, A. Tosi, G. Boso, T. Durduran, S. Arridge, F. Martelli, A. Farina, A. Torricelli, and A. Pifferi, "Fast silicon photomultiplier improves signal harvesting and reduces complexity in time-domain diffuse optics," Opt. Express **23**(11), 13937–13946 (2015).

31. E. Martinenghi, L. Di Sieno, D. Contini, M. Sanzaro, A. Pifferi, and A. Dalla Mora, "Time-resolved single-photon detection module based on silicon photomultiplier: A novel building block for time-correlated measurement systems," Rev. Sci. Instrum. **87**(7), 073101 (2016).

32. A. Farina, S. Tagliabue, L. Di Sieno, E. Martinenghi, T. Durduran, S. Arridge, F. Martelli, A. Torricelli, A. Pifferi, and A. Dalla Mora, "Time-Domain Functional Diffuse Optical Tomography System Based on Fiber-Free Silicon Photomultipliers," Appl. Sci. **7**(12), 1235 (2017).

33. S. Saha, Y. Lu, F. Lesage, and M. Sawan, "Wearable SiPM-Based NIRS Interface Integrated With Pulsed Laser Source," IEEE Trans. Biomed. Circuits Syst. **13**(6), 1313–1323 (2019).

34. L. Di Sieno, J. Zouaoui, L. Hervé, A. Pifferi, A. Farina, E. Martinenghi, J. Derouard, J.-M. Dinten, and A. Dalla Mora, "Time-domain diffuse optical tomography using silicon photomultipliers: a feasibility study," J. Biomed. Opt. **21**(11), 116002 (2016).

35. W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques* (Springer International Publishing, 2015), Vol. 111.

36. Becker & Hickl GmbH, "TCSPC Modules," https://www.becker-hickl.com/products/category/tcspc-devices/.

37. Picoquant GmbH, "TCSPC and time tagging electronics," http://www.picoquant.com/products/category/tcspc-and-time-tagging-modules.

38. C. Niclass, C. Favi, T. Kluter, M. Gersbach, and E. Charbon, "A 128 × 128 single-photon image sensor with column-level 10-bit time-to-digital converter array," IEEE J. Solid-State Circuits **43**(12), 2977–2989 (2008).

39. C. Niclass, M. Soga, H. Matsubara, and S. Kato, "A 100m-range 10-frame/s 340 × 96-pixel time-of-flight depth sensor in 0.18μm CMOS," Eur. Solid-State Circuits Conf. 107–110 (2011).

40. E. Ferocino, E. Martinenghi, A. Dalla Mora, A. Pifferi, R. Cubeddu, and P. Taroni, "High throughput detection chain for time domain optical mammography," Biomed. Opt. Express **9**(2), 755–770 (2018).

41. G. Boso, A. Tosi, A. Dalla Mora, and F. Zappa, "High-throughput gated photon counter with two detection windows programmable down to 70 ps width," Rev. Sci. Instrum. **85**(1), 013107 (2014).

42. L. Di Sieno, A. Dalla Mora, G. Boso, A. Tosi, A. Pifferi, R. Cubeddu, and D. Contini, "Diffuse optics using a dual window fast-gated counter," Appl. Opt. **53**(31), 7394–7401 (2014).

43. D. Tyndall, B. Rae, D. Li, J. Richardson, J. Arlt, and R. Henderson, "A 100Mphoton/s time-resolved mini-silicon photomultiplier with on-chip fluorescence lifetime estimation in 0.13μm CMOS imaging technology," Dig. Tech. Pap. - IEEE Int. Solid-State Circuits Conf. **55**, 122–123 (2012).

44. N. A. W. Dutton, S. Gnecchi, L. Parmesan, A. J. Holmes, B. Rae, L. A. Grant, and R. K. Henderson, "A time-correlated single-photon-counting sensor with 14GS/S histogramming time-to-digital converter," Dig. Tech. Pap. - IEEE Int. Solid-State Circuits Conf. **58**, 204–205 (2015).

45. N. Krstajić, S. Poland, J. Levitt, R. Walker, A. Erdogan, S. Ameer-Beg, and R. K. Henderson, "05 billion events per second time correlated single photon counting using CMOS SPAD arrays," Opt. Lett. **40**(18), 4305 (2015).

46. S. P. Poland, N. Krstajić, S. Coelho, D. Tyndall, R. J. Walker, V. Devauges, P. E. Morton, N. S. Nicholas, J. Richardson, D. D.-U. Li, K. Suhling, C. M. Wells, M. Parsons, R. K. Henderson, and S. M. Ameer-Beg, "Time-resolved multifocal multiphoton microscope for high speed FRET imaging in vivo," Opt. Lett. **39**(20), 6013 (2014).

47. H. Wabnitz, D. R. Taubert, M. Mazurenka, O. Steinkellner, A. Jelzow, R. Macdonald, D. Milej, P. Sawosz, M. Kacprzak, A. Liebert, R. Cooper, J. Hebden, A. Pifferi, A. Farina, I. Bargigia, D. Contini, M. Caffini, L. Zucchelli, L. Spinelli, R. Cubeddu, and A. Torricelli, "Performance assessment of time-domain optical brain imagers, part 1: basic instrumental performance protocol," J. Biomed. Opt. **19**(8), 086010 (2014).

48. A. Pifferi, A. Torricelli, A. Bassi, P. Taroni, R. Cubeddu, H. Wabnitz, D. Grosenick, M. Möller, R. Macdonald, J. Swartling, T. Svensson, S. Andersson-Engels, R. L. P. van Veen, H. J. C. M. Sterenborg, J.-M. Tualle, H. L. Nghiem, S. Avrillier, M. Whelan, and H. Stamm, "Performance assessment of photon migration instruments: the MEDPHOT protocol," Appl. Opt. **44**(11), 2104–2114 (2005).

49. H. Wabnitz, A. Jelzow, M. Mazurenka, O. Steinkellner, R. Macdonald, D. Milej, N. Zolek, M. Kacprzak, P. Sawosz, R. Maniewski, A. Liebert, S. Magazov, J. Hebden, F. Martelli, P. Di Ninni, G. Zaccanti, A. Torricelli, D. Contini, R. Re, L. Zucchelli, L. Spinelli, R. Cubeddu, and A. Pifferi, "Performance assessment of time-domain optical brain imagers, part 2: nEUROPt protocol," J. Biomed. Opt. **19**(8), 086012 (2014).

50. M. Wahl, T. Röhlicke, S. Kulisch, S. Rohilla, B. Krämer, and A. C. Hocke, "Photon arrival time tagging with many channels, sub-nanosecond deadtime, very high throughput, and fiber optic remote synchronization," Rev. Sci. Instrum. **91**(1), 013108 (2020).

51. D. Tyndall, B. R. Rae, D. D. U. Li, J. Arlt, A. Johnston, J. A. Richardson, and R. K. Henderson, "A high-throughput time-resolved mini-silicon photomultiplier with embedded fluorescence lifetime estimation in 0.13 μm CMOS," IEEE Trans. Biomed. Circuits Syst. **6**(6), 562–570 (2012).

52. W. Becker, *The Bh TCSPC Handbook* 6th Edition (Becker and Hickl GmbH, 2014).

53. D. V. O' Connor and D. Phillips, *Time-Correlated Single Photon Counting* (Academic Press, 1984).

54. A. Cominelli, G. Acconcia, P. Peronio, M. Ghioni, and I. Rech, "High-speed and low-distortion solution for time-correlated single photon counting measurements: A theoretical analysis," Rev. Sci. Instrum. **88**(12), 123701 (2017).

55. P. B. Coates, "The correction for photon "pile-up" in the measurement of radiative lifetimes," J. Phys. E: Sci. Instrum. **1**(8), 878–879 (1968).

56. R. Re, D. Contini, M. Turola, L. Spinelli, L. Zucchelli, M. Caffini, R. Cubeddu, and A. Torricelli, "Multi-channel medical device for time domain functional near infrared spectroscopy based on wavelength space multiplexing," Biomed. Opt. Express **4**(10), 2231–2246 (2013).

57. D. Contini, F. Martelli, and G. Zaccanti, "Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory," Appl. Opt. **36**(19), 4587–4599 (1997).

58. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical recipes in C," Cambridge Univ. Press 1, 3 (1988).

59. A. Pifferi, A. Torricelli, R. Cubeddu, G. Quarto, R. Re, S. Konugolu Venkata Sekar, L. Spinelli, A. Farina, F. Martelli, and H. Wabnitz, "Mechanically switchable solid inhomogeneous phantom for performance tests in diffuse imaging and spectroscopy," J. Biomed. Opt. **20**(12), 121304 (2015).

60. F. Martelli, P. Di Ninni, G. Zaccanti, D. Contini, L. Spinelli, A. Torricelli, R. Cubeddu, H. Wabnitz, M. Mazurenka, and R. Macdonald, "Phantoms for diffuse optical imaging based on totally absorbing objects, part 2: experimental implementation," J. Biomed. Opt. **19**(7), 076011 (2014).

61. A. Torricelli, D. Contini, A. Pifferi, M. Caffini, R. Re, L. Zucchelli, and L. Spinelli, "Time domain functional NIRS imaging for human brain mapping," NeuroImage **85**, 28–50 (2014).

62. L. Zucchelli, D. Contini, R. Re, A. Torricelli, and L. Spinelli, "Method for the discrimination of superficial and deep absorption variations by time domain fNIRS," Biomed. Opt. Express **4**(12), 2893–2910 (2013).

63. F. Acerbi and S. Gundacker, "Understanding and simulating SiPMs," Nucl. Instrum. Methods Phys. Res., Sect. A **926**, 16–35 (2019).

64. F. Zappa, S. Tisa, A. Tosi, and S. Cova, "Principles and features of single-photon avalanche diode arrays," Sens. Actuators, A **140**(1), 103–112 (2007).

65. L. Spinelli, M. Botwicz, N. Zolek, M. Kacprzak, D. Milej, P. Sawosz, A. Liebert, U. Weigel, T. Durduran, F. Foschum, A. Kienle, F. Baribeau, S. Leclair, J.-P. Bouchard, I. Noiseux, P. Gallant, O. Mermut, A. Farina, A. Pifferi, A. Torricelli, R. Cubeddu, H.-C. Ho, M. Mazurenka, H. Wabnitz, K. Klauenberg, O. Bodnar, C. Elster, M. Bénazech-Lavoué, Y. Bérubé-Lauzière, F. Lesage, D. Khoptyar, A. A. Subash, S. Andersson-Engels, P. Di Ninni, F. Martelli, and G. Zaccanti, "Determination of reference values for optical properties of liquid phantoms based on Intralipid and India ink," Biomed. Opt. Express **5**(7), 2037–2053 (2014).

66. A. Behera, L. Di Sieno, A. Pifferi, F. Martelli, and A. Dalla Mora, "Instrumental, optical and geometrical parameters affecting time-gated diffuse optical measurements: a systematic study," Biomed. Opt. Express **9**(11), 5524–5542 (2018).

67. "Dependence of contrast and CNR on the IRF full-width at half-maximum (FWHM)," Figshare, https://doi.org/10.6084/m9.figshare.6731279.

68. S. Isbaner, N. Karedla, D. Ruhlandt, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein, "Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging," Opt. Express **24**(9), 9429 (2016).

69. D. Orive-Miguel, L. Di Sieno, A. Behera, E. Ferocino, D. Contini, L. Condat, L. Hervé, J. Mars, A. Torricelli, A. Pifferi, and A. Dalla Mora, "Real-time dual-wavelength time-resolved diffuse optical tomography system for functional brain imaging based on probe-hosted silicon photomultipliers," Sensors (Switzerland) **20**, (2020).