

An Integrated-Circuit Node for High-Spatiotemporal Resolution Time-Domain Near-Infrared Diffuse Optical Tomography Imaging Arrays

Sajjad Moazeni^{1,2}, Member, IEEE, Kevin Renahan^{1,2}, Student Member, IEEE,
Eric H. Pollmann^{1,2}, Student Member, IEEE, and Kenneth L. Shepard^{1,2}, Fellow, IEEE

Abstract— Next-generation brain–computer interfaces (BCIs) for healthy individuals are expected to largely rely on noninvasive functional imaging methods to record cortex-wide neural activity because of the risk associated with surgically implanted devices. In this work, we present a fully integrated 1.8×1.8 mm single chip that can be arrayed on a wearable patch to perform noninvasive, functional brain imaging over large cortical areas. This chip node contains two bonded vertical-cavity surface-emitting lasers (VCSELs), an 8×8 single-photon avalanche diode (SPAD) array with event-driven time-to-digital converters (TDCs) per row, and a digital back-end for on-chip histogramming and time-gating. We achieved 70-ps resolution for time-of-flight (ToF) imaging at the record-high 100-MHz laser repetition rate with 80-mW total power. We showed that time-gating improves the imaging contrast by as much as 36% using a brain-skull phantom. The fully integrated and compact node presented here is the key enabler for future high-spatiotemporal-resolution time-domain diffuse optical tomography (TD-DOT) imaging arrays.

Index Terms— Diffuse optical tomography (DOT), noninvasive brain imaging, time-of-flight (ToF) imager, wearable patch.

I. INTRODUCTION

NON-INVASIVE functional brain imaging is a crucial tool for behavioral neuroscience studies, assessment of brain disease, and development of brain–computer interfaces (BCIs). Diffuse optical tomography (DOT) is a promising noninvasive imaging approach which relies on reduced optical scattering and absorption of the human skull and brain tissue in the near-infrared (NIR) spectrum [1]. Time-domain-DOT (TD-DOT) is a type of DOT in which the time-of-flight (ToF) of scattered photons is measured and used to extract

Manuscript received 12 March 2022; revised 26 July 2022 and 24 October 2022; accepted 15 November 2022. This article was approved by Associate Editor Shanthi Pavan. This work was supported in part by the DARPA Next-Generation Nonsurgical Neurotechnology (N3) Program under Contract N66001-19-C-4020 and in part by the National Science Foundation under Grant 1706207. (Sajjad Moazeni and Kevin Renahan are co-first authors.) (Corresponding authors: Sajjad Moazeni; Kenneth L. Shepard.)

Sajjad Moazeni was with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA. He is now with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: smoazeni@uw.edu).

Kevin Renahan, Eric H. Pollmann, and Kenneth L. Shepard are with the Department of Electrical Engineering, Columbia University, New York, NY 10027 USA (e-mail: krenahan@ee.columbia.edu; pollmaner@ee.columbia.edu; shepard@ee.columbia.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JSSC.2022.3223854>.

Digital Object Identifier 10.1109/JSSC.2022.3223854

information about penetration depth into the tissue, which improves image quality compared with continuous-wave approaches [2]. TD-DOT typically relies on arrays of optical sources and detectors and has the potential to allow for direct sensing of intracellular neural activity and hemodynamics at spatial resolutions that exceed those of electroencephalogram (EEG) and at temporal resolutions that exceed those of functional magnetic resonance imaging (fMRI) [3] while supporting more compact and cost-effective form factors. Recent work with benchtop TD-DOT hardware has even shown that spatial resolutions similar to fMRI can be achieved with dense source-detector configurations at a few mm-range distance [4], [5]. Despite many efforts in developing such a device in a wearable form factor [6], [7], [8], achieving high-spatiotemporal resolutions remains challenging [9]. Dense and compact co-packaging of optical and electronic components to enable sub-cm pitch for source and detector arrays [10], required to improve spatial resolution, is difficult to achieve with off-the-shelf components. Moreover, as the array size grows beyond hundreds of source/detectors, interfacing optical components and electronics requires complex and high-density wiring [8], [11]. At the same time, collecting raw ToF data in large-scale arrays can result in multi-Gb/s bandwidth requirements which can impose additional challenges [8].

In this work, we address all of these challenges by demonstrating the design of an integrated-circuit node that can be arrayed on a flexible printed-circuit board (PCB) and functions as a building block for an NIR TD-DOT wearable patch (see Fig. 1). The node contains not only light emitters and detectors, but all of the necessary electronics required to perform distributed TD-DOT imaging at laser repetition rates in excess of 100 MHz, which allows for frame rates in excess of 3 kHz while collecting sufficient photons per frame. Integration of CMOS electronics at each node enables us to reduce required data transfer bandwidths by histogramming ToF data.

This article is organized as follows. Section II describes the requirements for high spatiotemporal resolution TD-DOT imaging, and Section III describes a TD-DOT array that meets these requirements. System architecture and experimental results are presented in Sections IV and V, respectively. Section VI concludes the article and presents the comparison with prior work on chip-scale ToF imagers and state-of-the-art DOT experimental systems.

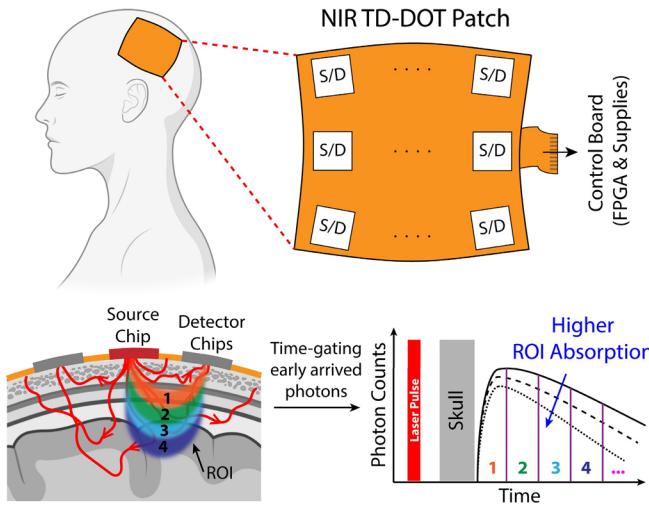


Fig. 1. NIR TD-DOT wearable patch for noninvasive brain functional imaging using an array of proposed integrated-circuit nodes (S and D stand for source and detector, respectively).

II. HIGH-SPATIOTEMPORAL RESOLUTION TD-DOT

Conventional DOT can be used to measure blood-oxygen-level-dependent (BOLD) signals by detecting changes in the optical properties of underlying tissue that arise due to the different absorption coefficients of oxygenated and deoxygenated hemoglobin (Hb/HbO_2) in the NIR spectrum. Changes in Hb/HbO_2 act as an indirect indicator of metabolism in the brain due to neuronal activity. TD-DOT is a method for the measurement of BOLD signals that utilizes complex timing electronics that can provide both intensity information and arrival times of incoming photons. Arrival-time histograms provide information about both the scattering and absorption of the tissue in the optical path [1], [2]. The presence of time-dependent absorption or fluorescence within underlying tissue (e.g., due to hemodynamics or voltage-/calcium-dependent absorbers or fluorophores) modulates the properties of the arrival-time histograms collected at the detector, where the primary effect of changes in absorption is reflected in the decay rate of the tail of the histogram (larger absorption corresponding to faster decay, and vice versa). By collecting histograms at many locations on the surface of the head, a time-dependent tomographic model of activity in the brain can be constructed using inverse imaging algorithms based on radiative transfer equations (RTE) [1], [2].

High spatial resolution in TD-DOT is dependent on creation of an array of optical sources and detectors at high density, where high density is normally taken to mean that the space between adjacent elements is no more than 15 mm [12]. However, recent work has shown that TD-DOT can realize spatial resolutions as high as 1 mm^3 through a 6.5-mm skull phantom when a 5×5 array of sources and detectors are positioned on a 4-mm pitch [4]. This result was obtained by scanning a single source and single detector around the phantom to 25 fixed positions separated by multiples of 4 mm, thereby emulating a $2 \times 2 \text{ cm}^2$, 5×5 array of sources and detectors with bulky optical equipment. While these results suggest great potential for TD-DOT hardware in rivaling fMRI as a noninvasive brain imaging technique, integrated TD-DOT

arrays with equivalent form factor have yet to be realized. This is likely because an array of this form factor would require the integration of optical sources and detectors at each node, as well as circuits that support precise timing resolution and a digital backend capable of histogramming photon arrival times so that data rates remain manageable as the number of nodes increases.

Beyond density, the form factor of a large array imposes additional challenges for detectors at each node. At source-detector separations as large as 2 cm, source power is attenuated by five orders of magnitude or more, which necessitates that detectors have very low noise-equivalent power (NEP), on the order of $20 \text{ fW}/\sqrt{\text{Hz}}$, to minimize contributions of read noise in measurements where photon count is very limited [12].

In contrast, at source-detector separations as small as 4 mm, detectors see large numbers of photons, with the majority of these early-arriving photons being non-informative as they are backscattered off of the skull [1], [10]. To help address this issue, detectors used in TD-DOT hardware should support time-gating so that early-arriving photons can be gated out of measurements. By rejecting skull-backscattered photons, which dominate in the case of sub-cm source-detector pitch configurations, the dynamic range of detection can be increased, and pile-up distortion can be minimized [10], [13]. This ultimately leads to better image contrast and contributes to higher spatial resolution during image reconstruction [14].

High temporal resolution in TD-DOT is dependent on the achievable frame rate. A frame rate of 10 Hz is considered optimal for the detection of BOLD signals [12], but higher

frame rates are needed for the direct recording of brain activity using red-shifted NIR genetically encoded calcium or voltage indicators [15] which require milliseconds temporal resolutions [5], [16]. Relative to benchtop systems, integrated TD-DOT hardware has the advantage that dedicated sources and detectors can be used at every location in the array, which eliminates the need for scanning and allows for detectors to operate in parallel so that integration time can be maximized for each source-detector pair. For example, in a benchtop scanning system with a single source and detector, a 10-Hz frame rate with a 5×5 array of sources and detectors implies only $160 \mu\text{s}$ of integration time for a given source-detector pair. When the time needed to change the position of the source and/or detector optics between measurements is considered, it rules out this type of system for the detection of BOLD signals *in vivo* [5]. In comparison, a fully integrated TD-DOT array could incorporate 25 unique detectors that operate in parallel, increasing available integration time to 4 ms per source-detector pair with no time lost to scanning. In addition, custom hardware can be designed to support laser repetition rates as high as 100 MHz, which is helpful for the collection of a greater number of photons within the available integration window.

The final aspect of TD-DOT hardware that must be considered is the impulse response function (IRF) of the overall system, as determined by the sources, the detectors, and the jitter of timing electronics. Long IRF tails can reduce contrast, limit the effectiveness of time-gating, and limit the resolution of ToF measurements, thereby negatively contributing to both

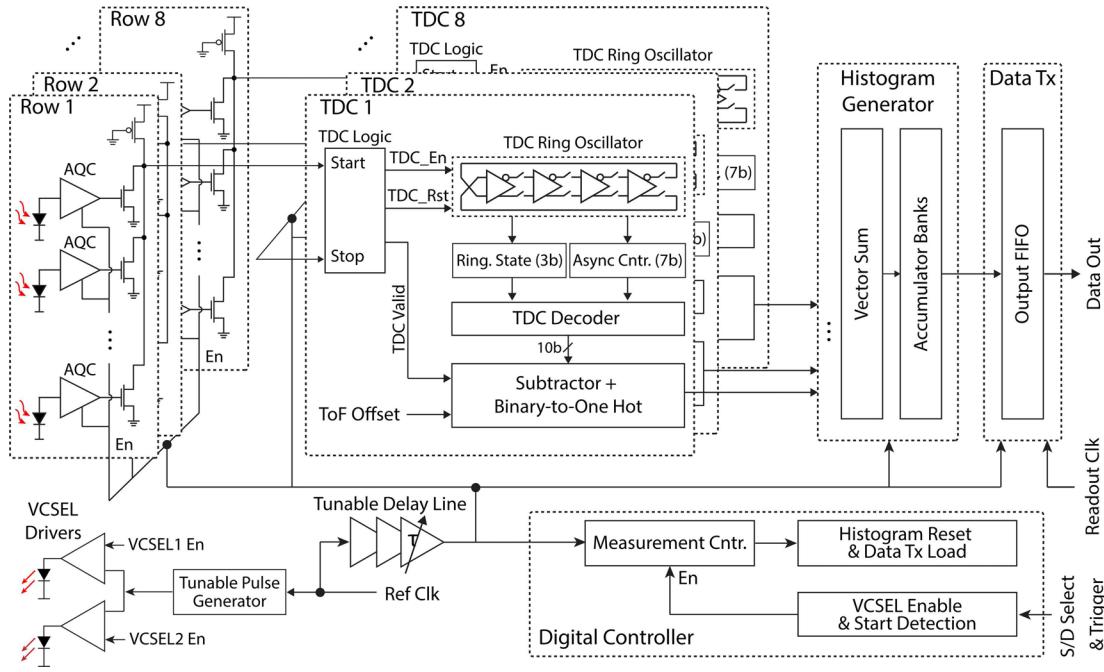


Fig. 2. Block diagram of the CMOS integrated circuit node.

spatial and temporal resolution [14], [17]. The selection of sources and detectors having sharp timing response is, therefore, critical.

III. TD-DOT IMAGING ARRAY

There is a need for a compact integrated element, herein referred to as a node, that could enable the realization of a TD-DOT array for brain imaging like the one proposed in [4] without having to rely on bulky optical components and timing electronics. Instead, this node would incorporate all necessary electronics and optics in a fully integrated form factor. The node would have dimensions less than 2×2 mm such that it could be tiled at a 4-mm pitch while leaving adequate space for circuit board routing within a larger array. The node would also have a digital backend capable of histogramming photon arrival times to keep data rates manageable as array size increases. Finally, the node would feature a source capable of generating sharp optical pulses, a detector with time-gating capabilities and NEP below $20 \text{ fW}/\sqrt{\text{Hz}}$, and timing electronics that minimize overall system IRF.

These needs are addressed by the node chip demonstrated here. Vertical-cavity surface-emitting lasers (VCSELs) are chosen as the source due to their sharp optical pulsedwidth and high electrical bandwidth, which support the need for minimal system IRF and high laser repetition rate. Single-photon avalanche diodes (SPADs) are chosen as detectors for their TCSPC capabilities, low NEP, and support for time-gating. Time-to-digital converters (TDCs) with ~70-ps timing resolution are designed to minimize overall system IRF, and a digital histogramming backend is incorporated to reduce off-chip data bandwidth. Through the integration of SPADs directly alongside timing and histogramming electronics, as well as direct mounting of VCSELs on the surface of the CMOS chip, the node is of a favorable form factor for the creation of dense

arrays. Fig. 1 illustrates a TD-DOT imaging array that can be realized using the node presented in this article. During each recording frame, one or more of these nodes are configured to act as light sources to send a programmable number of short laser pulses into the tissue, while all nodes simultaneously perform ToF measurement of detected photons.

IV. SYSTEM ARCHITECTURE

Fig. 2 shows the system-level block diagram of the node chip electronics. The chip uses two VCSELs with dedicated drivers to illuminate the imaging scene, which supports dual-wavelength operation. These VCSELs are stacked directly onto the chip surface with their cathode (bottom plate) in direct contact with a metal pad on the chip surface, while the anode pads are wire-bonded to another on-chip pad. With two wavelengths, we can measure two chromophores, in particular, Hb/HbO₂ in the case of BOLD measurements. From this, we can calculate total hemoglobin concentrations $[\text{THb}] = [\text{HbO}_2] + [\text{Hb}]$, and oxygen saturation $\text{StO}_2 = [\text{HbO}_2]/[\text{THb}]$. We can also use information at two different wavelengths to deduce changes in blood volume that would introduce artifacts in Hb/HbO₂ measurements performed at a single wavelength.

The node chip also contains electronics that handle detection and timestamping of incoming photons. Photon detection is accomplished through the use of SPADs that are operated by in-pixel active quench and reset circuits (AQC) with tunable delays to control time-gating. Timestamping is accomplished using TDCs.

The most basic implementation of photon detection and timestamping backend would involve just a single SPAD/AQC paired with a single TDC. However, in TD-DOT, where photon shot noise contributes to measurement noise, image contrast is improved by increasing the number of parallel photon detection paths to collect more photons for a given integration

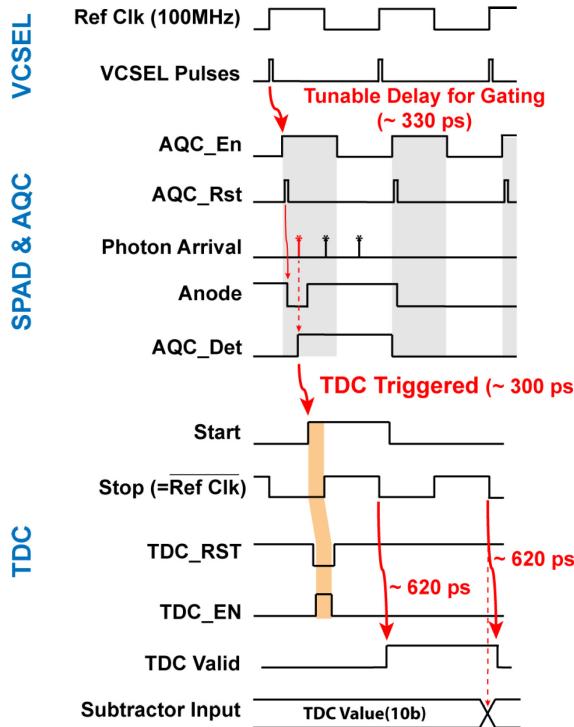


Fig. 3. Overall timing diagrams of the node operation.

window. With this in mind, the chip was designed to support as many parallel photon detection paths as possible given the chip form factor. With an allowable chip size of 2×2 mm, and after pre-allocating space for two VCSELs, two VCSEL drivers, and IO pads, we were able to fit a maximum of eight TDCs in a single column.

Further increasing the photon collection capabilities of the chip can be realized by replicating the column many times in a TDC-per-SPAD architecture. However, it is worth considering that with a SPAD-based imager, where light levels must be kept low so that pile-up distortion can be avoided, SPADs are expected to detect photons in less than 1% of the clock cycles in which they are enabled [18]. It is, therefore, possible to share a single TDC between multiple SPADs/AQCs in a winner-take-all (WTA) configuration with minimal loss of information due to overlapping detections. This yields a TDC-per-row architecture, which is more power efficient, especially when the relatively high static current draw of the TDC is considered. For this reason, we chose to incorporate as many SPADs as possible per TDC in a single row. SPADs were pitch matched in the horizontal and vertical direction with the TDCs, resulting in eight SPADs per TDC. The overall photon detection chain on the chip, therefore, consists of eight TDCs and an 8×8 array of SPADs/AQCs enabling detection and ToF measurement of up to eight photons simultaneously.

The overall timing diagram of node operation is shown in Fig. 3. In each imaging frame, one (or more) chip(s) can act as the source by emitting short VCSEL light pulses after each rising edge of the reference clock (Ref Clk). Thus, the Ref Clk frequency, which is set to 100 MHz, determines the laser repetition rate. Each node has a tunable inverter-based delay line to adjust the delay of the SPAD time gate (AQC_EN)

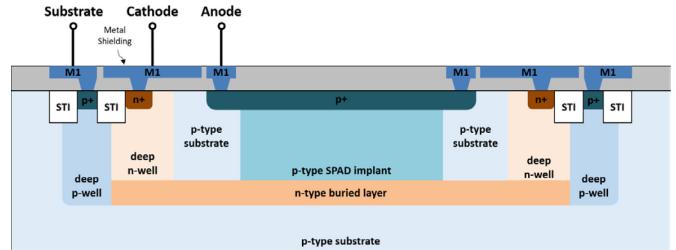


Fig. 4. Cross section of the SPAD.

relative to the Ref Clk, and VCSEL pulses, which allows for early photons to be gated out of the measurements. The delay line has a 200-ps least-significant bit (LSB) and eight bits with a maximum possible delay of approximately 50 ns. This allows for the SPAD time gate to be swept through the entire positive phase of the clock cycle at frequencies as low as 10 MHz, which sets the imager's minimum operating frequency. Once a photon has been detected (AQC_DET signal generated by the AQC circuit), it will trigger the Start signal for a TDC. The ToF will be indirectly measured with respect to the falling edge of reference clock. Hence, the stop signal for TDC is set by the inverted Ref Clk. Finally, once the TDC operation is complete (TDC_Valid signal), the result is sampled by the rising edge of Ref Clk and it goes through a digital backend for subtraction and histogramming. We will elaborate on the details of major chip components and blocks below.

A. NIR SPAD Design

SPADs are implemented with an $11\text{-}\mu\text{m}$ diameter active area using a custom doping implant. Pixels have a $60\text{-}\mu\text{m}$ pitch with 3% effective fill factor (FF). While the entire SPAD array acts as a single detector on a single chip, it is necessary to have an array of SPADs to increase the detection aperture and collect sufficient photons in each frame. The SPADs are optimized to have maximum photon detection probability (PDP) and minimal jitter in the NIR. This is accomplished using a deep p-type implant that forms an avalanche region with a deep n-type implant as illustrated in the cross section of the SPAD shown in Fig. 4. By moving the avalanche region of the SPAD deeper into the silicon, the PDP and jitter of the structure at NIR wavelengths can be improved at the expense of PDP and jitter at visible wavelengths. We have also used a custom p-type implant in this process in order to further improve the PDP. In addition, we have added metal shielding using the first back-end-of-line metal layer to prevent the absorption of photons by non-active regions of the SPAD such as guard rings and surrounding substrate. Carriers created in these regions by impinging photons have lifetimes on the scale of several nanoseconds and trigger delayed avalanches that occur long after the photon was first absorbed by the SPAD. This creates the “diffusion tail” commonly seen in the SPAD’s IRF. In TD-DOT systems, the delayed avalanches caused by these carriers cannot be distinguished from avalanches caused by late-arriving photons in the tail of the histogram. As a result, the carriers act to reduce contrast in the later time bins, which typically carry the most useful information. It is, therefore, crucial to suppress their effect as much as possible.

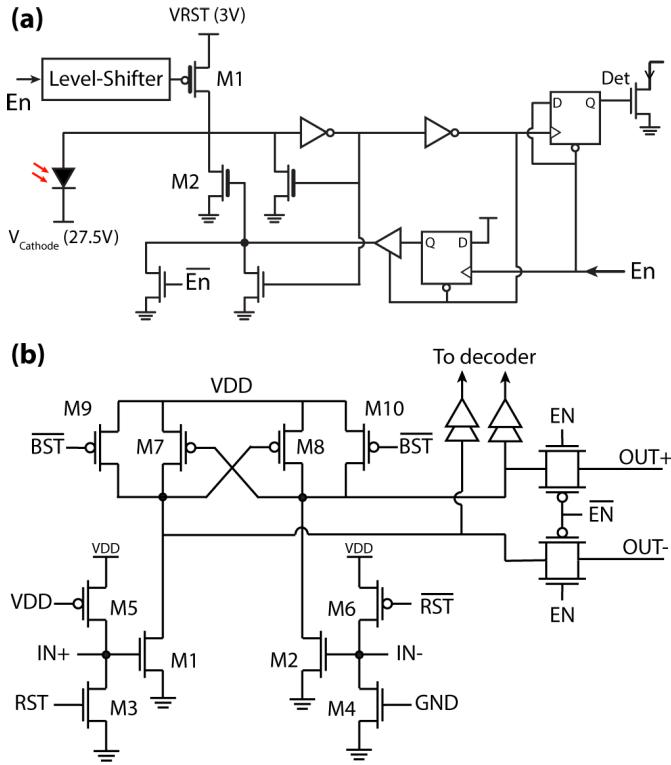


Fig. 5. Circuit diagrams of (a) AQC and (b) TDCs' inverter delay-cells (only one of two reset configurations is shown).

B. Active Quench and Reset Circuit

Fig. 5(a) shows the circuit diagram of the AQC, which is based on our previous work in [19]. To reduce SPAD dead time, the AQC is designed to gate the SPAD from an OFF state to an ON state in less than 200 ps. On the rising edge of EN, a level shifter converts this signal to the VRST domain, allowing it to turn off the PMOS quench transistor (M1). In parallel, a one-shot pulse is generated by latching VDD into a flip-flop on the rising edge of EN, turning on the NMOS reset transistor (M2), and discharging the SPAD anode past its breakdown voltage. Once the SPAD is in Geiger mode, the inverter-based comparator flips, causing the one-shot flip-flop to asynchronously reset low and turn off the NMOS reset transistor, leaving the SPAD anode in a high-impedance state. Upon detection of a photon, an e^-/h^+ pair gets accelerated by the large electric field in the depletion region pushing avalanche current into the high impedance anode causing the inverter-based comparator to flip. A second inverter uses thick-oxide devices and level-shifts down to the VDD domain, sending the rising edge of the photon detection signal into the event-detection flip-flop which passes the EN signal to the WTA NMOS pull-down transistor at the output. The SPAD remains in an OFF state until the next EN signal is generated based on the system clock. Since the AQC uses an internal monostable pulse generator to reset the SPAD, the circuit can be used with SPAD variants with extra capacitance at the anode. In addition, the VRST can be tuned between 1.5 and 3.3 V depending on the targeted excess bias voltage. The AQC's reset path delay supports operation at a 100-MHz repetition rate and supports excess-bias voltages of up to 3.3 V

with the use of thick-oxide reset and pre-charge switches at the anode.

Each pixel has its own enable signal to deactivate “hot” pixels [SPADs with abnormally high dark count rates (DCRs)]. This is implemented by adding a local AND gate between the global EN signal and a separate per-pixel enable signal.

C. VCSEL Driver

The VCSEL drivers are implemented using thick-oxide devices with up to 3.3-V drive capabilities and a level-shifter pre-driver. Voltage-mode VCSEL drivers have been used here to achieve higher power efficiencies over conventional current-based VCSEL drivers [20]. VCSEL pulses are generated by delaying the reference clock (see Fig. 3) with a NOR gate in the nominal 1.5-V digital voltage domain. A three-bit tunable delay with a 120-ps LSB is designed to generate tunable pulse widths for the VCSEL drivers. The NOR gate's output pulse is buffered and fed into a low-voltage level-shifter with a PMOS cross-coupled pair [21], followed by a series of pre-drivers and the final VCSEL driver in the high-voltage domain.

Since VCSELs have been shown to have more delay on the rising edge than the falling edge of a pulse if fully modulated from an OFF to an ON state [22], we have skewed the pulse generator to speed up the rising edge.

D. TDC and Histogramming Backend

TDCs are implemented using asynchronous control logic and a four-stage ring-oscillator (shown in Fig. 2) employing pseudo-differential inverter cells with a cross-coupled PMOS load [23]. Fig. 5(b) shows the TDC inverter cells' detailed circuit diagram. We have also added dummy reset switches [M4 and M5 in Fig. 5(b)] to the inputs for symmetry and to improve the differential nonlinearity (DNL). Enable switches are implemented using pass gates, and PMOS current-bleeders (M9 and M10) have been added to boost the oscillator's frequency and achieve the target LSB of ~ 70 ps. The 10-bit outputs of the eight TDCs are composed of three fine bits and seven coarse bits. The three fine bits are generated by encoding the eight possible internal states of the four-stage ring-oscillator as a three-bit binary number. The seven coarse bits come directly from the asynchronous counter, which is incremented at every full oscillation cycle of the ring oscillator. The TDC is event-driven: the rising edge of the detection signal from the AQC starts the ring oscillator, and the falling edge of the clock stops it. This ensures that the ring oscillator is only started when a photon is detected. Because the start signal can arrive at any time relative to the stop signal, it is important that the digital backend samples the TDC outputs at the correct time to guard against metastability, which is accomplished using asynchronous control logic. In the negative phase of the clock, which comes just after the stop signal, the control logic waits for a fixed amount of time to ensure that the asynchronous counter has time to settle to its final value before moving the fine and coarse bits into a set of output flops. After another fixed delay to ensure that the data has finished propagating into the output flops, the TDC Valid signal is asserted before the next rising edge of the clock. At the rising edge of the clock, the subtractor samples the

outputs flip-flops if the TDC Valid signal has been asserted. This mechanism allows for TDC data to move from the TDC to the backend logic with no possibility for metastability. While the ring oscillator and the counter are on the same VDD supply, we have added local decoupling caps to minimize the impacts of digital switching currents on the oscillator's frequency.

The TDC outputs (10 b total) are then fed into subtractor blocks. While the TDC can generate timestamps between 0 and 70 ns, the imager keeps data rates low by transmitting only a 10-ns window of the captured transient. Due to the event-driven architecture of the TDC, raw timestamps from the TDC represent the time at which a photon was detected relative to a half-cycle before the next laser pulse. The subtractor is typically configured at runtime with a value equal to half of the clock period divided by the TDC resolution. The output of the TDC is then subtracted from this value such that the output of the subtractor block represents the actual ToF between the VCSEL pulse and the detected photon. The subtractor can be configured with a value offset from this typical value such that TDC outputs can be shifted prior to entering the accumulator bank, allowing for any portion of the captured transient to be shifted into the 10-ns stream-out window. This facilitates a large reduction in the number of histogram bins while still allowing transients to be captured that occur more than 10 ns after the VCSEL pulse. The end result of limiting the stream-out window is a reduction in chip bandwidth by approximately a factor of seven. Furthermore, we can correct any ToF offset errors due to duty-cycle offset and chip-to-chip variations, for example, by using this subtractor block in a ToF measurement setup described in Section V.

The outputs of all eight subtractors flow into a bank of accumulators, where the ToF measurements collected by the eight data paths are aggregated into a single histogram per chip with 150 12-bits bins. Finally, the histogram is serialized in the data transmitter (Tx) block and transmitted to a central field programmable gate array (FPGA) unit at the end of each frame. The clock driving the Tx block is designed to be completely asynchronous to the clock driving the imager core. This allows the imager core to run at high frequencies to maximize the number of photons captured per unit time, while simultaneously allowing the data transmission block to run at very low clock frequencies. This reduces the need for complex clock routing of the Tx clock, which is favorable for densely arraying the node on a package.

The chip operates independently after initial configuration, with the number of laser pulses monitored with a 15-bit counter and histogram stream-out triggered when the counter reaches a programmable value. Histogramming on-chip keeps the data rate very low compared with chip architectures that transmit single SPAD counts and corresponding timestamps. The chip streams out 1800 bits, corresponding to 150 12-bit histogram bin values, every 320 μ s during normal operation at 100 MHz, which sets the data rate at 5.625 Mb/s. This implies a frame rate of 3125 frames/s is possible from a single chip. These low data rates help to ensure manageable aggregate data rates when the node is arrayed. The Tx block can support operation up to 25 MHz, which could allow for theoretical operation of a single node at frames rates as high

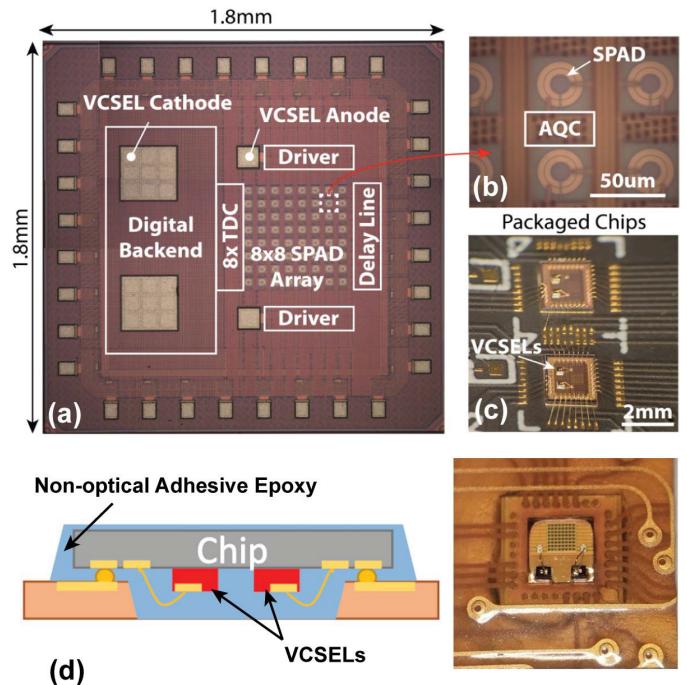


Fig. 6. Micrographs of (a) node chip and (b) pixel cell, a photograph of two chips at 4-mm node pitch with bonded VCSELs wire-bonded on (c) PCB board, and (d) flip-chip bonded packaging on a flexible substrate for ultimate wearable patch form factor.

as 13 888 frames/s, although this frame rate is likely not useful as it limits integration time within a frame to only 72 μ s.

Design choices regarding the ToF resolution, histogram bin sizes, and gating requirements are based on Monte-Carlo simulations verified by a phantom measurement from [14] and using a new sparse TD-DOT image reconstruction approach [4]. With an LSB of \sim 70 ps, the stream-out window of 150 bins allows for 10.5 ns of collection time within a single histogram. For clock frequencies between 50 and 100 MHz, 10.5 ns is sufficient time for all photon counts between subsequent VCSEL pulses to be captured and streamed out. Histogram bins were allocated only 12-bits with the understanding that for a large array, longer integration times more uniformly sample the field of view when the emitter is swept around the array a larger number of times at faster intervals rather than simply integrating for a longer time within a single sweep. With this approach, more histograms are generated during operation, but the number of counts within a histogram is kept small. This makes 12-bit counters adequate for providing the dynamic range needed for imaging with a large array.

V. SYSTEM CHARACTERIZATION AND MEASUREMENTS

A micrograph of the 1.8 \times 1.8 mm node chip, fabricated in a 130-nm high-voltage CMOS process, is shown in Fig. 6. The compact node area allows us to array these nodes with a tight pitch of 4 mm either on a rigid substrate with wire-bonding or on a flexible polyimide PCB substrate using flip-chip bump bonding. While the SPADs are monolithically designed and implemented in the CMOS chips, VCSELs have to be heterogeneously integrated. This has been done by bonding the VCSEL's bottom pad (cathode) via silver epoxy to a large

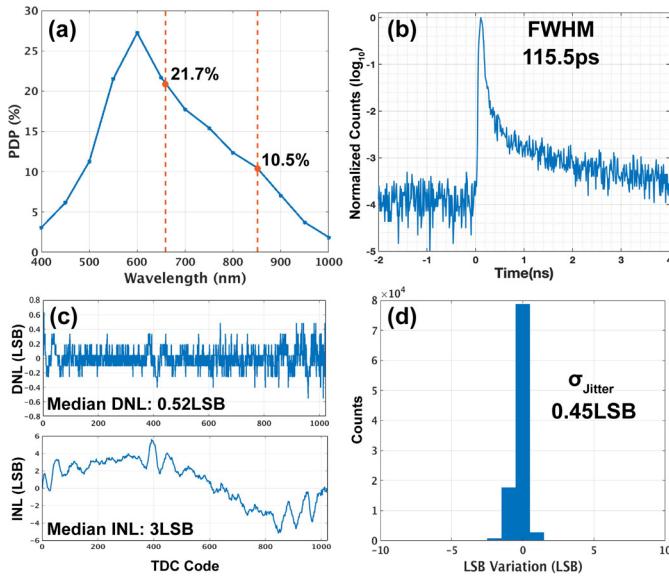


Fig. 7. (a) SPAD's PDP versus wavelength, (b) SPAD's jitter, (c) TDC's DNL and INL, and (d) TDC's jitter measurement.

CMOS pad ($250 \times 250 \mu\text{m}$) using a flip-chip bonder, and wire bonding the anode pad afterward, as shown in Fig. 6(c).

We have successfully managed to flip-chip bond the chip on a flexible PCB, as shown in Fig. 6(d). The flexible PCB has a cut-out to provide an optical window for VCSELs and SPADs. This approach allows us to package a large array of presented nodes in a wearable patch for future experiments on a curved brain/skull phantom or animal/human subjects. For our measurements, however, we chose to package the chip on a rigid board for ease of use in our experiment geometries. All supplies (VDD of 1.5 V, and HVDD of 3 V for VCSEL drivers) were sourced externally, and each supply was fit with discrete decoupling capacitors. Here we present the measurement results of chip blocks and in vitro experiments.

A. SPAD and TDC Characterization

The SPADs have a break-down voltage of 26.2 V and achieve PDPs of 21.7% and 10.5% at 670- and 850-nm wavelengths [see Fig. 7(a)], respectively, at a 1.5-V excess-bias voltage with a median DCR of 6.9 cps and a full-width half-maximum (FWHM) jitter of 115.5 ps [see Fig. 7(b)]. The NEP of the SPAD is $0.19 \text{ fW/}\sqrt{\text{Hz}}$ at 670-nm based on the formula in [24].

The TDCs have been characterized using an external pulse generator from the median of 50 samples, as shown in Fig. 5(c) and (d). The measurements show ~ 70 -ps timing resolution (LSB) over the entire range of ~ 70 ns, which allows the chip to support repetition rates as low as 10 MHz. The TDC's FWHM jitter is less than 0.45 LSB for a fixed delay of 10 ns from a total of 100-k measurements with a median DNL and integral non-linearity (INL) of 0.52 and 3 LSB across the entire TDC range, respectively. The TDC achieved similar or better than 0.45 LSB FWHM as the delay is varied. In addition, we have tested all TDCs simultaneously to make sure that potential supply droops will not impact the ToF measurements, and we have not noticed any changes in

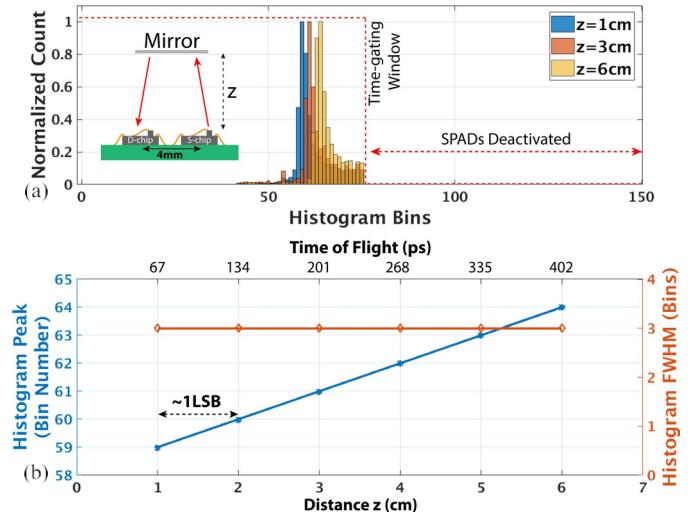


Fig. 8. ToF measurements: (a) sample histograms for multiple chip-mirror distances (z) and (b) histograms' peak and FWHM for 1–6-cm z -range.

their characteristics. Each TDC consumes $17.5\text{-}\mu\text{A}$ dynamic current at 500 kS/s with 3-mA static current drawn by PMOS current-bleeders.

B. VCSEL Characterization

We use 670-nm VCSELs from Vixar Inc., as our light sources, which feature a threshold current of $700 \mu\text{A}$, >5 GHz 3-dB electro-optical bandwidth, and 3-mW peak optical power in pulsed mode. We achieved optical pulse widths of ~ 210 ps with a -1 -V cathode bias and 3-V electrical peak drive voltage. These optical output power values are comparable with other TD-DOT devices [25]. While currently we have only used a single wavelength in our system, 850-nm VCSELs are also commercially available with higher optical power (>10 mW) and bandwidths. This higher power is an advantage despite the lower PDP at these wavelengths.

C. ToF Measurements

Initial device characterization and calibration have been performed by measuring the ToF using two nodes and a mirror at multiple distances (z) under low-photon intensity to avoid a pile-up. Fig. 8 shows the schematic of this test setup along with three sample histograms. The time-gating delay is set arbitrarily in this experiment to verify the gating functionality. Histograms are normalized to cancel the free-space propagation loss for various distances in this plot. The peak and FWHM of each histogram for the 1–6-cm z -range are plotted, showing the expected linear relationship with one LSB for every 2-cm ToF (~ 67 ps) and a 3 LSB (210 ps) total FWHM jitter. The accuracy of z -measurement is ~ 1 cm as determined by the ToF measurement LSB of ~ 67 ps. The end-to-end timing jitter includes contributions from all the critical elements of the system, including the SPAD's IRF, the VCSEL driver's optical pulsewidth, and all the jitter associated with the VCSEL driver, SPAD, AQC, and TDC. Total power consumption for each chip is 80 mW under ambient light condition with 32, 36, and 5 mW consumed by the digital backend, TDC current bleeders, and the VCSEL drivers at

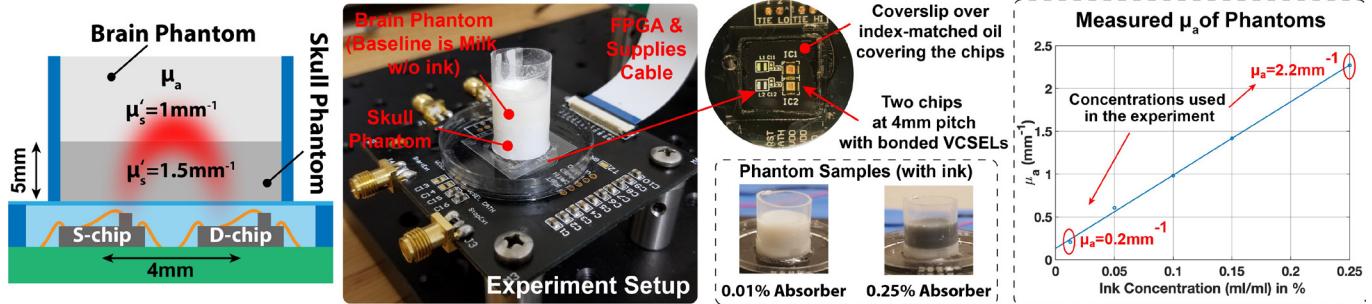


Fig. 9. Brain phantom experiment: photographs of the test setup, chip nodes, and samples of brain phantoms with different India ink concentrations.

a 100-MHz repetition rate, respectively. The rest of the power is consumed by the AQC and clocking blocks.

D. In Vitro Measurements

The imaging capabilities of the node along with the effectiveness of time-gating in improving the image contrast are evaluated by measuring the absorption coefficient changes in a brain-tissue phantom (a mixture of milk and India ink) through a 5-mm-thick skull phantom (a mixture of TiO₂ and epoxy resin), as shown in Fig. 9. In this work, we image a static scene where we expect no changes in solution volume within the optical path, and as a result, we show the results only at a single wavelength of 670 nm. Reduced scattering and absorption coefficients are denoted by μ'_s and μ_a , respectively. The values are chosen to match skull/brain and Hb/HbO₂ optical properties [26]. The concentration of the ink has been chosen to achieve $\mu_a = 0.2 \text{ mm}^{-1}$ and $\mu_a = 2.2 \text{ mm}^{-1}$, which correspond to estimated low and high μ_a for hemodynamics [26], [27].

Fig. 10 shows un-gated and gated histograms with an optimal gating delay of 1.05 ns for two different ink concentrations. The collection process for the data shown is described as follows. The VCSEL is pulsed for 25 000 repetitions, and timestamps for incoming photons are aggregated by the digital backend to form a single histogram. The effective integration time for this histogram is 250 μ s. The histogram is streamed off-chip by the Tx block, while the next histogram is collected in exactly the same way as was just described. This process is repeated until a total of 4000 histograms have been collected and streamed off-chip. These 4000 histograms are averaged together in backend software, which results in a 500-ms total integration time for the data shown. In this way, the histograms of Fig. 10 represent a denoised version of the information that is collected in 250 μ s with our system. As shown in this figure, the contrast in the tail of the histogram improves as the time-gating delay is increased. The image contrast is derived from $(N - N_0)/N_0$ equation, where N_0 and N denote unperturbed (no ink) and perturbed (with ink) photon counts, respectively. Larger gating delays reject non-informative early-arrived photons (from the skull) while still receiving enough useful scattered photons (from the brain) to maximize the contrast. Larger gating delays also help to reduce pile-up distortion in the tail of the measurement, which further improves contrast relative to ungated measurements. A more accurate μ_a estimation and gating delay can also be

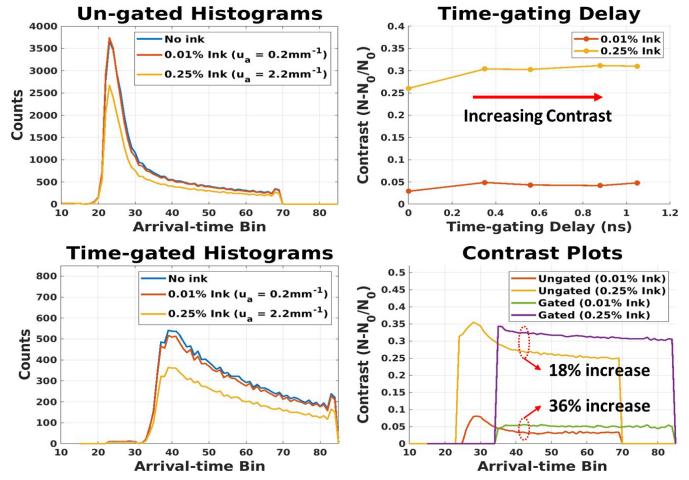


Fig. 10. NIR TD-DOT measurement using brain-skull phantoms: arrival-time histograms with and without gating, and the imaging contrast comparison plots.

determined from the decay rate of the histograms [2], [5]. Measurement results in Fig. 10 confirm that gating improves the contrast by as much as 36% in this experiment. Further improvements in contrast are possible by further enhancements to the system's IRF.

VI. CONCLUSION

The single-chip node with integrated light sources presented here is an important building block for scalable, wearable NIR TD-DOT imaging arrays. The performance summary and comparison with prior NIR ToF CMOS imagers [6], [7], [23], [28] are presented in Table I. While on-chip histogramming is key to significantly reduce the off-chip data rate, in the current implementation histogramming digital backend dominates the chip power similar to [6]. This power can be further reduced in future implementations by using multi- V th backend design and better application of clock gating. In addition, gating the SPADs also requires extra power in particular at high clock frequencies.

Table II compares our system-level results with state-of-the-art experimental TD-DOT systems [8], [29]. The system in [29] is table-top using fibers, a commercial SiPM chip, and a TCSPC module, while the work in [8] has fully arrayed S/D modules in a helmet form-factor. By fully integrating the light sources and detectors into a single chip, which is not done in [8], we are able to support higher pulse repetition rates, and

TABLE I
PERFORMANCE SUMMARY AND COMPARISON WITH PRIOR NIR TOF IMAGERS

	JSSC'15 [22]	ISSCC'11 [27]	TBiOCAS'19 [5]	JSSC'20 [4]	This Work
SPAD	Wavelength (nm)	800	637	850	780
	PDP (%)	12.2	19	5	8
	FF (%)	23.3	1	40	37
	DCR (Hz)	35k	50	10k	~1k
	Jitter FWHM (ps)	260	140	350	~50
TDC	Range (ns)	200	55	N/A (No on-chip TDC)	10
	Resolution (ps)	49.7	55		78
	DNL/INL (LSB)	0.44/0.47	0.3/2		<0.1/0.58
	Jitter FWHM (LSB)	1.21	N/R		0.45
	Dynamic Power⁺ (μW)	15	N/R		250
System	CMOS Technology	130nm (3D)	130nm	350nm	350nm
	Power (mW)	7 [*]	550 [*]	N/R	1300
	Array Size	1×400	160×128	8×16	25×70
	Chip Area (mm²)	0.77×5	11×12.3	2×2	6×7
	Repetition Rate (MHz)	40	16	100	80
	Total IRF FWHM (ps)	N/R	>200ps	430	1000
	Read-out Data-rate (GHz)	1.040	51.2	N/A	0.008⁺⁺
	Time-gating	No	No	Yes	Yes
	On-Chip Histogramming	No	No	No	Yes
	Integrated Light Sources	No	No	No	Yes

* Core only and under dark condition, + Measured @500kS/s rate, ++ Only one TDC has been used in this system, N/R: Not reported, N/A: Not applicable

TABLE II
PERFORMANCE SUMMARY AND COMPARISON
WITH PRIOR TD-DOT SYSTEMS

	Milan Probe SiPM [28]	Kernel Flow [6]	This Work
Detector	Hamamatsu SiPM	CMOS 130nm	CMOS 130nm
Technology			
Wavelength (nm)	670 & 830	690 & 850	670
Detector Responsivity (m².sr)	3.3×10 ⁻⁸	7.2×10 ^{-9*}	5×10^{-10†}
System IRF FWHM (ps)	308-556	290-350	210
Emitter	External Laser	Edge-emitting Laser	VCSEL
Technology			
Wavelength (nm)	800	637	670
ToF Resolution (ps)	80	N/R	70
Max Pulse Power (mW)	2	1	3
ToF DNL (LSB)	0.03	<0.5	0.52
System	External Electronics	Multi-chip on a PCB	Single chip
Technology			
Source-Detector Distance (mm)	30	10	<4
Module/chip Size (mm²)	900	1000 ⁺⁺	1.8×1.8
Repetition Rate (MHz)	40	20	100
Dynamic Range +	10 ²⁻³	10 ⁴⁻⁵	10³⁻⁴
Count Rate per Detector Node (Meps)	40	4.8	100

* Using external lenses, [†]Estimated, + For 5ms integration window, N/R: Not Reported.

more importantly, the smallest source-detector distance (4 mm) reported in a TD-DOT system.

This work sets the stage for mm-range spatial and sub-second temporal resolutions in future TD-DOT systems in a wearable form factor. Optimization of flip-chip bump bonding yield, control of power distribution within the array, and management of heat generated at the tissue will be critical to the eventual realization of a wearable devices with hundreds of sources and detectors. Beyond brain imaging, TD-DOT wearable devices have many other biomedical applications,

such as breast cancer detection and muscle oxygenation monitoring, which will continue to drive further development of this technology in the near future.

ACKNOWLEDGMENT

The authors gratefully acknowledge Taiwan Semiconductor Manufacturing Company (TSMC) for chip fabrication and support for the use of single-photon avalanche diodes (SPADs).

REFERENCES

- [1] D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: Approaches to optimizing image sensitivity, resolution, and accuracy," *NeuroImage*, vol. 23, pp. S275–S288, Jan. 2004, doi: [10.1016/j.neuroimage.2004.07.011](https://doi.org/10.1016/j.neuroimage.2004.07.011).
- [2] M. D. Y. Hoshi and Y. Yamada, "Overview of diffuse optical tomography and its clinical applications," *J. Biomed. Opt.*, vol. 21, no. 9, pp. 1–11, 2016, doi: [10.1117/1.JBO.21.9.091312](https://doi.org/10.1117/1.JBO.21.9.091312).
- [3] E. M. C. Hillman, "Optical brain imaging in vivo: Techniques and applications from animal to man," *J. Biomed. Opt.*, vol. 12, no. 5, p. 51402, 2007, doi: [10.1117/1.2789693](https://doi.org/10.1117/1.2789693).
- [4] H. K. Kim, Y. Zhao, A. Raghuram, A. Veeraraghavan, J. Robinson, and A. H. Hielscher, "Ultrafast and ultrahigh-resolution diffuse optical tomography for brain imaging with sensitivity equation based non-iterative sparse optical reconstruction (SENSOR)," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 276, Dec. 2021, Art. no. 107939, doi: [10.1016/j.jqsrt.2021.107939](https://doi.org/10.1016/j.jqsrt.2021.107939).
- [5] Y. Zhao, A. Raghuram, H. K. Kim, A. H. Hielscher, J. T. Robinson, and A. Veeraraghavan, "High resolution, deep imaging using confocal time-of-flight diffuse optical tomography," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 7, pp. 2206–2219, Jul. 2021, doi: [10.1109/TPAMI.2021.3075366](https://doi.org/10.1109/TPAMI.2021.3075366).
- [6] E. Conca et al., "Large-area, fast-gated digital SiPM with integrated TDC for portable and wearable time-domain NIRS," *IEEE J. Solid-State Circuits*, vol. 55, no. 11, pp. 3097–3111, Nov. 2020, doi: [10.1109/JSSC.2020.3006442](https://doi.org/10.1109/JSSC.2020.3006442).
- [7] S. Saha, Y. Lu, F. Lesage, and M. Sawan, "Wearable SiPM-based NIRS interface integrated with pulsed laser source," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1313–1323, Dec. 2019, doi: [10.1109/TBCAS.2019.2951539](https://doi.org/10.1109/TBCAS.2019.2951539).
- [8] H. Y. Ban et al., "Kernel flow: A high channel count scalable TD-fNIRS system," *Proc. SPIE*, vol. 11663, pp. 24–42, Mar. 2021, doi: [10.1117/12.2582888](https://doi.org/10.1117/12.2582888).

[9] M. Lacerenza et al., "Wearable and wireless time-domain near-infrared spectroscopy system for brain and muscle hemodynamic monitoring," *Biomed. Opt. Exp.*, vol. 11, no. 10, pp. 5934–5949, Oct. 2020, doi: [10.1364/BOE.403327](https://doi.org/10.1364/BOE.403327).

[10] A. Puszka et al., "Spatial resolution in depth for time-resolved diffuse optical tomography using short source-detector separations," *Biomed. Opt. Exp.*, vol. 6, no. 1, pp. 1–10, Jan. 2015, doi: [10.1364/BOE.6.000001](https://doi.org/10.1364/BOE.6.000001).

[11] P. Pinti et al., "The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience," *Ann. New York Acad. Sci.*, vol. 1464, no. 1, pp. 5–29, 2018.

[12] M. D. Wheelock, J. P. Culver, and A. T. Eggebrecht, "High-density diffuse optical tomography for imaging human brain function," *Rev. Sci. Instrum.*, vol. 90, no. 5, pp. 1–24, 2019, doi: [10.1063/1.5086809](https://doi.org/10.1063/1.5086809).

[13] A. Tosi et al., "Fast-gated single-photon counting technique widens dynamic range and speeds up acquisition time in time-resolved measurements," *Opt. Exp.*, vol. 19, no. 11, p. 10735, May 2011, doi: [10.1364/oe.19.010735](https://doi.org/10.1364/oe.19.010735).

[14] D. Contini et al., "Effects of time-gated detection in diffuse optical imaging at short source-detector separation," *J. Phys. D, Appl. Phys.*, vol. 48, no. 4, Feb. 2015, Art. no. 045401, doi: [10.1088/0022-3727/48/4/045401](https://doi.org/10.1088/0022-3727/48/4/045401).

[15] M. Kannan et al., "Fast, in vivo voltage imaging using a red fluorescent indicator," *Nature Methods*, vol. 15, no. 12, pp. 1108–1116, Dec. 2018, doi: [10.1038/s41592-018-0188-7](https://doi.org/10.1038/s41592-018-0188-7).

[16] P.-Y. Fortin, C. Genevois, A. Koenig, E. Heinrich, I. Texier, and F. Couillaud, "Detection of brain tumors using fluorescence diffuse optical tomography and nanoparticles as contrast agents," *J. Biomed. Opt.*, vol. 17, no. 12, Dec. 2012, Art. no. 126004, doi: [10.1117/1.JBO.17.12.126004](https://doi.org/10.1117/1.JBO.17.12.126004).

[17] J. J. Selb, D. K. Joseph, and D. A. Boas, "Time-gated optical system for depth-resolved functional brain imaging," *J. Biomed. Opt.*, vol. 11, no. 4, pp. 1–13, 2006, doi: [10.1117/1.2337320](https://doi.org/10.1117/1.2337320).

[18] D. E. Schwartz, E. Charbon, and K. L. Shepard, "A single-photon avalanche diode array for fluorescence lifetime imaging microscopy," *IEEE J. Solid-State Circuits*, vol. 43, no. 11, pp. 2546–2557, Nov. 2008, doi: [10.1109/JSSC.2008.2005818](https://doi.org/10.1109/JSSC.2008.2005818).

[19] S. Moazeni et al., "A mechanically flexible, implantable neural interface for computational imaging and optogenetic stimulation over 5.4 \times 5.4 mm² FoV," *IEEE Trans. Biomed. Circuits Syst.*, vol. 15, no. 6, pp. 1295–1305, Dec. 2021, doi: [10.1109/TBCAS.2021.3138334](https://doi.org/10.1109/TBCAS.2021.3138334).

[20] R. Inti et al., "A scalable 32-to-56Gb/s 0.56-to-1.28pJ/b voltage-mode VCSEL-based optical transmitter in 28 nm CMOS," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2021, pp. 1–2, doi: [10.1109/CICC51472.2021.9431548](https://doi.org/10.1109/CICC51472.2021.9431548).

[21] Y. Moghe, T. Lehmann, and T. Piessens, "Nanosecond delay floating high voltage level shifters in a 0.35 μ m HV-CMOS technology," *IEEE J. Solid-State Circuits*, vol. 46, no. 2, pp. 485–497, Feb. 2011, doi: [10.1109/JSSC.2010.2091322](https://doi.org/10.1109/JSSC.2010.2091322).

[22] J. A. Tatum, D. Smith, J. K. Guenter, and R. H. Johnson, "High-speed characteristics of VCSELs," *Proc. SPIE*, vol. 3004, pp. 151–159, May 1997, doi: [10.1117/12.273829](https://doi.org/10.1117/12.273829).

[23] J. M. Pavia, M. Scandini, S. Lindner, M. Wolf, and E. Charbon, "A 1 \times 400 backside-illuminated SPAD sensor with 49.7 ps resolution, 30 pJ/sample TDCs fabricated in 3D CMOS technology for near-infrared optical tomography," *IEEE J. Solid-State Circuits*, vol. 50, no. 10, pp. 2406–2418, Oct. 2015, doi: [10.1109/JSSC.2015.2467170](https://doi.org/10.1109/JSSC.2015.2467170).

[24] D. Bronzi, F. Villa, S. Tisa, A. Tosi, and F. Zappa, "SPAD figures of merit for photon-counting, photon-timing, and imaging applications: A review," *IEEE Sensors J.*, vol. 16, no. 1, pp. 3–12, Jan. 2016, doi: [10.1109/JSEN.2015.2483565](https://doi.org/10.1109/JSEN.2015.2483565).

[25] A. Torricelli et al., "Time domain functional NIRS imaging for human brain mapping," *NeuroImage*, vol. 85, pp. 28–50, Jan. 2014, doi: [10.1016/j.neuroimage.2013.05.106](https://doi.org/10.1016/j.neuroimage.2013.05.106).

[26] B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *J. Biomed. Opt.*, vol. 11, no. 4, pp. 1–16, 2006, doi: [10.1117/1.2335429](https://doi.org/10.1117/1.2335429).

[27] J.-P. Bouchard, I. Veilleux, R. Jedidi, I. Noiseux, M. Fortin, and O. Mermut, "Reference optical phantoms for diffuse optical spectroscopy. Part 1—Error analysis of a time resolved transmittance characterization method," *Opt. Exp.*, vol. 18, no. 11, pp. 11495–11507, May 2010, doi: [10.1364/OE.18.011495](https://doi.org/10.1364/OE.18.011495).

[28] C. Veerappan et al., "A 160 \times 128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter," in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 2011, pp. 312–314, doi: [10.1109/ISSCC.2011.5746333](https://doi.org/10.1109/ISSCC.2011.5746333).

[29] L. Di Sieno et al., "Probe-hosted large area silicon photomultiplier and high-throughput timing electronics for enhanced performance time-domain functional near-infrared spectroscopy," *Biomed. Opt. Exp.*, vol. 11, no. 11, pp. 6389–6412, Oct. 2020, doi: [10.1364/BOE.400868](https://doi.org/10.1364/BOE.400868).



Sajjad Moazeni (Member, IEEE) received the B.S. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2013, and the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of California at Berkeley, Berkeley, CA, USA, in 2016 and 2018, respectively.

From 2018 to 2020, he was a Post-Doctoral Research Scientist with the Bioelectronic Systems Laboratory, Columbia University, New York, NY, USA. He is currently an Assistant Professor with the Electrical and Computer Engineering Department, University of Washington, Seattle, WA, USA. His research interests are designing integrated systems using emerging technologies, integrated photonics, neurophotonics and biophotonics, and analog/mixed-signal integrated circuits.

Dr. Moazeni was a recipient of the 2022 NSF CAREER Award.



Kevin Renahan (Student Member, IEEE) received the B.S.E. degree in biomedical engineering from the University of Michigan, Ann Arbor, MI, USA, in 2016, and the M.S. degree in electrical engineering from Columbia University, New York, NY, USA, in 2018, where he is currently pursuing the Ph.D. degree in electrical engineering.

His research interests are in CMOS time-of-flight imagers for applications in diffuse optical tomography.



Eric H. Pollmann (Student Member, IEEE) received the B.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2017, and the M.S. degree in electrical engineering from Columbia University, New York, NY, USA, in 2018, where he is currently pursuing the Ph.D. degree in electrical engineering.

His research interests are in implantable CMOS fluorescence imagers for applications in biology and neuroscience.



Kenneth L. Shepard (Fellow, IEEE) received the B.S.E. degree from Princeton University, Princeton, NJ, USA, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1988 and 1992, respectively.

From 1992 to 1997, he was a Research Staff Member and the Manager with the VLSI Design Department, IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA, where he was responsible for the design methodology for IBM's G4S/390 microprocessors. He was the Chief Technology Officer with CadMOS Design Technology, San Jose, CA, USA, until its acquisition by Cadence Design Systems in 2001. Since 1997, he has been with Columbia University, New York, NY, USA, where he is currently the Lau Family Professor of electrical engineering and biomedical engineering, and the Co-Founder and the Chairperson of the Board of Ferric, Inc., New York, which commercializes technology for integrated voltage regulators. His current research interests include CMOS bioelectronics and power electronics.

Dr. Shepard has been an Associate Editor of the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, the IEEE JOURNAL OF SOLID-STATE CIRCUITS, and the IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.