

MIT Open Access Articles

*Adaptive optics wavefront sensors based
on photon-counting detector arrays*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Brian F. Aull, Daniel R. Schuette, Robert K. Reich and Robert L. Johnson, "Adaptive optics wavefront sensors based on photon-counting detector arrays", Proc. SPIE 7736, 773610 (2010). © 2010 COPYRIGHT SPIE

As Published: <http://dx.doi.org/10.1117/12.858006>

Publisher: SPIE

Persistent URL: <http://hdl.handle.net/1721.1/60971>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Adaptive optics wavefront sensors based on photon-counting detector arrays^{*}

Brian F. Aull^{*a}, Daniel R. Schuette^a, Robert K. Reich^a, Robert L. Johnson^b

^aMIT Lincoln Laboratory, 244 Wood St, Lexington, MA, USA 02420-9108; ^bStarfire Optical Range, Kirtland Air Force Base, Albuquerque NM, USA 87117

*aull@ll.mit.edu; phone 1 781 981-4676; fax 1 781 981-3433; www.ll.mit.edu

ABSTRACT

For adaptive optics systems, there is a growing demand for wavefront sensors that operate at higher frame rates and with more pixels while maintaining low readout noise. Lincoln Laboratory has been investigating Geiger-mode avalanche photodiode arrays integrated with CMOS readout circuits as a potential solution. This type of sensor counts photons digitally within the pixel, enabling data to be read out at high rates without the penalty of readout noise. After a brief overview of adaptive optics sensor development at Lincoln Laboratory, we will present the status of silicon Geiger-mode-APD technology along with future plans to improve performance.

Keywords: adaptive optics, wavefront sensing, photon counting, Geiger-mode avalanche photodiodes

1. INTRODUCTION

Adaptive optics systems for ground-based astronomy and space surveillance require sensors to measure the distortion of a wavefront due to atmospheric turbulence. The Shack-Hartmann technique uses arrays of lenslets with quad-cell detectors to measure wavefront tilt.¹ In scenarios with weak beacon signals or fast wavefront sampling rates, the performance of such sensors is often limited by detector readout noise. For example, if the wavefront needs to be sampled at 50 kHz, as in some surveillance applications, a charge-coupled device- or CMOS active pixel sensor-based wavefront sensor can have tens of electrons per pixel of readout noise at the required readout rates. The signal can be amplified by charge amplification chains² to overcome the readout noise, but the readout latency associated with such techniques limits the wavefront correction rate.

One way to eliminate readout noise at high readout rates is to use a photon-counting detector that digitizes photon detections by triggering a CMOS digital pixel circuit. This eliminates analog signals and circuits from the readout chain, enabling quantum-limited sensitivity. Furthermore the CMOS architecture is flexible and can be configured for low-latency readout and selective region-of-interest readout.

Lincoln Laboratory has developed arrays of Geiger-mode avalanche photodiodes hybridized to all-digital CMOS readout circuits for applications such as lidar³ and low-light passive imaging. In Geiger mode, the avalanche photodiode (APD) is biased several volts *above* the avalanche breakdown voltage. The absorption of a single photon initiates an avalanche discharge that produces a digital pulse directly from the detector. The APD can be connected directly to a CMOS logic element within a digital pixel circuit that counts or time stamps photon detection events. This “photon-to-digital conversion” technique has been used to build focal plane imagers for high-performance lidar systems.⁴

Recently, Lincoln Laboratory has been developing specialized Geiger-mode APD arrays for Shack-Hartmann wavefront sensors⁵ designed to operate at tens of kilohertz frame rates. The arrays consist of 2×2 subarrays or quad cells, each of which has a high fill factor resulting in low light loss in the central region. The array is hybridized to a CMOS readout circuit with a digital counter for each APD. At the end of an integration time, the counters are read out and centroiding computations are done off chip to determine the local wavefront tilt at each subarray. The following sections describe the architecture of the high-fill-factor APDs and give measured performance for the present APDs and integrated detectors.

^{*}This work was sponsored by the Air Force Research Laboratory Starfire Optical Range (AFRL SOR) under the Department of the Air Force contract number FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

2. APD ARRAYS

2.1 Device concept

Figure 1 shows a cross section of two neighboring Geiger-mode APDs designed for lidar applications. The detector is formed by selective-area ion implantation of dopants into a very lightly p-doped epitaxial layer grown on a silicon substrate. The substrate is ultimately removed for back-illuminated operation, and the back surface passivated with a thin ion-implanted p^+ contact layer. The APD has a p-i-p-i-n structure (reading from the bottom up and considering the epitaxial material to be close to intrinsic). When biased to detect light, the lower “i” layer functions as an absorber of signal photons; it has modest electric fields, sufficient to sweep photoelectrons into the upper “i” layer. The upper “i” layer functions as a multiplier; it has much stronger electric fields, sufficient to cause impact ionization that initiates the avalanche discharge.

In the structure shown in Figure 1, the n^+ implant on the front side extends well beyond the multiplier region, forming a guard-ring diode that collects dark current generated in the region between pixels. This dark current does not undergo avalanche multiplication and hence does not cause spurious detection events. The guard-ring structure also prevents breakdown at the periphery of the diode, since the fields there are much lower than in the multiplier. The price paid for these advantages is intrinsically low fill factor; photoelectrons generated in between pixels are unlikely to be detected. This is not a problem for applications that can use microlenses or other light concentrating techniques.

Figure 2 illustrates the specialized high-fill-factor design, showing a cross section through two APDs that comprise half of a quad cell. There is a stepped p^+ implant formed by ion implantation of boron through a stepped oxide mask. The shallow portion of this stepped implant separates the absorber and multiplier portions of each detector. The step lowers the electric field at the edges of the diode, preventing edge breakdown and forming a guard ring to collect surface-generated dark current without triggering Geiger-mode events. The deep portion of the implant, which is partially undepleted, prevents the guard ring from collecting photoelectrons generated in the absorber; as indicated in the figure, these photoelectrons reach a nearby multiplier region by a combination of diffusion and drift.

There are a number of technical challenges to be overcome with this detector architecture. First, a Geiger-mode APD emits hot-carrier radiation from the avalanche region during a discharge; some of the emitted photons trigger spurious detection events in neighboring APDs, or alternatively, the photoelectrons thus created can retrigger the emitting detector or neighbors later on. Second, dark current generated in the absorber is not collected by a guard ring structure as in the lidar design; the “fill factor” for thermally generated electrons is correspondingly higher. Third, the electric fields at the periphery of the n^+ implant, while low enough to avoid edge breakdown, may be high enough to cause tunneling current and parasitic avalanche multiplication, triggering spurious events from the periphery of the multiplier. Each of these non-idealities can cause spurious events that degrade the signal-to-noise ratio.

2.2 APD array fabrication, hybridization, and packaging

The APD quad-cell arrays are fabricated on six-inch silicon wafers in Lincoln’s Microelectronics Laboratory using standard VLSI processing techniques. Each quad cell occupies a $100 \times 100\text{-}\mu\text{m}^2$ footprint with a $50 \times 50\text{-}\mu\text{m}^2$ APD detector area. The quad cells are laid out on a $200\text{-}\mu\text{m}$ pitch (the inset in Figure 2 shows the quad cell layout). Both 16×16 and 32×32 quad-cell arrays have been fabricated. The n^+ regions on the front side of the diodes is formed by masked ion implantation of arsenic. Likewise, the stepped p^+ layer is formed by masked ion implantation of boron. The steps are created by implantation through mesas of oxide. In the original design, the step in the p^+ layer was relatively shallow ($0.35\text{ }\mu\text{m}$). As discussed below, this led to very high dark count rates due to periphery leakage currents. In a later design, the step was made deeper ($1.1\text{ }\mu\text{m}$) and a phosphorus implant was done around the periphery of the arsenic implant; this results in substantial improvement. In some arrays, metal-filled trenches are incorporated on the front side between APDs in order to suppress optical crosstalk by blocking emitted light. Contact to the back-surface p^+ layer is facilitated by implanting a p^+ “moat” implant around the periphery of each array. Planar metallization for later bump bond attachment is patterned over tungsten plug contacts. When hybridized to a CMOS chip, bump bonds are made both to the APDs and to the moat. In addition to the arrays, there are numerous test devices with various sizes and geometries.

Detector arrays using compound semiconductors (HgCdTe or InGaAs) have commonly been flip-chip bonded to CMOS readout chips and illuminated from the substrate side. In these cases, the detector is heteroepitaxial, grown on a higher-

bandgap substrate (CdZnTe or InP) that is optically transparent at the wavelength of operation. The arrays reported here are made on homoepitaxial silicon; the silicon substrate is opaque and therefore must be removed entirely, leaving a detector that is typically less than 20 μm in thickness. The thin silicon layer can greatly complicate handling during fabrication. Lincoln Laboratory has developed a technique⁶ for replacing the silicon substrate with a quartz substrate using oxide wafer bonding, avoiding the use of epoxy or other organic adhesives, which limits certain device processing. Fabrication concludes by opening contacts on the front side and then forming under-bump metallization and indium bumps. The wafer is then sawn into individual arrays which are then bump bonded to CMOS readout chips. The corresponding bump bonding pads on the CMOS chips are plated with an under-bump metal.

Figure 3 is a photomicrograph showing a quadrant of a thinned APD quad-cell array on quartz that has been bump bonded to a CMOS readout chip. The hybridized device is mounted in a hermetically sealed pin-grid-array package with a built-in thermoelectric cooler and thermistor temperature sensor.

3. CMOS READOUTS

3.1 Readout chip architecture

Readout chips for both 16×16 and 32×32 quad-cell arrays have been fabricated in the TSMC 0.35- μm digital CMOS process using the MOSIS foundry service. Figure 4 is a block diagram of the pixel in the 16×16 chip. The p terminals of all the APDs are biased at a negative voltage whose magnitude is just below avalanche breakdown (the figure shows -25V, which is a typical value). The n⁺ terminal of each APD is connected to a 5V logic circuit that “arms” the APD by pulling the terminal to logic high, biasing the APD above breakdown. When the APD detects a photon and begins to discharge, a sensing transistor (Q2) detects the voltage droop and turns on a pull-down transistor (Q3), rapidly discharging the n terminal to logic low. At this point, the APD is biased below breakdown, causing the avalanche to die out within a few nanoseconds. The APD is then rearmed by turning on Q1 after a minimum delay that is dictated by a synchronous timing circuit driven by a global polling clock. The use of a pull-down transistor to assist the APD discharge, known as active quenching, reduces the amount of charge that flows directly through the APD, correspondingly reducing carrier trapping and light emission.

The timing circuit includes a toggle flip-flop that changes state whenever the APD fires. This state then propagates down a chain of three D flip-flops that update on alternate edges of the polling clock. The outputs of the second and third flip-flops are input to an exclusive OR gate. If the APD fires, this gate produces a pulse that both turns on the rearm transistor Q1 and increments a 10-bit counter. The duration of the high portion of the polling clock is the minimum delay between the APD discharge and the rearm. This must be long enough to allow the avalanche to turn off and for any electrons or holes trapped in the multiplier to be released and collected, so that they do not trigger spurious events, known as afterpulses, upon rearming the APD. For silicon at room temperature, tens of nanoseconds is typically sufficient.

The readout circuitry multiplexes each 10-bit counter to a 10-bit parallel output. This scheme allows for low readout latency (20 μs), which facilitates high-frame-rate wavefront sensing.

The 32×32 chips use the same principles of operation. The most recent version, however, uses cascoded pull-up transistors to achieve higher APD voltage swings, up to 8V. This potentially offers higher photon detection efficiency and more rapid photoelectron collection. In this case, the 8V discharge from APD is ac coupled to produce a 3.3V logic signal that is compatible with the digital pixel circuitry.

3.2 Readout chip verification

Before being bump bonded to APD arrays, the CMOS chips are tested. As shown in Figure 4, each pixel has a pull-down transistor (Q4) that can be activated by a global disarm signal, simulating an APD discharge in every pixel. This feature is used to verify that the counting is functioning properly. In addition, a discharge can be simulated in selected pixels using a bright spot of light. Photocurrent in the drain-substrate junctions of the pull-down transistors is the most likely mechanism for this. This experiment verifies proper readout and display of the frames. Both the 16×16 and 32×32 readout chips have demonstrated proper function.

4. RESULTS

4.1 Packaged Device

A 16×16 quad cell array was packaged in a 68-pin hermetic package with a built-in TE cooler and temperature sensor. A printed circuit board was fabricated with a ZIFF socket for the package and circuitry for control and readout of the device. An Altera FPGA on the board has firmware that generates the clocking and control signals, and relays the pixel data to a set of Camera Link interface chips. A PC with a Camera Link interface card is used to acquire the data and perform display and processing with IDL-based software.

This was the first packaged device fabricated with the quartz wafer bonding technique. It has crosstalk suppression trenches on the front side, but is of the original design that has a shallow step in the boron implant and no phosphorus implant. Also, there are no antireflection coatings in the structure, so reflection losses are expected, particularly at the quartz/silicon interface.

An important measurement to determine centroiding performance is the intrapixel response, which is the partitioning of counts between two halves of a quad cell as a function of the position of a small light spot ($\sim 3\text{ }\mu\text{m}$). The response should be monotonic and change relatively quickly with spot position. For this measurement, a 684-nm Picoquant laser was attenuated to give roughly 1 detected photoelectron per light pulse. It was focused on the device and its position manually scanned up the middle of a quad cell in 5- μm increments. Figure 5 shows the normalized partition between the top and bottom halves as a function of spot position. The transition region is monotonic, indicating the device has a high fill factor, with a steep central region extending over 10 μm . Measurements have been made of dark count rate, photon detection efficiency, and correlation between the number of detection events of neighboring APDs within a quad cell. These parameters all vary with APD bias, integration time, and CMOS triggering threshold (the sensing transistor source voltage V_{ref}). Figure 6 shows dark count rate and photon detection efficiency versus p-terminal bias. This measurement was done at a temperature of -10C, with V_{ref} set to 4.35V. The detection efficiency rises to about 14% and saturates at low ($<1\text{V}$) overbias, likely because of blocking losses from the dark counts. The detection efficiency is limited by high ($\sim 40\%$) reflection losses at the quartz-silicon interface and by the low overbiases achievable due to the dark counts. The source of the high dark count rates is discussed in the next section.

4.2 Wafer probe dark count rates

Extensive wafer probing of APDs with both the high-fill-factor design and the lidar design was carried out on wafers. Most of this effort was aimed at investigating the dark count mechanisms in the high-fill-factor APDs, which as originally designed have very high dark count rates. Dark count rates were measured both with passive quenching, in which the APD is allowed to discharge itself and then recharge through a 200-k Ω ballast resistor, and with active quenching. During the course of this testing, heating or cooling was used to get temperatures as high as 40C and as low as -70C.

The original high-fill-factor design had a stepped p^+ implant that was formed by ion implantation over an oxide mask that was 0.5 μm thick; capacitance-voltage (CV) profiling shows that the deep part of the implant is only 0.35 μm deeper than the shallow part. In passive quenching measurements, these devices display a precipitous increase in dark count rate at small overbiases (1 V above breakdown), quickly reaching several MHz; this does not go away when the device is chilled. A 50- μm -diameter lidar-style device from the same wafer, in contrast, is well behaved, displaying a dark count rate of 6 kHz at room temperature and 4V overbias and 20 Hz at -28C and 6V overbias. The relatively shallow step is hypothesized to create a high electric field between the n^+ and p layers causing tunneling current and parasitic avalanching in the peripheral region of the junction. In addition, simulations have shown that electric field lines converge on the convex part of the junction at the edge of the n^+ implant, which is done with As.

An improved APD design incorporated two modifications to lower the fields at the periphery. First, the stepped boron implant was done at higher energy and through a thicker (1.5- μm) masking oxide, increasing the distance between the n^+ and p layers in the peripheral region of the device. CV profiling shows a 1.1- μm step in the implant depth. Second, a phosphorus implant was done at the edge of the arsenic implant; phosphorus diffuses much more than arsenic during implant activation, leading to decreased junction curvature and lower fields. With cooling to -50C, a test device with these improvements showed improved dark count rates compared to the shallow-step structures with rates less than 1 kHz at biases up to 4.5V above breakdown. However, there is still a rapid increase in dark count rate which occurs at 5V above breakdown. As the temperature is raised, the breakdown voltage increases as expected for silicon APDs (with a slope of 0.03V/C), but the voltage at which the rapid increase in dark count rates occurs does not shift significantly.

To investigate the rapid increase of dark count rates occurring at modest overbiases, test devices were measured that have a single high-fill-factor APD surrounded by simple p-i-n diodes made by blocking the boron implant; the pixel pitch in these arrays is 100 μm . This experiment was performed at room temperature, where the rapid increase in the dark count rate is observed at 2V above breakdown in improved-design devices. When the neighboring p-i-n diodes are reverse biased, they are expected to act as “scuppers,” that is, to collect thermally generated dark current that might otherwise trigger the APD. This is borne out by experiment. If the neighboring diodes are left open circuited, the dark count rate at 2.5V above breakdown is too high to measure with a passive quenching circuit; the APD never turns off long enough to recharge. A bias as low as 2V on the neighboring diodes decreases the dark count rate dramatically. The train of pulses observed, however, is bursty; there are quiet periods corresponding to a 20-kHz dark count rate interspersed with bursts of pulses.

In order to further investigate this bursty afterpulsing, high-fill-factor devices were measured with active quenching, varying the quenching time. These devices were also compared to a similar-sized lidar APD with a stepped boron implant with a diameter smaller than the As implant (the insets in Figure 7 show the two APD structures). Both devices had no scuppers and fairly large (28- μm) multiplier diameters. A LeCroy scope was used to histogram the delay between the end of the quenching time and the first subsequent detection; the inverse of the average value of this delay is a measure of dark count rate. The dark count rate was measured as a function of the quench time, and the results at room temperature are shown in Figure 7. The lidar device with the deep stepped implant had a dark count rate of 90 kHz, independent of quench times ranging from 10 μs to 100 μs . The high-fill-factor device, on the other hand, showed a pronounced dependence on quench time, varying from 6.6 MHz at 10 μs quench time to 180 kHz at 100 μs .

The high-fill-factor device requires a quenching time of several tens of μs in order for its dark count rate to become insensitive to further prolongation of quenching. This time scale is much too long for typical detrapping times. Furthermore, if the afterpulsing is mediated by trapped carriers in or adjacent to the multiplier region, the lidar device with the stepped implant would be expected to display the same phenomenon, as it has very similar doping and electric field profiles in that region. Tens of μs is a time scale that is commensurate with the minority carrier (electron) recombination time in the absorber and also with the time required for excess minority carriers to diffuse laterally out of the absorber region of devices with $\sim 30\text{-}\mu\text{m}$ diameter. The present hypothesis is that the light emitted from the multiplier region during an avalanche discharge is absorbed in the underlying absorber layer, creating photoelectrons that retrigger the APD. In the lidar device, this retriggering is mitigated, at least for smaller device diameters, by the proximity of the depletion region associated with the guard ring diode, which sweeps out and collects most of the photoelectrons.

In summary, the high dark count rates observed in high-fill-factor APDs come from three sources: thermally generated current, edge leakage due to tunneling, possibly worsened by parasitic avalanche gain, and optical emission leading to self retriggering. Quad-cell array designs now in fabrication, in addition to the deeper boron implant step and phosphorus guard ring implant, have built-in scupper diodes to collect both thermally generated current and photoelectrons created by light emission. It should also be noted that the APDs in the wafer probe experiments are discharging large (20 pF) capacitances associated with coaxial wires and scope probes. When the APD is bump bonded to a CMOS readout circuit, this capacitance is smaller by two orders of magnitude, which is expected to mitigate light emission by the same factor.

5. CONCLUSIONS

We have demonstrated a high-fill-factor, 16×16 quad cell array with monotonic centroiding response with a FWHM $< 10 \mu\text{m}$. Devices with a 32×32 format are in fabrication, and the higher-voltage CMOS chips for these have been fabricated and their functionality verified.

The dark count rate of the high-fill-factor devices is high compared with other detector technologies, and the mechanisms for this have been investigated. The dark count rate is hypothesized to arise from a combination of tunneling current at the junction periphery, unscuppered thermal dark current, and optical self retriggering. We expect substantial improvement from the design features in devices currently in fabrication. Decreasing the dark count rate will enable operation at higher biases, resulting in improved detection efficiency.

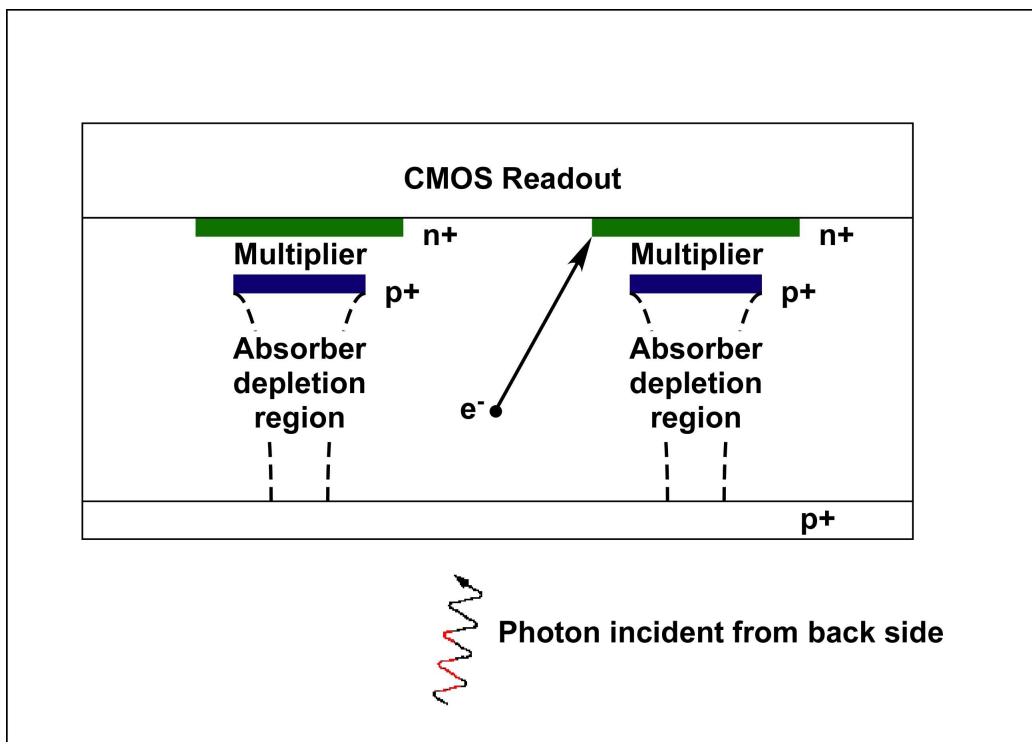


Figure 1. Cross section through two neighboring lidar-style APDs.

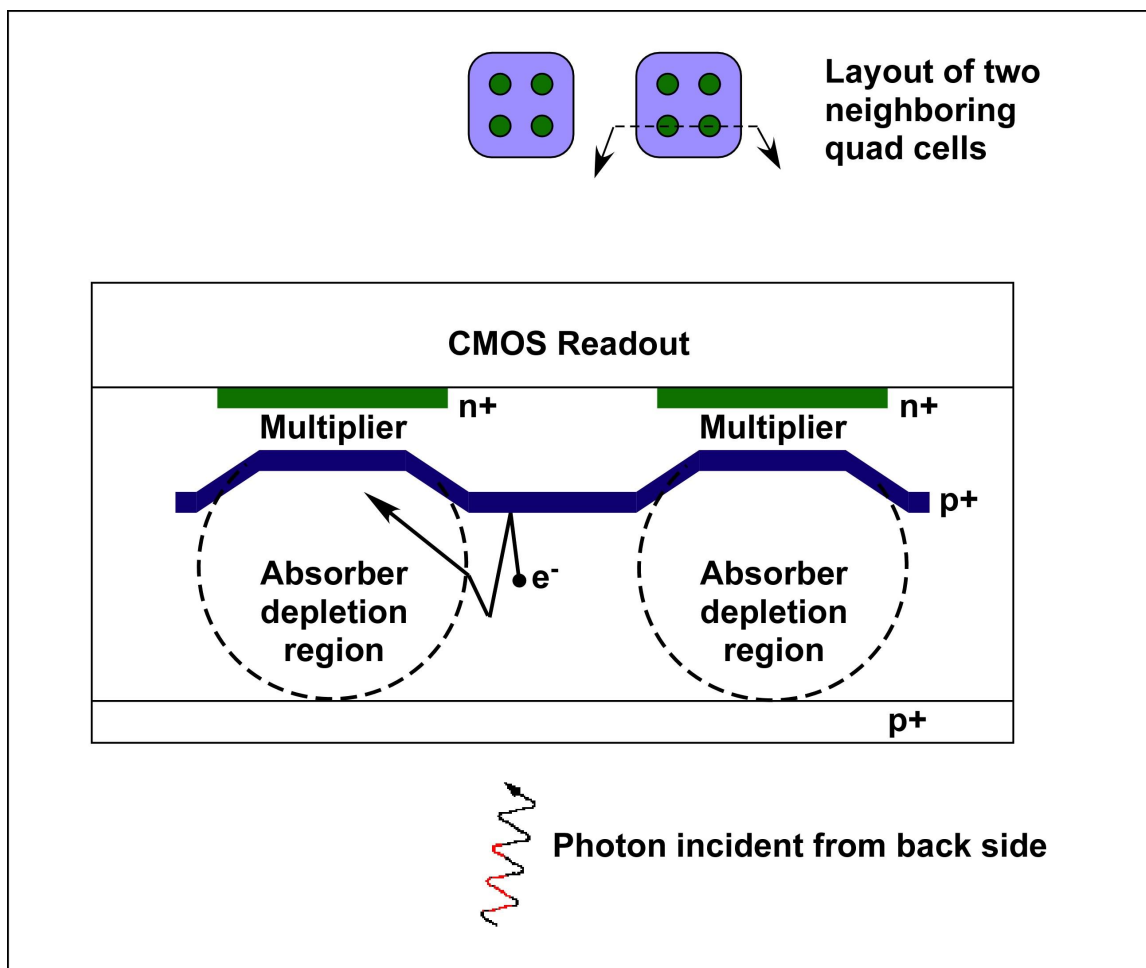


Figure 2. Cross section through two neighboring high-fill-factor APDs within a quad cell. Upper inset shows the layout of two neighboring quad cells, with a dashed line indicating the location of the cross section.

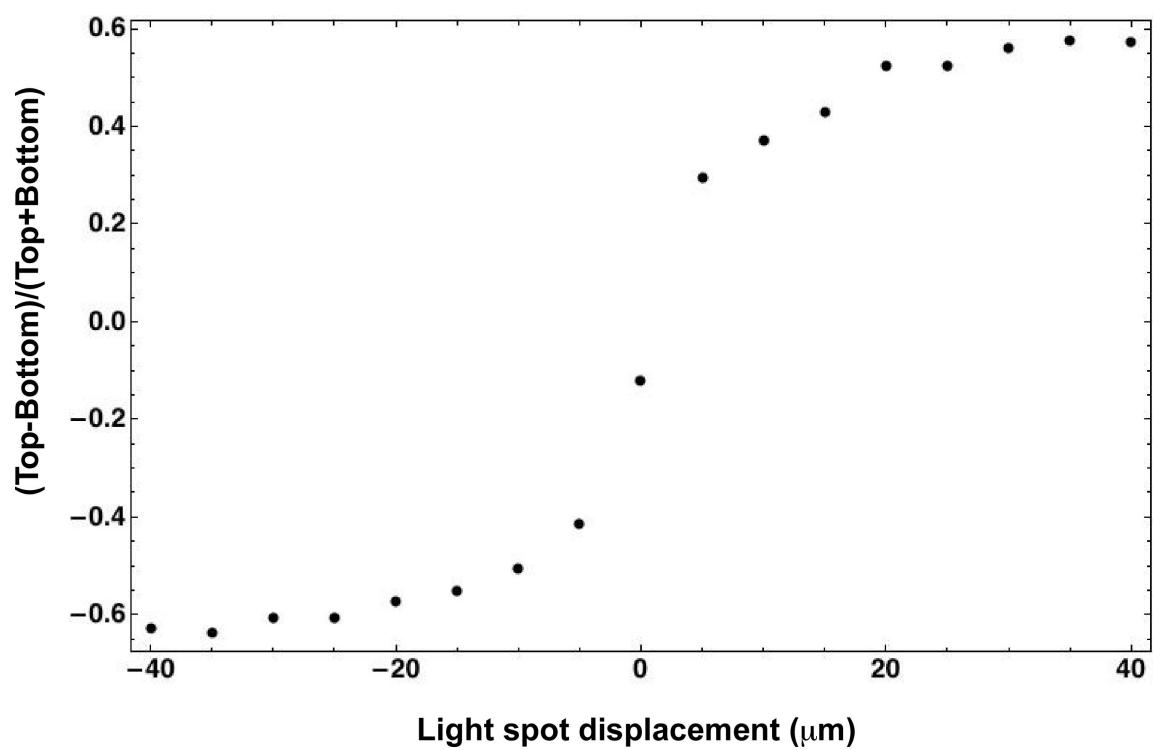


Figure 5. Intrapixel response.

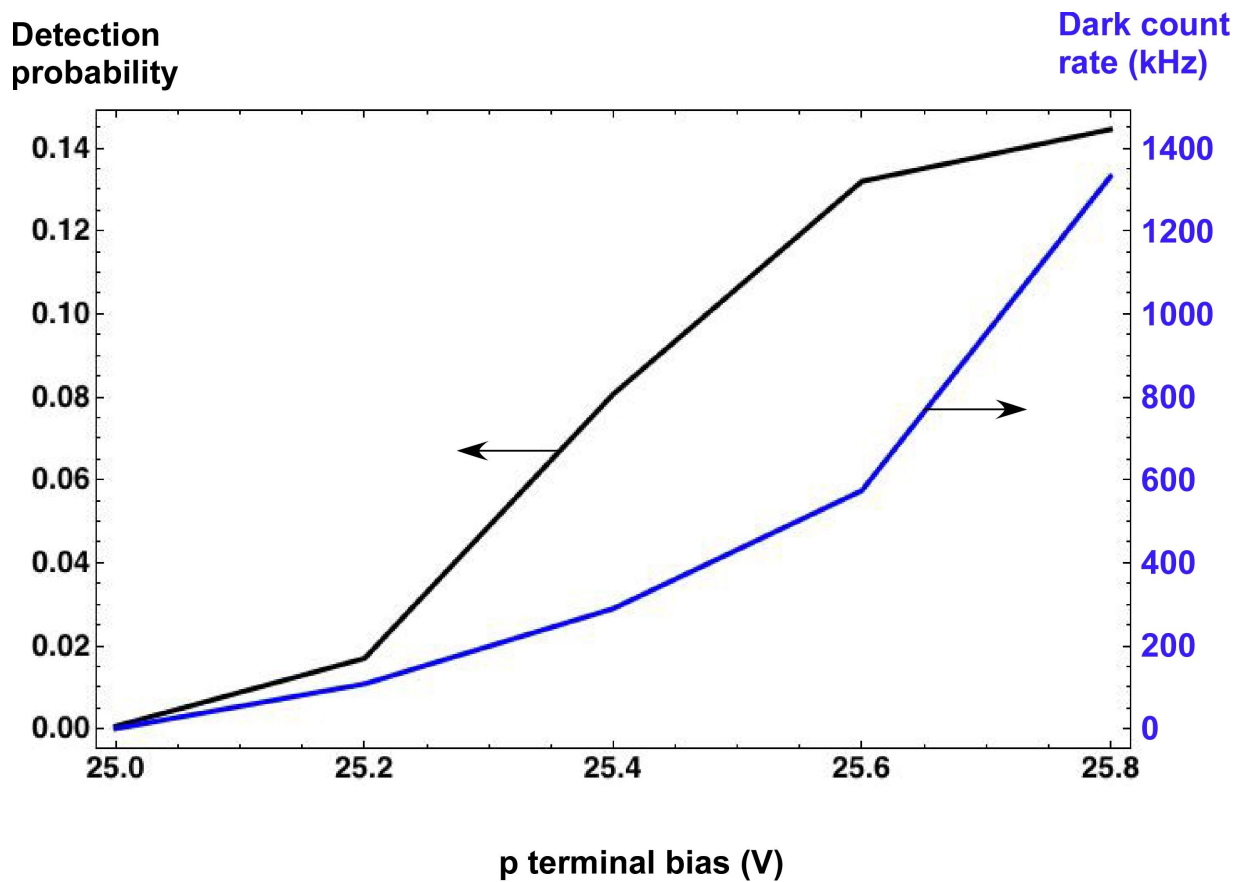


Figure 6. Detection efficiency and dark count rate versus bias.

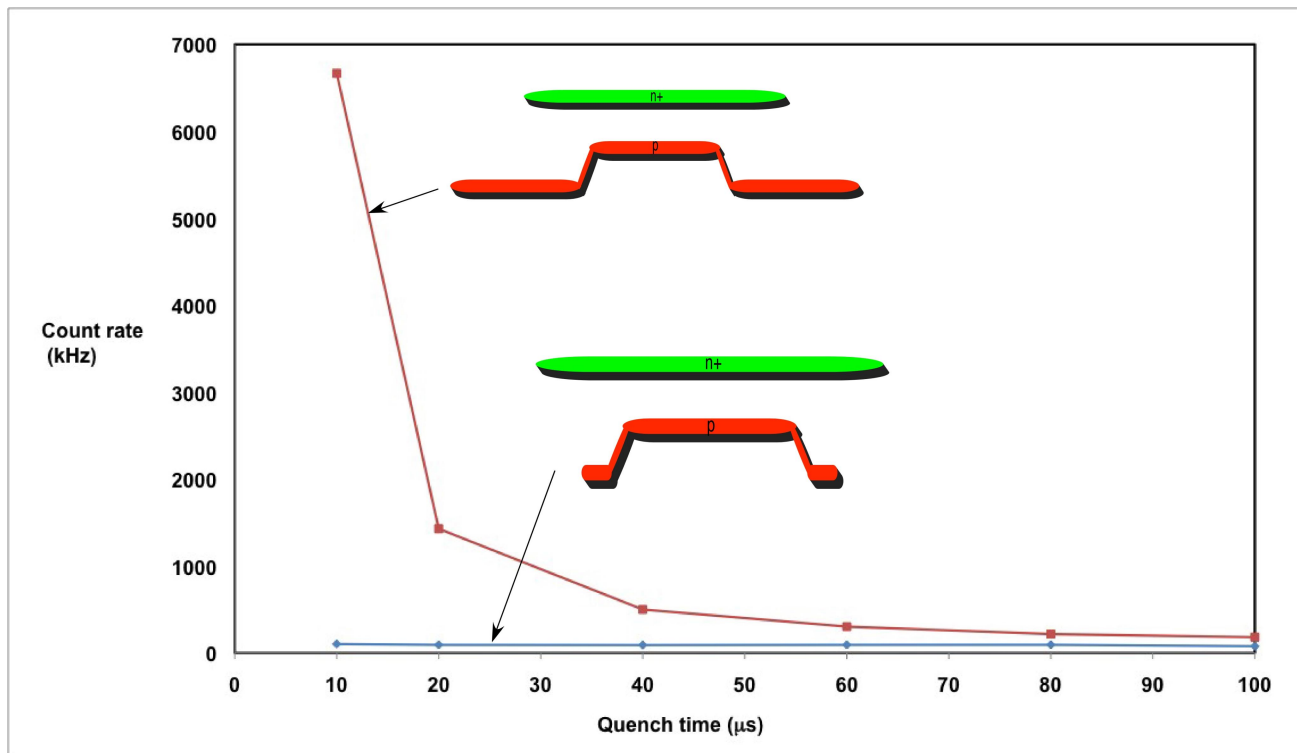


Figure 7. Dark count rate, as measured by the inverse of the mean time to first detection event, versus quench time. The blue curve is for the lidar-style APD with the truncated stepped boron implant, and the red curve is for the high-fill-factor APD.

REFERENCES

- [1] Shack, R.B. and Platt, B.C., "Production and use of a lenticular Hartmann screen," *J. Opt. Soc. Am.* 61(5), 656 (1971).
- [2] Downing, M., Hubin, N., Kasper, M., Jorden, P., Pool, P., Denney, S., Suske, W., Burt, D., Wheeler, P., Hadfield, K., Feautrier, P., Gach, J., Reyes, J., Meyer, M., Baade, D., Balard, P., Guillaume, C., Stadler, E., Boissin, O., Fusco, T., Diaz, J., "A dedicated L3vision CCD for adaptive optics applications," in J.E. Beletic et al. (eds.), *Scientific Detectors for Astronomy 2005*, Springer, The Netherlands, 321-328 (2006).
- [3] Aull, B. F., Loomis, A. H., Young, D. J., Heinrichs, R. M., Felton, B. J., Daniels, P. J., and Landers, D. J., "Geiger-mode avalanche photodiodes for three-dimensional imaging," *Lincoln Laboratory Journal* 13(2), 335-350 (2002).
- [4] Marino, R. M. and Davis, W. R., "Jigsaw: A foliage-penetrating 3D imaging laser radar system," *Lincoln Laboratory Journal* 15(1), 27-36 (2005).
- [5] Aull, B. F., Renzi, M. J., Loomis, A. H., Young, D. J., Felton, B. J., Lind, T. A., Craig, D. M., and Johnson, R. L., "Geiger-mode quad-cell array for adaptive optics," in *Conference on Lasers and Electro-Optics*, 2008 Technical Digest, (Optical Society of America, Washington, DC, 2008), CWF2.
- [6] Schuette, D. R., Westhoff, R. C., Loomis, A. H., Young, D. J., Ciampi, J. S., Aull, B. F., and Reich, R. K., "Hybridization process for back-illuminated silicon Geiger-mode avalanche photodiode arrays," *Proc. SPIE* 7681, to be published (2010).