Geiger-Mode APD Single Photon Detectors

Mark A. Itzler, Xudong Jiang, Rafael Ben-Michael, Bruce Nyman, and Krystyna Slomkowski
Princeton Lightwave Inc., 2555 US Route 130 S., Cranbury, NJ 08512 USA
mitzler@princetonlightwave.com

Abstract: We present experimental and modeling results for InP-based avalanche photodiodes designed for single photon counting in the wavelength range of 0.95–1.65 μm. We describe performance trade-offs and underlying mechanisms governing key device characteristics.

©2008 Optical Society of America
OCIS codes: (040.1345) Avalanche photodiodes (APDs); (030.5260) Photon counting

1. Introduction

The emergence of photonics applications requiring single photon detection has been driving significant advances in photodetectors with single photon sensitivity [1]. The exploitation of the quantum properties of photons for quantum cryptography and other quantum information processing techniques is critically dependent on single photon detection. Likewise, single photon sensitivity is a key enabler for photon-starved applications such as long-range free-space communications and flash ladar three-dimensional imaging. These applications are often best served by single-photon avalanche diodes (SPADs), which can achieve high performance consistent with practical operating conditions and robust deployment.

SPADs are avalanche diodes that are operated above their breakdown voltage $V_b$ in the so-called “Geiger mode”, for which the absorption of a single photon can generate a macroscopic current discernible using electronic threshold detection. By reverse biasing beyond $V_b$ at $V = V_b + \Delta V$, where $\Delta V$ is the excess bias, the SPAD is put into a metastable state in which a single photoexcited carrier can create a self-sustaining avalanche. After detection, the avalanche must be quenched by lowering the bias. Since threshold detection is digital, it is essentially noiseless; however, false counts (i.e., dark counts) can be caused by dark carrier generation via thermal or tunneling processes.

Though structurally related to more familiar “linear mode” APDs (for which the output photocurrent is linearly proportional to the input optical power), SPADs require different considerations for optimized designs [2,3]. They are generally operated at temperatures well below room temperature to reduce the thermally generated contribution to the dark count rate (DCR). The probability that an input photon is detected—the photon detection efficiency (PDE)—is intimately related to the avalanche breakdown probability, which is a strong function of the electric field amplitude in the avalanche region. PDE can be increased by imposing larger electric fields, but at the expense of increased DCR due to greater tunneling in the avalanche region. Significant performance enhancement can be realized for counting photons at shorter wavelengths by using an optimal stoichiometry for the InGaAsP absorption region. We demonstrate this optimization for shorter wavelength using both modeling and experimental results for single photon detection at 1.55 μm and 1.06 μm.

2. Photon detection efficiency versus dark count rate trade-off

The most fundamental trade-off in the operation of SPADs is that between DCR and PDE. Increasing the avalanche probability by operating at larger excess bias $\Delta V$ increases the probability for both photo-excited and dark carriers to generate detectable avalanches; therefore, both PDE and DCR increase. Moreover, if electric field-mediated dark carrier generation is significant at operating conditions of interest, the DCR will exhibit a faster increase with $\Delta V$ than PDE. We have modelled the behavior of PDE and DCR using a formalism described by Donnelly et al. [4]. Temperature-dependent impact ionization coefficients are used to obtain the bias dependence of the PDE by calculating appropriate avalanche breakdown probabilities. The DCR is modelled by taking into account thermal generation-recombination processes as well as band-to-band and trap-assisted tunneling in each layer of the structure (details of our SPAD structure can be found in [2,3]).

The DCR versus PDE trade-off is exhibited in Fig. 1 for SPADs optimized for operation at 1.55 μm [Fig. 1(a)] and 1.06 μm [Fig. 1(b)]; symbols signify experimental data and solid lines illustrate modeling results. For single photon detection at 1.55 μm, low dark count rates of ~10 kHz are achieved at a PDE of 20% at temperatures easily accessible using thermoelectric coolers (e.g., ~215 K) for a detector with an In$_{0.53}$Ga$_{0.47}$As absorber and a 25 μm optical diameter. PDE values beyond 40% can be reached with larger excess bias, at the expense of higher DCR (~100 kHz). For single photon detection at 1.06 μm, the use of a larger bandgap (~1.05 eV) InGaAsP absorption region dramatically reduces thermal contributions to DCR. For a detector with a relatively large optical diameter of 80 μm, DCR of ~1 kHz at a PDE of 30% is obtained for relatively modest cooling to ~240 K. This DCR vs. PDE performance at 1.06 μm is very attractive relative to commercially available Si-based SPAD performance: DCR
values are comparable, while the simultaneously obtained PDE values far outperform Si SPADs, for which 1.06 μm PDE does not exceed 1 – 2%.

![Graph showing DCR versus PDE at different temperatures for SPADs designed for detection at (a) 1.55 μm and (b) 1.06 μm. Symbols are experimental data and solid lines are modeling results. Detector optical diameters are (a) 25 μm and (b) 80 μm.]

Fig. 1. DCR versus PDE at different temperatures for SPADs designed for detection at (a) 1.55 μm and (b) 1.06 μm. Symbols are experimental data and solid lines are modeling results. Detector optical diameters are (a) 25 μm and (b) 80 μm.

The improvement in DCR vs. PDE performance that is achievable for shorter wavelength single photon detection is further illustrated by the modeling results in Fig. 2(a). These results were obtained for three well-optimized structures employing comparable multiplication widths of ~1.5 μm in which the absorption region was tailored to minimize thermal dark carrier generation (by maximizing the absorber bandgap) while guaranteeing high absorber quantum efficiency at the wavelength of interest (specifically, 1.55, 1.31, and 1.06 μm). Note that the 1.06 μm SPAD has DCR lower by at least two orders of magnitude than that of the 1.55 μm SPAD. For applications employing single photon detection at 1.31 μm (e.g., certain quantum key distribution architectures), a SPAD intermediate between the 1.55 and 1.06 μm SPAD performance can be realized.

![Graph showing modeled DCR versus PDE for different SPAD structures optimized for single photon detection at 1.55, 1.31, and 1.06 μm, as indicated. Graph showing modeled contribution to DCR from each of four principal layers in a 1.55 μm SPAD at 213 K.]

Fig. 2. (a) Modeled DCR versus PDE for different SPAD structures optimized for single photon detection at 1.55, 1.31, and 1.06 μm, as indicated. (b) Modeled contribution to DCR from each of four principal layers in a 1.55 μm SPAD at 213 K.

One goal of our modeling is to determine the dominant DCR mechanisms for various device designs and operating conditions. In Fig. 2(b), we exhibit the DCR contribution at 213 K as a function of excess bias for each of the four principal layers of the 1.55 μm SPAD structure. The modeling illustrates a competition between thermal generation in the absorption layer and trap-assisted tunneling in the multiplication region, with the tunneling component becoming more significant at higher bias. For higher temperatures (e.g., > 270 K), thermal generation dominates completely. A valuable output of the modeling is the potential identification of defect levels involved in the leakage mechanisms. While thermal generation is assumed to occur via mid-gap defect states (as in the canonical Shockley-Read-Hall description), our best fits for trap-assisted tunneling assume a trap level $E_t$ at an energy which is 0.78Eg above the InP valence band $E_v$ (i.e., $E_t = E_v + 0.78E_g$), where $E_g$ is the InP bandgap energy. This trap level is found for both the 1.55 and 1.06 μm SPAD structures (as expected, since both employ an InP multiplication region), and is comparable to the result $E_t \sim E_v + 0.75E_g$ reported in [4]. References in the literature on InP materials properties have reported a defect associated with P vacancies in the InP lattice at energies close to the values for $E_t$ found in InP-based SPADs, but a possible relationship between P vacancies and our DCR defects requires further study.
3. Afterpulsing effects and counting rate limitations

For applications required high counting rates (>> 1 MHz), a primary limitation of current InP-based SPADs is an effect known as afterpulsing. Avalanche events can create large instantaneous currents (e.g., ~ 1 mA), and even with relatively fast quenching (~1 – 10 ns), the number of electrical carriers flowing through the SPAD multiplication region is large (e.g., ~10^7 – 10^8). A small fraction of these carriers are trapped at defects within the multiplication region and are detrapped at a later time with an exponentially decaying behavior described by a detrapping time constant \( \tau_d \) that is often on the scale of microseconds. If a carrier is detrapped after the SPAD has been re-armed, this carrier can cause a dark count referred to as an “afterpulse”. Afterpulsing can be mitigated by imposing a sufficiently long “hold-off” time before rearming the SPAD, but this approach limits the photon counting rate.

Afterpulsing effects are apparent in various types of SPAD measurements and are highly dependent on operating conditions. Figure 3(a) illustrates the impact of afterpulsing for measurements obtained at repetition rates spanning 0.5 to 10 MHz. These data were obtained using gated mode operation in which gating pulses of 1 ns duration are applied at the indicated repetition frequency. Single photons were incident only on “even” gates, and the impact of afterpulsing is seen as a rise in DCR on the “odd” gates (relative to DCR in the absence of incident photons). For rates up to 5 MHz, afterpulsing is acceptable (e.g., < 1%) for PDE as large as 20%.

![Figure 3(a): Afterpulsing probability per 1 ns gate vs. PDE for gating frequencies of 0.5, 1, 5, and 10 MHz.](image)

![Figure 3(b): DCR vs. hold-off time illustrating sharp increase in DCR at short hold-off times due to afterpulsing.](image)

The impact of gate length, and the associated duration of charge flow through the device, is evident in Fig. 3(b), which includes experimental data [5] and afterpulse modeling based on the approach described in [6]. The hold-off time required to avoid excessive afterpulsing (evident as a sharp rise in DCR) varies approximately linearly with the amount of charge flowing through the device, as determined by the gate duration. From various measurement techniques, the amount of charge flowing through the multiplication region emerges as a primary factor in determining the severity of afterpulsing.

For many single photon applications, the DCR vs. DE performance of these detectors is satisfactory, and the counting rate limitation imposed by afterpulsing is the most crucial issue to be tackled. Eliminating trap defects is the most fundamental approach to reducing afterpulsing, but the pursuit of more favorable operating conditions may prove to be more expedient for achieving near-term progress in afterpulsing performance. Chief among these prospects for improved afterpulsing is the development of device designs and bias circuitry that effectively reduce the charge flow associated with avalanche events.

References


[5] Experimental data in Fig. 3(b) provided by A. Tosi, S. Cova, and F. Zappa, unpublished.