Thinning of Cascaded Multiplication Layers in Avalanche Photodiodes for High-Speed and High-Power-Tolerant Performance

Yen-Kun Wu, Chao-Chuan Kuo, Pei-Syuan Lin, Sean Yang, H.-S. Chen, Jack Jia-Sheng Huang, and Jin-Wei Shi, *Senior Member, IEEE*

***Abstract*—The thinning of the multiplication (M) layers in high-speed avalanche photodiodes (APDs) is an effective way to boost the gain-bandwidth product (GBP) and reduce excess noise. However, such downscaling usually comes at the price of a huge leakage current ($>$1 μA) induced by direct tunneling through the final thin M-layer ($<$100 nm) and degradation of the RC-limited bandwidth due to the decrease in the thickness of the depletion layer. In this work, we demonstrate how a p-side up top-illuminated APD structure with a thick InP collector layer buried below the thin and cascaded $In_{0.52}Al_{0.48}As$ based multiplication layer can effectively relax the trade-offs among the dark current, RC-limited bandwidth, multiplication gain, and avalanche delay time. This advanced device structure is then applied and the performance of APDs with different thin M-layer thicknesses ($<$50 nm) explored. Under 1.55 μm wavelength excitation, a device fabricated with a large active window (mesa) diameter of 10 (20) μm and an optimized M-layer thickness exhibits a dark current as low as $\sim$0.4 μA and a high responsivity (2.8 A/W; gain = 9.3) at 0.9 V_{br}. This device also exhibits excellent dynamic performance, including a wide optical-to-electrical bandwidth (44 GHz at 0.84 A/W), an extremely large GBP of 1.03 THz, and a high saturation current (12 mA), which corresponds to a large millimeter-wave (MMW) output power ($\sim$0 dBm) at 45 GHz. The exceptional speed performance coupled with the wide dynamic range and simple top-illuminated structure opens up new possibilities to further enhance the sensitivity of 50 G passive optical networks (PONs).**

***Index Terms*—Avalanche photodiode, p-i-n photodiode.**

I. Introduction

OVER the last 20 years, high-speed avalanche photodiodes (APDs) have played a vital role in the development of the Ethernet Passive Optical Network (EPON) and 10 G-EPON. The commercially available $In_{0.52}Al_{0.48}As$ based 10 G APDs can usually provide an 8 dB higher sensitivity than that of their p-i-n photodiode (PD) counterparts [1], [2], [3]. However, the 3-dB bandwidth demonstrated by such APDs under moderate gain operation ($\sim$10) is insufficient to meet the bandwidth requirements for 50 G operation ($>$30 GHz) in the next generation of 50 G-PON, offering less pronounced benefits for the sensitivity of 50 G APD based receivers [4], [5]. It has been demonstrated that the sensitivity at the receiver-end can be improved by the incorporation of complex equalizer integrated circuits (ICs) and high-power PDs, which are hybrid [6] or monolithically integrated with the semiconductor optical amplifier (SOA) [7], [8], [9]. However, both the additional ICs and pre-amplified SOAs require large extra bias currents which will lead to a significant increase of overall power consumption in the receiver-end [9], [10]. Moreover, the additional amplified spontaneous emission (ASE) noise in the SOAs may result in only a marginal improvement in receiver-end sensitivity [9].

Differently from III-V semiconductor based APDs, the Silicon (Si)/Germanium (Ge) counterparts, have demonstrated superior dynamic and static performance [4], [5], [11], [12], [13], [14], [15], [16] for $>$106 Gbit/sec transmissions per lane. Such excellent performances of Si/Ge APDs can be attributed to the fact that the Si-based multiplication (M-)layer has a larger ratio of electron over hole ionization coefficients than those of the III-V M-layers, such as $In_{0.52}Al_{0.48}As$ [5]. Nevertheless, interface defects between the Ge and Si active layers in this kind of APD can have an adverse effect on their reliability in harsh environments, e.g., under high optical power illumination ($\sim$mW). APDs for 50 G PON applications should be capable of handling high optical powers since the optical launch power in the 50 G fiber channel is 5 to 6 dB larger than that in a 10 G-EGPON channel ($\sim$11 vs. $\sim$6 dBm) [17]. A larger optical power around the 1.3 μm wavelength is essential to compensate for the degradation in sensitivity in the receiver-end and the higher fiber propagation loss than in the 10 G PON at the 1.55 μm wavelength. In other words, the 50 G APD requires high-power performance superior to that of the 10 G APD to satisfy the requirements for the afore-mentioned applications. In order to sustain the bandwidth necessary for 50 G PON applications while improving the saturation power, a thinning of both the absorber and depleted M-layers in the APDs is necessary, but this comes at a cost, a lower responsivity, poorer RC-limited

Received 7 June 2024; revised 1 August 2024 and 23 August 2024; accepted 30 August 2024. Date of publication 3 September 2024; date of current version 16 January 2025. This work was supported by the Ministry of Science and Technology in Taiwan under Grant 109-2221-E-008 -081 -MY3 and Grant 110-2622-E-008-022-CC2. *(Corresponding author: Jin-Wei Shi.)*

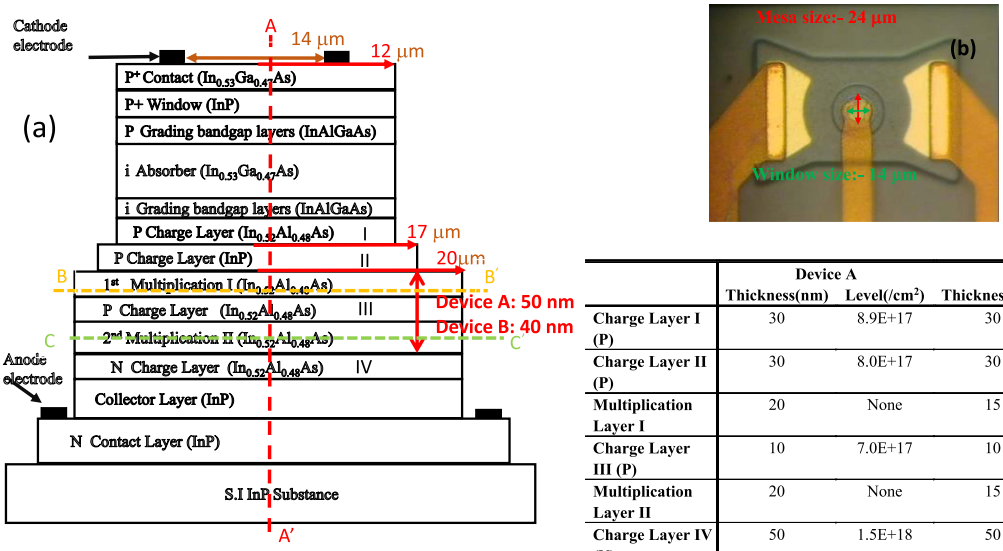
Yen-Kun Wu, Chao-Chuan Kuo, Pei-Syuan Lin, and Jin-Wei Shi are with the Department of Electrical Engineering, National Central University, Taoyuan 320, Taiwan (e-mail: jwshi@ee.ncu.edu.tw).

Sean Yang and H.-S. Chen are with Source Photonics, Hsinchu 300, Taiwan.

Jack Jia-Sheng Huang is with Source Photonics, Hsinchu 300, Taiwan, and also with Source Photonics, West Hills, CA 91304 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2024.3453851.

Digital Object Identifier 10.1109/JLT.2024.3453851



	Device A		Device B	
	Thickness(nm)	Level(/cm²)	Thickness(nm)	Level(/cm²)
Charge Layer I (P)	30	8.9E+17	30	8.6E+17
Charge Layer II (P)	30	8.0E+17	30	8.0E+17
Multiplication Layer I	20	None	15	None
Charge Layer III (P)	10	7.0E+17	10	7.0E+17
Multiplication Layer II	20	None	15	None
Charge Layer IV (N)	50	1.5E+18	50	1.5E+18

Fig. 1. (a) Conceptual cross-sectional view of the demonstrated APD. The vertical (AA') and two horizontal dotted lines (BB' and CC') are specified for the electric (E)-field simulation. (b) Top-view of the fabricated device. The inserted table shows details of the thickness and doping density of each charge and multiplication layer.

bandwidth, and higher dark current. In our previous work, we relaxed these trade-offs by the incorporation of a cascaded M-layer into thin APDs [15], [18], [19]. We demonstrated that such a design can suppress the leakage current and enhance the multiplication gain, compensating for the low unity-gain responsivity (0.3 A/W), thereby producing a device that can sustain a wider bandwidth performance under moderate gain operation ($M_G = \sim 7$). In this work, we explore the thinnest possible M-layer in an APD of this design. The maximum bandwidths (44 vs. 30 GHz) and GBP values (1.03 vs. 0.76 THz) [15] can both be further boosted, with an acceptable low dark current, by aggressively downscaling the thickness of the M-layer from 90 [15] to 50 nm. Moreover, a high saturation current (12 mA), which corresponds to a large millimeter-wave (MMW) output power (∼0 dBm) can be realized at a higher operating frequency, 45 vs. 40 GHz. The remarkable performance of our APD structure with its thin M-layer paves the way to further enhance the speed and receiver sensitivity performance of the photoreceivers needed for 50 G PON systems.

II. Device Structure Design and Fabrication

Fig. 1(a) shows a conceptual cross-sectional view of the structure of the demonstrated device. Note that these figures are not drawn to scale. The epi-layer structure, from top to bottom, is composed of a p+-$In_{0.53}Ga_{0.47}As$ contact layer, p+-InP window layer, p-type partially depleted $In_{0.53}Ga_{0.47}As$ absorber layer, composite p-type $In_{0.52}Al_{0.48}As$/InP charge layers, dual intrinsic $In_{0.52}Al_{0.48}As$ multiplication (M-) layers, n-type $In_{0.52}Al_{0.48}As$ charge layer, thick InP collector layer, and bottommost InP n+ contact layer, grown on a semi-insulating (S.I.) InP substrate. In order to further improve the APD bandwidths while sustaining the same responsivity performance as obtained in our previous work [15], we reduce the total thickness of the M-layers but keep the $In_{0.53}Ga_{0.47}As$ absorber layers unchanged. In the traditional APD structure, with a single M-layer inside, enhancement of the bandwidth through M-layer downscaling usually comes at the cost of an increase in the dark (tunneling) current [4], [5]. Here, a cascaded (dual) M-layer structure is utilized to release the trade-off between the speed, dark current, and responsivity [15], [18], [19]. The M-layer is divided into two parts, specified as the 1st and 2nd M-layer in Fig. 1(a), with an extra charge layer in between. This design results in a stepped electric field profile with most of the avalanche process confined to the extremely thin 2nd M-layer, which has the highest E-field across the whole epi-structure. The use of a stepped E-field can effectively reduce the k-factor (hole over electron impact ionization coefficient) in the APD. Here, the hole impact ionizations are suppressed because of the lower E-field encountered as they travel back toward the absorption layer. The smaller k-factor offered by this dual M-layer design encourages a higher multiplication gain and a larger GBP [20] than those obtained with the traditional APD design with the same M-layer thickness. In our previous work [21], we used dual $In_{0.52}Al_{0.48}As$ M-layers with different thicknesses of 60 and 90 nm. That APD exhibited bandwidth-responsivity product (∼180 GHz × A/W) and k-factor (∼0.14) values close to those of traditional high-performance $In_{0.52}Al_{0.48}As$ APDs with a single M-layer, with a thickness of only 90 nm [4]. This is similar to the working principles utilized in the impact-ionization-engineered (I_2E) APD structures [22]. The impact ionization process is localized in the materials with the narrowest bandgap within the M-layer, which is composed of hetero-junctions between several different bandgap materials. Unlike the I_2E structure, our M-layer adopts a homo-junction design but takes advantage of several charge layers and different doping densities to localize the avalanche process in the region which has the highest E-field.

Furthermore, compared to directly downscaling the single M-layer in the traditional APD to the same thickness as our 2nd M-layer, the incorporation of the additional 1st M-layer can effectively suppress tunneling leakage thereby providing a lower overall dark current. Here, we investigate an APD with two M-layers with different total thicknesses (1st + 2nd M-layer thickness). The total thicknesses of the M-layers in Devices A and B are 50 and 40 nm, respectively, as shown in Fig. 1(a). The thicknesses and doping densities of the 1st and 2nd M-layer and all the charge layers in these two devices are detailed in the table inserted in Fig. 1.

A composite charge layer ($In_{0.52}Al_{0.48}As$/InP) design is adopted to minimize the electric field at the side wall and suppress edge breakdown [23]. Selective chemical wet etching is used between these two layers to precisely etch away the charge layer above the M-layers and zero the E-field at the side wall. The simulated E-field distribution in such a structure will be discussed later; see Fig. 1(b) to (d).

In addition, the thick InP collector layer buried below the thin M-layer relaxes the burden imposed on the RC-limited bandwidth, allowing for a further enlargement of the APD's active diameter and alignment tolerance in the APD package. The InP layer also provides a larger overshoot electron drift-velocity than can be achieved using a $In_{0.52}Al_{0.48}As$ based collector layer [24], [25], which can further relax the trade-off between the device active area and the RC-limited bandwidth. Overall, in contrast to the traditional high-speed APD [1], [2], [3], [4] and uni-traveling carrier photodiodes, our novel APD structure can effectively relax the trade-offs among the RC-limited bandwidth, multiplication gain, and avalanche delay time when operated at the avalanche region [25]. The insertion of an n-type charge layer between the thin $In_{0.52}Al_{0.48}As$ cascaded M-layers and thick InP collector layer acts to manipulate the E-field distribution and maximize the O-E bandwidth of our device, which will be discussed later.

Fig. 1(a) shows the triple mesa structure used in our devices. As can be seen in the figure, three kinds of devices with three different first p-mesa diameters (20, 22, and 24 μm) were fabricated. The extraction of the internal carrier response time for the demonstrated APD epi-layer structure [15] will also be discussed in more detail later. Take the 24 μm device for example. During device fabrication, the 1st mesa with a 24 μm diameter was etched through the p-type absorber layer, followed by selective wet etching which stopped at the InP field control (charge) layer. A thin 34 μm diameter 2nd mesa was formed by the InP charge layer just above the first M-layer. Finally, we etched through the M-regions down to the N^+ contact layer to construct a 40 μm diameter 3rd mesa. A detailed view of the mesa structure is shown in Fig. 1(a). The electric-(E-) field distribution within the device, under the breakdown (V_{br}) and punch through bias voltage (V_{pt}), was simulated using the Silvaco Technology Computer Aided Design (TCAD) tools.[1] As shown in Fig. 2(a) and (b), proper selection of the doping density in the charge layers excludes an undesired breakdown process for the E-field on the AA' axis of devices A and B. This indicates

[1]Silvaco, 2811 Mission College Boulevard, 6th floor, Santa Clara, CA 95054.

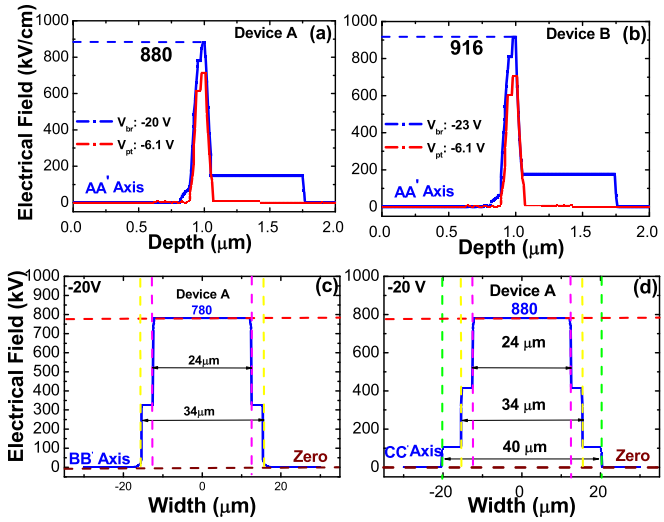


Fig. 2. Simulated E-field distributions of (a) devices A and (b) B in the vertical (AA') direction. The horizontal E-field distributions on the (c) BB' and (d) CC' axes of device A are also specified. The dotted lines in (c) and (d) correspond to the edges of the mesa shown in Fig. 1.

that the E-field in the absorption and collector layers is much less than their corresponding critical fields ($In_{0.53}Ga_{0.47}As$: 150 kV/cm; InP: 500 kV/cm [26]). The calculated E-fields in the 2 bottom M-layers in the horizontal direction (BB' and CC') at the breakdown voltage (V_{br}: -20.2 V) are depicted in Fig. 2(c) and (d), respectively. We can clearly see that, in the 20 μm diameter device, the horizontal E-field (840 kV/cm) in the bottom M-layer is well confined within the range of the top p-mesa, and it can be reduced to nearly zero at the edge of the bottommost M-layer, as discussed above. Fig. 1(e) shows a top view of the fabricated device with a 24 (14) μm diameter active mesa (window).

III. Device Measurement Results

The bias-dependent dark current, photocurrent, and operation gain of APDs with two different M-layer designs, labelled device A and device B, measured under various optical pumping powers at the optical wavelength of 1.55 μm, are shown in Figs. 3 and 4, respectively. Each device structure has 3 different window (mesa) size diameters, as discussed above, and the measured values of the V_{pt} and V_{br} are all specified in the figures. Here, the value of V_{pt} is defined as the data point at around the "knee" of the I-V curve, where the slope of the tangent is equal to 1. The V_{br} value corresponds to the reverse bias voltage when the leakage current reaches 10 μA.

As can be seen, the devices fabricated with our cascaded M-layer design, exhibited a dark current of around ~400 nA at 0.9 V_{br}, lower than that reported for the APDs used for 106 Gbit/sec applications (>1 μA at 0.9 Vbr) [4], [5], [11], [14]. In the case of device A, under an optical pumping power of 10 μW and bias of 0.9 V_{br}, the measured responsivity of all three devices was around 2.8 A/W. The corresponding gain was around 9.3. Here, the chosen unity-gain responsivity of our fabricated device is around 0.3 A/W, which is close to the

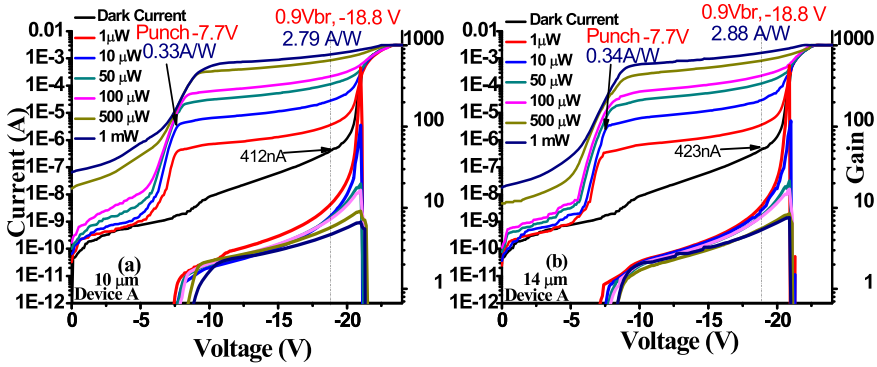


Fig. 3. Dark current, photocurrent, and gain versus bias voltage measured under different optical pumping powers at a wavelength of 1.55 μm for device A with different active window diameters: (a) 10 and (b) 14 μm.

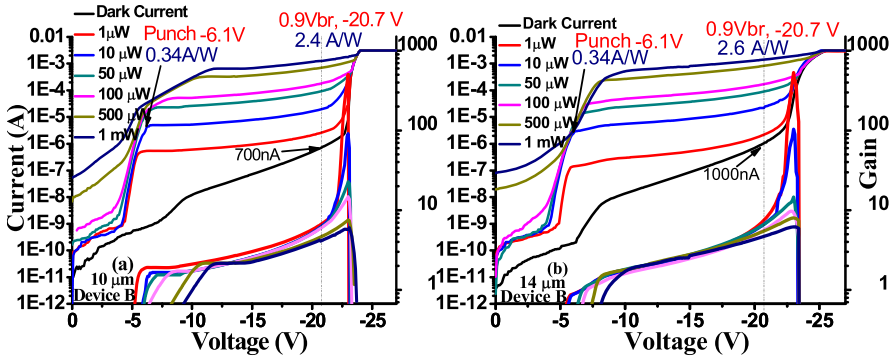


Fig. 4. Dark current, photocurrent, and gain versus bias voltage measured under different optical pumping powers at a wavelength of 1.55 μm for device B with different active window diameters: (a) 10 and (b) 14 μm.

responsivity (0.33 A/W) measured at around V_{pt}, assuming nearly zero optical coupling loss for a single pass of injected light into the absorption layer [15]. For reference, the figures also show the gain versus bias voltage for different optical powers from 1 to 1000 μW. When the reverse bias voltage is greater than the V_{br}, there is a significant reduction in the measured operation gains due to a massive increase in the dark current. Here, under the same optical pumping power values ($\sim$10 μW), the measured maximum gain and responsivity of our demonstrated ones are much larger than those reported for their III-V [14] and Si-Ge [11], [12], [13] counterparts.

This is due to the lower dark current and more pronounced avalanche process induced by our cascaded M-layer design, accompanied by the superior GBP performance to that of traditional APDs, as will be discussed later. However, the responsivity and gain of device B, measured under the same test conditions, i.e., same normalized bias (to V_{br}) and optical pumping power, are slightly less than those of device A, in all three devices. Take the devices with a 14 (24) μm window (mesa) diameter for example. Under a 10 μW optical pumping power, the maximum gains for device A and B are 105 and 90, respectively. Moreover, compared with the performance of high-speed APDs reported in our previous work [15], which had a similar cascaded M-layer design but with a greater M-layer thickness (90 vs. 50 nm), there is a degradation in the maximum gain (105 vs. 150). These results indicate that there is a decrease in the overall multiplication gain under the same effective bias (normalized to V_{br}) with the thinning of the total thickness of the cascaded M-layer. On the other hand, a thinner M-layer usually produces a shorter avalanche delay time and a wider bandwidth. Optimization of the M-layer thickness to maximize the GBP performance of our APDs will be discussed later.

Figs. 5 and 6 show the O-E frequency responses of devices A and B, respectively, measured at the 1.55 μm wavelength, under low (0.01 mW) optical pumping powers, and different bias voltages. The corresponding responsivity of each trace is specified in the figures. It can be clearly seen that devices A and B exhibit close 3-dB bandwidth values under the same normalized bias to V_{br}. When the diameter of the active mesa is 20 μm, the maximum 3-dB bandwidth under a low gain operation ($M_G = 2.8$; responsivity of around 0.84 A/W) is 44 GHz for device A and 41 GHz for device B. The slightly faster speed performance of device A can be attributed to the thicker depletion layer and wider RC-limited bandwidth than those of device B, as will be discussed later. On the other hand, when the bias reaches V_{br}, the avalanche delay time rather than the RC becomes the major bandwidth limiting factor in both devices, in which case, device B has a wider bandwidth (11 vs. 10 GHz) than device A. This reversal can be attributed to the thinner M-layer in device B, which results in a shorter avalanche delay time and faster speed performance, as expected. Here, the bandwidth and responsivity demonstrated by device A, with its

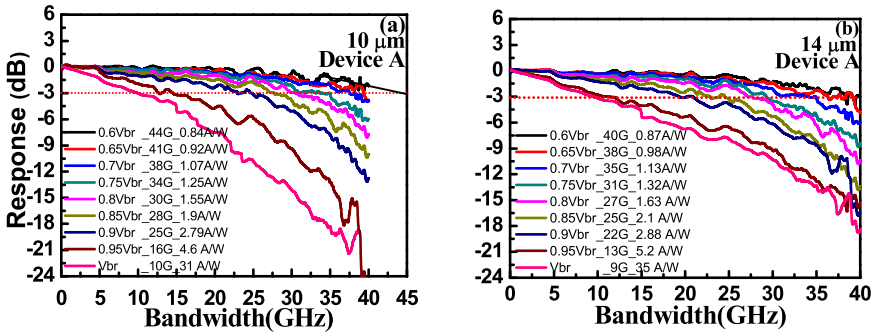


Fig. 5. Bias dependent O-E frequency responses of device A measured under a low 10 μW optical pumping power at a wavelength of 1.55 μm with different active window diameters: (a) 10 and (b) 14 μm.

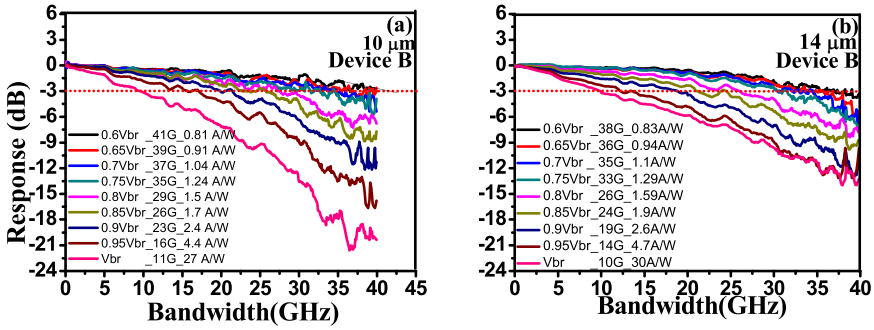


Fig. 6. Bias dependent O-E frequency responses of device B measured under a low 10 μW optical pumping power at a wavelength of 1.55 μm with different active window diameters: (a) 10 and (b) 14 μm.

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT TYPES OF APDS

	NTT	SiFotonics Technology	Albis(APD20E1)	MACOM (MARP_ BA56)	This work	Our Previous work
Reference	4,14,27	11	28	29		30
Type	Backside-illuminated	Topside-illuminated	Topside- illuminated	Backside-illuminated	Topside-illuminated	Backside-illuminated
Active Mesa Diameter	14μm	20μm	20μm	14μm	20μm	14μm
M-Layer Thickness	90nm				50nm	90nm
Dark Current(0.9V_{br})	2μA	0.9μA			410nA	175nA
Responsivity	0.75(1.75)A/W	3.53(6.5)A/W	4.5A/W	4.2A/W	0.81(1.9A/W)	3.4(6.52)A/W
Bandwidth	42(26)GHz	28(22)GHz	20GHz	30GHz	44(28)GHz	36(28)GHz
Input optical or output electrical saturation power	Input: -6.3dBm	Input: <0dBm	Input: <+5dBm		Input:+8.5dBm Output: 0dBm at 45GHz	Input: +8.8dBm Output: 0dBm at 40GHz

simple top-illuminated structure, is even better than that of its back-side illuminated III-V counterpart, which has a smaller (14 vs. 20 μm) active mesa diameter [14]. The responsivity performance of a back-side illuminated structure is usually better than a top-side illuminated structure, because of the folding of the optical absorption path caused by the topmost contact metal which acts as a reflector. The superior bandwidth-responsivity product of our APD confirms that our unique cascaded M-layer design can provide an intrinsically larger GBP value than is possible with the traditional single M-layer structure [1], [2], [3], [4]. Not only does this compensate for the lower unit gain responsivity (0.3 vs. 0.5 A/W) but it also produces a wider O-E bandwidth (44 vs. 42 GHz) under a larger operation gain (2.8 vs. 1.5). Table I shows the benchmark values reported for the high-performance APDs used for 28 or 56 Gbaud applications [11], [14], [27], [28], [29], [30]. Here, only APDs with a vertical-illuminated structure are listed. This type of structure is more popular with industry because it allows for a much larger optical alignment tolerance during device packaging than edge-coupled (waveguide) structures do [12], [13], [31]. The numbers in brackets in this table indicate the testing of APDs under different bias voltages.

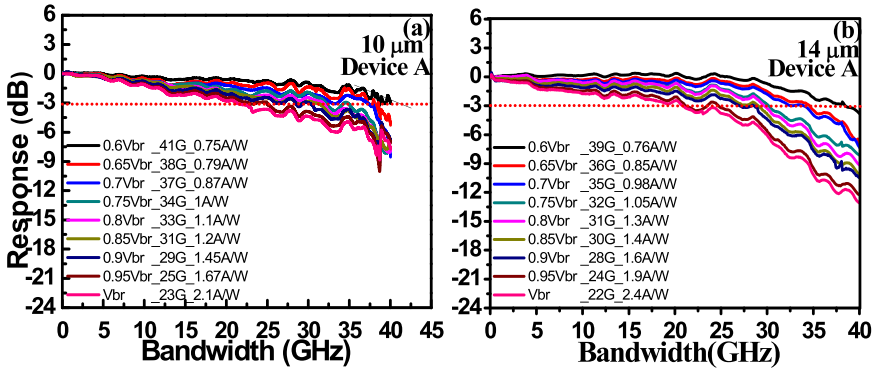


Fig. 7. Bias dependent O-E frequency responses of device A measured under an optical pumping power of 1 mW at the 1.55 μm wavelength with different active window diameters: (a) 10 and (b) 14 μm.

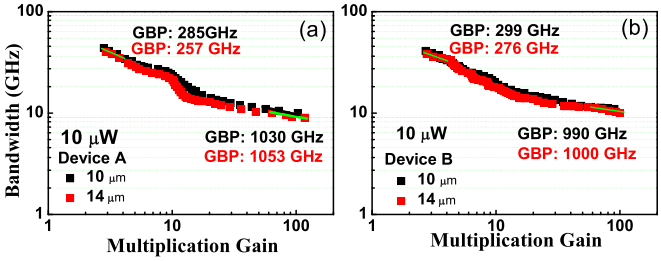


Fig. 8. Measured 3-dB O-E bandwidth versus multiplication gain for: (a) device A and (b) device B.

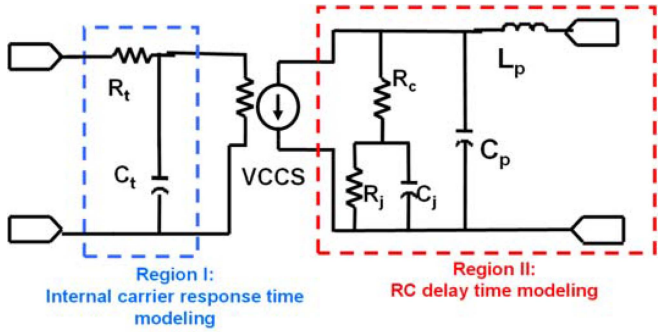


Fig. 9. Blue box: artificial RC-network mimicking the internal carrier response time limited frequency response. Red box: equivalent-circuit-model representing the physical structure of the APD. VCCS: Voltage controlled current source.

The measured responsivity of our device is smaller (1.9 vs. 3.53 A/W) than that of its top-illuminated Si-Ge counterparts with backside reflectors under the same operation speed (~28 GHz). This result suggests that, to compete with Si-Ge APDs, whose Silicon based M-layers are intrinsically superior to those of most III-V materials, it is preferable for our dual M-layer APD design to be combined with a back-side illuminated structure [29], [30]. Comparison with the benchmark values shows that our cascaded M-layer APD design combined with an advanced flip-chip bonding package for backside-illumination, as described in a previous work [30], demonstrates the largest bandwidth-responsivity products among all those reported for III-V or SiGe APDs [11], [14], [27], [28], [29], [30]. As previously discussed, in addition to the bandwidth and the responsivity, the high-power performance of an APD is another important parameter for 50 G PON network applications because of the increase in the optical launch power (>10 dBm) in the fiber channel. The high-power measurement results (at 1 mW) for device A are given in Fig. 7. As can be seen, device A can sustain a maximum 3-dB bandwidth over 40 GHz with the same active mesa diameter of 20 μm. The consistent high-speed performance indicates that our novel APD structure has good linearity across a wide range of optical pumping powers. Figs. 8(a) and (b) depict the gain bandwidth products (GBP) of devices A and B, respectively, measured at a low optical pumping power (10 μW). The devices have the same active mesa diameters of 20 or 24 μm, respectively. As can be seen from the results, very high GBPs of 1053 and 1000 GHz can be achieved for devices A and B, respectively, with the same unity-gain responsivity of 0.3 A/W. The obtained GBP values (1053 GHz) are much larger than those reported for high-speed $In_{0.52}Al_{0.48}As$ based APDs with a single M-layer (~320 GHz) [4], [32] or for cascaded M-layer designs (~750 GHz) [15], and even better than that of Si/Ge APDs (1033 GHz) [13] under the same extremely high-gain operating conditions. This excellent performance confirms that the GBP limitation encountered in traditional high-speed III-V APDs can be relaxed by aggressively downscaling the thickness of the cascaded M-layer to 50 nm, while retaining an acceptable low dark current (<1 μA). Note that, rather than directly reducing the thickness of the single M-layer as in the traditional APD structure, an extra 1st M-layer is added in our deign to provide a lower tunneling leakage current.

Such high GBP values also imply a small k-factor and low excess noise in our APDs. Detailed measurements of excess noise for devices A and B should be carried out in future [21].

Furthermore, in contrast to most high-speed APDs [1], [4], [15], which usually report a plateau in the bandwidth vs. gain curve under low gain operation (<10), our traces show a monotonous decrease of bandwidth with the increase of operation gain. This can be attributed to the large increase in the thickness of the collector layer which balances the RC-limited and internal carrier response time and leads to further improvement of the net O-E bandwidth (44 vs. 30 GHz [15]). However, as the

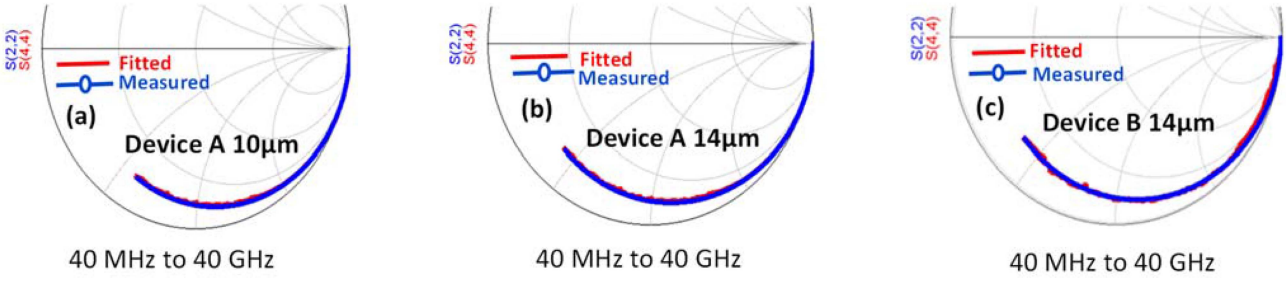


Fig. 10. Measured (blue line) and fitted (red line) S_{22} parameters from near dc to 40 GHz under a fixed dc bias (0.6 V_{br}) for device A with different active window diameters: (a) 10 and (b) 14 μm and (c) for device B with a 14 μm active window diameter.

multiplication gain increases, the multiplied electrons with high excess energy injected from M-layer tend to induce an undesired avalanche process in the thick collector layer, which diminishes the plateau region.

Comparison with device A shows that further downscaling of the thickness of the M-layer in device B (50 vs. 40 nm) does not lead to improvement of the GBP. As can be seen in Figs. 5 and 6, although device B has a slightly larger bandwidth than that of device A under V_{br} bias, it exhibits a smaller GBP. This can be attributed to the fact that the multiplication gain is less pronounced in device B, with its extremely thin M-layer, than in device A, as shown in Figs. 2 and 3. On the other hand, as demonstrated in a previous work [15], a high-speed APD with the same $In_{0.53}Ga_{0.47}As$ absorption layer design but greater total cascaded M-layer thickness (M_1+M_2) of 90 nm could exhibit a larger maximum gain (150 vs. 105) but a smaller GBP value (760 vs. 1053 GHz) than those of its 50 nm counterparts. Overall, the dynamic and static measurement results for APDs with 3 different cascaded M-layer thicknesses suggest that the optimum value for maximizing GBP performance would be at around 50 nm.

The net 3-dB O-E bandwidth (f_{3dB}) of APDs is determined by the internal carrier response time $1/f_t$ and the RC time constant $1/f_{RC}$. Its value can then be calculated using the following equation:

$$\frac{1}{f_{3dB}^2} = \frac{1}{f_{RC}^2} + \frac{1}{f_t^2} \tag{1}$$

Here, $1/f_t$ includes the carrier transit, diffusion, and build up time in the avalanche process [33]. Regarding with their RC-limited bandwidths (f_{RC}), they are extracted by performing the equivalent circuit modeling technique [15]. During such modeling process, the simulated microwave reflection coefficients (S_{22}) of established circuit model are used to fit the measured S_{22} of our APD device [15]. Fig. 9 shows the details of adopted circuit model. In region 1 (within the blue box), an artificial RC-network (R_t, C_t) is applied to mimic the internal carrier response time ($1/f_t$) limited frequency response. Region 2 in the red box represents an equivalent circuit model of the physical structure of the APD. During RC modeling, the two circuit elements, R_t and C_t, in region 1 are removed for extraction of the extrinsic f_{RC} of the APD, as described in detail in our previous work [30]. Table II shows the fitted values for each circuit element. The fitted and measured frequency responses

TABLE II
VALUES OF THE CIRCUIT ELEMENTS

	Physical Meaning	Devices			
		A			B
		10μm	12μm	14μm	14μm
C_j	Junction Capacitance (fF)	34	41	49	50
R_j	Junction Resistance (kΩ)	9	9	9	8.8
R_s	Contact Resistance (Ω)	22	22	22	22
C_p	Parasitic Capacitance (fF)	26	33	41	42
L_p	Parasitic Inductance (pH)	20	20	20	21

for the S_{22} parameters for device A with two different window diameters (10 and 14 μm) are shown in the Smith charts in Fig. 10(a) to (b). For comparison, the fitted and measured S_{22} frequency responses of device B with a 14 μm window diameter are also included in Fig. 10(c).

There is obviously a close match between the simulation and the measurement results. In addition, compared to device B, device A has a slightly shorter length of S_{22} trace, which implies that it has a smaller device capacitance due to thicker M-layer. However, it is possible to determine the internal carrier response time limited frequency responses of our device by carefully selecting the values of R_t and C_t in our established circuit model to fit the overall O-E frequency response.

Fig. 11(a) and (b) show the measured O-E, fitted O-E, extracted RC-limited, and internal response time limited frequency responses obtained under a bias of 0.6 V_{br} for device A with two different window diameters. Fig. 10(c) shows the same traces but for device B with a 14 μm window diameter for comparison. Here, the "RC delay" frequency responses are extracted from the proposed equivalent circuit model and the measured S_{22} traces, as illustrated in Figs. 9 and 10. The "Measured" frequency responses stand for the measured net O-E frequency responses. The "Internal" frequency responses, which include the carrier transit, diffusion, and build up time in the avalanche process, are obtained from the artificial RC network of R_t and C_t, as shown in Fig. 9. During the modeling process, the RC-delay and internal frequency responses are multiplied and the values of R_t and C_t tweaked to fit the measured net O-E responses. The multiplied traces (RC delay $\times$ Internal) represent the "Fitted" traces in these two figures. As can be seen, device B has a slightly poorer RC-limited bandwidth, as expected, which accounts for its smaller net O-E bandwidth than that of device A. Furthermore, under low gain operation, the RC-limited bandwidth of

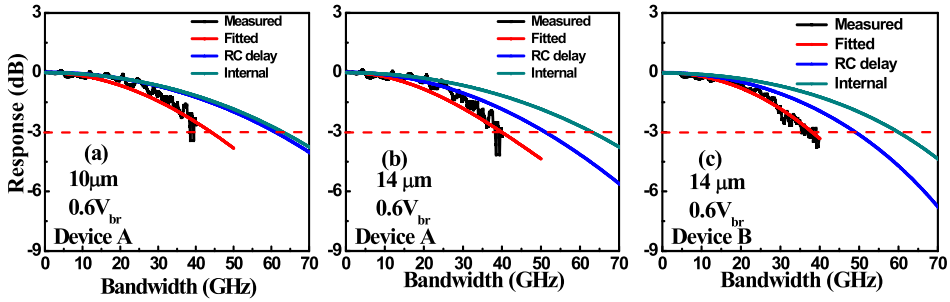


Fig. 11. Measured O-E, extracted RC-limited, internal response time, and fitted O-E frequency responses for device A with different active window diameters: (a) 10 and (b) 14 μm and (c) for device B with a 14 μm active window diameter. Both devices are under the same bias of 0.6 V_{br}.

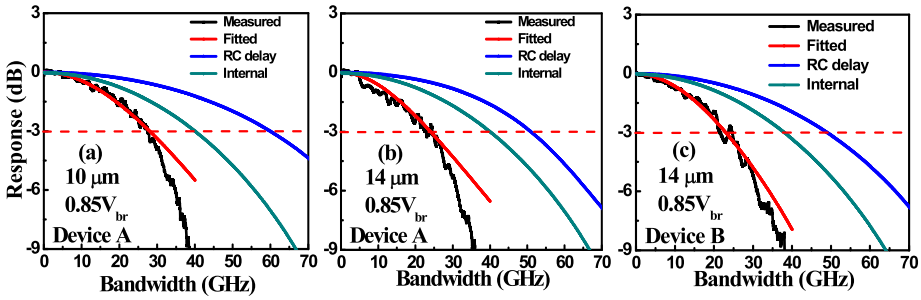


Fig. 12. Measured O-E, extracted RC-limited, internal response time, and fitted O-E frequency responses for device A with different active window diameters: (a) 10, (b) 14 μm and (c) for device B with a 14 μm active window diameter. Both devices are under the same bias voltage of 0.85 V_{br}.

device A with the 20 μm active mesa diameter is quite close to the internal limited bandwidth.

These results suggest that the total thickness of the depletion layers in the 20 μm active mesa diameter devices is optimal for balancing the internal carrier response time and external RC-limited bandwidth under low gain operation (0.6 V_{br}). Fig. 12 shows the results for the same cases as in Fig. 11 but with the application of a higher bias voltage reaching 0.85 V_{br} to devices A and B, for a moderate gain (6.3; 1.9 A/W) with a bandwidth of $\sim$30 GHz as is essential for 50 G PON applications.

We can clearly see that the dominant bandwidth limiting factor in these two devices has become the internal response time. The modeling results indicate that it might be preferable to shorten the width of the collector layer in the device to obtain a more balanced speed performance from low to moderate (3 to 7) gain operation [15]. However, this thick collector design provides a smaller junction capacitance, which can usually lower the noise current from the integrated trans-impedance amplifier (TIA) and benefit the sensitivity performance of the whole receiver module [34]. The next step in future work is a more detailed investigation of the sensitivity performance of this APD design with different collector layer thicknesses.

Moreover, as shown in Fig. 12, when the bias voltage is 0.85 V_{br}, the shapes of the measured and fitted O-E frequency responses do not match perfectly. This can be attributed to the fact that under high gain ($>$10) operation, the avalanche build up time becomes more pronounced and its frequency response cannot be described by a single time constant (STC) R_t, C_t network, as shown in Fig. 9 (region 1). A 3rd order network

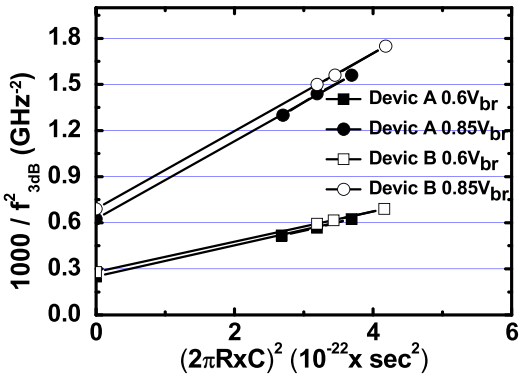


Fig. 13. Extracted $(1/f_{RC})^2$ versus measured $(10^3/f_{3dB}^2)$ for devices A and B under 2 different bias voltages (0.6 and 0.85 V_{br}).

describing the carrier response time, as discussed in a previous work [33], is necessary to get better fitting results.

Fig. 13 shows the values of $(1/f_{RC})^2$ versus $(10^3/f_{3dB}^2)$ for both kinds of APDs (A and B) fabricated with different mesa diameters. Here, to have a neat x-axis label, the $(1/f_{RC})^2$ is written as $(2\pi R \times C)^2$. The internal carrier response time can then be obtained using Equation 1 from the fitted RC-limited, measured net O-E bandwidths for each device (see Figs. 10 and 11), and the intercepts along the y-axis in Fig. 12. Here, the factor of 1000 scaling of the y-axis has been chosen to bring its values into the range between 0 and 1. Take device A for example. The f_t values obtained at bias voltages of 0.6 and 0.85 V_{br} are approximately 63 and 40 GHz, respectively.

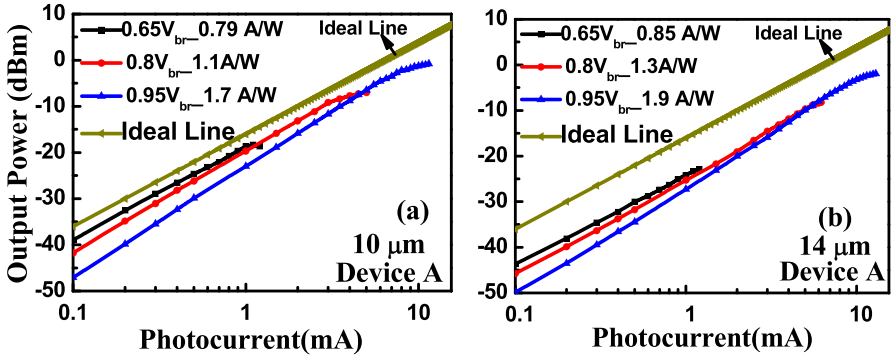


Fig. 14. Measured photo-generated MMW power versus photocurrent for device A with different active window diameters: (a) 10 μm and (b) 14 μm, at a frequency of 45 GHz under different bias voltages.

These values closely match the internal response time limited bandwidth obtained using the equivalent circuit modeling technique, as can be seen in Figs. 11 and 12.

Fig. 14 depicts the photo-generated RF output power measured with a heterodyne beating setup at a frequency of 45 GHz for device A with two different active mesa diameters (10 and 14 μm) at several different bias voltages.

The heterodyne-beating system is composed of two 1.55 μm wavelength free-running lasers, which have a 45 GHz (0.36 nm) spacing in their central wavelengths. The pink line in the figure shows the ideal relationship between the microwave power and the average photocurrent with a 100% optical modulation depth under a 50 Ω load, for reference. It can be seen that the value of the saturation photocurrent exhibited by these two devices is almost the same, 12 mA, for maximum output millimeter wave power at a bias of around 0.95 V_{br}. For example, the measured responsivity of device A with a 10 μm active window diameter under a saturation output RF power of ($-$0.77 dBm), which corresponds to a high optical launch power of $+$8.5 dBm, is around 1.7 A/W.

Compared with the high-power performance of APDs reported in our previous work [15], our new APD device A design, which has the same active mesa diameter of 24 μm, not only exhibits a higher saturation current (12 vs. 11 mA at 0.95 V_{br}) but at a higher operating frequency (45 vs. 30 GHz). This significant improvement in the saturation current-bandwidth product is mainly due to the aggressive downscaling of the thickness of the cascaded M-layer and the lower V_{br} (20.2 vs. 26 V). These two characteristics enhance the speed performance and reduce the device heating originating from the product of the operating bias voltage and output photocurrent, respectively. These improvements give the demonstrated APD structure strong potential to meet the requirements of the 50 G PON photo-receivers which need to handle strong light illumination. From Table I, we can clearly see that our device with its simple top-illuminated structure exhibits a state-of-the-art GBP, good capability for handling high optical power illumination, large mesa size, small capacitance, and has a record-high saturation current-bandwidth product among all those reported for high-speed APDs [16], [27], [35].

IV. Conclusion

In this work, we describe our unique $In_{0.52}Al_{0.48}As$ based APDs with a cascaded M-layer and composite charge layer inside. The M-layer thickness is downscaled to around 50 nm to improve the high-speed performance while minimizing the increase in the tunneling dark current. Moreover, in order to allow a large active window (mesa) size while maintaining a small junction capacitance for high-sensitivity performance, a thick and fully depleted InP collector layer is buried below the thin M-layer. Greater optical alignment tolerance and better device heat sinking can be obtained with this simple p-side up, top-illuminated structure, even with a large active mesa diameter of 20 μm. This APD exhibits a wide 3-dB O-E bandwidth (44 to 28 GHz) from a low to moderate multiplication gain (responsivity of 0.84 to 1.9 A/W), low dark current ($\sim$400 nA), low capacitance (34 fF), an extremely large GBP of 1.03 THz, and a high saturation current ($\sim$12 mA), which corresponds to a large millimeter-wave (MMW) output power ($-$0.77 dBm) at 45 GHz. Furthermore, the measurement results suggest that further downscaling of the cascaded M-layer thickness ($<$50 nm) would not benefit the GBP of the device given that the multiplication gain becomes less pronounced under the same normalized bias (to V_{br}). Overall, the demonstrated APD with its optimal M-layer thickness and remarkable dynamic/static performance opens up new possibilities to further enhance the sensitivity of the photoreceivers in a 50 G PON system, with a reduction in power consumption.

References

[1] J. C. Campbell, "Recent advances in avalanche photodiodes," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 278–285, Jan. 2016, doi: 10.1109/JLT.2015.2453092.

[2] E. Yagyu, E. Ishimura, M. Nakaji, T. Aoyagi, and Y. Tokuda, "Simple planar structure for high-performance AlInAs avalanche photodiodes," *IEEE Photon. Tech. Lett.*, vol. 18, no. 1, pp. 76–78, Jan. 2006.

[3] B. F. Levine et al., "A new planar InGaAs–InAlAs avalanche photodiode," *IEEE Photon. Technol. Lett.*, vol. 18, no. 18, pp. 1898–1900, Sep. 2006.

[4] M. Nada, Y. Yamada, and H. Matsuzaki, "Responsivity-bandwidth limit of avalanche photodiodes: Toward future ethernet systems," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 2, Mar./Apr. 2018, Art. no. 3800811.

[5] M. Huang et al., "Germanium on silicon avalanche photodiode," *IEEE J. Sel. Top. Quantum Electron.*, vol. 24, no. 2, Mar./Apr. 2018, Art. no. 3800911.
[6] Y. T. Han et al., "A compact 100G-ER4 ROSA Realized by hybrid integration of SOA and lensed PIN-PDs for QSFP28 transceivers," in *2019 Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, 2019, pp. 1–3.
[7] J. W. Raring et al., "40 Gbit/s photonic receivers integrating UTC photodiodes with high- and low-confinement SOAs using quantum well intermixing and MOCVD regrowth," *Electron. Lett.*, vol. 42, no. 16, pp. 942–943, Aug. 2006.
[8] M. Anagnosti et al., "Record gain x bandwidth (6.1 THz) monolithically integrated SOA-UTC photoreceiver for 100-Gbit/s applications," *J. Lightw. Technol.*, vol. 33, no. 6, pp. 1186–1190, Mar. 2015.
[9] L. Breyne et al., "50G burst-mode receiver using monolithic SOA-UTC and burst-mode TIA," in *2024 Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, 2024, pp. 1–3.
[10] G. Coudyzer et al., "100 Gbit/s PAM-4 linear burst-mode transimpedance amplifier for upstream flexible passive optical networks," *J. Lightw. Technol.*, vol. 41, no. 12, pp. 3652–3659, Jun. 2023, doi: 10.1109/JLT.2023.3262319.
[11] B. Shi et al., "106 Gb/s normal-incidence Ge/Si avalanche photodiode with high sensitivity," in *2020 Opt. Fiber Commun. Conf. Exhib.*, 2020, pp. 1–3.
[12] B. Wang et al., "64 Gb/s low-voltage waveguide SiGe avalanche photodiodes with distributed Bragg reflectors," *Photon. Res.*, vol. 8, no. 7, pp. 1118–1123, Jul. 2020.
[13] Y. Shi et al., "Avalanche photodiode with ultrahigh gain–bandwidth product of 1,033 GHz," *Nature Photon.*, vol. 18, pp. 610–616, 2024, doi: 10.1038/s41566-024-01421-2.
[14] M. Nada, T. Yoshimatsu, F. Nakajima, K. Sano, and H. Matsuzaki, "A 42-GHz bandwidth avalanche photodiodes based on III-V compounds for 106-Gbit/s PAM4 applications," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 260–265, Jan. 2019.
[15] Nassem et al., "Top-illuminated avalanche photodiodes with cascaded multiplication layers for high-speed and wide dynamic range performance," *J. Lightw. Technol.*, vol. 40, no. 24, pp. 7893–7900, Dec. 2022, doi: 10.1109/JLT.2022.3204743.
[16] T. Beckerwerth, R. Behrends, F. Ganzer, P. Runge, and M. Schell, "Linearity characteristics of avalanche photodiodes for InP based PICs," *IEEE J. Sel. Top. Quantum Electron.*, vol. 28, no. 2, Mar./Apr. 2022, Art. no. 3803408, doi: 10.1109/JSTQE.2021.3127853.
[17] S. Nishikawa et al., "SOA-integrated high-power EML-CAN for 50G-PON downstream," in *2024 Opt. Fiber Commun. Conf. Exhib.*, San Diego, CA, USA, 2024, pp. 1–3.
[18] Z. Ahmad et al., "Avalanche photodiodes with dual multiplication layers and ultra-high responsivity-bandwidth products for FMCW lidar system applications," *IEEE J. Sel. Top. Quantum Electron.*, vol. 28, no. 2, Mar./Apr. 2022, Art. no. 3800709, doi: 10.1109/JSTQE.2021.3062637.
[19] Z. A. Naseem et al., "Avalanche photodiodes with dual multiplication layers for high-speed and wide dynamic range performances," *Photonics*, vol. 8, no. 4, Mar. 2021, Art. no. 98.
[20] M. A. Saleh et al., "Impact-ionization and noise characteristics of thin III-V avalanche photodiodes," *IEEE Trans. Electron Devices*, vol. 48, no. 12, pp. 2722–2731, Dec. 2001.
[21] H.-Y. Zhao et al., "High-speed avalanche photodiodes with wide dynamic range performance," *J. Lightw. Technol.*, vol. 37, no. 23, pp. 5945–5952, Dec. 2019.
[22] N. Duan et al., "High-speed and low-noise SACM avalanche photodiodes with an impact-ionization-engineered multiplication region," *IEEE Photon. Technol. Lett.*, vol. 17, no. 8, pp. 1719–1721, Aug. 2005.
[23] Naseem et al., "Avalanche photodiodes with composite charge-layers for low dark current, high-speed, and high-power performance," *IEEE J. Sel. Top. Quantum Electron.*, vol. 28, no. 2, Mar./Apr. 2022, Art. no. 3801910, doi: 10.1109/JSTQE.2021.3111895.
[24] N. Li et al., "High-saturation-current InGaAs/InAlAs charge-compensated uni-traveling-carrier photodiode," *Phys. Statist. Sol.*, vol. 201, no. 13, pp. 3037–3041, Aug. 2004.
[25] Y. Yi, T. Umezawa, K. Akahane, and T. Kawanishi, "High-responsivity unitraveling carrier–type photodetector using avalanche inalas as a collector at a high-frequency range," *IEEE Access*, vol. 12, pp. 22813–22820, 2024, doi: 10.1109/ACCESS.2024.3363877.
[26] Y. L. Goh, J. S. Ng, C. H. Tan, W. K. Ng, and J. P. R. David, "Excess noise measurement in In0.53Ga0.47As," *IEEE Photon. Technol. Lett.*, vol. 17, no. 11, pp. 2412–2414, Nov. 2005.
[27] M. Nada, Y. Muramoto, H. Yokoyama, and H. Matsuzaki, "High-speed high-power-tolerant avalanche photodiode for 100-Gb/s applications," in *2014 IEEE Photon. Conf.*, San Diego, CA, USA, 2014, pp. 172–173, doi: 10.1109/IPCon.2014.6995303.
[28] *Albis Optoelectronics AG, Moosstrasse 2a, 8803 Rueschlikon, Switzerland. (Product: APD20E1)*. [Online]. Available: https://www.albisopto.com/albis_product/apd20e1-28-gbd-top-side-illuminated-apd/
[29] *MACOM, 100 Chelmsford Street, Lowell, MA, 0185, United States. (Product: MARP-BA56)*. [Online]. Available: https://www.macom.com/products/product-detail/MARP-BA56
[30] Naseem et al., "Simultaneous enhancement of the bandwidth and responsivity in high-speed avalanche photodiodes with an optimized flip-chip bonding package," *Opt. Exp.*, vol. 31, pp. 26463–26473, Jul. 2023.
[31] T. Okimoto et al., "106-Gb/s waveguide AlInAs/GaInAs avalanche photodiode with butt-joint coupling structure," in *2022 Opt. Fiber Commun. Conf. Exhib.*, 2022, pp. 01–03.
[32] G. S. Kinsey, J. C. Campbell, and A. G. Dentai, "Waveguide avalanche photodiode operating at 1.55 μm with a gain-bandwidth product of 320 GHz," *IEEE Photon. Technol. Lett.*, vol. 13, no. 8, pp. 842–844, Aug. 2001.
[33] B. Wang et al., "A compact model for Si-Ge avalanche photodiodes over a wide range of multiplication gain," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3229–3235, Jul. 2019, doi: 10.1109/JLT.2019.2913179.
[34] E. Jan, "Low noise transimpedance amplifier design using berkeley analog generator," University of California, Berkeley, CA, USA, Tech. Rep. UCB/EECS-2020-146, Aug. 2020. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2020/EECS-2020-146.html
[35] M. Nada, Y. Yamada, and H. Matsuzaki, "A high-linearity avalanche photodiodes with a dual-carrier injection structure," *IEEE Photon. Technol. Lett.*, vol. 29, no. 21, pp. 1828–1831, Nov. 2017.

Yen-Kun Wu was born in Taichung City, Taiwan, in 1997. He received the bachelor's degree in semiconductor engineering from the Department of Semiconductor Engineering, Kaohsiung University of Science and Technology, Kaohsiung City, Taiwan. He is currently working toward the master's degree with the Department of Electrical Engineering, National Central University, Taoyuan City, Taiwan. His research interests include high-speed photodiodes and avalanche photodiodes for optical receivers.

Chao-Chuan Kuo was born in Chiayi, Taiwan, in 1999. He is currently working toward the master's degree with the Department of Electrical Engineering, National Central University, Taoyuan, Taiwan. His research interests include high speed photodiodes and APDs (Avalanche Photodiodes) for optical receivers and instruments.

Pei-Syuan Lin was born in Taoyuan, Taiwan, in 2000. He is currently working toward the master's degree with the Department of Electrical Engineering, National Central University, Taoyuan, Taiwan. His research focuses on high-speed avalanche photodiodes.

Sean Yang received the B.S. degree in electrical engineering from National Taiwan Ocean University (NTOU), Taiwan, in 2006. From 2006 to 2017. he was a Photo Engineer with E Ink Holdings Inc. He is currently with Source Photonics Inc., as an Engineer of advanced tools.

H.-S. Chen received the Ph.D. degree from the Graduate Institute of Photonics and Optoelectronics, National Taiwan University, Taipei, Taiwan, in 2006. Since June 2007, he has been a Postdoctoral Research Fellow with the Institute of Physics, Academia Sinica. From 2011 to 2013, he was with the Institute of Photonics and Optoelectronics, National Taiwan University, as a Postdoctoral Fellow. In 2014, he joined Source Photonics Inc., as a Manager of advanced process development. His research interests include surface plasmon, near field optics, ultrafast lasers, nanorod array LEDs, and optical sensors.

Jack Jia-Sheng Huang received the B.S. degree in physics from National Taiwan University, Taipei, Taiwan, in 1992, and the M.S. and Ph.D. degrees in materials science from University of California, Los Angeles (UCLA), Los Angeles, CA, USA, in 1996 and 1997, respectively. He is currently a Senior R&D Scientist/Manager with Source Photonics, working on Advanced Photonics Devices including 100G DML, 40G CWDM DFB, 10G DFB, and 10G FP lasers as well as 10G and 25G APD photodetectors. His R&D projects also involve device reliability physics and device characterization. From 2000 to 2015, he was a Wafer Fab R&D/Operations Manager with Emcore, working on device design/process/characterization, reliability, electrostatic discharge (ESD), and failure analysis of analog BH lasers and digital ridge lasers for cooled and uncooled applications. From 1997 to 2000, he was a Member of Technical Staff with Lucent Technologies, Bell Labs, Orlando, FL, USA, working on electromigration, stress migration and failure analysis of 0.3, 0.25, 0.2, and 0.16mm ASIC and FPGA devices using CMOS technology. From 1992 to 1993, he was a Research Assistant with the Institute of Atomic and Molecular Sciences, Academia Sinica, Taiwan, studying the surface physics of gallium ion beams in an ultra-high vacuum. He has more than 90 publications in international journals and conferences in the areas of optoelectronics and ICs. He holds six U.S. patents and one U.K. patent. In 1997, he was awarded the Outstanding Ph.D. Award and Dissertation Year Fellowship Award from the Henry Samueli School of Engineering, UCLA.

Jin-Wei Shi (Senior Member, IEEE) was born in Kaohsiung, Taiwan, on January 22, 1976. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1998 and the Ph.D. degree from the Graduate Institute of Electro-Optical Engineering, National Taiwan University, in 2002. In 2003, he joined the Department of Electrical Engineering, National Central University, Taoyuan, Taiwan, where he is currently a Professor. In 2011, he joined the ECE Department of UCSB as a Visiting Scholar. He has authored or coauthored more than 160 Journal papers, 160 conference papers and hold 20 patents. His research interests include ultra-high speed/power optoelectronic devices, such as photodetectors, electro-absorption modulators, sub-millimeter wave photonic transmitters, and semiconductor lasers. In 2007, he was the recipient of the Excellent Young Researcher Award from the Association of Chinese IEEE and in 2010, he was also the recipient of the Da-You Wu Memorial Award. He has been an Optica fellow since 2023 and senior member of IEEE since 2012. He has been an Associate Editor for *Optics Express* since 2017.