Performance and Modeling of the MWIR HgCdTe Electron Avalanche Photodiode

JEFFREY BECK,[1,2] RICHARD SCRITCHFIELD,[1] BILLY SULLIVAN,[1] JAMIE TEHERANI,[1] CHANG-FENG WAN,[1] MIKE KINCH,[1] MARTHA OHLSON,[1] MARK SKOKAN,[1] LEWIS WOOD,[1] PRADIP MITRA,[1] MIKE GOODWIN,[1] and JIM ROBINSON[1]

1.—DRS Technologies – Infrared Technologies Division, P.O. Box 740188, Dallas, TX 75374, USA. 2.—e-mail: jdbeck@drs-irtech.com

The operation of the mid-wave infrared (MWIR) HgCdTe cylindrical electron injection avalanche photodiode (e-APD) is described. The measured gain and excess noise factor are related to the collection region fill factor. A two-dimensional diffusion model calculates the time-dependent response and steady-state pixel point spread function for cylindrical diodes, and predicts bandwidths near 1 GHz for small geometries. A 2 μm diameter spot scan system was developed for point spread function and crosstalk measurements at 80 K. An electron diffusion length of 13.4 μm was extracted from spot scan data. Bandwidth data are shown that indicate bandwidths in excess of 300 MHz for small unit cells geometries. Dark current data, at high gain levels, indicate an effective gain normalized dark density count as low as 1000 counts/μs/cm^2 at an APD gain of 444. A junction doping profile was determined from capacitance–voltage data. Spectral response data shows a gain-independent characteristic.

INTRODUCTION

The HgCdTe mid-wave infrared (MWIR) electron injection avalanche photodiode (e-APD) has been demonstrated to provide uniform noiseless gain out to gains in the 10^2 to near 10^4 range.[1–3] Photon-level sensitivity has been demonstrated at DRS with high noise equivalent photon operabilities in 128 × 128[4] and 320 × 240[5] gated arrays using e-APD in the *p*-around-*n*, high-density vertically integrated photodiode (HDVIP®) cylindrical diode configuration. Given this performance, there is increasing interest in using the e-APD in a variety of applications. Therefore, there is a need to develop quantitative characterization data and models that can be used to predict performance. This paper will address the performance and fundamental operation of the cylindrical geometry HDVIP e-APD, and models that have been developed for use in predicting performance. In particular, this paper will present new updated analysis of the dependence of gain and excess noise factor on the geometry of the unit cell. A two-dimensional (2D) numerical model has been developed to permit the calculation of the time-dependent response for a cylindrical *p*-around-*n* diode or APD. This model permits the prediction of bandwidth and collection efficiency as a function of unit cell geometry, minority carrier mobility, and lifetime. The 2D model also permits the calculation of the point spread function (PSF) and crosstalk. The concept of effective gain normalized dark current is reviewed. These performance factors are inserted into expressions for the noise equivalent power (NEP) and noise equivalent photon (NEPh). The modeling section of the paper will conclude with

a discussion of the tradeoffs between cylindrical and planar geometries. In the e-APD characterization section of the paper, spot scan data is shown along with crosstalk data. New bandwidth and dark current measurement results on cylindrical e-APDs are shown. Experimental data on high-performance gated imaging FPAs compare the effective gain normalized dark current, $(gndc_{\mathrm{eff}})$, determined from measured noise, to the gain normalized dark current $(gndc)$. The data will show that the lower effective gain normalized dark current is the better predictor of FPA noise performance in agreement with the observations of Perrias et al.[2] Low detector capacitance is key to achieving high bandwidth and low power dissipation in the readout integrated circuit (ROIC). Capacitance–voltage data for 25 μm pitch cylindrical APD unit cells is measured and analyzed and shown to be close to prediction.

THE CYLINDRICAL HgCdTe e-APD

The HgCdTe e-APD is a top-side illuminated, cylindrical, p-around-n homojunction. Its geometry, process, and operation are described below.

Geometry

The HgCdTe avalanche photodiode (APD) detector design is based on the highly successful HDVIP architecture. The HDVIP device is a front-side illuminated, cylindrical, p-around-n photodiode that is formed around a small via in the HgCdTe, shown in Fig. 1, where the via serves as the interconnect conduit between the n-side of the photodiode and the input to the readout circuit. The HDVIP structure is currently employed in MWIR and LWIR staring arrays in production at DRS. The keys to the success of this architecture are: (1) interdiffused CdTe passivation of both surfaces for low $1/f$ noise, (2) thermal cycle reliability that is detector and array size independent, (3) low defects due to diode junction orientation with respect to threading dislocations, and (4) front-side illumination for high quantum efficiency and MTF.

Process

The MWIR HDVIP APD process begins with a p-type HgCdTe film, typically 5 μm to 7 μm thick,

which is passivated on both surfaces with interdiffused CdTe. After the film is epoxied to the readout, vias are etched through the membrane down to preamp input pads on the readout. The process of etching the vias creates an n-region around the via, and a cylindrical p-around-n junction photodiode is created. The via is implanted to create an n^+ surface, then metal is deposited in the via to provide the electrical interconnect from the n region to the preamp input. The result is an $n^+/n^-/p$ cylindrical junction.[6] An antireflection (AR) coating is applied to maximize the quantum efficiency.

APD Operation

The front-side illuminated $n^+/n^-/p$ HDVIP offers a rather unique APD architecture. With increased reverse bias, the depletion region in the n^- region grows until the n^- region becomes completely depleted. In addition, as the bias is increased, the electric field becomes large enough to cause avalanche multiplication. The depleted n^- region is the multiplication region for the APD. The p-region around this depletion region is the absorption region. Photogenerated minority carrier electrons in the p-region diffuse to the multiplication region. The p-around-n geometry favors electron injection.

Exponential gain voltage characteristics with gains over 1000 at 13 V bias with gain-independent excess noise factors of around 1.3 have been measured in 4.3 μm cutoff HDVIP e-APDs.[1] At cutoffs of 5.1 μm the required biased to obtain gains of 1000 is reduced to about 10 V. The discussion in the next section shows how the measured gain and excess noise factor are related to the geometry of the unit cell.

Gain, Fill Factor, and Excess Noise Factor

The optical gain of the e-APD is defined as the ratio of the signal at low bias, where the gain is unity, and the signal at high bias. In the non-APD mode of operation (gain = 1), the optically active area is the total pixel area (i.e., the pixel pitch squared) minus the area of the via and the area of grid metal and corner contact, if present. The device is typically designed so that the lateral widths of the n- and p-regions in the unit cell are less than the diffusion lengths of the minority carrier holes and
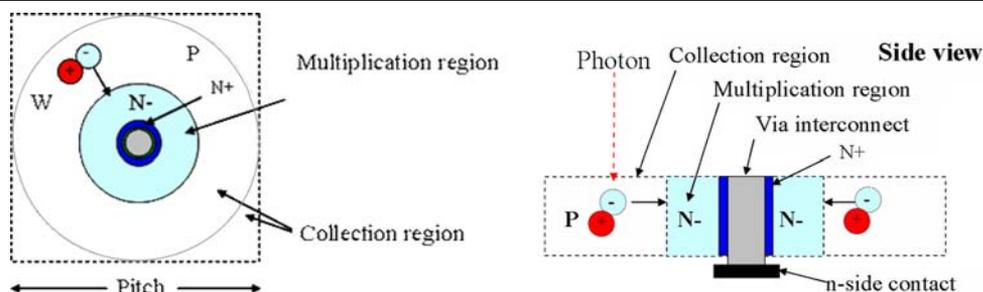

Fig. 1. HgCdTe electron avalanche photodiode unit cell: top view and cross section.

electrons in each region, respectively. The expression for the fill factor at unity gain, FF_{ug}, is given by:

$$FF_{ug} = \frac{p^2 - \left(Dia_{via}^2 \cdot \frac{\pi}{4} + A_{grid\_cc}\right)}{p^2},$$ (1)

where p is the pitch of the unit cell, Dia_{via} is the diameter of the via, and A_{grid_cc} is the area taken up by the distributed grid and substrate corner contact (described in Ref. 4), if present; for example, neglecting distributed grids and corner contact area, the unity gain fill factor for a 25 μm pitch unit cell with a 6 μm diameter via is 95.5%.

The measured optical gain of the APD (M_{measd}) is determined relative to the unity gain signal, which is set by this fill factor and the quantum efficiency. As the gain increases, the contribution from the carriers generated in the multiplication region becomes insignificant and the optical fill factor becomes approximately the area of the p-region within the pitch of the unit cell. The result is that the measured gain is less than the injected electron gain, which we define here as the true gain.

This loss in fill factor in the APD mode is reflected in the measured gain that is determined from the ratio of the signal at high gain, where the fill factor is lower, to the signal at unity gain, where the fill factor is high. Thus, the reported APD gains include the loss in fill factor. We define the true gain M as the number of electrons generated per electron injected from the collection region. These injected electrons receive full gain M as they must traverse the entire multiplication region. The true gain (M) is higher than the measured gain that is determined from the ratio of signals. The measured gain is related to the true gain by the following expression:

$$M_{measd} = \frac{FF_n + FF_p \cdot M}{FF_{ug}},$$ (2)

where the contribution of gained signal for carriers generated in the multiplication region is neglected; FF_n is the fill factor of the multiplication region (n region) and FF_p is the fill factor of the collection region (p region):

$$FF_p = \frac{p^2 - Dia_n^2 \cdot \frac{\pi}{4} - A_{grid\_cc}}{p^2},$$ (3)

$$FF_n = \frac{(Dia_n^2 - Dia_{via}^2) \cdot \frac{\pi}{4}}{p^2},$$ (4)

where Dia_n is the diameter of the n-region (gain region). The unity gain fill factor is the sum of the collection region fill factor and the multiplication region fill factor.

At high gains the measured gain will approach the ratio of the fill factor of the collection region to the unity gain fill factor:

$$M_{measd} = \frac{FF_p \cdot M}{FF_{ug}}.$$ (5)

We now address the fill factor effect on the measured excess noise factor. The excess noise factor is measured with an incident flux current I_ϕ that dominates the dark current. The measured excess noise factor, $F(M)_{measd}$, is calculated from:

$$F(M)_{measd} = \frac{I_n^2}{2 \cdot q \cdot I_\phi \cdot M_{measd}^2},$$ (6)

where I_n is the measured shot noise and I_ϕ is the measured input flux current (before gain).

Note that the expression for $F(M)$ uses the measured gain, which underestimates the true gain of the APD. This would imply that the measured excess noise factor is overestimated. The following analysis shows that the degree to which the measured excess noise factor exceeds the true excess noise factor is directly related to the APD fill factor.

The measured noise current given the true gain M and true excess noise factor of $F(M)$ is given by:

$$I_n^2 = 2 \cdot q \cdot I_{\phi p} \cdot M_p^2 \cdot F(M) + 2 \cdot q \cdot I_{\phi n} \cdot M_n^2 \cdot F(M),$$ (7)

where $I_{\phi p}$ is the flux generated current in the p-region, and M_p is the gain of the injected carriers from the p region (that we are defining as the "true" gain), $I_{\phi n}$ is the flux current generated in the multiplication region, and M_n is the average gain for optical carriers generated in the multiplication region. Note that hole multiplication has been neglected. The gain for carriers generated the multiplication region, M_n, will be considerably less than the gain of injected carriers, M_p, especially at high gains. Also $I_{\phi n}$ is much less than $I_{\phi p}$ because of the relatively small area of the multiplication region compared with the collection region. Therefore, at high gains the contribution from the n side is neglected and the expression for the current noise approaches:

$$I_n^2 = 2 \cdot q \cdot I_{\phi p} \cdot M^2 \cdot F(M),$$ (8)

where $M = M_p$, and solving for $F(M)$:

$$F(M) = \frac{I_n^2}{2 \cdot q \cdot I_{\phi p} \cdot M^2}.$$ (9)

At high gains the measured photocurrent I_ϕ will approach the multiplied current $I_{\phi p}$ from the p-region and the true excess noise factor in terms of the measured gain and input flux current becomes:

$$F(M) = \frac{I_n^2}{2 \cdot q \cdot I_\phi \cdot M_{measd}^2} \cdot \frac{FF_p}{FF_{ug}}.$$ (10)

Note that the first term in Eq. 10 is the measured excess noise factor defined by Eq. 6. Therefore, in

terms of the measured excess noise factor, the true excess noise factor is given by:

$$F(M) = F(M)_{\text{measd}} \cdot \frac{FF_{\text{p}}}{FF_{\text{ug}}}. \qquad (11)$$

The measured excess noise factor is inflated by the ratio of the unity gain fill factor to the high gain fill factor. Thus, the typical measured excess noise factor value of 1.3 actually reflects a true excess noise factor that is closer to 1.0.

Assuming unlimited gain, the fill factor only becomes important in that it determines the minimum achievable NEP at the point where the detector noise becomes dominant. It can be shown that the impact of APD fill factor on the ultimate detector noise limited NEP is $(FF_{\text{ug}}/FF_{\text{p}})^{1/2}$.

In the system noise limited regime the SNR will be directly proportional to the APD fill factor at a fixed true gain. If the gain at a fixed bias can be increased by increasing FF_{p} without degradation in performance due to decreased bandwidth or collection efficiency, then SNR and NEP will be improved. However, if bandwidth is critical a low FF_{p} can be compensated for by increasing the gain with additional bias.

Bandwidth, APD Fill Factor, and Collection Efficiency

The time it takes photogenerated electrons to diffuse to the multiplication region will determine the bandwidth of the device assuming that the bandwidth of the multiplication region and RC effects are not limiting. Because of the cylindrical geometry, the bandwidth is a function of the radial distance from the edge of the n^- multiplication region. As shown in Fig. 1, for a given pitch there is a parameter W_{p} that is the radial distance between the largest circle that fits in the unit cell and the multiplication region. The collection region width, W_{p}, and the unit cell pitch are critical dimensions that determine both the bandwidth and the APD fill factor (FF_{p}) of the cylindrical APD. In fact, for the cylindrical architecture there will always be a tradeoff between bandwidth and fill factor, as discussed below. The bandwidth of the multiplication process itself is greater than 1 GHz and is not a limitation for present device configurations.[1]

For optimal cases where the diffusion length of the electron is much greater than W_{p}, the width W_{p}, and the diffusion coefficient (D_{e}) of the electron, will determine this time. The pitch also plays a role in the cylindrical diode case, as will be shown later. If the diffusion length is much greater than W_{p}, then the response time constant in the one-dimensional case is roughly $W_{\text{p}}^2/D_{\text{e}}$ where D_{e} is the diffusion coefficient of the election. If the diffusion length is short compared with W_{p}, then the response time constant is approximately the lifetime of the minority carrier. Since D_{e} in MWIR and LWIR HgCdTe is relatively large due to the high electron mobility, the diffusion rate limited bandwidths were shown to reach levels of 2 GHz for W_{p} of 2 μm using a one-dimensional (1D) model.[1]

The 1D approximation has an obvious deficiency in the cylindrical case. In the cylindrical case the number of carriers generated per differential area and the diffusion time both increase with radius. In addition, in consideration of 2D square pixel array formats, contributions from the corner regions, which constitute 21% of the total area, between pixels are not included. Therefore, the carriers with a longer diffusion time weight the results more strongly when compared with the 1D model.

We have developed a time-dependent and steady-state 2D diffusion model for the calculation of the bandwidth and point spread response function of APDs (and cylindrical diodes) in a 2D array format. The 2D model allows the specification of the HDVIP FPA configuration: array size ($N \times M$ pixels), pixel pitch, pixel configuration (1×1, 2×2, etc. paralleled diodes), diode pitch, n^- region diameter, electron mobility, and electron lifetime. The spatial and temporal configuration of the optical input from flooded field down to a subpixel spot is also specified. The surrounding diodes are reverse biased. The model is a finite difference time-domain diffusion model written in C with a MATLAB interface.

The model produces a quantitative pulse response from which the 10–90% response times, τ_{rise}, can be calculated. These rise times are converted to equivalent bandwidths (BW) using an exponential approximation: $BW = 2.2/(2\pi\tau_{\text{rise}})$. In the flooded illumination case, the amplitude of the calculated pulse response reflects the product of the electron collection efficiency, $e\text{-}CE$, the fill factor of the pixel, and the bandwidth attenuation factor. The electron collection efficiency reflects loss of electrons due to recombination. The fill factor reflects the area of the collection region relative to the total pixel area, where it is assumed that carriers generated in the multiplication region contribute insignificantly to the signal.

The 1D model is compared in Fig. 2 to three cases to which the 2D model was applied: case 1 with a unit cell pitch of 35 μm, case 2 with a unit cell pitch of 20 μm, and case 3 with a unit cell pitch of 12 μm. An electron mobility of 3×10^4 cm^2/V/s and a lifetime of 8.67 ns, which are consistent with the 13.4 μm diffusion length determined from spot scan data (see "Spot Scan Measurement Results"), were used in the model calculations. A pixel in the center of an 8×8 array was selected. The surrounding pixels were reverse biased. An example of the 2D model square pulse time response calculation for the 35 μm pitch case with $W_{\text{p}} = 6.5$ μm is shown in Fig. 3. For this case the rise time is 3.87 ns (bandwidth of 90 MHz) with very close to 100% amplitude response to a 10-ns pulse.
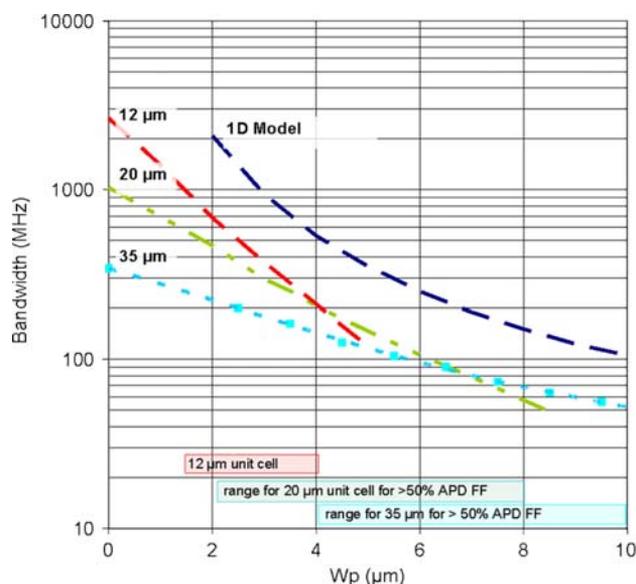
Fig. 2. Modeled 1D and 2D bandwidth as a function of diffusion distance parameter W_p and unit cell pitch. W_p range for APD FF >50% and diode diameter >4 μm are indicated for each unit cell pitch.
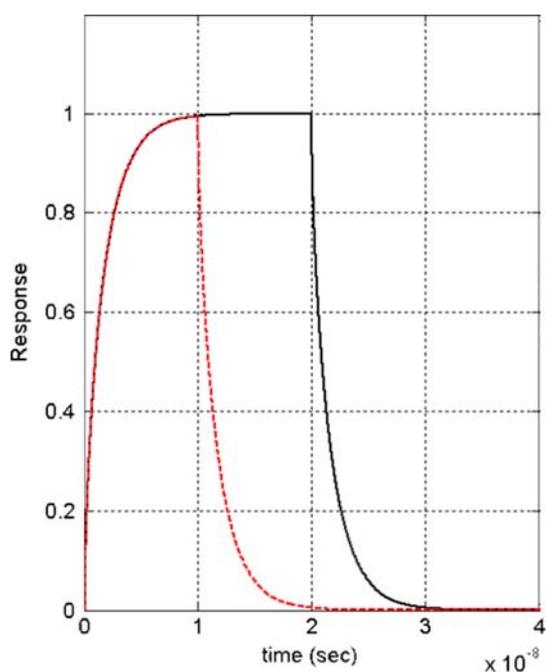

Fig. 3. Normalized 10 ns and 20 ns square pulse response of 35 μm unit cell with a 22 μm diameter diode (W_p = 6.5 μm).

The 2D model results in lower bandwidth predictions compared with the 1D model, as expected. The departure of the 2D model predictions from the 1D predictions is more pronounced as W_p approaches zero, since when $W_p = 0$, the corner regions are all that is left of the absorption region. For example in the 35 μm unit cell case the corner is 7.2 μm from the edge of the multiplication region

while for the 20 μm unit cell this distance drops to 4.1 μm. At this extreme, it is clear that as the pitch gets larger, the size of the corner region increases, causing the response time to increase even though $W_p = 0$.

Figure 2 shows that the MWIR HgCdTe e-APD cylindrical diode is predicted by the 2D model to reach bandwidths in excess of 1 GHz with APD fill factors over 50% for unit cell pitches of 12 μm. A 100 MHz bandwidth, by the way, is more than adequate for the detection of pulses down to 4 ns, and much shorter if there is enough gain margin; for example, in the case of $W_p = 6.5$ μm and pitch = 35 μm, the attenuation factor for a 4 ns square pulse is 0.90 (−0.46 dB), or a 10% loss in amplitude. In APDs with gain margin, this is easily made up for by a small increase in the APD bias. The same analysis has been applied to the detection of Gaussian pulses which require a higher bandwidth to maintain a given response amplitude.

Bandwidths above 1 GHz with reasonable APD fill factor in the cylindrical case will require smaller pitches, for example, 12 μm, as shown in Fig. 2. MWIR HDVIP FPAs have been successfully demonstrated with 12 μm diode pitch (however, APD operation in pitches this small has not been evaluated). Another possible approach to >1 GHz bandwidth with reasonable fill factor would be to design the optics to concentrate the flux in a doughnut shape in the high bandwidth central region of the unit cell (outside the multiplication region). Eventually the bandwidth will be limited by RC time constant and multiplication region transit time effects. These issues are being addressed in ongoing work at LETI.[7,8]

There are two other factors besides bandwidth in the geometry tradeoff between pitch and multiplication region diameter: (1) electron collection efficiency, e-CE, and (2) APD fill factor. The electron diffusion length, L_e, at 80 K for MW HgCdTe and the p-type material currently in use at DRS is about 13 μm (see "Spot Scan Measurement Results"). Care must be taken in the design of the unit cell to minimize the diffusion distances in the unit cell with respect to L_e to prevent the loss of carriers to recombination, which degrades e-CE. The APD fill factor, previously defined as FF_p will be degraded as the diode diameter increases; however, the bandwidth electron collection efficiency will increase. There is an optimum geometry for the detection of a laser pulse with a particular pulse width. If the application requires a pixel that is larger than the optimum unit cell size, then the pixel can be composed of an $N \times N$ array of unit cells connected together in parallel, as discussed in a previous paper.[1]

The tradeoff for optimum pulse detection is illustrated in Fig. 4 for a 35 μm unit cell designed for the hypothetical detection of a 10 ns wide square laser pulse. The pulse attenuation factor is the loss factor in peak amplitude relative to the wide pulse
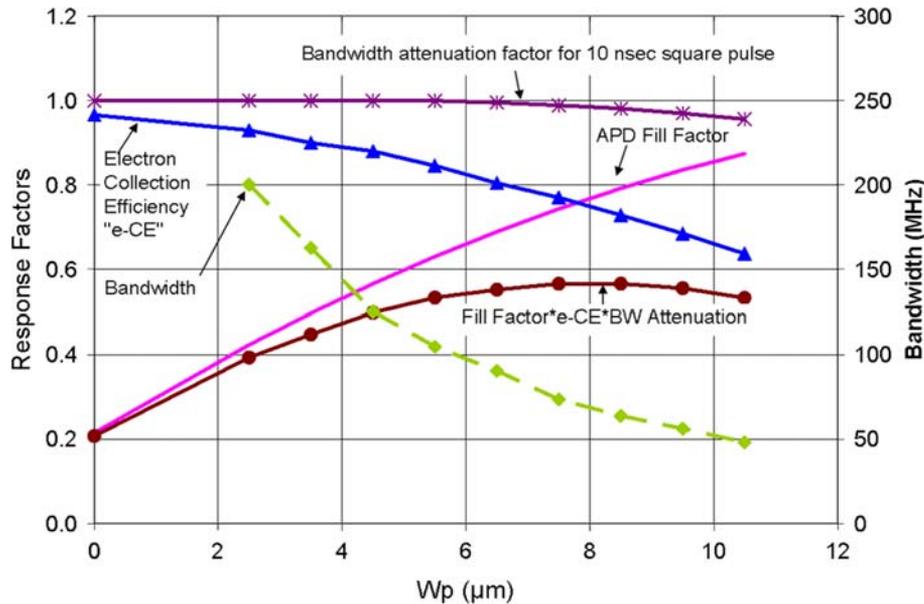
Fig. 4. Tradeoff analysis for optimal detection of 10 ns square pulse for 35 μm unit cell e-APD using 2D model. Optimum response occurs at $W_p = 8$ μm (diode diameter = 19 μm).

response limit. This is of course intimately related to the bandwidth (rise time) of the APD. As can be seen there is an optimum value of W_p of 8 μm for maximum response to a 10 ns pulse that reflects these tradeoffs. In this example, the corresponding optimum diode diameter would be 19 μm.

Point Spread Function

We define the point spread function (PSF) as the spatial response function of a pixel that would be measured as a photocurrent as an infinitely small spot is scanned across the pixel. Response from beyond the geometrical area of the pixel represents crosstalk. The Fourier transform of the PSF is the 2D modulation transfer function (MTF) for the pixel.

The 2D diffusion model (in the steady-state response mode) was used to calculate the theoretical PSF of a cylindrical diode pixel in an $M \times M$ array of reverse-biased diodes. In the case of an APD array the multiplication region is considered optically inactive. As stated previously, the model also allows the calculation of the PSF of $N \times N$ diode pixel configurations. In Fig. 5 we plot the calculated $\log_{10}$ PSF for a 2×2 64 μm pixel with four 18 μm diameter 32 μm pitch diodes hooked together in parallel. The $\log_{10}$ PSF is plotted to show the response behavior outside the pixel. The assumed electron diffusion length, L_e, was 13.4 μm. This value of L_e was obtained from the measurement described in the section "Spot Scan Measurement Results". The assumed spot size was a 0.78 μm diameter Ares disk which closely approximates the infinitely small spot for the geometry and diffusion length used.

If one defines the crosstalk as the ratio of the integrated signal from an immediately adjacent pixel to the integrated signal in the pixel itself, then a 5.1% carrier crosstalk is calculated from the PSF shown in Fig. 5.

Dark Current

Dark current limits the ultimate gain and performance of an APD, assuming background flux is negligible. With a significant background flux present, dark current is still important in as much as you want the dark current to be insignificant compared with the flux current to achieve highest possible performance. As the bias and gain of an APD are increased, there will be a point where the dark current starts increasing significantly faster than the gain and the device becomes dark current noise limited. For this reason, measurements of dark current as a function of bias or gain are important in determining the maximum useful gain and in predicting APD performance for a given ROIC noise level.

A question arises: What is the expression for the shot noise associated with the dark current that is measured? There are a couple of illustrative scenarios. If the dominant dark current is a diffusion current from the p side of the junction, then the dark current is gained up by the full gain M of the APD, discussed above. In this case the measured current noise is given by:

$$I_n = \sqrt{2 \cdot q \cdot gndc \cdot M_{measd}^2 \cdot F(M)_{measd}}, \quad (12)$$

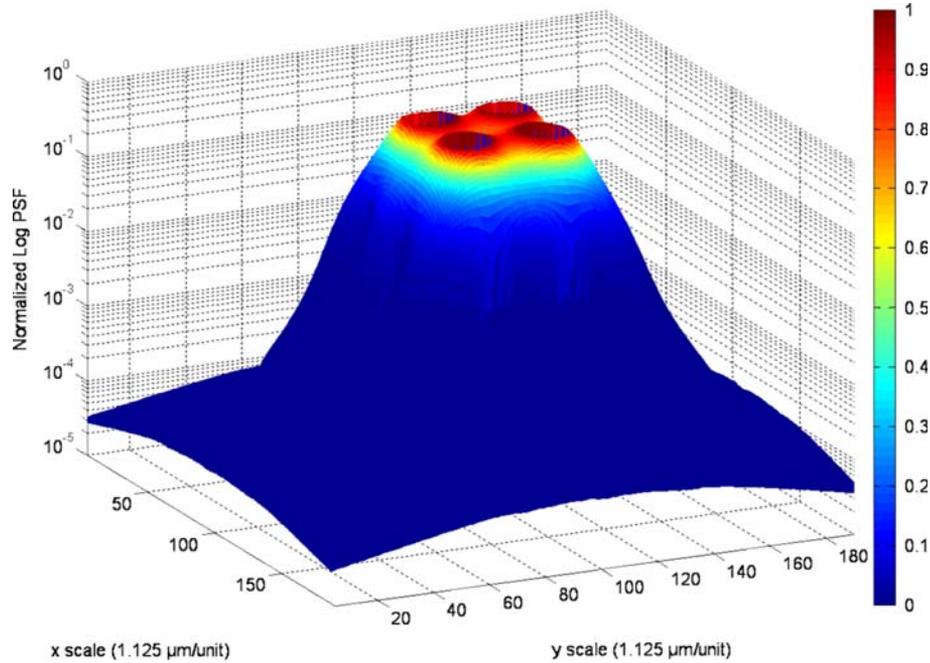where $gndc$ is the measured gain normalized dark current (equal to the measured dark current I_{dark}

Fig. 5. Log$_{10}$ PSF, side view, of a 64 μm pixel with 2 × 2 18 μm diameter diodes at 77 K (scale factor 1.125 μm per step).

divided by the measured gain), M_{measd} is the measured gain, and $F(M)_{\text{measd}}$ is the measured excess noise factor. The emphasis on "measured" is to point out the self-consistency of the analysis. In terms of the measured dark current, I_{dark}, which is assumed to be fully multiplied, the noise is given by:

$$I_{\text{n}} = \sqrt{2 \cdot q \cdot I_{\text{dark}} \cdot M_{\text{measd}} \cdot F(M)_{\text{measd}}}. \qquad (13)$$

On the other hand, assume the dark current is not avalanche gain multiplied, for example, dark current created in the n^+ contact region. We could argue that this dark current is not multiplied due to the lack of hole multiplication, in which case the noise would be given by:

$$I_{\text{n}} = \sqrt{2 \cdot q \cdot I_{\text{dark}}}. \qquad (14)$$

The difference in predicted noise between multiplied and non-multiplied noise is a factor of $[M \times F(M)]^{1/2}$, which can be very significant at high gains.

Another more likely case would be dark current generated in the depletion region. Electrons in this case would experience some average gain $\langle M_{\text{gr}} \rangle$ less than M depending on where they are generated and the applied bias:

$$I_{\text{n}} = \sqrt{2 \cdot q \cdot I_{\text{dark}} \cdot \langle M_{\text{gr}} \rangle \cdot F(M)_{\text{measd}}}. \qquad (15)$$

Since the gain for the dark current in this case is not easily found, we need to define an effective

gain normalized dark current based on the measured noise. The effective gain normalized dark current would be the dark current producing the measured noise, if the dark current were fully multiplied. This issue for the e-APD was first treated theoretically and experimentally by Perrias et al.[2]

$$gndc_{\text{eff}} = \frac{I_{\text{n}}^2}{2 \cdot q \cdot M_{\text{measd}}^2 \cdot F(M)_{\text{measd}}}. \qquad (16)$$

Note that I_{n} is the measured noise. The above equation has been used to obtain $gndc_{\text{eff}}$ from measured FPA noise data.

This analysis indicates that the measured dark current, when normalized to the measured optical gain, will predict a worst-case detector dark current limited noise. Researchers at CEA-LETI have found from noise measurements that the noise is one-third to one-fifth of the noise predicted from the gain normalized dark current on the MWIR e-APD at 80 K.[2] These finding have been corroborated in our work as discussed in the section "Effective Gain Normalized Dark Current".

Narrow Pulse NEP$_\lambda$ and NEPh

We take these concepts to obtain an expression for the narrow pulse noise equivalent power of the FPA, NEP_λ, which is defined as the peak pulse peak signal power that produces a signal equal to the root-mean-square (rms) noise. NEP at the wavelength λ can be expressed in terms of the rms current noise In_{total} (rms A) and the current responsivity R_λ (A/W) at wavelength λ.

$$NEP_\lambda = \frac{In_{\text{total}}}{R_\lambda}, \qquad (17)$$

where R_λ is given by:

$$R_\lambda = \frac{q \cdot \eta_{\text{CE}} \cdot \eta_{\text{opt}} \cdot \eta_{\text{bw}} \cdot FF_{\text{ug}} \cdot \lambda \cdot M_{\text{measd}}}{h \cdot c}, \qquad (18)$$

where η_{CE} is the electron collection efficiency, η_{opt} is the optical quantum efficiency of the active region, and η_{bw} is the pulse response amplitude attenuation factor discussed in the section "Bandwidth, APD Fill Factor, and Collection Efficiency".

The quiescent rms current noise, In_{tot} is given by:

the wavelength of interest. For the detection of short wavelengths a 2 μm thick absorption layer could be used and very high bandwidths should be achieved. However, the detection of longer wavelength photons requires a thicker absorption region and hence lower bandwidth; for example, to detect 4 μm photons with high efficiency you would want an absorption region thickness of around 6 μm for maximum quantum efficiency. Note that there is no wavelength (photon absorption depth) dependence of bandwidth (or gain) in the cylindrical diode case (see "Spectral Response Versus Gain"). On the other hand, bandgap grading in the absorption region will improve the planar

$$In_{\text{tot}} = \sqrt{2 \cdot q \cdot (gndcd_{\text{eff}} \cdot A_{\text{det}} + I_\phi) \cdot M_{\text{measd}}^2 \cdot F(M) \cdot NBW + In_{\text{ROIC}}^2}, \qquad (19)$$

where A_{det} is the area of the pixel (i.e., the pitch squared), I_ϕ is the background photon flux induced current, NBW is the noise bandwidth, and In_{ROIC} is the input referred rms current noise of the ROIC. To calculate the signal shot noise limited noise, the signal current is substituted for I_ϕ. The broadband NEP_λ for pulse detection is calculated from Eqs. 18 and 19.

For gated detection, the noise equivalent photon, $NEPh$, is defined as the number of integrated photons that create a signal equal to the quiescent noise. The $NEPh$ is related to the gate time (τ_{gate}). This expression is used to predict $NEPh$ performance for our fated FPAs:

device bandwidth as will planar designs that allow the imposition of a drift field in the absorption region.

The electron collection efficiency of the planar diode should be high ($> 90\%$) for layer thicknesses $< L_e$. In spite of these apparent advantages, there are a couple of important factors in favor of the cylindrical diode: high operability at higher gain levels in array formats, lower dark currents, higher quantum efficiency (due to front-side illuminated architecture). The capacitance per unit pixel area of the cylindrical diode will be comparable to or lower than that of the planar diode.

At the present time it is strongly suspected that the cylindrical HDVIP APD may provide higher

$$NEPh = \frac{\sqrt{\left[M_{\text{measd}}^2 \cdot F(M)_{\text{measd}} \cdot (gndcd_{\text{eff}} + J_\phi) \cdot A_{\text{det}} \cdot \frac{\tau_{\text{gate}}}{q} \right] + N_{\text{ee}}^2}}{M_{\text{measd}} \cdot \eta \cdot FF_{\text{ug}}}, \qquad (20)$$

where J_ϕ is the background flux generated current density and N_{ee} is the noise equivalent input of the readout in terms of rms electrons.

Cylindrical e-APD Architecture Versus Planar

We now compare the cylindrical e-APD diode to the planar e-APD. In terms of fill factor, the backside illuminated planar architecture will have an APD mode fill factor near 100%, assuming that all the photons are absorbed in p-type absorption region. Second, based on the I-D model calculations shown in Fig. 2, it would appear that a planar backside illuminated n-on-p APD[3] can achieve greater bandwidth. However, the thickness of the p-type absorption layer must be adjusted to accommodate

array operabilities at a given gain level than the planar architecture. This supposition is based on the greater immunity of the cylindrical diode to threading dislocations that result from the liquid-phase epitaxy (LPE) or molecular-beam epitaxy (MBE) growth processes currently employed. This greater immunity arises from the lower cross section of the p–n junction for such defects that intersect the p–n junction in the cylindrical diode case. The cylindrical junction is parallel to the direction of the threading dislocation. In the planar diode the junction is perpendicular to this direction and much more likely to be intersected by threading dislocations. The effect of these differences will be seen in comparisons of the maximum useful gain that can be achieved with low dark current. We believe that

the dark current density levels reported in this paper for cylindrical e-APD FPAs at high gains are the lowest reported to date in the realm of MWIR APD FPAs. As evidence for these claims we note that cylindrical e-APDs have demonstrated 128×128 and 340×240 gated imaging FPAs with SNR operabilities, at high gain levels, of greater than 99% at detection levels in the single-photon region.[4,5]

As material growth defects are reduced, the planar e-APDs should have an advantage. However, the gain margin due to the intrinsic advantage of the DRS cylindrical, low-defect, LPE-based, e-APD in terms of dark current defects and the low dark current levels, means that the current NEP sensitivities are likely to be competitive, if not superior, to structures claiming 100% fill factors, especially when operability is taken into account.

e-APD Characterization

Gain and excess noise factor have been reported extensively on e-APD single pixels and FPAs.[1,4,5] Here we will present additional experimental characterization results on the MWIR cylindrical e-APD. In particular, we will show spot scan measurements on e-APDs which are analyzed to determine crosstalk. We also present recent gain normalized dark current (*gndc*) data measured on test diodes and FPAs, and effective gain normalized dark current (*gndc*$_{\text{eff}}$) data measured on FPAs.

Spot Scan Measurements

Spot scan data on back-side illuminated planar e-APDs were first reported by Reine et al.[3] Here we present spot scan data on top-side illuminated MWIR cylindrical APDs operating at 77 K.

Laser Spot Scanner Design and Operation

We have developed a long focal length laser spot scan apparatus that will perform PSF and crosstalk measurements on pixels and test diodes cooled in a cryogenic Dewar. The spot scan system, in the nonpulsed (DC) mode allows the measurement of the pixel PSF and interpixel carrier crosstalk due to photogenerated carrier diffusion. A software and hardware interface is being created that will allow us to perform pulsed and DC spot scan measurement on an operating FPA. In pulsed laser mode, the apparatus will be capable of measuring the dynamic signal crosstalk in the FPA. The optical configuration is shown in Fig. 6. The optical source is a 635 nm laser diode.

For pixel PSF and carrier cross talk measurements, the system scans a small 2 μm spot over a designated area of the FPA or test structure in 2 μm steps while the device under test is connected to a Keithley 236 source measurement unit (SMU) that provides the detector bias and measures the photo signal. The test device is placed in a down-looking
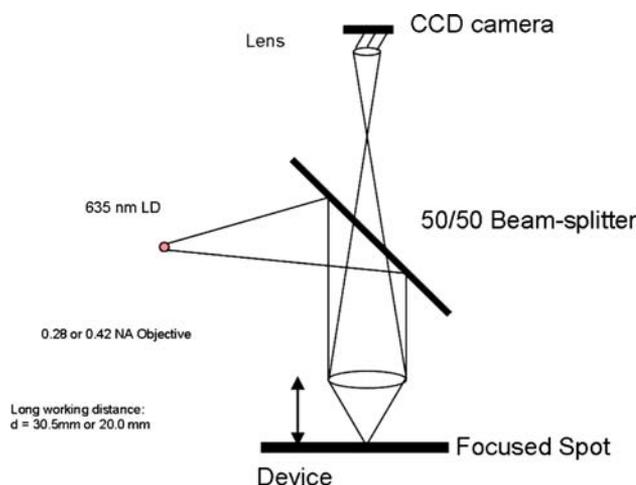

Fig. 6. Spot scan optical setup.

(or side-looking) Dewar. A three-axis stepping motor permits focusing and scanning of the spot, controlled by a graphical user interface (GUI). The GUI is used to set the spot position and scanning area. The GUI also controls the step delay, the number of averages, and the data display and output. A real-time image of the spot scan data is displayed in the GUI. A charge-coupled device (CCD) camera (Fig. 6) displays a real-time image of the test region on a separate monitor. An autofocus capability is planned as one of the next improvements. The system is easily converted from down-looking to side-looking Dewar configurations.

Spot Scan Measurements Results

Pixel Point Spread Function and Crosstalk The spot scan apparatus was used to measure the response profile and crosstalk of a MWIR 64 μm APD pixel with a 2×2 configuration of 32 μm pitch diodes with estimated junction diameters of 16 μm. The pixel was in an 8×8 fanout array configuration, operated at 80 K. A centrally located pixel was connected to a current monitor. We had the option of reverse biasing or floating the surrounding diodes. The spot scan was measured at reverse biases of 0.1 V bias and 4.6 V where the APD gain is about 6. A plot of the spot scan results at 4.6 V (APD gain = 6), with the surrounding diodes reverse biased, is shown in Fig. 7. The surface plots at low and high bias show the response around the vias. An enlargement of the reduced response region around the via at 4.6 V (compared with the 0.1 V bias case) was seen in the normalized response plots that reflects the relative nonresponsiveness of the *n*-region in the APD mode. The small spots in the 4.6 V case appear to be due to retroreflections from the vias.

The crosstalk, defined as the ratio of the integrated signal over the 64 μm $\times$ 64 μm area of the neighbor to the integrated signal over the
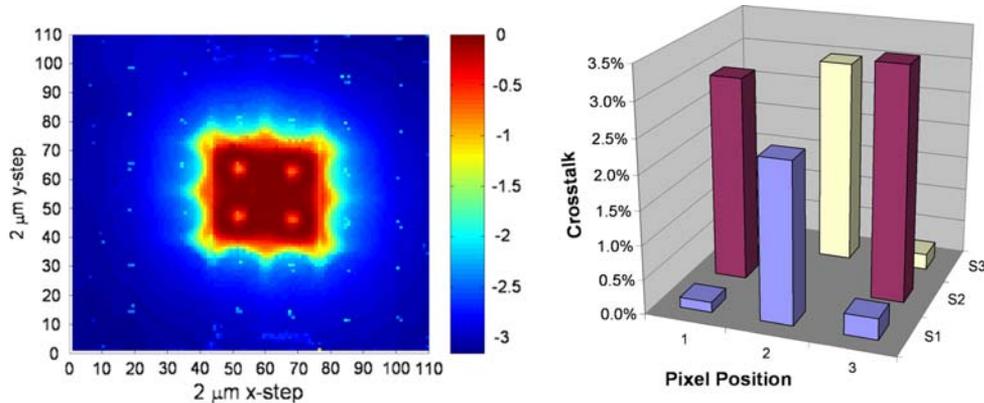
Fig. 7. Normalized spot scan profile of 64 × 64 μm pixel at gain of 6 at 4.6 V bias (Log$_{10}$ z-axis) and crosstalk in nearest neighbors.

64 μm × 64 μm area of the device under test, was calculated for each of the eight nearest neighbors and is also shown in Fig. 7. The crosstalk was not corrected for the finite size of the spot. The mean orthogonal nearest neighbor crosstalk in the low bias test was 2.4%. In the high bias test the calculated mean orthogonal nearest neighbor crosstalk was 2.9%. The crosstalk should be independent of gain, so the rough agreement between the two measurements is expected. The crosstalk measurement was repeated on a 64 μm pixel in another 8 × 8 array from another lot which had a slightly smaller estimated diode diameter of 14 μm. This time the measured nearest neighbor crosstalk was 3.7%. The larger crosstalk in this case may be due to the smaller diode diameters.

The 2D model predicted 4.5% to 5.6% crosstalk for the 18 μm diameter junction case. The lower measured crosstalk is likely due to the presence of 2 μm wide substrate contact grid lines around the pixel, not included in the model.

Diffusion Length Measurements Critical in modeling the performance of the APD is knowledge of the diffusion length as it determines the carrier collection efficiency and the crosstalk of the pixel. If the diffusion length is too short compared with the junction separation, then a significant fraction of the photogenerated carriers will recombine before reaching the junction and the collection efficiency will be degraded. Diffusion length also has an obvious effect on interpixel carrier crosstalk. By measuring the falloff in signal with distance, the diffusion length can be estimated. It helps if the geometry is simple in order to derive the diffusion length accurately. The long guard diode ring around the 8 × 8 fanout provided a simple geometry for analysis that can be compared with the 2D model. The guard diode is a string of diodes about 512 μm long on each side. Since 512 μm is long compared with the estimated electron diffusion length, the signal falloff from the guard diode outward to the substrate contact gives us the desired 2D geometry.

The substrate contact is about 39 μm away from the guard diode, which should be far enough to allow the measurement of the diffusion length.

The spot scan method we are using in these measurements is known as the optical beam induced current (OBIC) technique. OBIC was used to measure the electron diffusion length of SWIR HgCdTe at room temperature.[9] We applied the spot scan technique (OBIC), in conjunction with the 2D model, to determine the electron diffusion length at 77 K in MWIR HgCdTe.

The data, plotted in Fig. 8, show an exponential falloff in the signal going toward the contact with an exponential slope of $(11.6 \ \mu\text{m})^{-1}$. The 2D diffusion model fit reveals a diffusion length of 13.4 μm. At 77 K, for an electron mobility of 3×10^4 cm^2/V s, the corresponding electron lifetime is 8.67 ns.

Bandwidth

We measured the pulse response and bandwidth of APDs as a function of unit cell pitch. A high-speed InGaAs detector was used to calibrate the pulse output of the laser. The laser pulse rise time was 0.8 ns (bandwidth = 440 MHz). The bandwidth data were measured in a high-bandwidth Dewar at a temperature of 132 K on a number of test devices with pixel pitch from 21.3 μm to 40 μm and with varying junction diameters is shown in Fig. 9, which plots the bandwidth (pulse response rise time determined as defined in the section "Bandwidth, APD Fill Factor, and Collection Efficiency") as a function of estimate values for W_{p}. Even though the device temperature was 132 K, the bandwidths are in rough agreement with the model that assumes an 80 K electron mobility of 3×10^4 cm^2/V s and a 8.67 ns lifetime. At 132 K the bandwidth will be lower than at 80 K, since the electron mobility will be lower and the lifetime will be longer. At long W_{p} values the bandwidth seems to be approaching the limit set by the minority carrier lifetime. At the short W_{p} limit the measured bandwidths are approaching the limit of the laser.
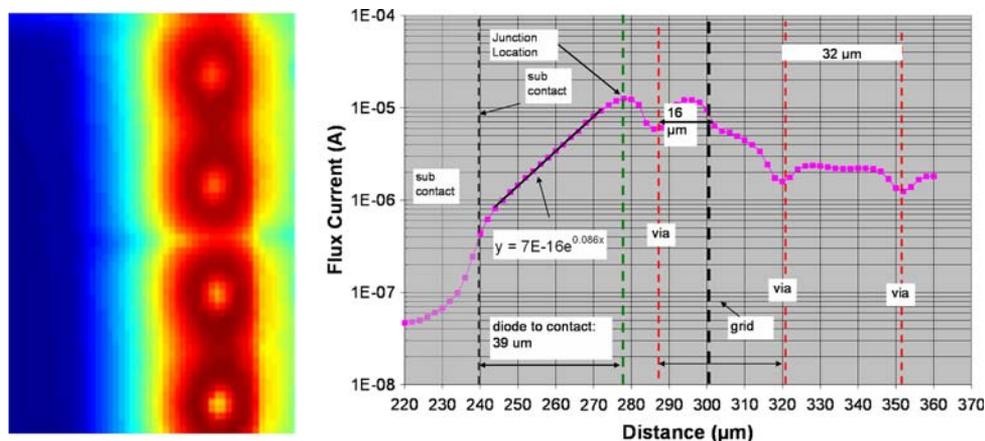
Fig. 8. Spot scan image of guard diode and flux signal current versus position through center of guard diode and adjacent pixel.
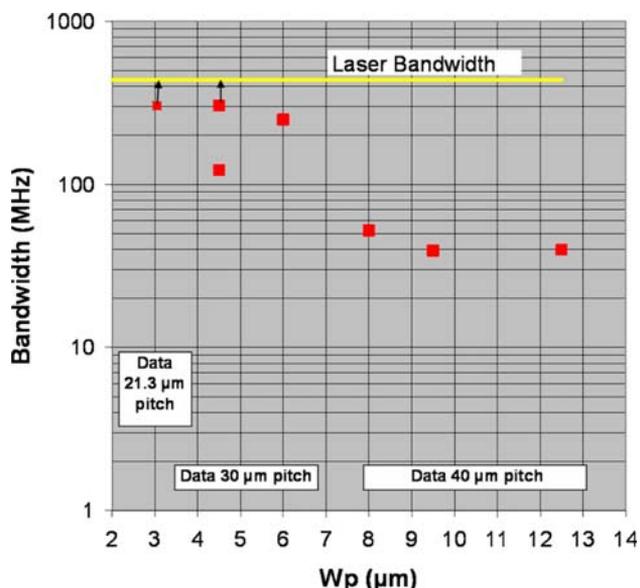

Fig. 9. Bandwidth at 132 K as measured by laser pulse rise time. The W_p values are rough estimates based on a junction diameter that is assumed to be three times the via diameter.

Dark Current

The first high bias dark current data on MWIR HDVIP e-APDs was reported in 2004. Since then we have improved our measurement setup and, as a consequence, obtained a better picture of the dark current behavior. The improvement reduced the system limited leakage currents measured in our 77 K test setup. In addition, we have improved our passivation process with the result that very low high bias dark currents have been reproducibly realized. An example of gain and gain normalized dark current density (*gndcd*) measurements out to 12 V on a 4.4 μm cutoff 64 $\times$ 64 μm e-APD with the new passivation process is shown in Fig. 10. Notice that the gain normalized dark currents are very low (<1 nA/cm^2 or <0.26 e/μs). Note that below 8 V the current is system "noise" limited. Above 8 V the gain normalized dark current increases faster than the gain, implying that the dark current is not a diffusion current from the *p*-side of the junction. Thus we would expect that this dark current is not fully gain multiplied. Note also that one device (5,5) shows the abrupt breakdown phenomena described by Reine et al.[10]
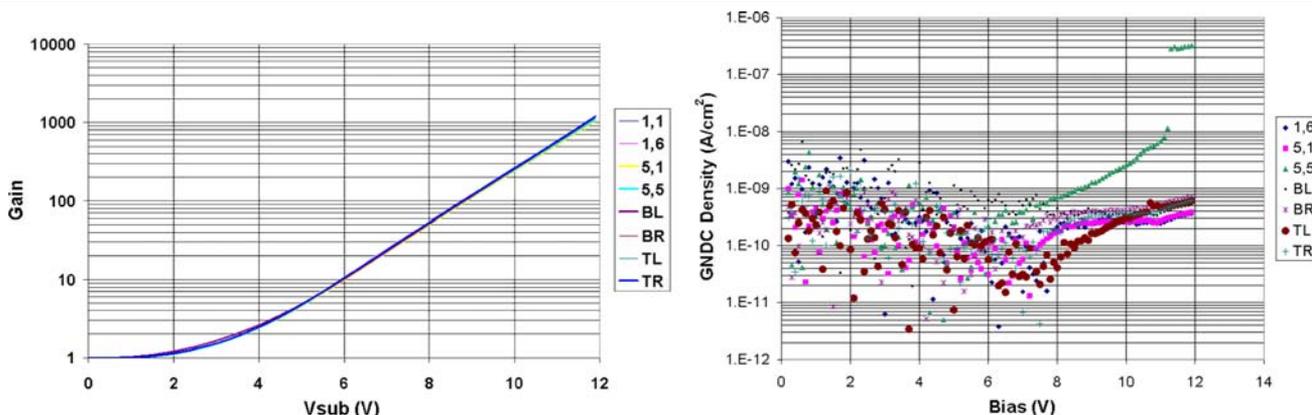

Fig. 10. Gain versus bias and gain normalized dark current density versus bias at 77 K on D4989-4A test pixels.
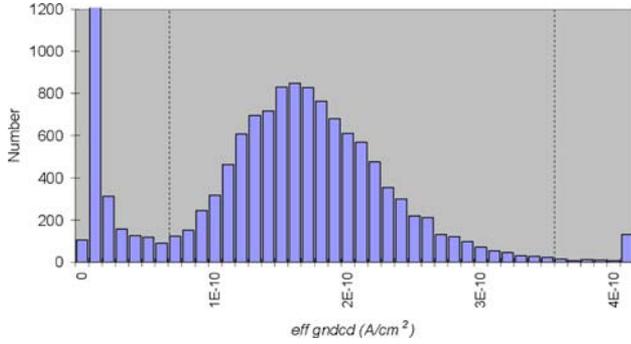
Fig. 11. Effective gain normalized dark current for A5087-4A at 10 V showing boundaries defining "good".

Effective Gain Normalized Dark Current

We determined the effective gain normalized dark current from noise measurements on gated APD FPAs. For a gated readout with an integration capacitor C_{int}, it can be shown that the effective gain normalized dark current density $gndcd_{\text{eff}}$ is related to the rms noise voltage at high bias low flux, $VnHBLF$ through the relationship:

$$gndcd_{\text{eff}}$$

$$= C_{\text{int}}^2 \cdot \frac{VnHBLF^2 - Vn_{\text{sys}}^2}{q \cdot M_{\text{measd}}^2 \cdot \tau_{\text{gate}} \cdot A_{\text{det}} \cdot F(M)_{\text{measd}} \cdot ROICGAIN}, \tag{21}$$

where Vn_{sys} is the system noise, and $ROICGAIN$ is voltage gain of the ROIC.

We recently measured the effective gain normalized dark current density from noise measurements and compared this to gain normalized dark current density obtained from dark current measurements.

Our noise measurements on a 128×128 gated imaging FPAs[4] with 40 μm pitch pixels (Fig. 11) revealed a good median $gndcd_{\text{eff}}$ of 0.164 nA/cm^2 at

10 V bias where the median APD gain was 444. The median $gndcd$ at 10 V was 0.447 nA/cm^2, which is 2.7 times the measured $gndcd_{\text{eff}}$. These low effective dark currents at high gain levels translate to sub-photon detection sensitivities. The device in Fig. 11 had a median $NEPh$ of 0.55 photons at a gain of 444 at 77 K for a 1 μs gate time. This result concurs with results obtained at LETI.[2]

Capacitance–Voltage Characterization

Capacitance–voltage data were measured on cylindrical diodes that were formed around 6 μm vias. The estimated junction diameter was 18 μm. A 6 μm via and 18 μm diameter n^- region is a fairly typical geometry for our e-APDs. The measurements were made on multidiode test structures composed of 16, 96, 384, and 864 diodes connected together in parallel. The tests were conducted at 77 K on a sample with a 4.3 μm cutoff.

The analysis assumed a cylindrical geometry and an abrupt n^-/p^+ junction. In this case, it can be shown that the radial position of the depletion region with respect to the center of the unit cell r_{dn} can be calculated from the measured capacitance, C:

$$r_{\text{dn}} = r_{\text{j}} \cdot e^{-\frac{2 \cdot \pi \cdot h \cdot \varepsilon_{\text{s}}}{C}}, \tag{22}$$

where r_{j} is the radius of the n^- region and h is the thickness of the semiconductor (height of cylinder). The doping concentration, N_{D}, at this position is giving by:

$$N_{\text{D}} = \frac{1}{q} \cdot \frac{1}{(2 \cdot \pi^2 \cdot h^2 \cdot \varepsilon_{\text{s}})} \cdot \frac{1}{r_{\text{dn}}^2} \cdot \frac{1}{\frac{d}{dV}\left(\frac{1}{C^2}\right)}. \tag{23}$$

The capacitance per diode was obtained from the slope of the capacitance versus number of diodes data measured at each bias voltage. The capacitance per diode versus bias data with the external stray capacitance removed is shown in Fig. 12. At 5 V
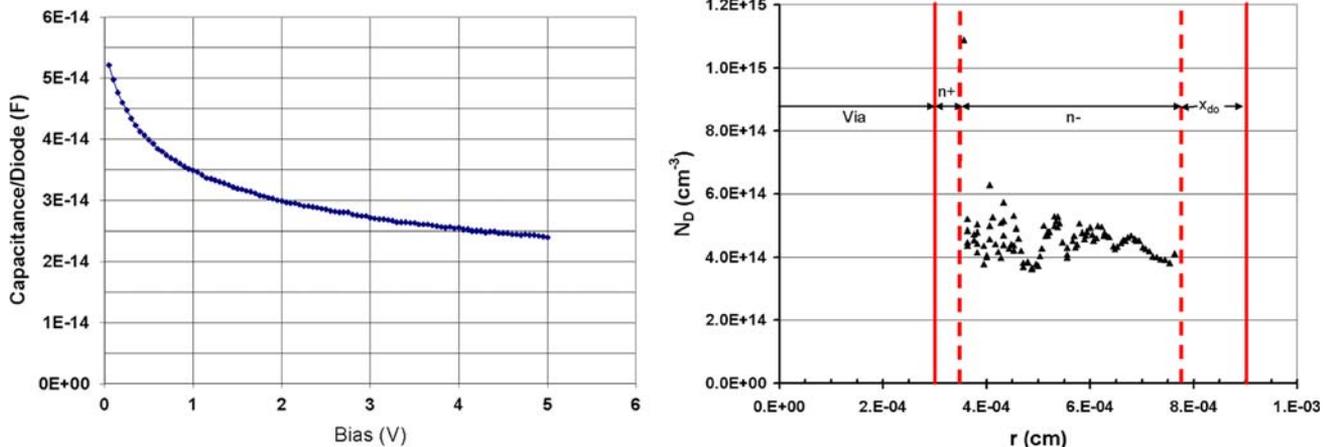

Fig. 12. Capacitance per diode versus bias measured on MWIR cylindrical diodes at 77 K and calculated doping profile (N_{D} versus radial position) assuming 18 fF of internal stray and 6 μm junction spread.

bias the capacitance appears to be approaching a limiting value as the depletion region approaches the narrow n^+ region around the via. The junction profile, also shown in Fig. 12, was calculated from the C–V data using the above equations assuming the estimated junction radius, r_j, of 9 μm and an internal stray capacitance per diode of 18 fF. The profile indicates the positions of the via, n^+ region, n^- region, depletion region at zero bias x_{do}, and the junction position. This analysis revealed that an internal stray capacitance correction of about 18 fF per junction (due to stray capacitance between the substrate and fanout metallization) was needed to obtain physically realistic data. The capacitance per diode (with the internal and external stray capacitance subtracted) is about 6 fF at 5 V, and the n^- doping level is 4×10^{14} cm^{-3} to 5×10^{14} cm^{-3}. The inferred capacitance of 6 fF is close to the 6.01 fF prediction for a 6 μm thick cylindrical diode with a junction radius of 9 μm and 5.5 μm wide n^- region that is fully depleted, and a dielectric constant of 17.

Spectral Response Versus Gain

The cylindrical geometry of the HDVIP photodiode results in a wavelength-independent gain and bandwidth. This is due to the fact that the lateral radial position with respect to the multiplication region is independent of how deep in the material the hole–electron pair is photogenerated. We verified this behavior by measuring spectral response as a function of gain on one of our MWIR e-APDs at 80 K. The spectral response data are plotted in Fig. 13 for biases from 0.5 V (unity gain) to 5 V (gain of about 10). The plots show a uniform increase in response versus wavelength as the APD gain is increased. The normalized spectral data (Fig. 13) show no evidence of a significant wavelength dependence of gain, as expected. The wavelength dependence of APD gain versus APD geometry is discussed by Stillman and Wolfe (on pp. 357–360).[11]

SUMMARY AND CONCLUSIONS

The key aspects of cylindrical p-around-n e-APD operation and performance have been described. One of the key features is the APD mode fill factor, defined as the ratio of the collection region area to the total area of the pixel. It has been shown that, in the high gain region at fixed bias voltage, the measured gain is degraded by the ratio of the APD mode fill factor to the unity gain fill factor. The excess noise factor is increased by the reciprocal of the same ratio.

In ROIC or system noise limited scenarios, gain is the key factor, and gain margin due to defect immunity is a highly desirable attribute that the cylindrical e-APD has demonstrated. In the detector noise limited region, the maximization of the APD collection region fill factor results in optimum SNR performance.

A 2D time-dependent diffusion model has been developed to achieve more accurate bandwidth and collection efficiency predictions for cylindrical diodes as a function of unit cell geometry. Optimum performance in terms of maximized response for a given unit cell pitch involves the determination of the optimum junction diameter that maximizes the product of the electron collection efficiency, bandwidth attenuation factor, and APD mode optical fill factor. An example was shown that illustrates how the 2D model can be used to obtain the optimum unit cell design.

In steady-state mode, the 2D model is being used to predict point spread function, MTF, and crosstalk of cylindrical diodes and APDs.

Planar geometries were compared with cylindrical geometries and found to offer the possibility of greater bandwidth and close to 100% APD mode fill factor. However, the large junction area, which is positioned perpendicular to the epitaxial growth direction, is a cause for concern as far as intersecting defects and capacitance are concerned in comparison with the cylindrical APD.
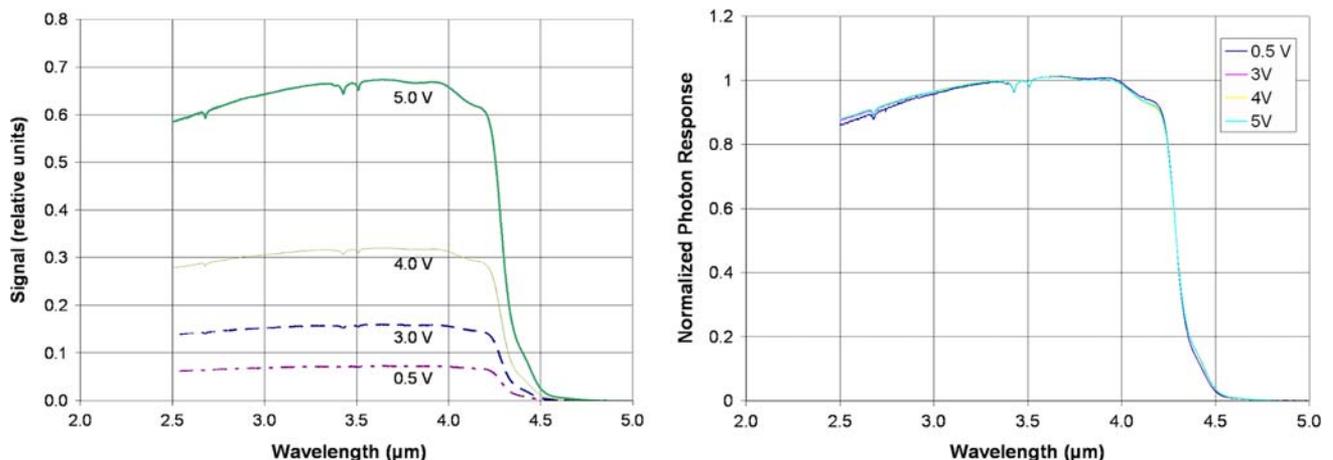

Fig. 13. Relative spectral response versus bias and normalized spectral response versus bias.

A microspot scan system has been developed and demonstrated, showing great value in determining important APD characteristics such as the point spread function, crosstalk, and MTF. The spot scan system has also been used to determine an experimental value for the electron diffusion length, which is a key parameter in performance modeling. In pulsed mode it will also be possible to use this system for measurements of total crosstalk in FPAs.

The effective gain normalized dark current ($gndc_{\text{eff}}$) is defined as the fully multiplied dark current that would be required to produce the measured noise. The $gndcd_{\text{eff}}$ is lower than the gain normalized dark current ($gndc$) because the dark current is not fully gain multiplied. A median effective gain normalized dark current density of 0.16 nA/cm^2 (equal to 2.56 fA for the 40 μm $\times$ 40 μm pixel) was determined from noise measurements in MWIR 128 $\times$ 128 FPAs at an APD gain of 444 while the measured gain normalized dark current density was 0.447 nA/cm^2. The $gndc_{\text{eff}}$ was about one-third of the $gndc$.

Because of the very low dark currents in the cylindrical e-APDs, subphoton *NEPhs* have been demonstrated in gated mode FPAs. Very low dark current levels at gains of 1000 have been measured on MWIR devices with an improved passivation process. Gain normalized dark current densities of less than 0.7 nA/cm^2 at a gain of 1000 were measured on these devices.

Capacitance–voltage measurements have been used to determine a high bias junction capacitance of 6 fF per diode for a cylindrical diode 18 μm in diameter. This data was analyzed to obtain a junction profile in the n^- region that showed a fairly uniform doping level in the vicinity of 4×10^{14} cm^{-3}.

Finally, gain-independent spectral response characteristics, predicted for the cylindrical architecture, have been demonstrated.

ACKNOWLEDGEMENTS

REFERENCES

1. J. Beck, C.-F. Wan, M. Kinch, J. Robinson, P. Mitra, R. Scritchfield, F. Ma, and J. Campbell, *J. Electron. Mater.* 35, 1166 (2006). doi:10.1007/s11664-006-0237-3.
2. G. Perrais, O. Gravrand, J. Baylet, G. Destefanis, and J. Rothman, *J. Electron. Mater.* 36, 963 (2007). doi:10.1007/s11664-007-0147-z.
3. M.B. Reine, J.W. Marciniec, K.K. Wong, T. Parados, J.D. Mullarkey, P.A. Lamarre, S.P. Tobin, K.A. Gustavsen, and G.M. Willliams, *J. Electron. Mater.* 36, 1059 (2007). doi:10.1007/s11664-007-0172-y.
4. J. Beck, M. Woodall, R. Scritchfield, M. Ohlson, L. Wood, P. Mitra, and J. Robinson, *J. Electron. Mater.* 37, 1334 (2008). doi:10.1007/s11664-008-0433-4.
5. M. Goodwin, J. Beck, L. Wood, C. Rau, R. Scritchfield, M. Skokan, M. Ohlson, M. Woodall, P. Mitra, and J. Robinson (2008 MSS Specialty Group Meeting on Detectors, Nellis AFB, NV, 4–7 February 2008).
6. M.A. Kinch, *Proc. SPIE* 4369, 566 (2001). doi:10.1117/12.445316.
7. G. Perrais, J. Rothman, G. Destefanis, and J.-P. Chamonal, *J. Electron. Mater.* 37, 1261 (2008). doi:10.1007/s11664-008-0459-7.
8. G. Perrais, L. Mollard, J.-P. Chamonal, G. Destefanis, and J. Rothman, *J. Electron. Mater.* (2009, to be published).
9. M. Liu, S. Wang, J.C. Campbell, J.D. Beck, C. Wan, and M.A. Kinch, *J. Appl. Phys.* 98, 074509 (2005). doi:10.1063/1.2060948.
10. M.B. Reine, J.W. Marciniec, K.K. Wong, T. Parados, J.D. Mullarkey, P.A. Lamarre, S.P. Tobin, R.W. Minich, K.A. Gustavsen, M. Compton, and G.M. Williams, *J. Electron. Mater.* 37, 1376 (2008). doi:10.1007/s11664-008-0420-9.
11. G.E. Stillman and C.M. Wolfe, *Semicond. Semimet.* 12, 291 (1981). doi:10.1016/S0080-8784(08)60150-7.