Study of the transit time limitations of the impuse response in HgCdTe avalanche photodiodes

Gwladys Perrais, Sophie Derelle, Laurent Mollard, Jean-Paul Chamonal, Gerard Destefanis, Gilbert Vincent, Sylvie Bernhardt, Johan Rothman

Study of the Transit-Time Limitations of the Impulse Response in Mid-Wave Infrared HgCdTe Avalanche Photodiodes

GWLADYS PERRAIS,[1] SOPHIE DERELLE,[2] LAURENT MOLLARD,[1] JEAN-PAUL CHAMONAL,[1] GERARD DESTEFANIS,[1] GILBERT VINCENT,[3] SYLVIE BERNHARDT,[2] and JOHAN ROTHMAN[1,4]

1.—CEA-LETI-MINATEC/DOPT, 17, rue des Martyrs, 38054 Grenoble Cedex 9, France.
2.—ONERA/DOTA/CIO, Chemin de la Hunière, 91761 Palaiseau Cedex, France.
3.—CNRS-LTM, 17, rue des Martyrs, 38054 Grenoble Cedex 9, France. 4.—e-mail: johan.rothman@cea.fr

The impulse response in frontside-illuminated mid-wave infrared HgCdTe electron avalanche photodiodes (APDs) has been measured with localized photoexcitation at varying positions in the depletion layer. Gain measure-ments have shown an exponential gain, with a maximum value of M = 5000 for the diffusion current at a reverse bias of V_b = 12 V. When the light was injected in the depletion layer, the gain was reduced as the injection approached the N+ edge of the junction. The impulse response was limited by the diode series resistance–capacitance product, RC, due to the large capacitance of the diode metallization. Hence, the fall time is given by the RC con-stant, estimated as RC = 270 ps, and the rise time is due to the charging of the diode capacitance via the transit and multiplication of carriers in the depletion layer. The latter varies between $t_{10–90}$ = 20 ps (at intermediate gains M < 500) and $t_{10–90}$ = 70 ps (at M = 3500). The corresponding RC-limited bandwidth is BW = 600 MHz, which yields a new absolute record in gain–bandwidth product of GBW = 2.1 THz. The increase in rise time at high gains indicates the existence of a limit in the transit-time-limited gain–bandwidth product, GBW = 19 THz. The impulse response was modeled using a one-dimensional deterministic model, which allowed a quantitative analysis of the data in terms of the average velocity of electrons and holes. The fitting of the data yielded a saturation of the electron and hole velocity of $v_e = 2.3 \times 10^7$ cm/s and $v_h = 1.0 \times 10^7$ cm/s at electric fields E > 1.5 kV/cm. The increase in rise time at high bias is consistent with the results of Monte Carlo simulations and can be partly explained by a reduction of the electron saturation velocity due to frequent impact ionization. Finally, the model was used to predict the bandwidth in diodes with shorter RC = 5 ps, giving BW = 16 GHz and BW = 21 GHz for x_j =4 lm and x_j =2 lm, respectively, for a gain of M = 100.

Key words: HgCdTe, APD, impulse response, bandwidth, Monte Carlo simulation, velocity, saturation

INTRODUCTION

In recent years, HgCdTe avalanche photodiodes (APDs) have been demonstrated to be a promising path to focal plane arrays (FPA) for low-flux and high-speed applications such as active and hyperspectral imaging. Several groups[1–12] have reported multiplication gains of M = 100 to 1000 at low values of reverse bias, around 10 V, associated with a quasideterministic multiplication which yields a conserved signal-to-noise ratio (SNR). These exceptional characteristics are due to an exclusive impact ionization of the electrons, which is why these devices

have been termed electron-initiated avalanche photodiodes (e-APDs).[1] The exclusive electron multiplication also implies that the response time due to the junction transit time of the carriers and the corresponding electrical bandwidth (BW) should be close to independent of gain. This behavior contrasts with that observed in APDs in which both carriers multiply, for which a constant gain–bandwidth product is observed due to the increased number of carrier junction transitions of the carriers at high gain.[13] In a recent publication[12] we presented experimental data consistent with the independence of the carrier transit time from the gain. This property allowed us to evidence a record high gain–bandwidth product, GBW = 1.1 THz, for the total diffusion-limited response time. The transit-time-limited gain–bandwidth product, estimated from the rise time of the response, was even higher: GBW = 16.3 THz. Although diffusion contributed to both values, they were the highest ever achieved in solid-state photodetectors and gave another indication of the exceptional potential of amplified photodetection in HgCdTe e-APDs.

In the present communication we report impulse response time measurements realized with test diodes designed to eliminate the diffusion component that contributed to both the rise and fall time in our previous measurements.[12] The diffusion was eliminated by localizing the photogeneration of carriers at variable positions in the junction through micrometer-sized holes at the front side of the APDs. By this variation, we could generate an impulse response which was dominated by the carrier transit time and multiplication. Hence, the impulse response is directly related to the position of the light injection in the junction. In the following section, we detail the APD device technology, the parameters of the test structures used to generate the localized photogeneration, and the experimental setup used for the impulse response measurement. Next, we report the experimental results of the photogeneration-position-dependent gain and impulse response time measurement. The impulse response measurement is then compared with a one-dimensional (1D) model, which calculates the impulse response as a function of the position of light injection, gain, the electric-field-dependent average velocity of the carriers, and the constant of the APDs. The adjustment of the calculated impulse response allowed us to estimate the electric-field dependence of the drift velocity of the carriers and to compare the latter with saturation velocities calculated using Monte Carlo simulations. Finally, we have used the estimated parameters to predict the performance in optimized transit-time-limited APDs.

DEVICE STRUCTURE AND EXPERIMENTAL SETUP

The typical device structure used for the localized photogeneration impulse response measurements is illustrated in Fig. 1. The APDs are similar to our standard N+ n–p planar diode, formed in a p-type liquid-phase epitaxy (LPE), $Hg_{1-x}Cd_xTe$ substrate with a Cd composition of $x_{Cd} = 0.3$, corresponding to a cutoff wavelength of $\lambda_c = 5.2$ μm at $T = 80$ K. In the present samples, the p-type doping was close to $N_a = 8 \times 10^{15}$ cm^{-3}. The junction is formed by generating an N+ region in a shallow hole patterned through the passivation layer. The formation of the N+ region will generate a Hg flux which results in an n-region, with a doping level given by residual impurities close to $N_d = 5 \times 10^{14}$ cm^{-3}. In the present devices, the P doping in the substrate is low and the width of the n-region is expected to be larger than that observed in our standard devices. In addition, as the p-type doping is low, the junction is expected to extend slightly on the p-side at high reverse bias.

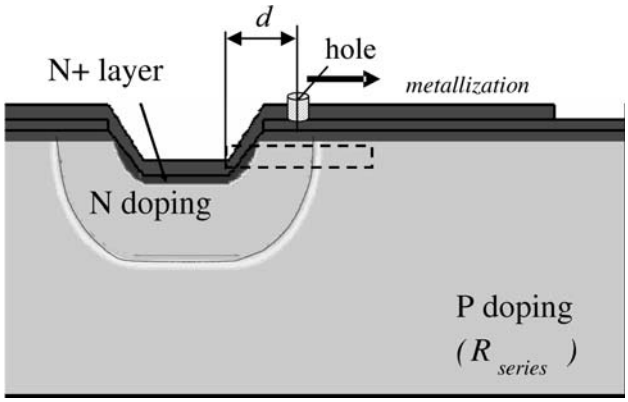

Fig. 1. Schematic illustration of the device structure.

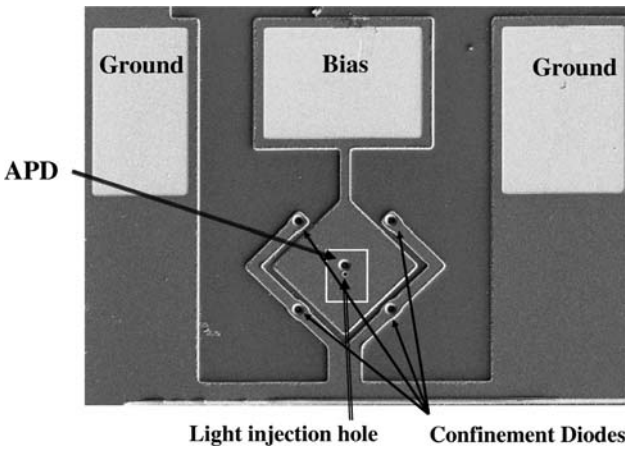

Fig. 2. Scanning electron micrograph of a diode structure.

Light is injected through a hole patterned into the diode contact metallization, as illustrated in Fig. 1. Figure 2 shows a scanning electron micrograph of a diode structure. In this particular device, the diode is surrounded by four confinement diodes, which are short-circuited to the substrate. These diodes originate from another experiment and do not have any function in the present measurements. The APD N+

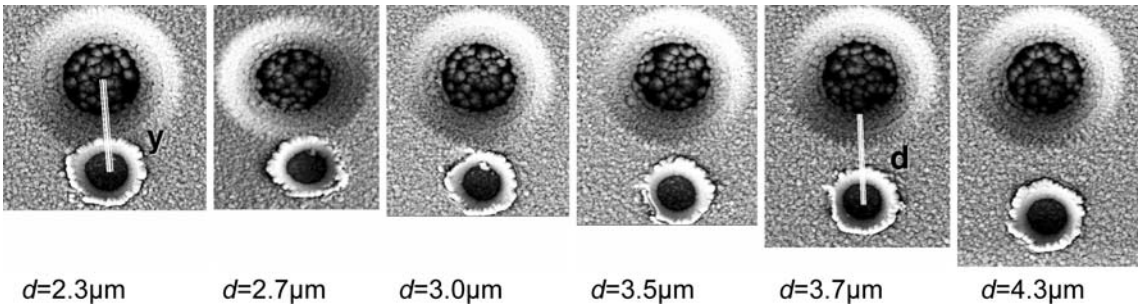

Fig. 3. Scanning electron micrographs of the different APDs with varying light injection distance, *d*.

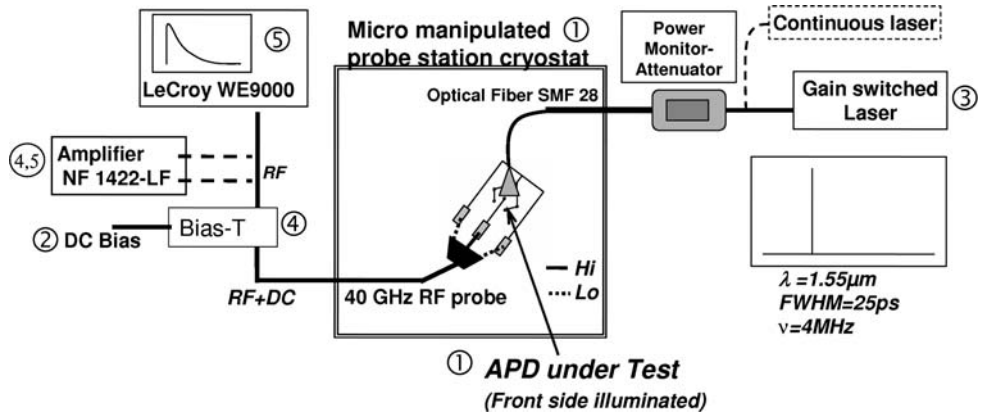

Fig. 4. Impulse response measurement setup.

formation area and the hole for light injection are labeled in the figure. The average diameter of the hole was $\phi = 1.5\ \mu$m. The inner diameter of the junction was close to $\phi_J = 4\ \mu$m. Figure 3 shows a series of close-up views of APDs with varying distance between the N+ region and the light injection hole. In the following experiments, we have used the distance d between the innermost diameter of the APD contact and the centre of the light injection hole to compare the evolution of the response time of the APDs, as illustrated in Fig. 1. The distance to the N+ junction edge might however be smaller, as the diameter of the APD contact hole is larger at the top of the structure.

The impulse response was measured using a micromanipulated probe system, equipped with two 40-GHz-bandwidth radio-frequency probes and a fiber-optic entrance. The experimental setup is illustrated in Fig. 4. The impulse response was induced by a gain-switched semiconductor laser (CALMAR optcom), capable of generating pulses with a full-width at half-maximum (FWHM) of 37 ps at a wavelength of 1.55 μm. Attenuation of the optical pulse was provided by an EigenLight Power Monitor-Attenuator M410. The APDs were biased using a bias-T and the impulse response was recorded with a LeCroy WE9000 sampling oscilloscope using a 20-GHz-BW sampling module. The RF signal from the bias-T was amplified with a 20 GHz BW, 18 dB gain New Focus amplifier model 1422. The normalized responses presented in this communication have not been deconvoluted by the laser impulse shape nor corrected for the bandwidth of the chain of detection.

GAIN MEASUREMENTS

$I(V)$ and gain curves were characterized in APDs with different distance d at $T = 80$ K. Figure 5 reports gain curves measured in APDs with

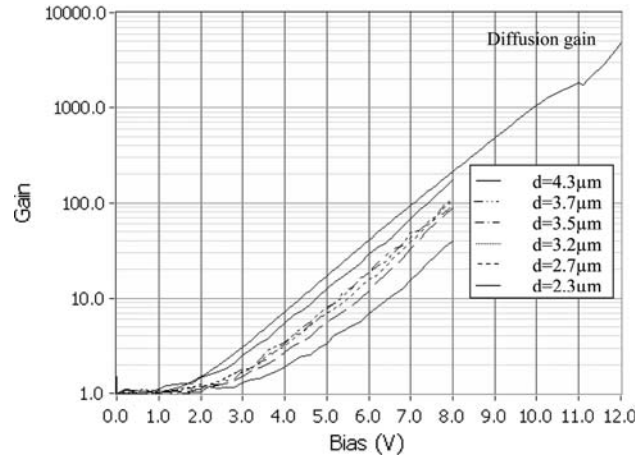

Fig. 5. Gain curves measured at $T = 80$ K, for the APDs with different position of light injection, compared with the gain of the diffusion component.

continuous-wave (CW) laser light injected through the holes positioned at $d = 2.3$ μm to 4.3 μm. The curve labeled "diffusion gain" corresponds to the gain measured with the light injected through the gap in the metallization between the APDs and the confinement diodes. This curve shows the average gain of the diffusion current generated outside the junction and was identical in all the measured APDs. At a reverse bias of 5 V, the gain is $M = 20$, which is a standard value in wide-depletion-region APDs. At high reverse biases, $V_b > 10$ V, the gain displayed instabilities which could be due to electrical charging of the passivation layer. A gain of $M = 5000$ was nevertheless measured at a reverse bias of 12 V. When the light is injected through the holes the gain is reduced. This reduction indicates that the carriers are predominantly generated inside the depletion layer and that the lateral depletion width is larger than the largest injection distance, i.e., $x_j > 4$ μm. Hence, the measured response time will be dominated by the dynamics of the carriers generated inside the junction, at least for low gains. At high gains, the carriers generated outside the junction might start to contribute to the response time. In any case, the diffusion component of the response time due to electron diffusion in the p-type material should be negligible. The gain with light injected through the holes at high bias, $V_b > 8$ V, was estimated by extrapolating the gain curves measured at low bias. A maximum gain of $M = 3500$ and $M = 2000$ was estimated at $V_b = 11.7$ V for the APDs with $d = 4.3$ μm and $d = 3.7$ μm, respectively.

IMPULSE RESPONSE MEASUREMENTS

The impulse response was measured as a function of reverse bias for each distance of light injection, d. Figures 6 and 7 report impulse responses measured as a function of the reverse bias at a distance $d = 2.7$ μm and 3.7 μm, respectively. At high biases

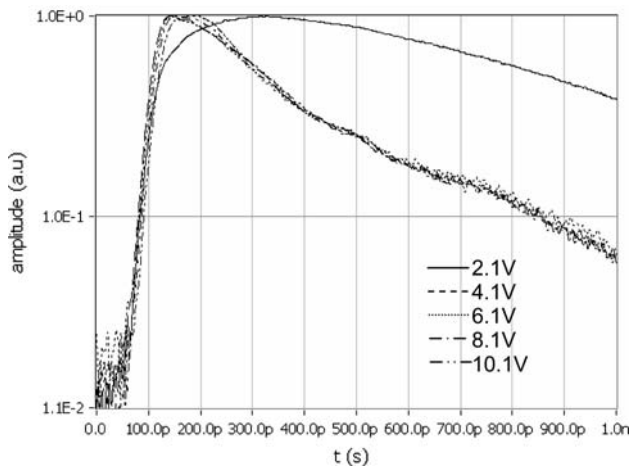

Fig. 6. Impulse responses measured at different bias for the APD with $d = 2.7$ μm.

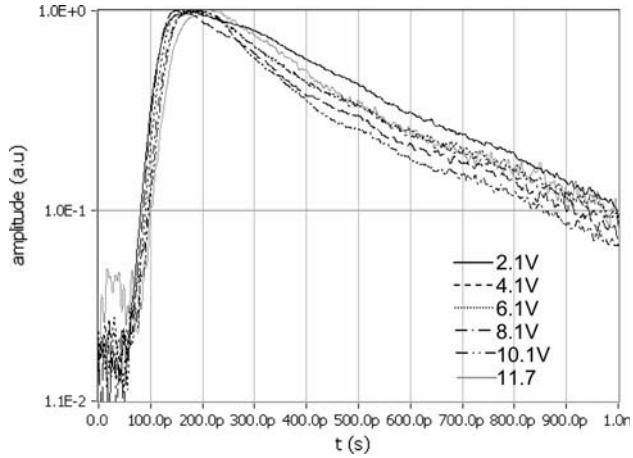

Fig. 7. Impulse responses measured at different bias for the APD with $d = 3.7$ μm.

$V_b > 4$ V, the impulse response is characterized by a fast rise time, $t_{10-90} = 50$ ps to 70 ps, and a slower exponential fall time, $t_{90-10} = 600$ ps. The latter was found to be close to independent of the bias and the localization of the photogeneration. The observation of a fall time which is independent of the localization of the injected light confirms that the diffusion contributes little to the response time. As a consequence, the RC constant limits the bandwidth of the diodes at high bias, and the rise time is determined by the depletion layer transit time of the electrons and holes.

A qualitatively different behavior is observed at low reverse bias, $V_b < 4$ V, depending on whether the injection is localized close to the N+ or P side of the junction:

- When the light is injected close to the p-side of the junction (Fig. 6), the fall time of the impulse response is close to constant up to reverse biases of $V_b = 10$ V, indicating that the impulse response is in the RC regime also for these measurements.
- When the light is injected close to the N+ side of the junction (Fig. 7), both the rise and fall times are bias dependent. As the gain is close to unity for these biases, the variation of the fall time is related to the depletion of the n-region, which reduces the junction capacitance and suppresses a hypothetical diffusion component of the holes generated in the nondepleted n-region. As the diffusion gain curves measured for the different APDs are the same, the junction profile and capacitance can be supposed to be similar. As the RC constant is independent of bias for diodes with large injection distances d, this indicates that the diode capacitance is dominated by the large diode metallization. Hence, the modulation of the junction capacitance should have a small influence on the response time and be independent on the localization of the photogeneration. For this reason, we believe that the variation of the fall time is due to the suppression of the slow

diffusion of holes generated in the n-region. As a consequence, the variation of the rise time can be explained by both the variation of carrier velocity and the onset of the diffusion collection. In the following, we will however neglect the influence of the diffusion onset in the quantitative analysis of the response. Therefore, the carrier transit time will be overestimated and drift velocities underestimated at low biases, when light is injected close to the N+ side of the depletion layer. The presence of diffusion of holes at low biases also confirms that the junction formed in the p-material is large, $x_j > 4$ μm, and indicates that the n-doping in the depletion layer is higher than $N_d > 5 \times 10^{14}$ cm^{-3}.

The RC limitation of the bandwidth is due to the large diode capacitance, induced by the large contact metallization used in the present study. The typical fall time $t_{90-10} = 600$ ps corresponds to $V = 270$ ps and yields an estimation of the total diode capacitance $C = 533$ fF, supposing $R_S = 500$ Ω, and a fall-time-limited bandwidth of BW = $1/(2\pi RC) = 600$ MHz. The highest gain used in this study was $M = 3500$, yielding a new record high gain–bandwidth product of GBW = 2.1 THz.

It should be emphasized that the RC-limited bandwidth is not the intrinsic limit of the response time of the detectors. This limit can be estimated from the rise time of the response, given by the depletion layer transit times of the carriers for a diode in the RC limit with negligible contribution from carriers collected by diffusion.

IMPULSE RESPONSE MODEL

In order to estimate the carrier transit time and response time limitations of the APDs, we have developed a one-dimensional (1D) model to calculate the average impulse response of an exclusive electron multiplication APD as a function of the gain and the position of the photogeneration, x_i, in the junction. The model is based on an initial unitary electron–hole pair generation at the position x_i from the p-side of a p-i-n diode with a depleted intrinsic layer of width x_j, as illustrated in Fig. 8. The avalanche multiplication is calculated supposing continuous smooth generation of electrons and holes through the depletion layer.

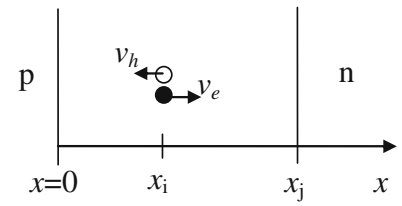

Fig. 8. Schematic illustration of the initial position of the original electron–hole pair generated at position x_i in the depletion layer.

$$M(x, x_i, x_j) = M_{x_i}(V_b) \exp\left(\frac{x}{x_j - x_i}\right), \quad (1)$$

$$\Delta q(x) = \frac{dM}{dx} \Delta x_e q_e, \quad (2)$$

where q_e is the elementary charge. The impulse response is calculated by determining the spatial distribution of deterministically generated and drifted carriers through the junction as a function of time with a time step Δt. At each time step, every electron and hole partial charge will travel a distance Δx_e and Δx_h until they reach the end of the depletion layer:

$$\Delta x_e = v_e \Delta t, \quad (3)$$

$$\Delta x_h = v_h \Delta t. \quad (4)$$

The total difference in electric potential generated across the depletion layer is calculated by summing the potential difference of each partial charge pair, supposing that the partial charges are uniformly distributed on a surface A:

$$\Delta V_i = \frac{\Delta q_i}{\varepsilon A} d_i, \quad (5)$$

$$V_q(t) = \sum_i \Delta V_i. \quad (6)$$

At each time step, the new partial charges are added at the position of the electrons in the structure, according to Eq. 2, and the separation of the previously generated carriers will increase according to Eqs. 3 and 4. The generation of the output current is obtained by calculating the decharging of the capacitance through the series resistance at each time step Δt:

$$\Delta V_{RC}(t) = \frac{V(t)}{RC} \Delta t. \quad (7)$$

This allows estimating the instantaneous output current from the following expression:

$$i(t + \Delta t) = V(t + \Delta t)/R = \big(V(t) + \Delta V_q - \Delta V_{RC}\big)/R, \quad (8)$$

where ΔV_q is the increase in the potential over the junction due to generation and displacement of charges between t and $t + \Delta t$.

Finally, the width of the laser impulse has been included by convoluting the calculated response with a Gaussian impulse with a FWHM equal to that measured for the laser impulse, FWHM = 37 ps. Figure 9a and b illustrates the impulse response of a junction having a depletion layer width of $x_j = 4$ μm, gain of $M = 100$, electron velocity of $v_e = 2 \times 10^7$ cm/s, and hole velocity of $v_h = 1 \times 10^7$ cm/s, calculated in the short-circuit limit ($RC = 0.5$ ps) and in the RC limit ($RC =$ 270 ps), respectively. The calculated impulse

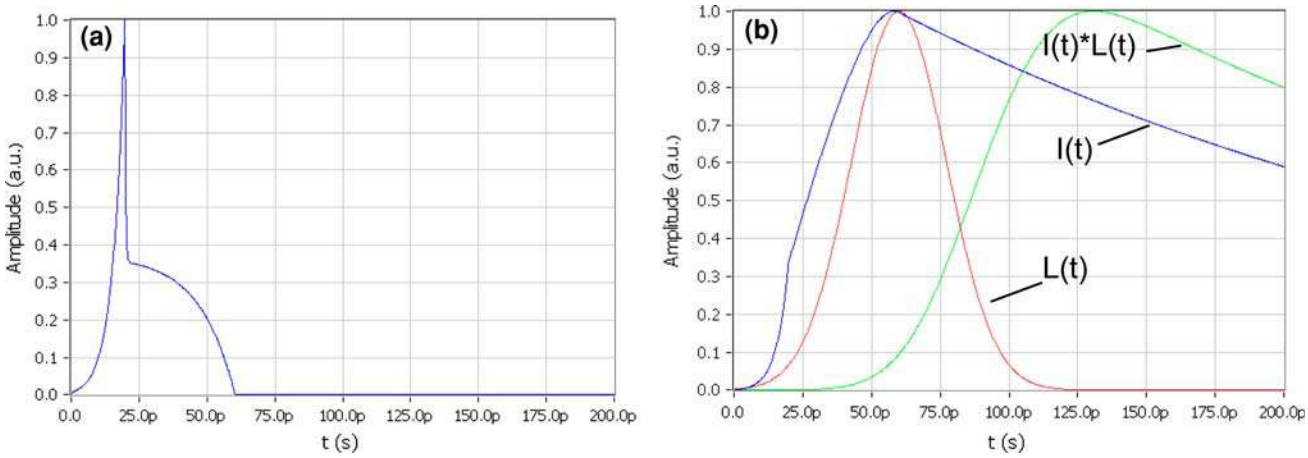

Fig. 9. APD impulse response $I(t)$ in (a) the short-circuit limit (RC = 0.5 ps) and (b) the RC limit (RC = 270 ps) for a x_j = 4 μm-wide depletion layer with $v_e = 2 \times 10^7$ cm/s and $v_h = 1 \times 10^7$ cm/s, compared with a Gaussian laser impulse $L(t)$ and the convoluted impulse response $I(t)*L(t)$.

response is consistent with other theoretical predictions.[14] In the short-circuit limit, the rise time is mainly given by the multiplication and transit of electrons, and the fall time is due to the slow evacuation of generated holes. In the RC limit, both the transit of electrons and holes build up the potential over the junction, and the rise time is a function of the transit time of both carriers. When the holes are slower, they will make the main contribution to the rise time. As for the present measurements, the fall time is given by the RC constant. The influence of the laser impulse width is also illustrated in Fig. 9b, where both the Gaussian laser impulse and the resulting convoluted impulse response are illustrated. It can be seen that the width of the laser impulse will prevent one from distinguishing between the regimes dominated by the electron and hole transits. This shows that the use of a shorter laser pulse and faster electronics should improve the estimations of the electron and hole contributions to the impulse response.

TRANSIT-TIME ESTIMATIONS

The parameters of the impulse response time model were manually adjusted to fit the measured impulse response as a function of bias and photogeneration position, d. All of the curves have been fitted with the electron and hole velocities and the zero-level offset as the only adjustable parameters. The RC constant was kept fixed at RC = 270 ps and the gain at each bias and position of injection were estimated from the CW gain measurements in Fig. 5. Figure 10a–f compares the adjusted and measured impulse response for d = 2.7 μm and for d = 3.7 μm, measured at reverse biases of V_b = 3.1 V, 4.1 V, and 10.1 V. The total junction width in the adjustments was taken to be x_j = 4.0 μm, except for the measurements at low bias, $V_b < 4$ V, for which the junction width was varied to take into account the depletion of the junction towards the N+ side of the n-layer: x_j = 2.5 μm at V_b = 2.1 V, x_j = 3.0 μm at V_b = 3.1 V, and x_j = 3.4 μm at V_b = 4.0 V. A close fit to the rise time of the impulse response was obtained at all biases. In particular, the shift of the onset of the signal at higher gains, observable in Figs. 6 and 7 is predominantly due to the increase in gain, which shifts the maximum of the impulse response even for constant velocities. The quality of the fit of the falling edge to the RC decay varies as a function of bias and the position of injection. In the case of injection close to the N+ side of the depletion layer at low reverse bias, we observe an excess of signal on the falling edge, which we believe is due to the finite size of the generation and the contribution from diffusion of holes generated on the nondepleted n-side of the junction. At high biases, $V_b > 8$ V, the maximum of the response displays a trailing maximum before the exponential decay. This behavior cannot be fitted by the model, which is why other phenomena should be considered to obtain a good fit of the data, such as unbiasing due to generated charges that reduces the gain during carrier transit, contributions from electrons generated outside the junction and collected with high gain after slow diffusion, and/or a first indication of hole multiplication in HgCdTe e-APDs.

The measured and calculated rise times are reported as a function of gain for different photogeneration positions d in Fig. 11. The corresponding electron and hole velocities are reported as a function of the electric field in Fig. 12. Good agreement between the calculated and measured rise times is observed for intermediate gains. The difference in the adjustment observed at low and high gains results in an underestimation of the calculated rise times and in an inaccurate estimation of the carrier velocities. The origin of the differences has been discussed above. It should however be emphasized that the onset of the impulse response is very sensitive to the electron velocity, which is why the

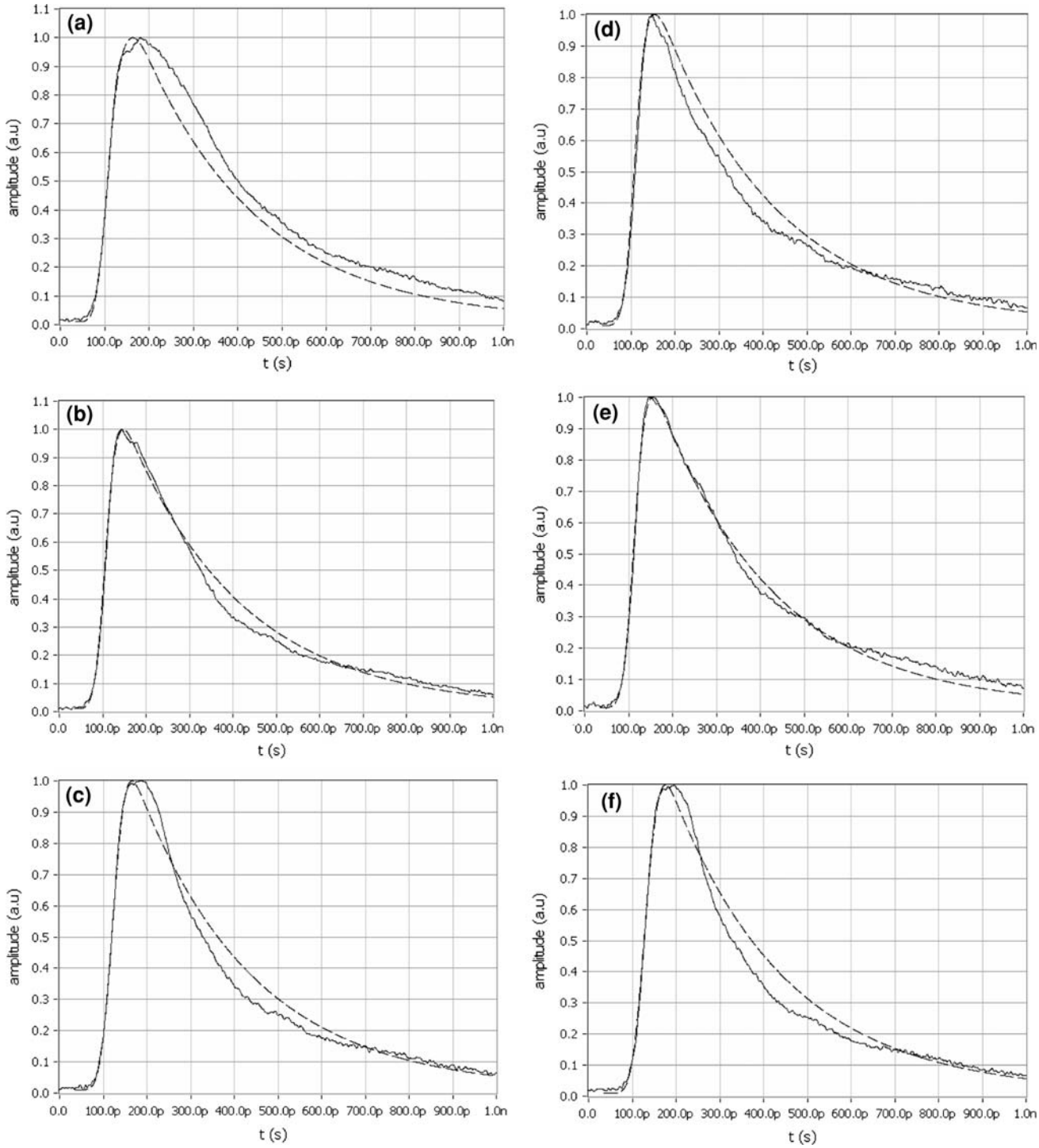

Fig. 10. Measured and calculated impulse response for APDs with d = 2.3 μm (a–c) and d = 4.7 μm (d–f), at V_b = 3.1 V (a, d), 4.1 V (b, e), and 10.1 V (c, f).

inaccuracy of this parameter should be limited. The velocity of both carriers shows a qualitatively similar behavior. At a field $E = 1.5 \times 10^4$ V/cm, the electron and hole velocities saturate at values of $v_e = 2.3 \times 10^7$ cm/s and $v_h = 1 \times 10^7$ cm/s. The hole velocity saturation value is close to the phonon scattering velocity limitation $v_h = 8 \times 10^6$ cm/s, estimated by Kinch.[15] At higher fields and gains, the increased rise times translates into a reduction of both carrier velocities.

The saturation and reduction of electron velocity has been reproduced in Monte Carlo simulations

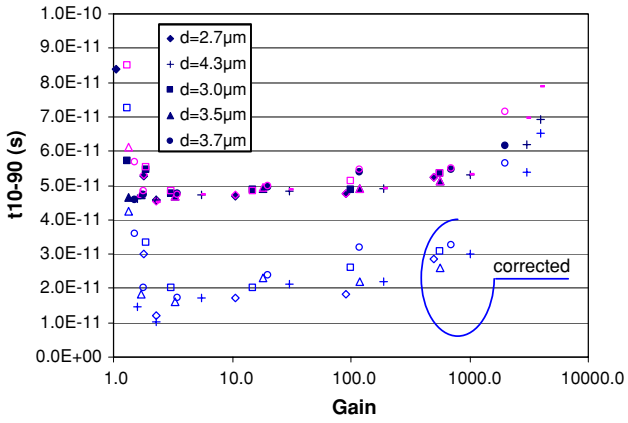

Fig. 11. Comparison between the measured (open symbols), calculated (full symbols), and laser pulse FWHM corrected measured (lower open symbols).

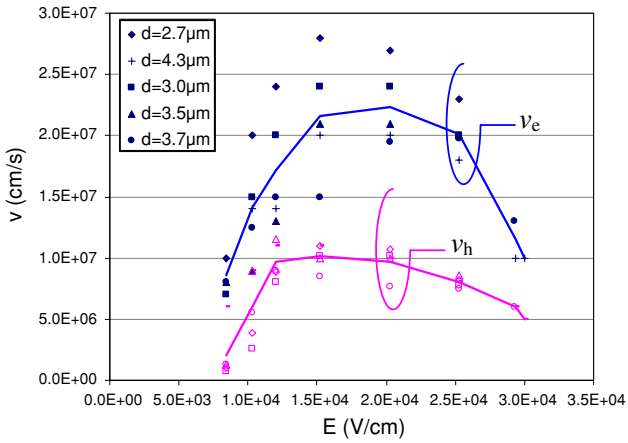

Fig. 12. Estimated electron velocity (full symbols), v_e, and hole velocity (open symbols), v_h.

at constant electric field. Figure 13 reports the electron transit velocity, calculated from the average transit time of the electrons in a junction with constant electric field. The band structure and scattering parameters were the same as those used in Refs. 16 and 17, and a more detailed study of the dependence of transit time on gain and junction width will be published elsewhere.[18] The qualitative agreement with the estimated electron velocity in Fig. 12 shows that the observed saturation and reduction of the electron velocity is consistent with the combined effect of the nonparabolic conduction band (which limits the speed of the electrons to $v_e = 2 \times 10^8$ cm/s at high energies[15]) and the scattering mechanisms included in the simulations, i.e., phonon scattering, alloy scattering, and impact ionization. At high gains, impact ionization will frequently reset the energy and speed of the electrons, which should contribute to the reduction of the average transit velocity of the electrons.

The decrease in hole velocity at high electric fields is not expected from phonon and alloy scattering in

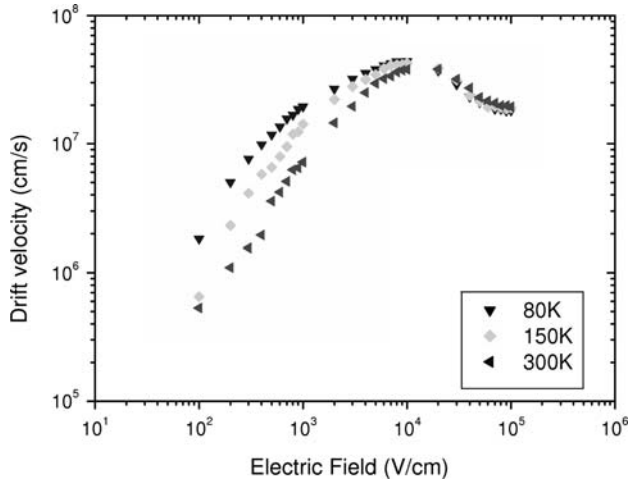

Fig. 13. Electron drift velocity calculated as a function of electric field using Monte Carlo simulations.

a parabolic band. The estimated reduction might be induced by the same phenomena that induces the trailing edge in the impulse response at high fields, i.e., unbiasing, multicarrier scattering or hole impact ionization, although the latter has not been evidenced by the gain measurements. Further Monte Carlo simulations are needed to understand the origin of the trailing edge at high electric fields.

The measured rise times corrected for the widths of the laser impulse are also reported in Fig. 11. The laser-width-corrected rise times can be used to estimate the transition-time-limited bandwidth of the APDs:

$$\mathrm{BW} = \frac{0.35}{\left(\sqrt{t_{10-90}^2 - t_{\mathrm{Laser}}^2}\right)},$$

where $t_{\mathrm{Laser}} = 44$ ps corresponds to the resulting rise time of a laser impulse with FWHM of 37 ps convoluted with a step function. The transition-time-limited BW in the present $x_j = 4$ μm-wide depletion layer is reported in Fig. 14. At gains up to $M = 1000$, the bandwidth is BW > 10 GHz, which is sufficient for application in high-speed free-space optical telecommunication systems. In particular, at $M = 200$, the laser impulse width corrected rise time was found to be close to $t_{10-90} = 20$ ps, corresponding to a transit-time-limited bandwidth of BW = 15 GHz (GBW = 3 THz). The increase in rise time at high gains shows a transit-time limit in the gain–bandwidth product of GBW = 3500 $\times$ 5.3 GHz = 19 THz. The impact of a variation in depletion layer width on the gain–bandwidth product limit will depend on the origins of the saturation of the bandwidth evidenced by the present measurements. However, at moderate gains, $M < 1000$, the bandwidth is limited by the slow transit of the holes, which is why the bandwidth should scale as the reverse of the junction width. Higher bandwidths

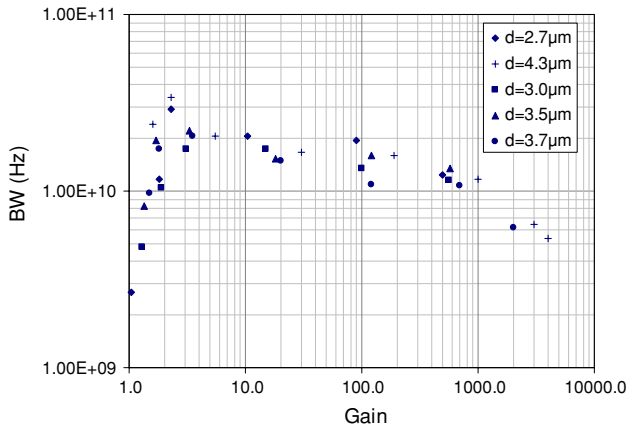

Fig. 14. Short-circuit limit BW estimated from the measured corrected lifetimes (BW estimation).

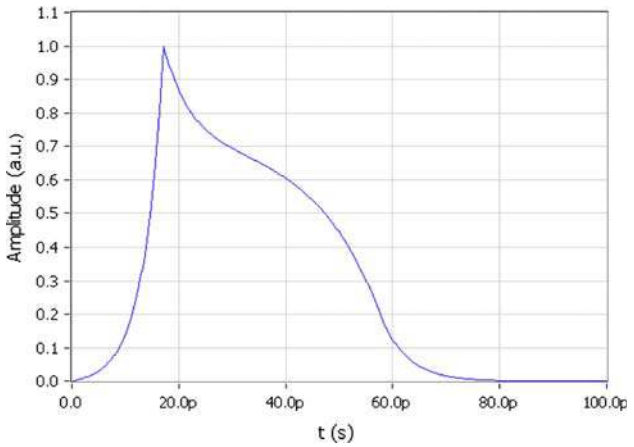

Fig. 15. Predicted impulse response in an APD with junction width $x_j = 4$ μm, $v_e = 2.3 \times 10^7$ cm/s, $v_h = 1.0 \times 10^7$ cm/s, and $RC = 5$ ps at $M = 100$.

and gain–bandwidth products should therefore be expected in narrower junctions.

To conclude, we have used the estimated electron and hole saturation velocities to predict the transit-time-limited impulse response at gain $M = 100$ in a detector with reduced RC constant of $RC = 5$ ps (corresponding to $R_s = 500$ Ω and $C = 10$ fF). The calculated impulse response, illustrated in Fig. 15, has a FWHM of 32 ps, corresponding to BW = 0.44/FWHM = 13.8 GHz. In a $x_j = 2$ μm-wide device, the impulse response starts to be limited by the RC constant, resulting in FWHM = 21 ps and BW = 21 GHz.

CONCLUSIONS

We have reported gain and impulse response measurements on HgCdTe N+ n–p e-APDs with photon injection localized at a distance d from the N+ side of the depletion layer. At all positions of light injection, the gain was lower than the total diffusion current gain, implying that the light injection was localized inside the junction for all the measured devices and that the junction width is larger than $x_j > 4$ μm. As expected, the gain was also found to decrease as the position of injection approached the N+ side of the junction. A maximum gain of $M = 5000$ was measured for the diffusion current at $V_b = 12$ V, yielding a maximum gain for the light injected in the holes of $M = 3500$ for $d = 4.3$ μm.

The bias dependence of the impulse response showed that the APDs were in the RC limit, due to the large capacitance of the diode contact metallization used to generate the localized light injection. The RC constant was estimated to be $RC = 270$ ps, corresponding to BW = 600 MHz and a record high gain–bandwidth product, GBW = 2.1 THz. However, the bandwidth of the APDs is ultimately limited by the transit time of the carriers, which determines the rise time of the impulse response in the present RC-limited measurements. The rise time corrected for the laser impulse width was found to be close to $t_{10–90} = 20$ ps for gains up to $M = 200$, corresponding to a transit-time-limited bandwidth of BW = 15 GHz (GBW = 3 THz). At higher gains, the rise time was found to increase, $t_{10–90} = 70$ ps at $M = 3500$, indicating the presence of a gain–bandwidth product limitation in the present device of about GBW = 19 THz.

The impulse response was compared with that calculated using a 1D model, taking into account the gain, the position of light injection in the junction, and the average electron and hole velocities. Manual adjustment of the carrier velocities yielded a good fit between the measured and calculated impulse response for all biases and injection positions. Hence, the fit allowed us to estimate the electron and hole velocities as a function of the electric field in the junction. The drift velocities of the electrons and holes were found to saturate at $v_e = 2.3 \times 10^7$ cm/s and $v_h = 1 \times 10^7$ cm/s at $E = 1.5 \times 10^4$ V/cm, consistent with Monte Carlo simulation of the electron velocity and the phonon diffusion limitation of the hole velocity. At high gains, the increase in transit time is, within the limitation of the model, translated into a reduction of the electron and hole saturation velocities at high fields and gains. This behavior has also been observed for the electrons in the Monte Carlo simulations, and is attributed to the high frequency of impact ionizations. As no clear evidence of impact ionization initiated by the holes has been observed, the reduction of the velocity of holes has been related to second-order effects, such as an onset of particle-to-particle interactions.

Finally, the model has been used to predict the transit-time-limited bandwidth of detectors with a shorter RC constant, $RC = 5$ ps, $v_e = 2.3 \times 10^7$ cm/s, $v_h = 1 \times 10^7$ cm/s, and $M = 100$. Bandwidths of BW = 13 GHz and 21 GHz were predicted for junctions widths of $x_j = 4$ μm and $x_j = 2$ μm, respectively.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Jean-Cristophe Peyrard, at the French General Delegation of Arms (DGA), for supporting this study.

REFERENCES

1. J.D. Beck, C.-F. Wan, M.A. Kinch, and J.E. Robinson, *Proc. SPIE* 4454, 188 (2001). doi:10.1117/12.448174.
2. J.D. Beck, C.-F. Wan, M.A. Kinch, J.E. Robinson, P. Mitra, R. Scritchfield, F. Ma, and J. Campbell, *Proc. SPIE* 5564, 44 (2004). doi:10.1117/12.565142.
3. J.D. Beck, C.-F. Wan, M.A. Kinch, J.E. Robinson, P. Mitra, R. Scritchfield, F. Ma, and J. Campbell, *J. Electron. Mater.* 35, 1166 (2006). doi:10.1007/s11664-006-0237-3.
4. J.D. Beck, M. Woodall, R. Scritchfield, M. Ohlson, L. Wood, P. Mitra, and J.E. Robinson, *J. Electron. Mater.* 37, 1334 (2008). doi:10.1007/s11664-008-0433-4.
5. I. Baker, S. Duncan, and J. Copley, *Proc. SPIE* 5406, 133 (2004). doi:10.1117/12.541484.
6. M. Vaidyanathan, A. Joshi, S. Xue, B. Hanyaloglu, M. Thomas, M. Zandian, D. Edwall, G. Williams, J. Blackwell, W. Tennant, and G. Hughes, *2004 IEEE Aerospace Conference Proceedings* (2004), p. 1776.
7. R.S. Hall, N.T. Gordon, J. Giess, J.E. Hails, A. Graham, D.C. Herbert, D.J. Hall, P. Southern, J.W. Cairns, D.J. Lees, and T. Ashley, *Proc. SPIE* 5783, 412 (2005). doi:10.1117/12.603386.
8. M.B. Reine, J.W. Marciniec, K.K. Wong, T. Parodos, J.D. Mullarkey, P.A. Lamarre, S.P. Tobin, and K.A. Gustavsen, *Proc. SPIE* 6294, 629401 (2006). doi:10.1117/12.674137.
9. M.B. Reine, J.W. Marciniec, K.K. Wong, T. Parodos, J.D. Mullarkey, P.A. Lamarre, S.P. Tobin, and K.A. Gustavsen, *J. Electron. Mater.* 36, 1059 (2007). doi:10.1007/s11664-007-0172-y.
10. G. Perrais, J. Rothman, G. Destefanis, J. Baylet, P. Castelein, J.-P. Chamonal, and P. Tribolet, *Proc. SPIE* 6935, 69350H (2006).
11. G. Perrais, O. Gravrand, J. Baylet, G.L. Destefanis, and J. Rothman, *J. Electron. Mater.* 36, 963 (2007). doi:10.1007/s11664-007-0147-z.
12. G. Perrais, J. Rothman, G. Destefanis, and J.-P. Chamonal, *J. Electron. Mater.* 37, 1261 (2008). doi:10.1007/s11664-008-0459-7.
13. R.B. Emmons, *J. Appl. Phys.* 38, 3705 (1967). doi:10.1063/1.1710199.
14. B.E.A. Saleh, M.M. Hayat, and M.C. Teich, *IEEE Trans. Electron. Dev.* 37, 1976 (1990). doi:10.1109/16.57159.
15. M.A. Kinch, J.D. Beck, C.-F. Wan, and J. Campbell, *J. Electron. Mater.* 33, 630 (2004). doi:10.1007/s11664-004-0058-1.
16. S. Derelle, S. Bernhardt, R. Haidar, J. Primot, J. Deschamps, J. Rothman, and G. Perrais, *Proc. SPIE* 7003, 70031P (2008). doi:10.1117/12.780501.
17. S. Derelle, S. Bernhardt, R. Haidar, J. Primot, J. Deschamps, and J. Rothman, *IEEE Trans. Electron. Dev.* 56, 569 (2009).
18. S. Derelle, to be published (2009).