Self-Consistent 2-D Monte Carlo Simulations of InSb APD


Abstract—Self-consistent Monte Carlo simulations are used to study the low noise and high gain potential of InSb avalanche photodiodes. It is found that for an electron-initiated avalanche, excess noise factors well below the minimum McIntyre value persist up to gain values of around 60 for a 3.2 μm avalanche region. For these very low noise values, it is found that multiplication has a very unusual voltage dependence which may be exploited for highly efficient novel low noise planar arrays operating at low voltage.

Index Terms—Avalanche photodetector, dead space, detector arrays, impact ionization, InSb, low noise, Monte Carlo (MC) simulation.

I. INTRODUCTION

InSb devices are of interest for low-power, high-frequency applications. The low bandgap with very small electron mass and very high mobility leads one to expect that very high speed response is possible for cooled structures. At room temperature, the high energy tail of the thermal carrier distribution extends above the bandgap energy where ionization processes are possible. These processes occurring within a thermal distribution are sometimes referred to as Auger generation and can lead to high leakage currents. By using minority carrier exclusion techniques and carrier extraction, it is possible to greatly reduce these leakage currents making room-temperature operation feasible [1], [2]. It is also possible to operate at low temperature or fabricate ohmic contacts from wider gap InAlSb in order to reduce the leakage from Auger generation. High-speed devices e.g., InSb FETs operating at room temperature have been successfully fabricated for large gate structures [2]–[4] and Monte Carlo (MC) simulations have indicated that very high speed performance is possible for short gate structures [5]. The unusually high electron drift velocity of InSb results from impact ionization cooling the electrons within the Γ valley and leads to the exceptional high speed potential [5]. In this paper, we explore the possibilities for avalanche photodiode (APD) devices.

There is ongoing interest in low-noise avalanche detection for military applications and a growing interest in single photon avalanche diodes (SPAD) for very low light level detection with applications in, for example, quantum cryptography and astronomy. Conventional SPAD detectors using wide gap materials such as Si or Ge operate at high voltage, which can be inconvenient. Narrow gap materials offer the possibility of low voltage operation and very low avalanche noise when operated below breakdown. Recent results for Hg0.7Cd0.3Te have shown very low excess noise, well below the minimum McIntyre value, at multiplication values up to 10^5 with no tendency to break down [6]–[8]. The extremely low noise appears to be related to the electron ionization dominating the hole ionization and a quasi-ballistic transport of electrons to an effective ionization threshold followed by rapid impact ionization and a repeat of this process throughout the diode. The related dead space phenomena are not included in the classical McIntyre formula for excess noise [9] and act to make the ionization process more deterministic [10]. Much work has been done on the wider gap materials to determine the role of dead space and we refer to Plummer et al. [11] for a recent review of this topic. It seems that the narrow gap materials HgCdTe with Cd fractions less than 0.4 are in a different regime as far as noise is concerned. They appear to exhibit a quasi-ballistic behavior which in the presence of negligible hole ionization leads to excess noise factors approaching unity at high multiplication.

In this paper, we show that this very low-noise avalanching is also a property of InSb. Two-dimensional (2-D) simulations are performed on structures which have the potential to form imaging arrays. The results indicate that high gain at low applied voltage should be possible in diodes with large depletion widths. The electric field inhomogeneity in APD arrays can be a source of noise unless special precautions are taken to detect light in regions with uniform electric field. The voltage dependence for multiplication observed in this work would eliminate this noise source and suggests the possibility of highly efficient low noise low voltage planar arrays.

The simulations use the self consistent SLURPS 2-D Monte Carlo (MC) code [12]. We have added suitable subroutines to handle impact ionization and photoionjection. A Keldysh form for the ionization scattering rate \( I \) was used in the form

\[
I = S \left( \frac{E - E_{th}}{E_{th}} \right)^2.
\]  

Here, \( E \) is the carrier energy and \( E_{th} \) is the threshold energy. The values for \( S \) are taken as \( 1 \times 10^{12} \text{s}^{-1} \) for electrons and \( 2 \times 10^{11} \text{s}^{-1} \) for holes. The threshold energies are estimated from energy and wave vector conservation and are taken as 1.08 Eg for electrons, 1.5 Eg for light holes (LH) and 2 Eg for heavy holes (HH), where Eg, the direct band gap, is taken as 0.174 eV at room temperature. Material parameters, ionization

Manuscript received April 25, 2005; revised July 22, 2005. The review of this paper was arranged by Editor L. Lunardi.

D. C. Herbert and P. A. Childs are with the Semiconductor Device Research Group, Department of Electronic, Electrical and Computer Engineering, Emerging Device Technology Research Center, The University of Birmingham, Edgbaston, Birmingham, B15 2TT U.K. (dcwherbert@supanet.com).

R. A. Abram, G. C. Crow, and M. Walmsley are with the Department of Physics, University of Durham, Durham, DH1 3LE U.K.

Digital Object Identifier 10.1109/TED.2005.856802

0018-9383/$20.00 © 2005 IEEE
coefficients and velocity field characteristics for InSb together with details of the simulation techniques have been presented previously [5] and are used unchanged in this paper. For convenience, the ionization coefficients are reproduced in Fig. 1 and the velocity field characteristics in Fig. 2. Unfortunately the ionization coefficients have not been measured directly at room temperature and we have used the work of Devreese et al. [13] who derive the Keldysh form in (1) specifically for InSb. By fitting photo quantum efficiency data at 77 K they also determine a value for the rate constant $S$ of about $3 \times 10^{13} \text{ cm}^2 \text{s}^{-1}$ [4]. The lower value of $S = 3 \times 10^{13}$ estimated from first principles [14] leads to a strong negative differential mobility caused by transfer of electrons to the $X$ valley and a much lower saturated velocity.

band gaps should not differ appreciably for these narrow gap materials where only the LH and HH bands are involved in the ionization. The simulations included $\Gamma$, $X$, and $L$ electron valleys and HH, LH, and spin orbit split off (SSO) valence bands, although in practice the $L$ valley and SSO valence band play a negligible role in the transport. All of the standard phonon scattering mechanisms were included in the simulations and an optical phonon energy of 28 meV was used.

II. MODEL

The structure used for the simulations in this paper is sketched in Fig. 3. It provides a useful tool for exploring 2-D aspects of avalanching, and may also have potential applications for low noise arrays in the narrow bandgap materials. In practical APD arrays where the pic-cell contacts on the top surface are say n-type with electron photocurrent flowing to this surface, it is necessary to electrically isolate these contacts. One approach is to etch trenches into the semiconductor which penetrate the n-type contact layers, sometimes reaching through the intrinsic avalanche region as far as the first p-type layer. Penetration of the avalanche region will expose surfaces in the high field region. These surfaces can be sources of leakage current, even after passivation. If the optical signal also enters through the substrate, then in general, some signal will be lost unless the light is channelled into the detectors, for example, by shaping the etched regions to guide the light by internal reflection. The structure considered here, by using both n- and p-type contacts at the top surface, avoids exposing surfaces to high field regions and consequently should be capable of low leakage. It also offers high efficiency as there are no dead regions where signal is wasted.
The device shown in Fig. 3 will exhibit significant field nonuniformity and with wider bandgap materials this can be an additional source of avalanche noise, as photoelectrons generated at different locations experience different fields while traversing different trajectories toward the contacts and therefore multiply to different extents. The bunching of field lines at contacts introduces one type of field variation, even if guard rings are employed, and etching may also introduce further field variations. Results presented in this paper indicate that InSb and probably midwave infrared (MWIR) HgCdTe, will not be subject to this noise source if the fields are sufficiently low to exclude hole ionization. As a consequence, the field variations in the proposed structure should not introduce additional avalanche noise when operated at sufficiently low electric fields.

In this paper, we are aiming to elucidate the avalanching properties of InSb and have greatly simplified the structure shown in Fig. 3 for this purpose. A practical implementation would have to consider details of contacts and light injection. The simulated structure consists of InSb material for the avalanche zone. Wider gap InAlSb is used for the contact regions in order to minimize leakage currents. Ion implantation could be used to obtain N and P ohmic contacts in the same layer at the top surface. The buffer P layer could be extended beyond the active area to form the substrate contact metallization and allow light to enter through the substrate. In this case, it would be necessary to design the buffer layer composition, doping and thickness to ensure that the light is absorbed in the P buffer or at the lower edge of the avalanche region. For our purpose here, we simply inject electron–hole pairs at random into the 0.05 µm thick injection layer. Typically, the p-type G contact can be set at the same bias as the substrate S contact. The G contact then provides a potential barrier separating the pic-cell n-type contacts. A typical potential profile is shown in Fig. 4 and it can be seen that photoelectrons generated in the injection region are channelled toward the D contact. The structure represents one half of a pic-cell for a one-dimensional (1-D) array, the other half is obtained by reflecting about the left hand vertical edge and would produce identical results if included in the simulation.

The structure in Fig. 3 used 192 mesh points to model the 3.0-µm width and 128 mesh points to model the 3.2-µm height of the avalanche region. The MC model generates a weak leakage current at room temperature from impact ionization in the contact doping layers and to compute the excess noise from avalanching of optically injected carriers we labeled the optical carriers and all carriers generated from a given injected carrier. This allows separation of the optical current and associated noise in the simulations. When all of the carriers associated with a particular optical absorption event have exited the device, the contribution to the multiplication M and \( M^2 \) are computed and a new electron–hole pair is injected with random location into the injection layer. In this way the mean optical current is maintained at a constant value. At the end of the simulation the excess noise factor \( F(M) \) is calculated as

\[
F(M) = \frac{\langle M^2 \rangle}{\langle M \rangle^2},
\]  

III. RESULTS

From the computed ionization coefficients in Fig. 1 it is seen that for fields below \( 6 \times 10^3 \) V cm\(^{-1} \) the ratio of electron to hole ionization coefficients (\( \alpha/\beta \)) is greater than 10\(^2 \). This implies that very low avalanche noise can be expected for fields in this range. Figs. 5 and 6 show the optical multiplication and excess multiplication noise for a range of applied bias. For a multiplication close to one, the photocurrent density is about 2 amps cm\(^{-2} \). It is remarkable that the excess noise factor for the 3.2-µm avalanche region remains well below the minimum...
McIntyre value, up to a multiplication value of 60. This suggests the possibility for very high performance detectors if leakage currents can be controlled. The extremely low noise is thought to be caused by the very high electron mobility and small threshold energy. The electrons are accelerated to an effective threshold for ionization very rapidly in a quasi-ballistic manner after which ionization occurs. In our MC model, the initiating electron for an ionization event loses kinetic energy equal to one threshold energy and a new electron–hole pair is introduced. For simplicity we have set the initial kinetic energy of the injected hole to zero and set the energy of the initiating and generated electron to be equal. This process is repeated and leads to a definite number of ionization events per initiating particle with only small fluctuations in this number at low bias. As the bias increases the hole ionization gradually increases and the excess noise levels start to rise.

Fig. 5 shows the multiplication plotted on a log scale. The curve labeled “3.2 µm no hole” was obtained by switching off the hole ionization and yields an approximate straight line, demonstrating the expected exponential dependence on voltage for β/α = 0. The “3.2 µm” curve deviates from the straight line for voltages exceeding about 2 V or fields exceeding about 6 kV/cm. The “1.6 µm” curve was obtained by reducing the avalanche region from 3.2 to 1.6 µm. At low applied voltage the curves approximately superimpose showing that multiplication depends on voltage rather than field. This confirms our interpretation of the low noise in which an electron accelerates quasi-ballistically to an effective threshold energy and then rapidly loses its kinetic energy by ionization. In this way the number of ionization events for the initiating particle is given approximately by \( V / E' \), where \( V \) is the applied voltage between \( S \) and \( D \), and \( E' \) is the mean energy supplied by the field to generate a new electron hole pair by ionization. From energy conservation \( E' \) can be written as the sum \( E' = E_g + \Delta + E_l \), where \( E_g \) is the energy gap, is the minimum energy required to create a new electron–hole pair, \( \Delta \) is the kinetic energy of the created electron–hole pair and \( E_l \) is the energy dissipated to the lattice while heating the initiating electron to the required energy for impact ionization. The standard expressions for polar optic phonon scattering show that the rate of emission of energy to the lattice (phonon emission rate minus phonon absorption rate) varies from about \( 3.9 \times 10^{12} \text{ s}^{-1} \) at low energy to about \( 2.6 \times 10^{12} \text{ s}^{-1} \) at the threshold for electron ionization. If we take the saturated drift velocity to be \( 5 \times 10^7 \text{ cm s}^{-1} \) then we can expect around \( 5 L \) phonon energies to dissipate while traversing an avalanche region of width \( L \mu \text{m} \). The distance required for an electron to ionize will be reduced as the multiplication \( (M) \) increases. If \( M = 2^n \), then the number of phonons per ionization event will be of the order \( 5 L / n \). For the first ionization event the electron will also experience velocity overshoot [17] so that the number of phonon emissions will be further reduced in this case.

For this quasi-ballistic model, the multiplication becomes \( M = 2^{V/E'_{l}} \), and from the slopes in Fig. 5, we estimate \( E'_{l} \approx 2.55 E_g \). The small number of phonon collisions per ionization event and the small angle scattering obtained from polar optic phonons ensures that the distance moved “down stream” for each ionization event is roughly constant. For a material like Si, the randomizing inter-valley scattering and the larger threshold energy leads to a much larger number of phonon collisions per ionization event with strong lateral diffusion. This causes much greater variation in the “down stream” distance per ionization event and the multiplication then depends more on field than voltage.

To gain further insight into the value of \( E_l \) we performed simulations for the structure in Fig. 3 with the phonon scattering set to zero. The results are shown in Fig. 7. It is interesting to note that the multiplication curves with and without phonon scattering are quite similar with the “ballistic” curve being roughly obtained from the “non-ballistic” curve by a small displacement toward lower voltage. This displacement is expected to be close to the value of \( E_l \) in the quasi-ballistic model and appears to be reducing as the curves converge for high multiplication values. This is consistent with \( E_l \) decreasing with multiplication. For \( M = 3^2 = 9 \), the displacement is 0.098 V which corresponds to 3.5 phonon energies. In the quasi-ballistic model we
demonstrates plateau formation which is a sign of Si p-i-n which emphasizes that the quasi-ballistic effects are strong and depletion regions are forming. Reference [7] does not give details for simulation parameters and refers to a different material, but we surmise that those simulations may have used a harder threshold with a higher value for S in (1).

It is clear that at low fields the multiplication varies exponentially with applied voltage. To achieve high multiplication while maintaining low electric field and ultra low excess noise, it is necessary to use large depletion widths so that relatively high voltage with low electric field can be obtained. In Fig. 6 it is seen that a depletion width of 3.2 μm gives a multiplication of 60 before the noise approaches the minimum McIntyre value while for a depletion of 1.6 μm a multiplication of 29 is achieved. To obtain very large depletion widths requires very low doping levels within the avalanche zone and in practice the background doping will impose constraints on achievable performance. In the simulations reported here we have used a background doping of $10^{14}$ cm$^{-3}$ and at room temperature this is lower than the intrinsic carrier density of about $10^{15}$ cm$^{-3}$. Under reverse bias, however, the intrinsic carriers are extracted from the avalanche region and the ionized impurity charge determines the achievable depletion width. In practice, in high-quality material with no defect assisted tunnelling, impact ionization is expected to occur before Zener tunnelling sets in and the tunnelling current has not been considered in this paper.

Fig. 4 shows the potential profile for the structure where $V_{\text{f}} = 0$, $V_{\text{i}} = V_{\text{g}} = -1.9$ V. It can be seen that setting the two p-type contacts to the same bias gives a simple way of splitting the photocurrent between pic-cells and could form the basis for a novel detector array technology. The vertical right hand edge in Fig. 3 represents the boundary between pic-cells and the field normal to this boundary is zero. As a consequence the drift current connecting adjacent cells will vanish and due to the quasi-ballistic motion the diffusion current is also expected to be small leading to negligible cross talk currents. The structure however, shows large variations in field and for a material like Si where the quasi-ballistic model fails, this field variation can be a source of noise as electron-hole pairs injected at different locations experience different fields and different multiplication values. To confirm this, further simulations were performed using Si but setting the hole ionization to zero. Results for the array geometry in Fig. 3 were compared with a p-i-n structure and the results are shown in Fig. 8 together with the corresponding McIntyre result. The potential profile for the array had a very similar shape to that in Fig. 4, though much higher voltages are used. As expected the array shows high noise levels compared with the p-i-n due to the field inhomogeneity. It is also interesting to observe that the p-i-n diode gives lower noise than predicted by the McIntyre theory. It is thought that this is a consequence of dead space which is neglected in the McIntyre formula. It should be noted that corrections from dead space are quite small for the 3.2-μm Si p-i-n which emphasizes that the quasi-ballistic effects observed in InSb have a much larger effect than dead space alone.

expect about 16 phonon emissions while traversing a distance of 3.2 μm and the displacement agrees well with the estimate $E_l \approx 16/5$. The displacement is greatest at low multiplication but decreases as the multiplication drops below 2. This decrease is thought to be caused by velocity overshoot.

The excess noise from fluctuations in phonon emission is expected to be small due to the low number of emissions. If we assume that the central limit theorem [18] applies then the fluctuation in phonon emission is roughly $\sqrt{(5LM/n)}$, where $M = 2^n$ is the multiplication and L is the length of the avalanche region. For $M = 64$ this is about 13 phonons which would give $M \approx 64 \pm 0.8$. This is consistent with the sub McIntyre noise observed in Fig. 6. It is difficult to give analytical estimates for the excess noise arising from fluctuations in $\Delta$ but the MC simulations for no hole ionization with and without phonon scattering would indicate that this is small at high multiplication. For $M \approx 43$ the excess noise factors for the array were calculated as 1.05 with phonon scattering and 1.04 without phonon scattering suggesting that both mechanisms give very small contributions to noise with the fluctuations in $\Delta$ being larger at high multiplication.

It is also worth noting that the resonance behavior in the excess noise seen in MC simulations for CdHgTe [7] was not observed in our simulations. For a hard threshold where the electrons accelerate to a precise energy and then ionize instantly, one might expect to see plateaus in the multiplication versus voltage curve corresponding to different values of n in the expression $M = 2^n$ and peaks in the excess noise close to voltages corresponding to the midpoints in the plateaus. Reference [7] gave some evidence for this in CdHgTe at low multiplication and referred to the phenomenon as a resonance effect. To check whether the absence of resonance in our simulation was due to 2-D effects we simulated a simple p-i-n diode. The results in Fig. 7 show that the p-i-n and array results are very close so that field nonuniformity cannot explain the result. The effect of threshold hardness was then considered by increasing the value of $S$ in (1) by a factor of 200 to $2 \times 10^{14}$ s$^{-1}$. The results in Fig. 7 were obtained by stepping the voltage in units of 0.1 V and setting the hole ionization and phonon scattering to zero. The formation of plateaus is evident though there is smearing of the effect at the lowest multiplications where velocity overshoot effects are strong and depletion regions are forming.

Fig. 7. Optical multiplication for 3.2-μm InSb diodes at room temperature. A p-i-n with no hole ionization and an array with no hole ionization and no optical phonon scattering are compared with the 3.2-μm array result with no hole ionization. The latter is also shown in Fig. 5. The curve with the rate constant $S$ increased to $2 \times 10^{14}$ demonstrates plateau formation which is a consequence of the increased threshold hardness.
For InSb and similar small bandgap materials where multiplication depends on voltage more than field, the injected electrons experience a constant voltage drop between the $S$ and $D$ contacts and noise arising from field inhomogeneity is not present. Extremely low noise levels with the excess noise factor approaching unity as shown in Fig. 6 are then possible.

**IV. CONCLUSION**

Self-consistent MC simulations have been performed for a simple 2-D model of an InSb APD array. The results show that the ultralow excess noise, well below the minimum McIntyre value, which has been reported in the literature for CdHgTe detectors, should also be observable for InSb devices. This low noise is available up to high multiplication levels if large depletion widths can be achieved for the avalanche zone. It is believed that the extremely low noise is caused by the impact ionization occurring almost entirely within the $\Gamma$ valley where the phonon scattering is dominated by small angle polar optic processes. In addition to dead space effects it requires a quasi-ballistic transport to an effective threshold energy followed by rapid ionization. As a consequence the results reported here for arrays are also expected to apply to CdHgTe at MWIR and possibly to other narrow gap materials. In these cases, the simple structure indicated in Fig. 3 may lead to very high performance imaging arrays. The results in this paper all refer to room temperature but similar effects are also obtained at low temperature.

**ACKNOWLEDGMENT**

D. C. Herbert would like to thank N. T. Gordon and R. S. Hall at QinetiQ, Malvern, U.K. and J. P. R. David and his group at Sheffield University, Sheffield, U.K. for valuable discussions.

**REFERENCES**


P. A. Childs received the B.Sc. degree in physics and the Ph.D. degree for experimental and theoretical work on light emission and hot-carrier effects in MOS and bipolar transistors from the University of Liverpool, Liverpool, U.K. in 1984.

He subsequently joined the staff at Plessey Research Ltd., Caswell, Northamptonshire, U.K., where he worked in the area of silicon device processing and simulation before moving to the University of Surrey, Guildford, U.K. where, as an Academic Member of Staff, he reestablished his research on hot-carrier effects in semiconductor devices. In 1988, he joined the academic staff at the University of Birmingham, Birmingham, U.K., where he currently undertakes research on semiconductor devices and nanoelectronics. Initial work at Birmingham involved fundamental studies of hot-carrier transport and electron–electron interactions in MOS transistors. He has also undertaken experimental work on self-assembly of single electron transistors in conjunction with the Nanoscale Physics Research Laboratory, Birmingham. Current research interests include hot-carrier effects in silicon, III-V and II-VI compounds, nanoscale electronics, and theoretical and experimental studies on carbon nanotubes.

Richard A. Abram received the Ph.D. degree at the University of Manchester, Manchester, U.K. in 1971. He is a Professor of Physics, Department of Physics, University of Durham, Durham, U.K. He worked as a Theoretical Researcher for nearly seven years at the central laboratories of the Plessey Company, Caswell, Northamptonshire, U.K. In 1978, he moved to Durham University, to join the Department of Applied Physics and Electronics and transferred in 1991 to the Department of Physics, where he is currently Head of the department. He is also leader of the Condensed Matter Theory Research Group, which has interests covering a number of fundamental and applied topics, including the physics of semiconductors materials and devices, molecular solids, liquid crystals, semiconductor quantum structures, and photonic microstructures.

Dr. Abram became a Fellow of the Institute of Physics in 1989 and is both a Chartered Engineer and Chartered Physicist.

G. C. Crow, photograph and biography not available at the time of publication.

M. Walmsley, photograph and biography not available at the time of publication.