Quantum-dot based avalanche photodiodes for mid-infrared sensing

Majeed M. Hayat\textsuperscript{a}*, Oh-Hyun Kwon\textsuperscript{b}, J. Shao\textsuperscript{c}, and Sanjay Krishna\textsuperscript{c}

\textsuperscript{a}Electrical & Computer Engineering Department
University of New Mexico, Albuquerque, NM, 87131
\textsuperscript{b}Air Force Research Laboratory, Ohio 45433
\textsuperscript{c}Center for High Technology Materials and Electrical & Computer Engineering Department
University of New Mexico, Albuquerque, NM, 87106

ABSTRACT

A mid-infrared sensor is proposed in which an intersubband quantum-dot (QD) detector is integrated with an avalanche photodiode (APD) through a tunnel barrier. In the proposed three-terminal device, the applied biases of the QD and the APD are controlled separately; this feature permits the control of the QD’s responsivity and dark current independently of the operational gain of the APD. It is shown theoretically that the proposed device can achieve a higher signal-to-noise ratio (SNR) over the QD detector without the APD component. Indeed, prior studies have revealed that although a heterostructure barrier lowers both the dark current and the photocurrent of the QD detector, the barrier has a greater impact on the dark current. Thus the dark-current-limited SNR is enhanced in the presence of the barrier. However, due to the reduced photocurrent, the SNR may not achieve its potential in the presence of Johnson noise, which may become dominant, for example, at low integration times or when detecting ultra-weak signals. In the proposed device, the APD component provides the necessary photocurrent gain required to elevate the SNR to the dark-current limit. This improvement, however, comes at a slight penalty in the SNR, due to the excess noise introduced by the APD. In this paper, guidelines for the SNR improvement are discussed in terms of the QD’s operational bias voltage and the required APD gain. The higher SNR could be used to obtain a higher sensitivity at the same temperature, or to achieve a comparable performance at higher operating temperatures.

Keywords: Quantum dots, avalanche photodiodes, infrared, sensors, mid-wave infrared, avalanche gain, SNR, dark current, potential barrier, DWELL

I. Introduction

Many of today’s sensing applications rely on high performance sensors and detectors in the mid-infrared (MIR) spectral regime. Such applications include laser-aided ranging (LADAR), remote sensing of toxic chemical agents, chemical spectroscopy, and vegetation and geological monitoring\textsuperscript{1}. The important wavelength regimes for operation of terrestrial sensors is in the two transmission

\footnote{E-mail: hayat@ece.unm.edu; phone: +1 505 277 0297; fax: +1 505 277 1439}
windows of the atmosphere in the mid-wave infrared (MWIR) regime (3-5 µm) and the long-wave infrared (LWIR) regime (8-14 µm). Space-based sensors, such as those mounted on satellites and used for thermal imaging, usually operate in the very long-wave infrared (VLWIR) regime (λ >14 µm) since it is easy to determine the local temperature of a blackbody by viewing its VLWIR emission.

Intersubband quantum-dot (QD) infrared detectors are currently viewed as a promising and emerging technology for LWIR detection. This is because they offer many advantages: (1) they are based on a mature GaAs technology, (2) they are sensitive to normal incidence radiation, (3) they can exhibit a large quantum confined stark effect, which can be exploited for bias-controlled spectral tuning2, (4) and they have lower dark currents than their quantum-well counterparts. In fact, it had been known for a long time that the discrete atom-like density of states present in QDs would have the potential to lead to a breakthrough in LWIR device technology. Nonetheless, the technology to fabricate these nanoscale objects did not exist in the laboratory. For room temperature operation of these devices the largest dimension of the dots has to approach 15-20 nm. Conventional lithography techniques are not considered as effective tools for fabricating QDs since the process-induced defects render the QDs optically dead. Fortunately, a highly promising approach that has emerged over the last decade facilitated the fabrication of QDs with demonstrated success. The technique is based on the creation of self-assembled quantum-dots by means of using highly strained epitaxial growth1. Under certain growth conditions, the epitaxial film thermodynamically prefers to minimize its free energy by assembling into 3-dimensional islands instead of forming a 2-dimensional layer. This process is akin to the process of formation of oil droplets on a wet surface: The oil molecules prefer the formation of droplets rather than uniformly covering the wet surface since it is “thermodynamically easier.” Such self-assembled QDs are typically pyramid to lens shaped with lateral dimensions of 15-20nm and vertical dimension of 7-8nm (see Fig. 1) and can be formed from important semiconductor systems such as InGaAs/GaAs, SiGe/Si and InGaAs/InP.

Figure 2 shows the calculated energy levels in an In_{0.4}Ga_{0.6}As/GaAs pyramidal shaped quantum dot with dimensions of 18.1nm base and 4.5nm height, calculated by Jiang et al.3 The energy spacing between these levels has been experimentally verified by photoluminescence, capacitance voltage spectroscopy and absorption measurements3. The interesting point is that since the intersubband energy spacing in the QDs lies in the MIR range (i.e., 5-20 µm), QDs can be used to fabricate MWIR and LWIR sources and detectors. Indeed, self-assembled QDs have been used in making optoelectronic devices such as interband edge- and surface-emitting lasers4-6 and intersubband long wavelength detectors7-12. Moreover, it is expected, by virtue of the singular density of states of ideal QDs, that characteristics of QD devices will surpass those of quantum well (QW) devices13-14.
In a QD infrared photodetector (QDIP), an incident photon promotes an electron from the ground state of the dot to an excited state. This electron is then swept out as photocurrent due to an applied external bias. If the excited state is a bound state, then this transition is known as a bound-to-bound transition. If the excited state lies in the continuum close to the edge of the conduction band of the barrier, this transition is known as a bound-to-continuum transition. Professor Krishna has already fabricated mid-IR sources and detectors based on intersubband transitions in QDs. Some researchers have proposed that the intersubband relaxation time in QDs increases with increase in temperature. The slow capture time in QDs is attributed to the presence of the “phonon-bottleneck,” which refers to the suppression of phonon scattering in QDs since the intersubband energy spacing (\(\Delta E \approx 60-100\, \text{meV}\)) is larger than the longitudinal optical (LO) phonon energy (\(E_{\text{LO}} = 36\, \text{meV}\)). Thus, the electrons in the excited state of the dot cannot relax to the ground state by scattering with an atom in the crystal lattice. This is expected to lead to a long relaxation time, which in turn ensures that the photogenerated carriers stay at the excited state for a longer time and contribute more efficiently to the photocurrent.

Since their first demonstration in 1997, QDIPs have demonstrated (a) normal incidence operation in the MWIR, LWIR and VLWIR (>14 \(\mu\)m); (b) high temperature operation (up to 150K at ~4 \(\mu\)m) with a cut-off at 8.2 \(\mu\)m at \(T=78\,\text{K}\); and (d) large gain (~25) and conversion efficiency (57\%) The fact that QD detectors offer normal incidence operation (which alleviates the need for gratings in sensors) on a mature GaAs technology makes them highly attractive material system for large area FPAs. However, one of the biggest problems plaguing QDIPs is their low quantum efficiency, which leads to a lower detectivity, responsivity and limits their operating temperature to about 70-80K. If the operating temperature of QD detectors can be increased to 150K-200K, fairly inexpensive Peltier coolers can be used. Presently, there are no photonic detectors that can operate in this temperature range. A 100K increase in the operating temperature would lead to a dramatic decrease in the cost and complexity of infrared imaging systems and would represent a major technological breakthrough. Moreover, as the operating wavelength is increased, the performance of a sensor deteriorates and, hence, the operating temperature must be decreased. It is to be noted that the only photonic detectors that operate in the VLWIR regime do require cooling to 4.2K. This cooling requirement places an enormous constraint on an infrared sensors and imaging systems, thereby leading to an increase in the complexity and cost.

We have recently shown that dark current in a QDIP can be reduced by using the novel concept of dots-in-well (DWELL) design in which InAs dots are placed in a thin InGaAs well, which, in turn, is surrounded by a GaAs barrier. This lowers the ground state of the dot with respect to the GaAs band.
edge, thereby decreasing thermionic emission and reducing dark current. Some results for QDIPs fabricated and characterized by Professor Krishna’s group at the University of New Mexico are shown in Fig. 3.

![Graph showing experimental verification of the expected improvement in the detectivity (SNR) as a result of the insertion of the AlGaAs potential barrier. Blue (●): Without the barrier; Red (▲): with the barrier. From Ref. [27].](image)

A highly promising approach for reducing the cooling requirement is based on coupling of an avalanche layer with an intersubband QD detector through a tunnel barrier. The tunnel barrier reduces the dark current while the avalanche section supplies the photocurrent with internal gain. It has been demonstrated that a barrier (either homojunction or heterojunction) can simultaneously lower the dark current and the photocurrent of the QD detector; however, the impact on dark current is larger. Thus, the dark-current-limited signal-to-noise ratio (SNR) is larger in the presence of the barrier.

However, due to the reduced photocurrent, the SNR may not achieve its dark-current limit in the presence of Johnson noise. If we were to somehow make up for the reduced photocurrent, possibly by introducing avalanche-multiplication-induced optoelectronic gain, then the overall SNR would approach its dark-current limit. The benefit of the barrier would therefore be available in the presence of Johnson noise, which would be significant in sensing ultraweak signals. The aforementioned device is termed the quantum-dot avalanche diode (QDAP), which is schematically shown in Fig. 4 (a); the avalanche layer provides the necessary photocurrent gain required to restore the photocurrent and elevate the SNR to almost its dark-current limit. There is a slight penalty, however, due to the excess noise factor associated with the avalanche multiplication gain mechanism. Of course, the mode of operation of the QDAP (i.e., linear mode versus Geiger mode) would be dictated by the specific application the sensor is used for. In addition to linear-mode operation, the QDAP can be operated in Geiger mode to provide a single-photon counting in the LWIR regime. Despite the recent demonstration of single-photon detectors for the LWIR regime (e.g., by Astafiev et al., Lipatoc et al., and by Toshiba Industries), the state of such devices is presently in its infancy as the detection efficiency is very low (<12%). Combined with a potential barrier, bandgap engineering of the APD
section of the QDAP, and with the use of an DWELL structures, the QDAP would offer a promising approach for photon counting for the LWIR regime.

**Fig. 4.** (a) Simplified schematic of the QDAP structure and (b) its principle of operation. The photogenerated carriers in the n-i-n section tunnel through the thin p-doped layer and undergo impact ionization in the i layer of the p-i-n section.

**II. Detailed Description of the Quantum-Dot Avalanche Diode (QDAP)**

Consider the three-terminal QDAP device schematically shown in Fig. 4(a). In this device, an n-i-n intersubband DWELL detector is grown on top of a p-i-n avalanche-photodiode (APD) structure in a single-step epitaxy. For proper operation, the n-i-n and the p-i-n segments of the QDAP are both reverse biased, viz., $V_{\text{top}} < V_{\text{mid}} < V_{\text{bot}}$. Photons are absorbed in the n-i-n structure and the photogenerated electrons are swept by the applied electric field in the direction of the APD region. The thickness of the p-region must be kept low to allow the photoexcited electrons to tunnel into the intrinsic region of the reverse biased p-i-n region. The bias applied to the p-i-n structure must be
sufficiently large to guarantee punch-through, enabling impact ionization of carriers and thus providing avalanche gain, which results in an amplified photocurrent in the external circuit. An attractive feature of this device is that only electrons are injected into the depletion layer (no hole injection) and there are no primary holes triggering the avalanche multiplication process. This feature is similar to the low-noise separate-absorption-multiplication (SAM) heterostructure APDs, which are specifically designed to achieve this property. Consequently, the excess noise produced in the APD section of the QDAP is expected to be significantly low. The applied biases $V_{\text{top}}$, $V_{\text{mid}}$, and $V_{\text{bot}}$ of the QD and the avalanche layer, respectively, are controlled separately. This feature permits the control of the QD’s responsivity and dark current independently of the operating avalanche gain. This is an important feature as it allows us to freely select the operational avalanche gain to maximize the photocurrent’s SNR for each given level of Johnson noise.

Additionally, if a heterojunction barrier is employed, the thickness of the avalanche layer may also be optimized to minimize the excess noise factor. Figure 5 shows three possibilities for the APD section of the QDAP: (a) a pin GaAs homojunction, the type considered in the QDAP shown in Fig. (4); (b) a pin AlGaAs/GaAs heterojunction; and (c) a pin AlGaAs/AlGaAs/GaAs two-layer multiplication region heterostructure. The yellow layer(s) represent the multiplication region (electrons are injected from the top). These structures are expected to significantly reduce the excess noise factor due to the so-called initial-energy effect. Consequently, we expect achieving higher SNRs by incorporating and optimizing these heterojunction APDs.

Theoretical analysis of the QDAP was undertaken using a commercially available simulation software (MEDICI). The QDs were simulated by N-traps with activation energy of 300 meV (in accordance with the position of ground state of the QD). In particular, the bandstructure and transport properties of the device were investigated for different doping levels and thickness of the p-type layer. Figure 6 shows the calculated bandstructure of the QDAP along with the electronic quasi-Fermi levels for $V_{\text{top}}=-1V$, $V_{\text{mid}}=0V$ and $V_{\text{bot}}=5V$ for $p = 5 \times 10^{18} \text{cm}^{-3}$. Note that the calculated bandstructure is in general agreement with the heuristic schematic shown in Fig. 4(b). We have also calculated the barrier height, measured from the conduction band edge of GaAs, as a function of the doping of the p-type region. It was seen that as the doping in the p-type region is increased the maximum of the barrier shifts from the intrinsic region of the p-i-n section to the p-type region. This is because the voltage drop across a highly doped layer becomes very small. Our simulations indicate that for a p-type doping in the range of $5 \times 10^{18}, 10 \times 10^{18} \text{cm}^{-3}$, the potential barrier is approximately 400 meV.

![Fig. 5. The APD structures to be integrated with the QDAPs: (a) a pin GaAs homojunction; (b) a pin AlGaAs/GaAs heterojunction; and (c) a pin AlGaAs/AlGaAs/GaAs two-layer multiplication region heterostructure. The yellow layer(s) represent the multiplication region. Electrons are injected from the top.](image-url)
The transport properties of the structure were studied by simulating the IV characteristics of the device. Figure 7 shows the current collected at the bottom contact (which is indicative of the tunneling probability through the p-type barrier) as a function of the voltage, parameterized by the thickness of the p-type layer, doped at $1 \times 10^{18}$ cm$^{-3}$. As expected, the tunneling probability increases with a decrease in the thickness of the barrier layer. Our modeling suggests that there is a good amount of tolerance to the specified doping and thickness value, making the design fairly robust to fluctuations during growth.

The performance of the device can be further improved by replacing the p-type GaAs barrier with a p-type AlGaAs. This has two advantages. First, the barrier height can be controlled by varying the composition of the AlGaAs layer while keeping the doping fixed. Secondly, the AlGaAs barrier can serve to improve both the excess-noise and breakdown characteristics of the APD due to the fact that its ionization threshold energy is higher than that of GaAs. This leads to the occurrence of the so-called initial-energy effect, whereby the carriers that are tunnelled into the intrinsic GaAs layer enter the layer with a substantial energy, which they acquire as they traverse the high bandgap AlGaAs layer without ionizing (since the electron’s dead space in the AlGaAs layer is comparable to its width). It has been argued recently that when the primary electrons that initiate the avalanche multiplication in the GaAs layer enter the layer hot, the excess noise factor is reduced and the breakdown probability is enhanced as a function of the excess bias. For example, for 0.6 Al composition, a possible scenario for enhanced performance is when the thickness of the Al$_{0.6}$Ga$_{0.4}$As and GaAs layers are approximately 30 nm and 100 nm, respectively. In this structure a gain of 10, which is accompanied by an excess noise factor of approximately 3, can be achieved at a reverse bias of 8.2 V across the multiplication region (corresponding to an electric field of 630 kV/cm). A simple calculation shows that an avalanche gain of 10 would improve the noise-equivalent power (NEP) by a factor of 2 (assuming that the sum of the readout, $1/f$ and pre-amplification noise variances is ten-times that of the dark-current noise). Additionally, for a Geiger-mode APD design (used in photon counting), using a thin (<200 nm) intrinsic region of the diode can lead to a reduced value of the operating voltage. We believe that the value of the voltage applied to the bottom contact could be reduced to ~3-5 V, thereby making the device compatible with standard readout circuits.

**Fig. 6.** Bandstructure of the device for $V_{top} = -0.5$ V, $V_{mid} = 0$ V and $V_{bot} = 5$ V. The dotted line represents the quasi-Fermi level in the device.
III. Rationale for signal-to-noise ratio (SNR) improvement in QDAPs

A. Dark current reduction due to the barrier layer

Previously we argued that the dark current can be reduced as a result of the barrier layer, and that this is accompanied by a reduction in the responsivity. However, the reduction in the dark currents is more significant. It is, therefore, anticipated that the presence of a potential barrier layer results in an improvement in the SNR at least in the dark-current-noise limit (i.e., when the Johnson noise is negligible). We introduce a formalism that can support the rationale for the SNR improvement as a result of the potential barrier layer. First, let us define the dark-current reduction, $1/r$, to be $i_d/i_{do}$ where $i_d$ is the dark current with the barrier layer and $i_{do}$ is the dark current without the barrier layer and the photocurrent reduction, $p$, to be $i_p/i_{po}$ where $i_p$ is the photocurrent with the barrier layer and $i_{po}$ is the photocurrent without the barrier layer. Now, the SNR without the barrier layer is

$$SNR_o = \frac{i_{po}^2}{2(i_{po} + i_{do})B},$$

where $B$ is the detector’s electrical bandwidth and the SNR with the barrier layer is

$$SNR = \frac{i_p^2}{2(i_p + i_d)B}.$$
Then, if we assume for simplicity that $i_{po} = i_{do}$ as in the case of weak signal detection, we carry out a simple calculation and conclude that the condition to satisfy $SNR > SNR_o$ leads to the following inequality:

$$2rp^2 - rp - 1 > 0$$

Simplifying the above inequality results in

$$\rho > \frac{1}{4}(1 + \sqrt{1 + 8/r})$$

(1)

Subsequently, we applied the experimental data, reported in Ref. 2, to verify that the SNR is indeed improved due to the potential barrier according to the derived equation (1). Rotella et al.\textsuperscript{27} fabricated the InAs/InGaAs DWELL detectors with and without the AlGaAs current blocking layer. The photocurrent reduction, $\rho$, is obtained from the ratio of responsivity curves (with and without the current blocking layer) as a function of applied bias at 78 K. The dark current reduction, $1/r$, is obtained from the ratio of dark current curves (with and without the current blocking layer) as a function of applied bias at 80 K. (We assumed that the effect of the small temperature difference is negligible.) Figure 8 illustrates how the predicted photocurrent reduction and dark current reduction are related in general and how those reductions affect the SNR. In the figure, the thick solid curve sets the boundary that $SNR$ is either improved (above the curve) or degraded (below the curve). Note that SNRs with and without the barrier layer are identical along the curve. The thin curves represent the

![Fig. 8. Relationship between dark-current reduction and photocurrent reduction due to the potential barrier layer and its effect on the SNR improvement. Actual data (circles) are mapped to the level of improvement for each bias.](image-url)
level of the SNR improvement (dashed: 10%, dashed-dotted: 20%, and dotted: 30%). From this figure we claim that the SNR can be improved as a result of the potential barrier layer as long as the reduction in the photocurrent and dark current satisfies the requirement [c.f. Eq. (1)]. The experimental data from Rotella et al.\textsuperscript{27} are mapped on the figure (circles connected with dotted curve). It shows that when the reverse bias is applied the SNR with the barrier layer is greater than the SNR without the barrier layer. For example, when a reverse bias of 0.5 V is applied, the SNR improvement is more than 30%.

**B. Johnson-noise-limited SNR improvement due to avalanche gain**

In applications for which Johnson noise is not negligible, the presence of the barrier, despite its beneficial effect in the dark-current limit, may result in a lower SNR, simply due to the combined effect of Johnson noise and the reduction in the responsivity. In order to preserve the SNR improvement provided by the presence of the barrier, the role of Johnson noise must be suppressed. The approach to solve this problem is to employ the multiplication of an APD in the QD detector.

The general expression for the SNR is given by

\[
\text{SNR} = \frac{I_p^2}{\sigma_S^2 + \sigma_T^2},
\]

where

\[
I_p^2 = (M R P_{in})^2
\]

\[
\sigma_S^2 = 2 q M^2 F_A (R P_{in} + I_d) \Delta f
\]

\[
\sigma_T^2 = 4 (k_B T / R_L) F_n \Delta f
\]

In our calculations we assumed the following parameters: \(R_L = 1 \text{k}\Omega\), \(T = 78 \text{ K}\), \(\Delta f = 50 \text{ MHz}\), \(F_n = 2\). The dark current data are obtained from Ref. 27. The dark current of a device with the current blocking layer (the barrier) at the reverse bias of 1 V and at the temperature of 80 K is 1.81 \(\mu\text{A}\) and without the current blocking layer at the same conditions is 23.0 \(\mu\text{A}\). We used the mean gain, \(M\), and the excess noise factor, \(F_A\), of a 200-nm GaAs APD data generated according to our dead-space-multiplication theory model (DSMT)\textsuperscript{33}. We also assumed an incident power for which the incident photocurrent is approximately a tenth of the dark current. Using the responsivity of the device which is 0.09 A/W in 1198, \(P_{in}\) is set to be 0.20 \(\mu\text{W}\) to make \(I_p = R P_{in} \approx 0.1 I_d\). The corresponding SNR is calculated using (2). Figure 9 shows the calculated SNR as a function of the APD mean gain for the different operating temperature. The SNR of a device without the barrier layer is higher when \(M = 1\). Once the APD starts the multiplication (i.e., \(M > 2\)), the SNR of the device with the barrier layer is higher than the one without the barrier layer. The figure also shows the optimum mean gain of the APD that maximizes the SNR.

In general, the SNR of a QD detector is higher when it is operated in low temperature. A QDAP with the optimum mean gain can be operated in substantially high temperature, yielding the equivalent SNR as a QD detector being operated in low temperature without the gain. For example, the SNR at \(T = 20 \text{ K}\) without the APD gain is approximately 25 dB in the given assumptions, including the same responsivity and \(I_p\). On the other hand, the peak SNR of the QDAP, when \(M = 4\), at \(T = 80 \text{ K}\) (Fig. 9) is over 25 dB when the device is equipped with the barrier layer. In the other words, the APD gain can increase the operating temperature by 60 K, yielding the equivalent SNR. However, if the device is not
incorporated with the barrier layer, the peak SNR with the optimum APD gain at $T = 80$ K would not reach the SNR without the APD gain at $T = 70$ K, indicating little temperature benefit in this case.

IV. Conclusion

A quantum-dot avalanche photodiode (QDAP), which combines an intersubband quantum-dot detector with an avalanche photodiode, can offer an SNR enhancement in the presence of Johnson noise. A key feature of the QDAP is that it is a three-terminal device for which the biases of the quantum-dot and avalanche-photodiode sections can be controlled independently. This feature (1) facilitates optimizing the avalanche gain to maximize the SNR when operated in a linear mode while maximizing the quantum-dot’s responsivity and (2) permits operating the device in Geiger-mode (for single-photon counting applications) by increasing the reverse-bias of the APD section beyond avalanche breakdown. We believe that the value of the voltage applied to the bottom contact could be reduced to $\sim 3-5$ V in thinner multiplication layers, thereby making the device compatible with standard readout circuits.

![Graph](image-url)

**Fig. 9.** Calculated signal-to-noise ratio (SNR) of QDAPs, as a function of the APD mean gain, $M$, with (solid) and without (dashed) a barrier layer for the set of system and operational parameters described in the text.
Acknowledgement
This work is supported by AFRL Contract Number F29601-01-C-0156 and NSF Grants IIS-0434102, ECS-0401154 and ECS-0428756.

References


