An Improved Non-Local History-Dependent Model for Gain and Noise in Avalanche Photodiodes Based on Energy Balance Equation


Abstract—We present a non-local history-dependent model for impact ionization gain and noise in avalanche photodiodes (APDs) especially suited for staircase APDs. The model uses a simple energy balance equation to define effective electric fields valid also in the presence of band discontinuities which are then used to express the ionization coefficients. The model parameters have been calibrated against literature data for gain and noise in GaAs and AlGaAsAs (x = 0.2, 0.6, 0.8) p-i-n diodes. Application to experimental data for gain and noise in heterojunction and staircase SAM-APDs is reported to demonstrate the ability of the model in describing complex APD structures. It is found that, in spite of conduction band discontinuities being much larger than valence band ones, hole impact ionization contributes a significant degradation of the noise metrics in GaAs/AlGaAs staircase APDs. These non-trivial insights demonstrate the usefulness of the model to steer device design and optimization.

Index Terms—Avalanche Photodiodes, Impact Ionization, Simulation

I. INTRODUCTION

Avalanche photodiodes (APDs) find applications as receivers in optical communication systems [1] as well as for detection of X-rays [2]-[6]. Internal multiplication is obtained via impact ionization. If the associated multiplication noise is low, significant advantages can be harvested compared to the use of external amplification. The trade-off at the basis of APD design is thus between gain and noise, and both these figures of merit should be modelled accurately for device optimization.

Such an endeavour would be well tackled by the solution of the Boltzmann Transport Equation (BTE), with accurate description of the band structure and of the phonon and impact ionization scattering rates up to and above the bandgap energy, which is very computationally demanding [7], [8], [9]. For this reason, simple local models for impact ionization have been used to investigate the trade-off between gain and noise [10] and between gain and bandwidth [11]. These analyses pointed out that having similar impact ionization coefficients for electrons and holes is detrimental in terms of noise and gain-bandwidth product. Unfortunately, this is the case for most III-V compounds, that on the other hand offer the advantages of large absorption coefficient, low leakage and high carrier velocity. Complex heterojunction structures (in the following referred to as staircase APDs) have been proposed [12], where the position of electron impact ionization process is better defined by the use of conduction band discontinuities making electron ionization stronger than hole ionization. Non-local history-dependent models [13], [14], where the ionization coefficients at a given position do not depend on the local electric field, but on the whole electric field profile experienced by the carrier, have been developed to overcome the limitations of local models. Consequently, carriers at a given position have different ionization probability depending on the location they were generated either by optical absorption or by ionization processes.

The non-local history-dependent models proposed so far in the literature differ in the way they relate the ionization coefficients to the electric field profile. For instance, the dead space model [13] assumes that the ionization coefficient is null within a distance from the generation position. This distance is the one required to the carriers to attain an energy equal to a suitable ionization threshold. The history-dependent (HD) model in [14], instead, transforms the electric field profile into an effective field by convolution with a Gaussian function in space that mimics a soft ionization threshold. These models have been proven to accurately describe gain and excess noise in homogeneous p-i-n APDs with different thickness [15]–[18].
Their application to staircase APDs, however, is not straightforward. For instance, in the dead space model, it is not obvious to decide if the dead space should be accounted for at every step or only at the first one encountered by the ionizing carrier [19]. The HD model in [14], instead, is not applicable to staircase APDs because it relates the parameters of the Gaussian function to the local electric field and becomes singular in the presence of band discontinuities.

Inspired by [20], [21], in this paper we present an improved non-local energy balance history-dependent model (EBHDM) applicable to interpret electrical measures on single photon staircase APDs.

The paper proceeds as follows. The proposed model is presented in Sec.II. Calibration against experiments in [15]-[17] and [22], [23] for p-i-n APDs is reported in Sec.III. The model is then applied to template staircase APDs in Sec.IV to demonstrate that it is consistent with the theory in [12]. Comparison with experimental literature data for heterojunction [4], [5] and staircase APDs [2] is provided in Sec.V. Conclusions are finally drawn in Sec.VI.

II. MODELLING FRAMEWORK

Following [13], [14], the model computes the multiplication gain for an e-h pair generated at the position $x$ as

$$ M(x) = \frac{N_e(x) + N_h(x)}{2} $$

and the corresponding excess noise as

$$ F(x) = \frac{(m(x)^2)}{[M(x)]^2} = \frac{(\langle m_e(x)^2 \rangle + \langle m_h(x)^2 \rangle + 2\langle n_e(x) \rangle \langle n_h(x) \rangle)}{4 \langle M(x) \rangle^2} $$

where $n_e(x)$ and $n_h(x)$ are the stochastic variables describing the number of pairs created respectively by an electron and a hole generated (optically or by impact ionization) at position $x$ and $N_e(x)$ and $N_h(x)$ are their ensemble averages. As in [13], [14], the ensemble averages and r.m.s. values of these variables are found by solving Eqs.3-6 at the bottom of this page, where we consider a depletion region extending from $x = 0$ (electron injection) to $x = W$ (hole injection).

Differently from [14], we do not solve the system of integral equations by iteration: after discretization on a spatial mesh, we obtain an algebraic system that is directly solved by matrix inversion (see Appendix A), guaranteeing numerical stability.

The probabilities for an electron and a hole to travel from position $x$ to $x'$ without suffering ionization events are:

$$ P_{se}(x|x') = \exp\left(-\int_x^{x'} \alpha(x|x'')dx''\right) $$

$$ P_{sh}(x|x') = \exp\left(-\int_x^{x'} \beta(x|x'')dx''\right) $$

They depend on the ionization probabilities per unit length $\alpha$ and $\beta$. The latter are related to suitable effective fields as [14]:

$$ \alpha(E_{eff,e}(x|x')) = A_e \cdot \exp\left(-\left(\frac{E_{eff,e}(x|x')}{{E_{ee} \gamma_e}}\right)^\gamma_e\right) $$

$$ \beta(E_{eff,h}(x|x')) = A_h \cdot \exp\left(-\left(\frac{E_{eff,h}(x|x')}{{E_{eh} \gamma_h}}\right)^\gamma_h\right) $$

where $A_e$, $A_h$, $E_{ee}$, $E_{eh}$, $\gamma_e$ and $\gamma_h$ are the adjustable model parameters to be calibrated against experimental data. The effective fields are obtained by convolving the quasi-electric-field (gradient of the bands, that is, including the band steps), obtained from TCAD simulations (which solve the drift–diffusion equations without considering impact ionization) under the assumption that the potential profile is not significantly modified by the generated charge, with an exponential function:

$$ E_{eff,e}(x|x') = \frac{1}{\lambda_e} \int_x^{x'} \frac{dE}{dx} \exp\left(x'' - \frac{x'}{\lambda_e}\right) dx'' $$

$$ E_{eff,h}(x|x') = \frac{1}{\lambda_h} \int_x^{x'} \frac{dE}{dx} \exp\left(x' - \frac{x''}{\lambda_h}\right) dx'' $$

This is different from what is done in [14], where a Gaussian function with $\lambda_e,h = V_{th} \gamma_e/h(x)$ is used ($V_{th}^e,h$ being the ionization threshold and $E(x)$ the electric field). Here, the parameters $\lambda_e$ for electrons and $\lambda_h$ for holes are constant (i.e. they neither depend on position nor on the bias), they represent a sort of mean free path and the ionization probability tends to that of the local model after travelling a large distance compared to $\lambda$ from the generation point. Consider for example a region of almost uniform electric field: regardless of the initial conditions at the generation point, the carrier distribution a few mean free paths from the generation point tends to a unique function of energy well defined by the electric field value. The use of exponential function in Eqs.11-12 is also consistent with the definition of effective field as derived from the energy balance equation [20], [21]. For this reason, in the following the new model will be then referred to as...
Energy Balance History-Dependent Model (EBHDM). In this paper we focus on APDs working in the linear regime below breakdown: Eqs.3-6 hold in this regime. However, Eqs.7-12 are more general and could be used to determine the breakdown probability as explained in Sec. V.I.C of [14].

III. MODEL VALIDATION AND CALIBRATION

As in [13], [14] the adjustable model parameters are eight; namely the mean free paths $\lambda_e$ and $\lambda_h$ for electrons and holes, respectively, and $A_{ee}, A_{eh}, E_{ce}, E_{ch}, \gamma_e, \gamma_h$ which relate the effective fields with the ionization coefficients (see Eqs.9-10). The model calibration firstly requires a reliable estimate of the electric field profiles. To this end, we have reproduced with TCAD [25] the same doping concentrations and geometrical structure for the p-i-n diodes reported in [15]-[17] and [22], [23] and then applied the EBHDM in post-processing, taking as input the electric field profiles from drift-diffusion TCAD simulations. The good agreement at high gains validates the post-processing approach, indicating that the generated charge has little impact on the electric field profile.

Fig.1 compares the model results with the measured gain as a function of bias and with the measured excess noise as a function of gain. In both cases, good agreement between simulations and experiments is found over a wide range of intrinsic layer thicknesses with a unique set of material dependent model parameters summarized in Tab.I.

<table>
<thead>
<tr>
<th>Material</th>
<th>$A_{ee}$</th>
<th>$E_{ce}$</th>
<th>$\gamma_e$</th>
<th>$\lambda_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>7.30</td>
<td>1.75</td>
<td>1.00</td>
<td>18.00</td>
</tr>
<tr>
<td>Al$<em>{0.2}$Ga$</em>{0.8}$As</td>
<td>88.30</td>
<td>5.94</td>
<td>1.00</td>
<td>17.00</td>
</tr>
<tr>
<td>Al$<em>{0.6}$Ga$</em>{0.4}$As</td>
<td>88.80</td>
<td>4.13</td>
<td>1.00</td>
<td>18.00</td>
</tr>
<tr>
<td>Al$<em>{0.8}$Ga$</em>{0.2}$As</td>
<td>14.60</td>
<td>3.53</td>
<td>1.00</td>
<td>20.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Material</th>
<th>$A_{eh}$</th>
<th>$E_{ch}$</th>
<th>$\gamma_h$</th>
<th>$\lambda_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs</td>
<td>10.73</td>
<td>1.42</td>
<td>1.00</td>
<td>22.00</td>
</tr>
<tr>
<td>Al$<em>{0.2}$Ga$</em>{0.8}$As</td>
<td>10.81</td>
<td>2.99</td>
<td>1.00</td>
<td>21.50</td>
</tr>
<tr>
<td>Al$<em>{0.6}$Ga$</em>{0.4}$As</td>
<td>260.00</td>
<td>4.99</td>
<td>1.00</td>
<td>16.00</td>
</tr>
<tr>
<td>Al$<em>{0.8}$Ga$</em>{0.2}$As</td>
<td>0.14</td>
<td>0.67</td>
<td>5.12</td>
<td>35.00</td>
</tr>
</tbody>
</table>

TABLE I. MODEL PARAMETERS FOR GAAS, AL$_{0.2}$GA$_{0.8}$AS, AL$_{0.6}$GA$_{0.4}$AS AND AL$_{0.8}$GA$_{0.2}$AS USED THROUGHOUT THE PAPER.

Fig.1 shows that in GaAs and Al$_{0.2}$Ga$_{0.8}$As the excess noise factor increases proportionally to the gain (figures on the right) and that it is smaller for thinner intrinsic layers. As pointed out in [15], this behavior is in contrast with the predictions of a local model [24], but it is consistent with the measurements. In fact, when the extension of the intrinsic region becomes comparable to $\lambda_e$ and $\lambda_h$, the ionization probability of the secondary carriers decreases, making the ionization process somewhat more deterministic. The $F$ versus $M$ curves for Al$_{0.3}$Ga$_{0.7}$As put in evidence a more complex non-monotonic trend as a function of $W$: in fact, when $W$ decreases from 1024 nm to 312 nm, $F$ increases for given $M$, owing to the fact that at high electric field (as in the short diodes) $\beta$ approaches $\alpha$ [10]. However, as $W$ shortens, non-local effects come into play and reduce $F$ as in GaAs and Al$_{0.2}$Ga$_{0.8}$As.

It is worth mentioning that the good agreement between simulations and experiments from literature for the excess noise versus gain curve in Fig.1 is a distinctive result of history-dependent models. In fact, a non-local history-independent model (that is a model where $\alpha$ and $\beta$ are still expressed as a function of an effective field but at a given position all carriers have the same ionization probability regardless of the position they were generated [20], [21]) predicts that the excess noise versus gain curve only slightly depends on the thickness of the intrinsic layer. This can be easily confirmed setting $\alpha(x|x') = \alpha(0|x')$ and $\beta(x|x') = \beta(W|x')$ in our model, where $0$ and $W$ are the left and right end side boundaries of the simulation domain. The results (Fig.2) are easily explained by the observation that in this case the secondary carriers have the same ionization probability as the primary ones and this renders the ionization process more random compared with
history-dependent models. The results of the non-local but history-independent model are thus close to the ones of a local model in terms of excess noise versus gain, while predictions for the gain versus voltage curves are different [20].

![Excess Noise Factor (F) Diagram](image)

**Fig. 2.** Excess noise as a function of the gain for the 100-nm- and the 1600-nm-thick GaAs p-i-n diodes described in [15]. The EBHDM described in Sec.II (dashed lines) is compared with the results obtained using a non-local model that is history-independent (solid lines).

**IV. APPLICATION TO STAIRCASE STRUCTURES**

A sanity check of the proposed model and of its implementation is offered by a simple staircase APD where the model predictions should be consistent with the theory [12]:

$$F(0) = 1 + \frac{\delta [1 - (2 - \delta)^{-N}]}{(2 - \delta)}$$  \hspace{1cm} (13)

where $N$ is the number of steps of the staircase and $\delta$ is how much the gain per stage deviates from 2, i.e. $M(0) = (2 - \delta)^N$. Here $x = 0$ is located just before the first conduction band step along the direction of the electron flow.

All the calculations reported in this section have been performed on ideal staircase structures. The amplitude of the conduction band steps has been arbitrarily chosen and can be larger than the maximum achievable at GaAs/AIGaAs interfaces (0.55 eV [25]) or at the AlInAsSb/InAsSb interface (0.6 eV [26]). In reality, structures like the ones presented in this section would be possible only with a huge difference of electronic affinity among the materials that compose the ladder.

![Graph](image)

**Fig. 3.** Gain for a single conduction band step of amplitude $\Delta E_C$ obtained with the EBHDM. GaAs impact ionization parameters are used at the right side of the step, while the $\lambda_s$ of GaAs is used at both sides. The electric field is null and hole impact ionization is set to zero.

We now compute the excess noise factor versus gain for the single step structure used in Fig.3 and for five of such steps in sequence. Results are shown in Fig.4a, where the gain per stage reported on the x-axis has been obtained by varying $\Delta E_C$: the EBHDM is in perfect agreement with Eq.13.

We now apply a constant electric field to the conduction band discontinuities (set to 0.5 eV), while maintaining the hole ionization off. Fig.4b shows that now the gain per stage can be larger than 2, since impact ionization occurs also in the regions between the steps thanks to the energy given by the field that adds to the one provided by the steps. The excess noise factor starts deviating from Eq.13 and tends to saturate, which means that a higher gain partly comes with a higher noise. A large electric field is needed to obtain significant multiplication per stage. In fact, the step of 0.5 eV is much lower than the gap of GaAs. Experimental data in [26] show that steps of 0.6 eV can produce a multiplication close to 2 if impact ionization takes place in a low bandgap material (InAsSb, $E_g = 0.25$ eV).

**V. APPLICATION TO HETEROJUNCTION APDs**

In this section we apply the EBHDM model to two realistic structures: the *Separate Absorption and Multiplication* (SAM) GaAs/Al$_{0.8}$Ga$_{0.2}$As APD described in [4], [5] (Fig.5a) and the staircase SAM-APD presented in [2] (Fig.5b).

These devices have been experimentally characterized in [2], [4], [5] using X-ray sources. For the sake of a fair comparison with experiments, mixed injection has been taken into account in our model. Following [15], gain and excess noise factor in the case of mixed injection are computed as:

$$M_{mixed} = \int_0^W g_0(x)M(x)dx / \int_0^W g_0(x)dx$$ \hspace{1cm} (14)

$$F_{mixed} = \frac{\int_0^W g_0(x)M^2(x)F(x)dx}{M_{mixed}^2 \int_0^W g_0(x) dx}$$ \hspace{1cm} (15)
where the injection profile is given by the Beer-Lambert’s law:

\[
g_i(x) = \exp \left( - \int_0^x \frac{dx'}{L(x')} \right)
\]

(16)

A. Heterojunction SAM-APD in [4]

The nominal structure of the device reported in [4], [5] (Fig.5a) is reproduced with TCAD [25] and the comparison between the simulated and the experimental [4] C-V curves is shown in Fig.6. The device consists of a 400-nm i-GaAs absorption region, separated from the 100-nm i-Al₈₀Ga₂₀As multiplication region by a 130-nm p-doped GaAs/Al₀.₈Ga₀.₂As layer. A 300-nm n-doped Al₀.₅Ga₀.₅As/GaAs layer is placed between the multiplication region and the n⁺ GaAs substrate, to confine most of the applied bias voltage inside the multiplication region. The agreement between measurements and simulations is good, meaning that the TCAD provides a reliable estimate of the electric field profile. Note that the Thermionic Emission Current model for transport at the heterointerfaces has been used. This may be incorrect in indirect-to-direct band gap heterojunctions, as in this case [28]. However, this is expected to have a negligible impact on the device electrostatics. It may however result in a different value of the effective \( \Delta E_C \) to use in our post-processing model, but this can be assessed only via accurate Monte-Carlo simulations.

The EBHDM is then applied using as inputs the \( E_C \) and \( E_V \) profiles obtained from TCAD to compute \( M(x) \) and \( F(x) \). Model parameters are the ones listed in Tab.I for GaAs and Al₀.₈Ga₀.₂As. Finally, the application of Eqs.14 and 15 yields the gain (Fig.7a) and the excess noise factor (Fig.7b). The breakdown voltage is slightly underestimated by the model. We also see that assuming pure electron or mixed injection has only a small impact on the gain. On the other hand, see Fig.7b, mixed injection results in a slightly higher excess noise. Noise from the model is hardly comparable with the experiments in [5]. In such reference only the FWHM value is reported. Assuming that the only noise contribution comes from the multiplication noise of the APD, one extracts a \( F \approx 7 \) almost constant over a range of gains up to 5 (no data available for higher gains). This suggests that experiments (at such low gains) are mostly dominated by the noise of the read-out.

Fig. 6. Comparison between simulated (dashed line) and experimental (circles) C-V characteristics of the SAM-APD in [4] (Fig.5a).

Fig. 7. Simulated and measured gain versus voltage (a) and excess noise factor versus gain (b) for the SAM-APD reported in [4], [5]. In the model we consider either pure electron injection of mixed injection (Eqs.14, 15).
B. Staircase SAM-APD from [2]

For the APD fabricated and measured in [2], we start with the nominal structure, assuming abrupt doping profiles. The device consists of a 4.5-µm i-GaAs absorption region, separated from the multiplication region by a δ-p doping. The multiplication region is the periodic repetition of a structure composed by a 20-nm i-Al0.45Ga0.55As, a 25-nm i-Al0.1Ga1−xAs (x from 0.0 to 0.45), and a 35-nm i-GaAs layers. The total length of the this region is ≈ 1 µm. Finally, a 200-nm n-GaAs layer is placed between the n+ GaAs substrate. A sample band structure profile is shown in Fig.8a: we observe the conduction band steps, while the ones in the valence band are very small. Since the energy barriers seen by the electrons are quite thick (45 nm) and include a graded region, tunneling is not accounted for in our model.

Simulated C-V curves (from TCAD) are reported in Fig.8b: the out-diffusion of dopants from the bottom n+ region has to be assumed to match experiments with simulations. Defaults values provided by the simulator [25] are assumed for the dielectric constant of the different layers and for the band discontinuities amplitude.

![Band diagram of the multiplication region of the device in [2]. Case with doping diffusion. Applied voltage is 0 V.](image)

We apply the EBHDM using as inputs the $E_C$ and $E_V$ profiles obtained again from drift-diffusion TCAD simulations using the Thermionic Emission Current model for transport at the heterojunctions. The model parameters for an arbitrary mole fraction of Al in Al1-xGa1−xAs have been linearly interpolated from Tab.I. The gain versus voltage and excess noise versus gain curve are reported in Fig.9a and b. The agreement between simulated and measured gain versus voltage is not perfect, while the excess noise versus gain is reproduced quite well.

The strong dependence of the gain on the applied voltage points out that due to the large electric field (see Fig.8a) also the regions between the steps give an important contribution to the multiplication process, more than the energy steps alone, which are only 0.55 eV high [25].

Due to the large electric field, hole impact ionization has a substantial impact on the results as demonstrated by the comparison with the curves for $\beta(x|x') = 0$. If we turn off holes in the model, see again Fig.9b, the excess noise is always below 2 also at high gains. These results confirm that low noise performance requires high energy steps compared to the gap of the material where multiplication takes place [26].

![Simulated and measured gain versus voltage (a) and excess noise factor versus gain (b) for the staircase APD reported in [2]. Experimental results for Fig.b are taken from [3]. The structure in [3] slightly differs from the one in [2], but the EBHDM, with the model parameters of Tab.I, gives the same $F(M)$ curve for both of these devices.](image)

VI. CONCLUSIONS

We have proposed a non-local history-dependent model for impact ionization suitable for (but not limited to) APDs with staircase structure. The ionization coefficients for electrons and holes are related to effective field profiles obtained from a first-order energy balance equation rather than resorting to the definition of a dead space. This renders the model robust in the presence of abrupt heterojunctions. For a given material, eight parameters are needed. The model has been calibrated against literature experimental data for gain and noise in p-i-n diodes and then applied to more complex APDs based on heterojunctions between III-V compounds. It is found that hole impact ionization still plays a substantial role in these devices and degrades the noise performance.

The model is thus a powerful tool in support of the interpretation of single photon APDs electrical measurements and for the optimization of their gain and noise.

APPENDIX A

Instead of using the iterative procedure proposed in [14], we devise a method to calculate the matrix form for Eqs.3-6. We will focus on the solution of Eqs.3-4, but the same method has been used for Eqs.5-6.

After discretization on a spatial mesh (see Fig.10), Eqs.3-4 can be written as products between vectors and matrices:

$$\bar{N}_e = \bar{A} + \bar{B}_1 \bar{N}_e + \bar{B}_2 \bar{N}_h \quad (17)$$

$$\bar{N}_h = \bar{C} + \bar{D}_2 \bar{N}_h + \bar{D}_1 \bar{N}_e \quad (18)$$

where $\bar{N}_e, \bar{N}_h$ are column vectors which represent the values of $N_e, N_h$ on the simulation mesh; $\bar{A}, \bar{C}$ are column vectors that do not contain neither $N_e$ nor $N_h$, while $\bar{B}_i$ and $\bar{D}_i$, with $i \in \{1, 2\}$, are the constant matrices which will
multiply the unknown variables. We can further re-arrange the expressions obtaining the following matrix equation:

\[
\begin{bmatrix}
\vec{N}_e \\
\vec{N}_h \\
\end{bmatrix} = -\begin{bmatrix}
\vec{B}_1 - \vec{I} \\
\vec{B}_2 \\
\end{bmatrix}^{-1} \cdot \begin{bmatrix}
\vec{D}_1 \\
\vec{D}_2 - \vec{I} \\
\end{bmatrix} \cdot \begin{bmatrix}
\vec{A} \\
\vec{C} \\
\end{bmatrix}
\]

It has to be observed (see again Fig.10) that \( \vec{N}_e, \vec{N}_h \) include only the unknown samples of \( \vec{N}_e \) and \( \vec{N}_h \). The values imposed by the boundary conditions \( \{N_0(0) = 1, N_e(H') = 1\} \) are not part of these vectors (highlighted in red in Fig.10).

Fig. 10. Spatial discretization of the variables \( N_e \) and \( N_h \) and indication of the boundary conditions (in red).

REFERENCES


