PHYSICAL REVIEW APPLIED **17,** 034044 (2022)

Editors' Suggestion

Atomistic Transport Modeling, Design Principles, and Empirical Rules for Low-Noise *III-V* Digital-Alloy Avalanche Photodiodes

Sheikh Z. Ahmed,[1,*] Yaohua Tan,[2] Jiyuan Zheng,[3] Joe C. Campbell,[1] and Avik W. Ghosh[1,4]

[1] *Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, Virginia 22904, USA*
[2] *Synopsys Inc, Mountain View, California 94043, USA*
[3] *Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China*
[4] *Department of Physics, University of Virginia, Charlottesville, Virginia 22904, USA*

(Received 30 August 2021; revised 23 November 2021; accepted 11 February 2022; published 16 March 2022)

A series of *III-V* ternary and quarternary digital alloy avalanche photodiodes have recently been seen to exhibit very low excess noise. Using band inversion of an environment-dependent atomistic tight-binding description of short-period superlattices, we argue that a combination of increased effective mass, minigaps, and band split-off are primarily responsible for the observed superior performance. These properties significantly limit the ionization rate of one carrier type, either holes or electrons, making the avalanche multiplication process unipolar in nature. The unipolar behavior in turn reduces the stochasticity of the multiplication gain. The effects of band folding on carrier transport are studied using the nonequilibrium Green's function method that accounts for quantum tunneling, and the Boltzmann transport equation model for scattering. It is shown here that carrier transport by intraband tunneling and optical phonon scattering are reduced in materials with low excess noise. Based on our calculations, we propose five simple inequalities that can be used to approximately evaluate the suitability of digital alloys for designing low-noise photodetectors. We evaluate the performance of multiple digital alloys using these criteria and demonstrate their validity.

DOI: 10.1103/PhysRevApplied.17.034044

I. INTRODUCTION

The demand for efficient optical detectors is constantly growing due to rapid developments in telecommunication, light imaging, detection and ranging systems, and other military and research fields [1–7]. Photodetectors are increasingly being incorporated in photonic integrated circuits for the Internet of things and 5G communications [8–10]. These applications require higher sensitivity in comparison to traditional *p-i-n* photodiodes [11]. Avalanche photodiodes (APDs) are often deployed instead due to their higher sensitivity, enabled by their intrinsic gain mechanism. However, the stochastic nature of the impact ionization process of APDs adds an excess noise factor $F(M) = \langle m^2\rangle/\langle m\rangle^2 = kM + (1-k)(2-1/M)$ to the shot noise current, $\langle i^2_{\text{shot}}\rangle = 2qIM^2F(M)\Delta f$ [12–14]. Here, q is the electron charge, I is the total photo plus dark current, m is the per primary electron avalanche gain, $M = \langle m\rangle$ is the average multiplication gain, and Δf is the bandwidth. A low value of k, which is the ratio of the hole ionization coefficient β to the electron ionization coefficient, α, is desirable for designing low-noise n-type APDs. This ratio stipulates that, for pure electron injection, a significantly lower hole ionization than the electron ionization rate leads to reduced shot noise. If impact ionization is caused by pure hole injection, k in the equation will be replaced by $1/k$. This behavior is generally true for low electric fields, which is usually applicable for thick avalanche regions. Additionally, dead space effects can be exploited to attain low noise in thin structures.

Recently, several *III-V* digital alloys, i.e., short-period superlattices with binary components stacked alternately in a periodic manner, were found to exhibit extremely low-noise currents and a high gain-bandwidth product in the short-infrared wavelength spectrum [15–17]. Characterization of (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) digital alloy APDs have shown very small values of k [15–17], whereas other digital alloys, like (In,Ga)As and (Al,Ga)As, demonstrate much higher k values [18,19]. The k values of these materials were determined using an Agilent 8973A noise figure analyzer to obtain the excess noise factor, $F(M)$. The total noise was measured when the APDs were illuminated and in the dark. The dark noise was then subtracted to determine the photocurrent noise. Initially, the noise was measured at the APD unity gain point and then the bias was increased to obtain the gain-dependent

---

[*] sza9wz@virginia.edu

noise characteristic. Plots of the excess noise factor versus the average gain, $\langle M\rangle$, were fit to the expression $F(M) = kM + (1-k)(2-1/M)$ to find the effective k values. Based on previous full-band Monte Carlo simulations [20–22], the low k has been attributed to the presence of superlattice minigaps inside the valence band of the material bandstructure, along with an enhanced effective mass arising from the lower bandwidth available to the holes. Such valence band minigaps often coexist with similar (but not symmetrical) minigaps in the conduction band. However, electrons in the conduction band typically have very low effective mass, which allows quantum tunneling and enhanced phonon scattering to circumvent minigaps in the conduction band. Furthermore, certain digital alloys show that minigaps do not exhibit low noise, and the reason behind that has not yet been addressed. More recently, Sb-containing random alloy (Al,In)(As,Sb) and (Al,Ga)(As,Sb) APDs have demonstrated low excess noise, and the underlying mechanism is not properly understood [23,24]. Oğuzman *et al.* showed that at high electric fields the impact ionization rate for the light-hole and split-off bands for bulk Si and GaAs is much larger compared to the heavy-hole bands [25]. In Si, ionization events originating in the split-off band are comparable to that of the light-hole rate, while for GaAs, the split-off band rate clearly dominates the hole ionization process. The β value was shown to be inversely proportional to the spin-orbit splitting [26]. Liu *et al.* demonstrated that the excess noise in GaAs can be significantly reduced by alloying with small fractions of bismuth [27]. The strong spin-orbit coupling of the heavier Bi atoms results in a larger separation between light-hole and split-off bands that reduces the hole ionization coefficient. Our postulate is that a combination of valence band minigap, a large separation between tight-hole and split-off bands, and corresponding enhanced hole effective mass tend to limit the hole ionization coefficient in the digital alloys. A comprehensive analysis is clearly necessary to understand the carrier impact ionization in these materials.

In this paper, we employ a fully atomistic, environment-dependent tight-binding (EDTB) model [28], calibrated to density functional theory (DFT) bandstructure as well as wave functions, to compute the bandstructures of several *III-V* digital alloys. Using a full three-dimensional quantum kinetic nonequilibrium Green's method (NEGF) formalism with the EDTB Hamiltonian as input, we compute the ballistic transmission across these digital alloys that accounts for intraband quantum tunneling across minigaps and light-hole and split-off bands offset. Additionally, a full-band Boltzmann transport solver is employed to determine the energy-resolved carrier density distribution under the influence of an electric field in order to study the effect of optical phonon scattering in these short-period superlattices. The calculations are performed using computational resources at the University of Virginia and XSEDE [29]. Using these transport formalisms, we elucidate the impact of minigap sizes, light-hole and split-off bands offset, and effective masses on carrier transport in the valence band.

Our simulations demonstrate that the squashing of subbands into tighter bandwidths, such as arising from minigap formation, or the engineering of large light-hole and split-off bands offset lead to the suppression in transport of one carrier type, by resisting quantum tunneling or phonon-assisted thermal jumps. For (In,Al)As, the improved performance is primarily due to the minigaps generated by the digital alloy periodicity and the corresponding enhanced effective mass. For (Al,In)(As,Sb) and Al(As,Sb), the gain is a combination of minigaps, large effective mass, and LH-SO offset. The LH-SO offsets in these two alloy results arise from the strong spin-orbit coupling due to the Sb atoms, a characteristic which is also observed in their random alloy counterparts that exhibit low noise. A quantitative comparison of the various alloy gains measured is presented in the last two columns of Table IV in Sec. IV.

The unique superlattice structure of the digital alloys opens the possibility for designing alternative low-noise alloy combinations for detection of other frequency ranges. Ideally, it is easier and cheaper to at first computationally study the suitability of the alloys for achieving low noise before actually fabricating these. For this purpose, we need a set of design criteria for judging the alloy performance using theoretically calculated parameters. Based on our simulations, we propose five simple inequalities that can be used to judge the suitability of digital alloys for use in low-noise APDs. We judge the aptness of five existing digital alloys: (In,Al)As, (In,Ga)As, (Al,Ga)As, (Al,In)(As,Sb), and Al(As,Sb). We observe that the inequalities provide a good benchmark for gauging the applicability of digital alloys for use in low-noise APDs.

II. SIMULATION METHOD

A. Environment-dependent tight binding and band unfolding for the atomistic description

In order to understand the influence of minigap filtering in digital alloy structures, an accurate band structure over the entire Brillouin zone is required. The periodic structure of the (In,Al)As digital alloy shown in Figs. 1(a) and 1(b) shows the typical structure of a *p-i-n* APD. We have developed an EDTB model to accurately calculate the band structure of alloys [28,30]. Traditional tight-binding models are calibrated directly to bulk bandstructures near their high symmetry points and not to the underlying chemical orbital basis sets [30]. These models are not easily transferable to significantly strained surfaces and interfaces where the environment has a significant impact on their material chemistry. In other words, the tight-binding parameters work directly with the eigenvalues (E-k) and

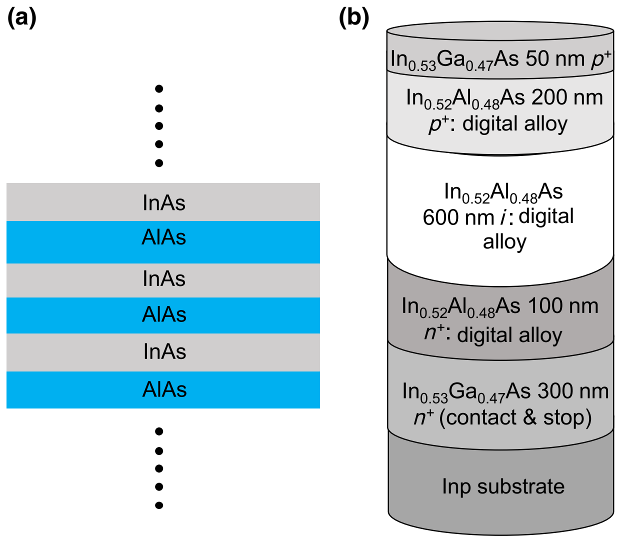

FIG. 1. (a) Digital alloy structure. (b) Typical structure of an APD.

not with the full eigenvectors. While the crystallographic point group symmetry is enforced by the angular transformations of the orbitals, the radial components of the Bloch wave functions, which determine bonding and tunneling properties, are left uncalibrated. Previously, in order to incorporate accuracy of radial components, an extended Hückel theory [31,32] was used that incorporated explicit Wannier basis sets created from nonorthogonal atomic orbitals that were fitted to density functional theory for the bulk Hamiltonian. The fitted basis sets were transferrable to other environments by simply recomputing the orbital matrix elements that the bonding terms were assumed to be proportional to. As an alternative, the EDTB model employs conventional orthogonal Wannier-like basis sets. The tight-binding parameters of this model are generated by fitting to both hybrid functional (HSE06) [33] band structures and orbital-resolved wave functions. Our tight-binding model can incorporate strain and interface-induced changes in the environment by tracking changes in the neighboring atomic coordinates, bond lengths, and bond angles. The onsite elements of each atom have contributions from all its neighboring atoms. The fitting targets include unstrained and strained bulk *III-V* materials as well as select alloys. We have shown in the past that our tight-binding model has the capability of matching the hybrid functional band structures for bulk, strained layers, and superlattices [28,34].

The band structures of the alloys contain a massive number of spaghettilike bands due to the large supercell of the system that translates to a small Brillouin zone with closely separated minibands and minigaps. In order to transform the complicated band structure into something tractable, we employ the technique of band unfolding [35–37]. This method involves projecting the eigenvalues back to the extended Brillouin zone of the primitive unit cell of either component, with weights set by decomposing individual eigenfunctions into multiple Bloch wave functions with different wavevectors in the Brillouin zone of the original primitive unit cell. The supercell eigenvector $|\vec{K}m\rangle$ is expressible in terms of the linear combination of primitive eigenvectors $|\vec{k}_i n\rangle$. The eigenstate E_p of an atom with wavevector k can be expressed as a linear combination of atomic-orbital wave functions. The supercell electron wave function $|\psi^{\rm SC}_{m\vec{K}}\rangle$ can be written as a linear combination of electron wave functions in the primitive cell as [15]

$$|\psi^{\rm SC}_{m\vec{K}}\rangle = \sum_n a(\vec{k}_i, n; \vec{K}, m)|\psi^{\rm PC}_{n\vec{k}_i}\rangle \qquad (1)$$

with $\vec{k}_i \in \{\vec{\tilde{k}}_i\}$, where $|\psi^{\rm PC}_{n\vec{k}_i}\rangle$ is the electron wave function for the wavevector $\vec{k}_i$ in the nth band of the primitive cell. Here, $\vec{\mathbf{K}}$ and $\vec{\mathbf{k}}$ denote the reciprocal vector in the supercell and primitive cell, respectively. The folding vector $\vec{\mathbf{G}}_{\vec{\mathbf{k}}\to\vec{\mathbf{K}}}$ contains the projection relationship and is expressed as

$$\vec{\mathbf{K}} = \vec{\mathbf{k}} - \vec{\mathbf{G}}_{\vec{\mathbf{k}}\to\vec{\mathbf{K}}}. \qquad (2)$$

The projection of the supercell wave function $|\psi^{\rm SC}_{m\vec{K}}\rangle$ into the primitive cell wave function $|\psi^{\rm PC}_{n\vec{k}_i}\rangle$ is given as

$$P_{m\vec{K}} = \sum_n |\langle\psi^{\rm SC}_{m\vec{K}}|\psi^{\rm PC}_{n\vec{k}_i}\rangle|^2. \qquad (3)$$

Plotting these projection coefficients gives a cleaner picture of the band evolution from the individual primitive components to the superlattice bands.

B. Nonequilibrium Green’s function method for coherent transmission

Under the influence of a large electric field, it is possible for carriers to move across minigaps by means of quantum tunneling. Such transport involves a sum of complex transmissions limited by wave function symmetry between several minibands. We make use of the nonequilibrium Green’s function formalism to compute the ballistic transmission and study the influence of minigaps on quantum tunneling in digital alloys. The digital alloys we are interested in studying are translationally invariant in the plane perpendicular to the growth direction and have finite nonperiodic hopping in the transport (growth) direction. Thus, we need a device Hamiltonian H whose basis is Fourier transformed into k space in the perpendicular x-y plane but is in real space in the z growth direction, i.e., $H(r_z, k_x, k_y)$. Conventionally, this can be done with a DFT Hamiltonian in real space, $H(r_z, r_x, r_y)$, which is Fourier transformed along the transverse axes to get $H(r_z, k_x, k_y)$. However, DFT Hamiltonians are complex and sometimes do not match with bulk material bandstructure. Thus, it is simpler to utilize a tight-binding Hamiltonian whose $E - \vec{k}$s are calibrated to the bulk bandstructure, and inverse transform along the growth direction.

The matrix elements of the three-dimensional (3D) EDTB Hamiltonian are given in the basis of symmetrically orthogonalized atomic orbitals $|nb\mathbf{R}\rangle$. Here $\mathbf{R}$ denotes the position of the atom, n is the orbital type (s, p, d or s^*), and b denotes the type of atom (cation or anion). The Hamiltonian can also be represented in the k-space basis $|nb\mathbf{k}\rangle$ by Fourier transforming the elements of the real-space Hamiltonian. The 3D Hamiltonian is then converted into a quasi-1D Hamiltonian [38]. The Hamiltonian elements can be represented in the basis $|nbj\,\mathbf{k}_{||}\rangle$ with “parallel” momentum $\mathbf{k}_{||} = (k_x, k_y)$ and “perpendicular” position $x_j = a_L/4$ as parameters. For a zinc-blende crystal, the distance between nearest-neighbor planes is one-fourth the lattice constant a_L. The 3D Hamiltonian is converted to the quasi-1D Hamiltonian by means of a partial Fourier transform [38,39]:

$$|nbj\,\mathbf{k}_{||}\rangle = L_{\mathrm{BZ}}^{-1/2}\int dk_z e^{-ik_z j a_L/4}|nb\mathbf{k}\rangle. \tag{4}$$

Here $L_{\mathrm{BZ}} = 8\pi/a_L$ is the length of the 1D Brillouin zone over which the k_z integral is taken. The quasi-1D Hamiltonian is position dependent in the growth direction. Thus, we are able to utilize the accurate bandstructure capability of the EDTB.

In the presence of contacts, the time-independent open boundary Schrödinger equation reads

$$(EI - H - \Sigma_1 - \Sigma_2)\Psi = S_1 + S_2, \tag{5}$$

where E represents energy, I denotes the identity matrix, and $\Sigma_{1,2}$ are the self-energies for the left and right contacts, respectively, describing electron outflow, while $S_{1,2}$ are the inflow wave functions. The solution to this equation is $\Psi = G(S_1 + S_2)$, where the Green’s function [40]

$$G(E) = [EI - H - \Sigma_1 - \Sigma_2]^{-1}. \tag{6}$$

Here H includes the applied potential, added to the onsite 1D elements. Assuming that the contacts are held in local equilibria with bias-separated quasi-Fermi levels $\mu_{1,2}$, we can write the bilinear thermal average $\langle S_i S_i^\dagger\rangle = \Gamma_i f\,(E - \mu_i)$, where f is the Fermi-Dirac distribution and $\Gamma_{1,2} = i(\Sigma_{1,2} - \Sigma_{1,2}^\dagger)$ denotes the broadening matrices of the two contacts. The equal time current $I = q(d/dt + d/dt')Tr\langle\Psi^\dagger(t)\Psi(t')\rangle|_{t=t'}$ then takes the Landauer form $I = (q/h)\int dET(f_1 - f_2)$, where the coherent transmission between the two contacts is set by the Fisher-Lee formula

$$T(E) = \mathrm{Tr}[\Gamma_1 G\Gamma_2 G^\dagger], \tag{7}$$

where Tr represents the trace operator. The energy-resolved net current density from layer m to layer $m+1$ is expressed as [38]

$$J_{m,m+1}(E) = -\frac{iq}{h}\int\frac{\mathbf{k}_{||}}{(2\pi)^2}\mathrm{Tr}[G^{n,p}_{m+1,m}H_{m,m+1} - G^{n,p}_{m,m+1}H_{m+1,m}], \tag{8}$$

where $G^n = \langle\psi^\dagger\psi\rangle$ and $G^p = \langle\psi\psi^\dagger\rangle$ respectively represent the electron (n) and hole density (p), and $H_{m,m+1}$ is the tight-binding hopping element between layers m and $m+1$ along the transport-growth direction.

C. Boltzmann transport model for incoherent scattering

The NEGF approach is particularly suited to ballistic transport where coherent quantum effects dominate. Incoherent scattering requires a self-consistent Born approximation that is computationally quite involved. We need a practical treatment of scattering. Under an external electric field, the carrier distributions in digital alloys no longer follow a local Fermi distribution, but redistributed over real-space and momentum space. To understand the carrier distribution under an electric field in digital alloys, we employed the multiband Boltzmann equation:

$$\vec{v}\cdot\nabla_{\mathbf{r}}f_n + \vec{F}\cdot\nabla_{\mathbf{k}}f_n = \sum_{m,\vec{p}'}S(\vec{p}',\vec{p})f_m(\vec{p}')[1 - f_n(\vec{p})] - \sum_{m,\vec{p}'}S(\vec{p},\vec{p}')f_n(\vec{p})[1 - f_m(\vec{p}')]. \tag{9}$$

Here, $f = f(\mathbf{r}, \mathbf{k})$ is the carrier distribution, n and m are band indices, $\vec{p}$ and $\vec{p}'$ are the momenta of the carriers, and $S(\vec{p}',\vec{p})$ is the scattering rate. The left-hand side of this equation alone describes the ballistic trajectory in the phase space of carriers under an electric field. The right-hand side of the equation corresponds to the scattering processes including intraband and interband scattering.

In a homogeneous system where the electric field is a constant, the distribution function is independent of position, $\nabla_{\mathbf{r}}f = 0$, and the equation is reduced to

$$\vec{F}\cdot\nabla_{\mathbf{k}}f_n = \sum_{m,\vec{p}'}S(\vec{p}',\vec{p})f_m(\vec{p}')[1 - f_n(\vec{p})] - \sum_{m,\vec{p}'}S(\vec{p},\vec{p}')f_n(\vec{p})[1 - f_m(\vec{p}')]. \tag{10}$$

For APDs, it is critical to consider optical phonon scattering, which is the dominant process besides tunneling that allows carriers to overcome the minigap arising in the band structures of digital alloys. The optical phonon has a nontrivial energy of $\hbar\omega_{opt}$ that can be absorbed or emitted by carriers. The scattering rate $S(\vec{p}\,',\vec{p})$ has the form set by Fermi's golden rule:

$$S(\vec{p}\,',\vec{p}) = \frac{2\pi}{\hbar}|H_{\vec{p},\vec{p}\,'}|^2\delta_{\vec{p}\,',\vec{p}\pm\vec{\beta}}\delta[E(\vec{p}\,') - E(\vec{p}) \pm \hbar\omega_{\mathrm{opt}}]. \tag{11}$$

The $E(\vec{p})$ and $E(\vec{p}\,')$ are band structures of the digital alloy calculated with the tight-binding model. We can calculate $H_{p,p'}$ by evaluating the electron-phonon coupling matrix elements explicitly. In this work, we extract a constant effective scattering strength $H_{\vec{p},\vec{p}\,'}$ from experimental mobility μ. The scattering lifetime τ, which is $1/S(\vec{p}\,',\vec{p})$, can be extracted from the mobility using $\mu = q\tau/m^*$. Because of lack of experimental mobilities of the digital alloys, we considered the average of the binary constituent room-temperature mobilities for extracting the lifetime. A simple average is done since the binary constituents in periods of most of the digital alloys considered here are equally divided. In using room-temperature values the underlying assumption is that the dominant scattering mechanism here is phonon scattering due to a large phonon population. Ionized impurity scattering is considered to be much lower due to digital alloys having clean interfaces [16]. It is then possible to extract $H_{\vec{p},\vec{p}\,'}$ from the scattering lifetime. To get the equilibrium solution, we solve Eq. (10) self-consistently, starting from an initial distribution $f = \delta_{\vec{k},0}$.

A detailed model of carrier transport in APDs also requires a NEGF treatment of impact ionization self-energies and a Blanter-Buttiker approach to extract shot noise, but we leave that to future work. Our focus here is on conductive near-ballistic transport, and the role of quantum tunneling and perturbative phonon scattering in circumventing this.

III. RESULTS AND DISCUSSION

There are three common ways to achieve a low noise and high gain-bandwidth product: selecting a semiconductor with favorable impact ionization coefficients, scaling the multiplication region to exploit the nonlocal aspect of impact ionization, and impact ionization engineering using appropriately designed heterojunctions [11]. Typically, the lower hole impact ionization coefficient in semiconductors is due to stronger scattering in the valence bands, as depicted in Fig. 2(a). Previously, the lowest noise with favorable impact ionization characteristics were realized with Si in the visible and near-infrared range [41–44], and InAs [45–49] and (Hg,Cd)Te [50,51] in the midinfrared spectrum. In comparison, (In,Ga)As/(In,Al)As [52,53] random alloy APDs exhibit significantly higher noise than Si, (Hg,Cd)Te, or InAs, which are the highest performance telecommunication APDs. In the recent past, digital alloy (In,Al)As APDs have demonstrated lower noise compared to their random alloy counterpart [15]. This seems a surprise, as the suppression of one carrier type (the opposite of ballistic flow expected in an ordered structure) is necessary for low excess noise. Initially, the low value of k in (In,Al)As was attributed to the presence of minigaps [22]. However, minigaps were also observed in (In,Ga)As digital alloy APDs that have higher excess noise [18,54]. So, a clearer understanding of the minigap physics was needed and hence a comprehensive study was required.

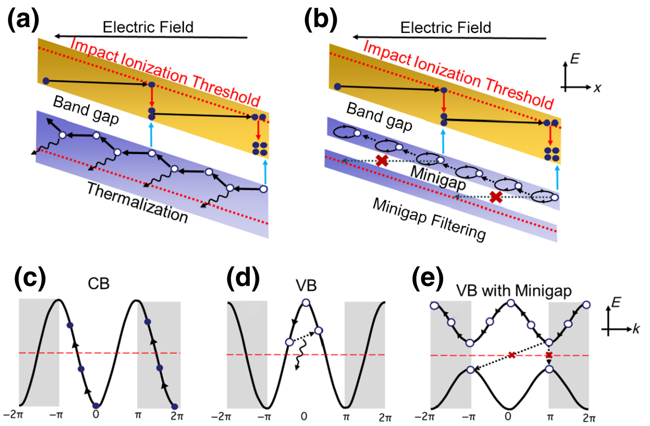

FIG. 2. Impact ionization process in a normal (random alloy) APD and superlattice APD. In both APDs, it is easier for electrons to gain energy and reach the impact ionization threshold (c). The low effective mass of the electrons in these materials allow easy acceleration under an applied electric field and overcome any minigaps or scattering processes present in the conduction band. In normal APDs (a), holes find it harder to gain high energy compared to electrons because of thermalization. The hole energy is reduced by thermalization due to various scattering processes, as shown in (d). In the superlattice APD (b), the existence of minigaps makes it harder for holes to reach higher energies. The minigaps act as barriers that prevent holes from moving to the lower valence bands. In the plots, the y axis E is the total energy (kinetic plus potential), meaning that in between inelastic scattering events the particles travel horizontally.

Our recent results suggest that well-defined minigaps introduced in the valence band of digital alloys suppress the density of high-energy holes and thereby reduce the impact ionization greatly, as shown in Fig. 2(b). In a regular low-noise electron-injected APD, the electron ionization coefficient is much higher than the hole ionization coefficient. The conduction band minigaps in the digital alloys can be bypassed by quantum tunneling due to the low effective masses of the electrons. Thus, electrons can easily climb to higher kinetic energies in the conduction band, depicted in Fig. 2(c), and participate in the impact ionization process by gaining the impact ionization threshold energy. Random and digital alloys have similar electron impact ionization coefficients [19] that verify that conduction band minigaps do not limit electron impact ionization. On the other hand, holes lose energy by various inelastic scattering processes [Fig. 2(d)], collectively known as thermalization. Thermalization prevents holes from reaching their secondary impact ionization threshold. In superlattice APDs, minigaps provide an additional filter mechanism that prevents holes from reaching the threshold energy required to initiate secondary impact ionization.

The effect of minigaps is shown in Fig. 2(e). However, not all digital alloy APDs exhibit low noise. The excess noise $F(M)$ versus multiplication gain characteristics of experimental (In,Ga)As, (Al,Ga)As, (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) digital alloy APDs are shown in Fig. 3 [15–19]. (In,Ga)As APDs have the highest excess noise while Al(As,Sb) has the lowest. A key observation from this compilation is that As-based APDs have higher excess noise than Sb-based APDs. The dotted lines represent the theoretical $F(M)$ versus M calculated using the well-known McIntyre's formula [12], introduced in the first paragraph of this paper. In order to understand the underlying physics in these digital alloys, an in-depth analysis of the material bandstructure and its effect on the carrier transport is required.

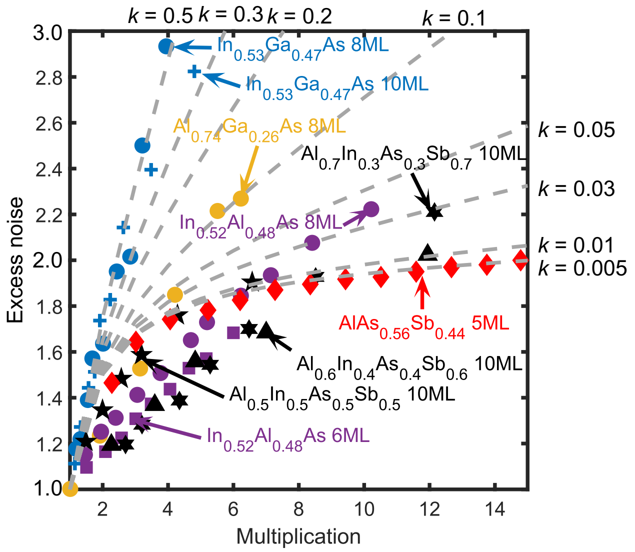

FIG. 3. Experimentally measured excess noise versus multiplication gain of (In,Ga)As, (Al,Ga)As, (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) digital alloys are shown here [15–19]. The dotted lines for the corresponding k values are plotted using McIntyre's formula [12].

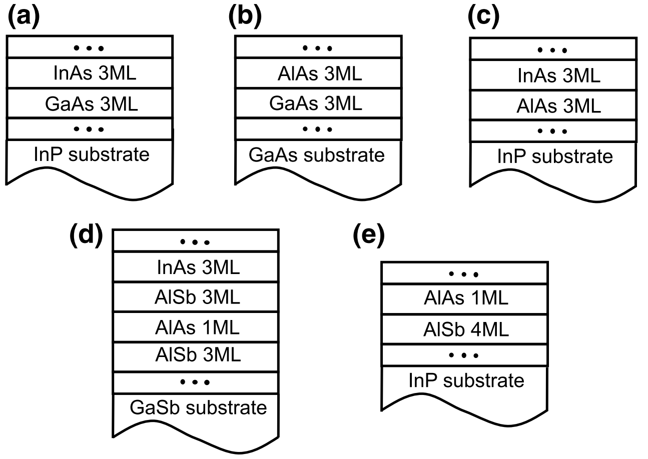

FIG. 4. Lattice structures of (a) (In,Ga)As, (b) (Al,Ga)As, (c) (In,Al)As, (d) (Al,In)(As,Sb), and (e) Al(As,Sb) digital alloys considered in this paper.

We calculate the atomistic DFT-calibrated EDTB bandstructure of these materials and unfold their bands using the techniques described in Sec. II A, to understand the underlying physics of their noise performance. In Fig. 4, we show the periods of the different digital alloys considered: (a) 6ML [monolayer (ML)] (In,Ga)As, (b) 6ML (Al,Ga)As, (c) 6ML (In,Al)As, (d) 10ML $Al_{0.7}In_{0.3}AsSb$, and (e) 5ML Al(As,Sb). Here, 6ML (In,Ga)As includes 3ML InAs and 3ML GaAs, 6ML (Al,Ga)As has 3ML AlAs and 3ML GaAs, and 6 ML (In,Al)As has 3ML InAs and 3ML AlAs. 10ML $Al_{0.7}In_{0.3}AsSb$ consists of 3ML AlSb, 1ML AlAs, 3ML AlAs, and 3ML InAs in its period. Al(As,Sb) has 4ML AlSb, and 1ML AlAs. The unfolded bandstructures of these alloys are shown in Fig. 5. We observe that minigaps exist in at least one of the valence bands (heavy hole, light hole, or split off) for all the material combinations. The (In,Al)As valence band structure is magnified in Fig. 6. The minigap between the LH and SO bands is denoted in the figure. Additionally, the large separation between the LH and SO bands at the Γ point is highlighted. In general, the minigap size shows a decreasing trend with increasing period thickness, as was observed for the (In,Al)As digital alloy [15]. However, the minigaps disappear for very short-period (≤ 4ML) structures, as was recently observed by Wang *et al.* for the 4ML (In,Al)As digital alloy [55]. This is primarily due to the increased edge roughness in these structures that result in a larger smearing of the bands around the minigap regions.

The role of the minigaps on hole localization is not identical across different alloys. For instance, the presence of minigaps in the material bandstructure is not sufficient to realize low noise in APDs. Taking a closer look at the bandstructures, we observe that the positions in energy of the minigaps with respect to the valence band edge differ from one material to another. Additionally, the minigap sizes of the different alloys vary in magnitude. A complimentary effect of the minigap size is the flattening of the energy bands, i.e., a large minigap size results in flatter bands around the gap. This in turn results in an increased effective mass that tends to inhibit carrier transport. Table I lists the energy location of the minigap with respect to the valence band edge ΔE_b, the minigap size ΔE_m, the light-hole (LH) and split-off (SO) band effective masses, and the energy difference between the LH and SO bands ΔE_{LS} at the Γ point for the digital alloys studied.

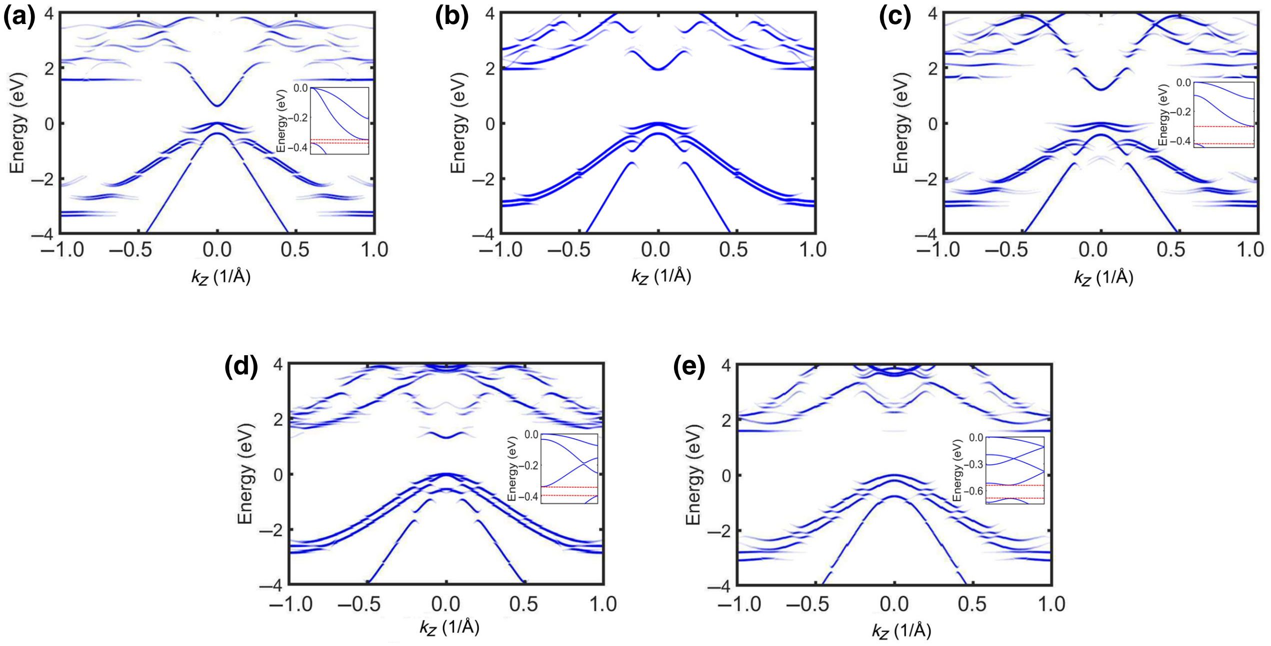

FIG. 5. Unfolded bandstructure of (a) 6ML (In,Ga)As, (b) 6ML (Al,Ga)As, (c) 6ML (In,Al)As, (d) 10ML (Al,In)(As,Sb), and (e) 5ML Al(As,Sb). The minigaps of (In,Ga)As, (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) real bandstructures are shown in the insets.

We can see in the table that there are significant variations in minigap size and position between different materials. At first glance, there seems to be no direct correlation between these variations and the excess noise, prompting us to do added transport analyses. Under a high electric field, a carrier must gain at least the threshold energy, E_{th}, in order to impact ionize. Typically, E_{th} is assumed to be approximately 1.5 times the material bandgap, E_G. Thus, in the presence of minigaps, electrons and holes must bypass these gaps by some transport mechanism in order to gain energy equivalent to E_{th}. Two such major transport mechanisms are quantum mechanical tunneling and optical phonon scattering. Our transport study must incorporate these two mechanisms to understand the effectiveness of minigaps on the APD excess noise.

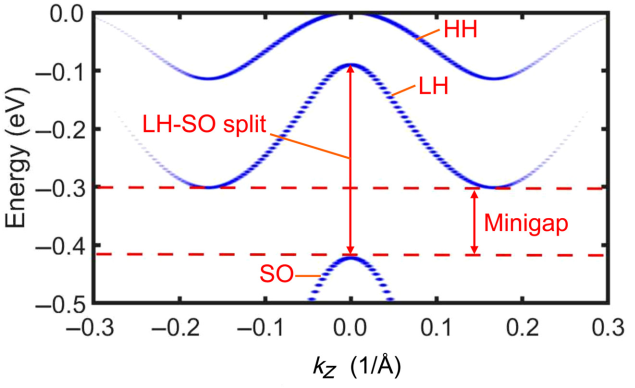

FIG. 6. A magnified picture of the (In,Al)As valence band shows the minigap closest to the valence band edge. The split between the LH and SO at the Γ point is also highlighted.

We employ the NEGF formalism described in Sec. II B to compute the ballistic transmission in the valence band as a function of energy, $T(E)$, dominated by tunneling processes. The effect of different minigap sizes is highlighted in Fig. 7. For our simulation, we set the quasi-Fermi level of the left contact at $-qV$ below the valence band edge and the quasi-Fermi level of the right contact at another $-qV$ below. This is done in order to only observe the intraband tunneling inside the valence band that is responsible for overcoming minigaps under ballistic conditions. In Fig. 7(a), we demonstrate that a small minigap in the valence band creates a small tunneling barrier for the holes. A hole with a small enough effective mass will be able to tunnel across this barrier and render it ineffective. That is the case for (In,Ga)As, which has a LH effective mass of $0.13m_0$ and $\Delta E_m = 0.03$ eV. The spectral current density for (In,Ga)As under a bias $V = 0.25$ V is shown in Fig. 7(b). We observe that the current spectrum in the valence band is continuous in the Fermi energy window and there is no drop in transmission due to the minigap. For a large minigap, the holes encounter a larger tunneling barrier, as shown in Fig. 7(c), preventing them from gaining the threshold energy E_{th} for secondary impact ionization. This case is operational in (In,Al)As digital alloys, as shown in the spectral density plot in Fig. 7(d). (In,Al)As has a minigap size of 0.12 eV and LH effective mass of $0.4m_0$. Within the Fermi window we see that

TABLE I. Material parameters of the different digital alloys simulated in this paper.

Material	E_G (eV)	ΔE_b (eV)	ΔE_m (eV)	HH m^*	LH m^*	SO m^*	ΔE_{LS} (eV)
(In,Ga)As	0.63	0.34	0.03	0.31	0.13	0.045	0.35
(Al,Ga)As	1.94	1.03	0.34	0.45	0.31	0.12	0.33
(In,Al)As	1.23	0.30	0.12	0.5	0.4	0.1	0.31
(Al,In)(As,Sb)	1.19	0.33	0.06	0.42	0.38	0.08	0.48
Al(As,Sb)	1.6	0.56	0.1	0.45	0.3	0.13	0.54

there are regions with extremely low current due to low tunneling probability across the minigap. This is further demonstrated by the $T(E)$ versus E plot in Fig. 7(e). Here, it is observed that there are regions of low transmission for (In,Al)As, whereas the (In,Ga)As transmission is continuous. This signifies that the minigaps in the (In,Al)As valence band are large enough to prevent holes from gaining in kinetic energy, resulting in a low hole ionization coefficient.

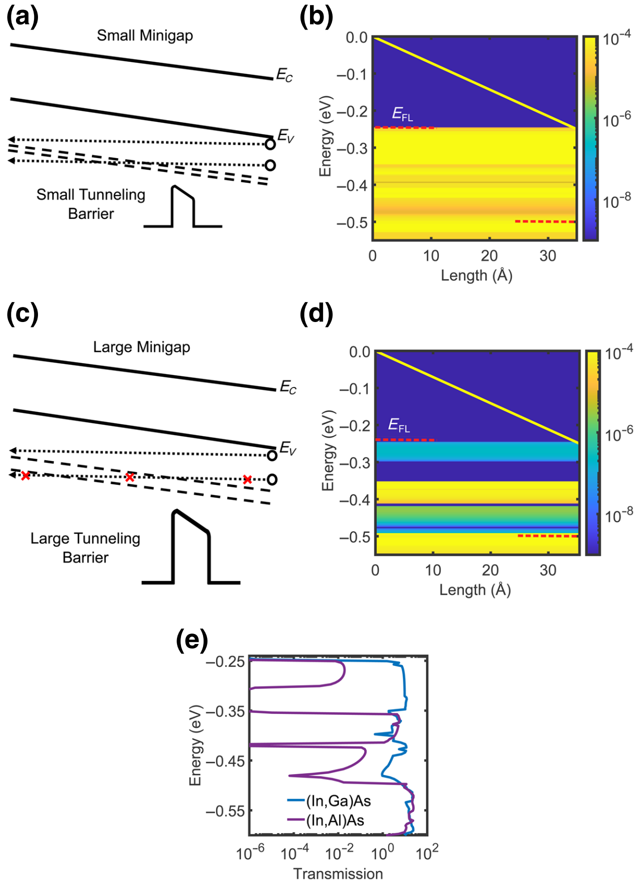

FIG. 7. Small minigaps in the valence band, as shown in (a), create a small tunneling barrier that can be overcome by holes with low mass. The spectral current density for (In,Ga)As, which has a small minigap and small LH effective mass, is shown in (b). The current spectrum for (In,Ga)As in the Fermi window is continuous. The creation of a large tunneling barrier by a larger minigap is shown in (c). This barrier prevents hole transmission. (In,Al)As has a larger minigap and LH m^*. Regions of low current density are observed within the Fermi window in the (In,Al)As spectral current density in (d). The large minigap in (In,Al)As results in reduced transmission, as shown in the $T(E)$ versus E plot of (e). The simulations for (b), (d) and (e) are conducted under a bias of $V = 0.25$ V. Here E_V is the valence band edge, E_C the conduction band edge and E_{FL} represents quas-Fermi level of the left contact.

In order to investigate the role of minigaps in the remaining digital alloys, we look at the transmission versus energy plots for all the alloys. The $T(E)$ versus E characteristics for the five digital alloys are shown in Fig. 8 for two bias conditions: (a) $V = 0.25$ V and (b) $V = 0.5$ V. We use a 21×21 grid for the transverse wavevectors (k_x, k_y) within the first Brillouin zone. For this simulation, the structure lengths for (In,Ga)As, (Al,Ga)As, (In,Al)As, and Al(As,Sb) are considered to be two periods. For (Al,In)(As,Sb) we consider one period length. This allows us to keep the structure lengths as close as possible. We consider lengths of 3.48 nm (In,Ga)As, 3.42 nm (Al,Ga)As, 3.54 nm (In,Al)As, 3.06 nm (Al,In)(As,Sb), and 3.08 nm Al(As,Sb) channels. The channel sizes chosen are small compared to actual device lengths in order to keep the computation tractable. For both the bias conditions in Fig. 8 we see that there are energy ranges for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) in which the transmission probability drops drastically. This low tunneling probability can be attributed to two factors. The first factor is the presence of a sizeable minigap in all directions in the material bandstructure. The other contributing factor is the separation between the LH and SO bands. To a large extent this factor is responsible for the low transmission

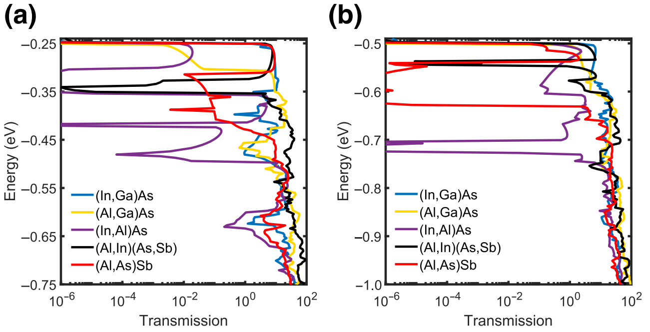

FIG. 8. The transmission $T(E)$ versus energy E for all the digital alloys at (a) $V = 0.25$ V and (b) $V = 0.5$ V. A 21×21 grid for transverse wavevectors is used.

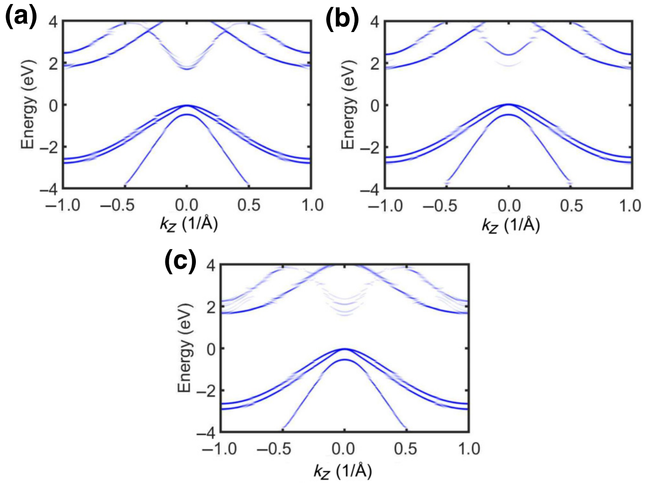

FIG. 9. Unfolded bandstructure of Sb-containing random alloys: (a) $Al_{0.79}In_{0.21}As_{0.74}Sb_{0.26}$, (b) $AlAs_{0.44}Sb_{0.56}$, and (c) $Al_{0.85}Ga_{0.15}As_{0.56}Sb_{0.44}$.

regions in (Al,In)(As,Sb) and Al(As,Sb), whose minigap sizes (from Table I) are smaller than (In,Al)As but also demonstrate lower excess noise. (In,Ga)As and (Al,Ga)As do not have any large drop in transmission for both biases. This characteristic implies that either the minigap size is too small to affect the carrier transport like in (In,Ga)As or there is no minigap at all, as in (Al,Ga)As.

To further underscore the role of the separation between the LH and SO bands, we looked at the bandstructure of Sb-containing random alloys in Fig. 9. Here, for $Al_{0.79}In_{0.21}As_{0.74}Sb_{0.26}$, $AlAs_{0.44}Sb_{0.56}$, and $Al_{0.85}Ga_{0.15}As_{0.56}Sb_{0.44}$, the ΔE_{LS} values are 0.44, 0.46, and 0.52 eV, respectively. The LH and SO offset values of the Al(As,Sb) and (Al,In)(As,Sb) random alloys are comparable to their digital alloy counterparts that are given in Table I. The different gap sizes of these alloys most likely originate from variations in bonding and antibonding interactions due to their different chemical compositions. The valence band effective masses of these random alloys are also similar to the digital alloy masses. However, there are no minigaps present in these random alloys. Thus, the low excess noise observed recently in the (Al,In)(As,Sb) and (Al,Ga)(As,Sb) random alloys can be attributed to their large LH/SO offsets. This also indicates that these offsets play a crucial role in achieving low noise in Al(As,Sb) and (Al,In)(As,Sb) digital alloys.

For further confirmation of these observations on the digital alloys, we compute the spectral current density for the case of constant total period length for all the structures. The period size of each unit cell stays the same, but the number of unit cells is increased to make the total period length the same for all alloys. We consider the case with total period of 30MLs and voltage bias of 0.25 V. The current spectral density plots for the five digital alloys using a 15 × 15 transverse wavevector grid are shown in Fig. 10. A smaller number of grid points is used here to save computation time. In the figure, a very small minigap is observed for (In,Ga)As within the Fermi window and a continuous spectrum is seen for (Al,Ga)As. Regions of low transmission and current are observed for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb). These observations are consistent with our previous calculations. We can thus infer that at least under fully coherent transport including tunneling, holes will not be able to gain sufficient kinetic energy to achieve impact ionization.

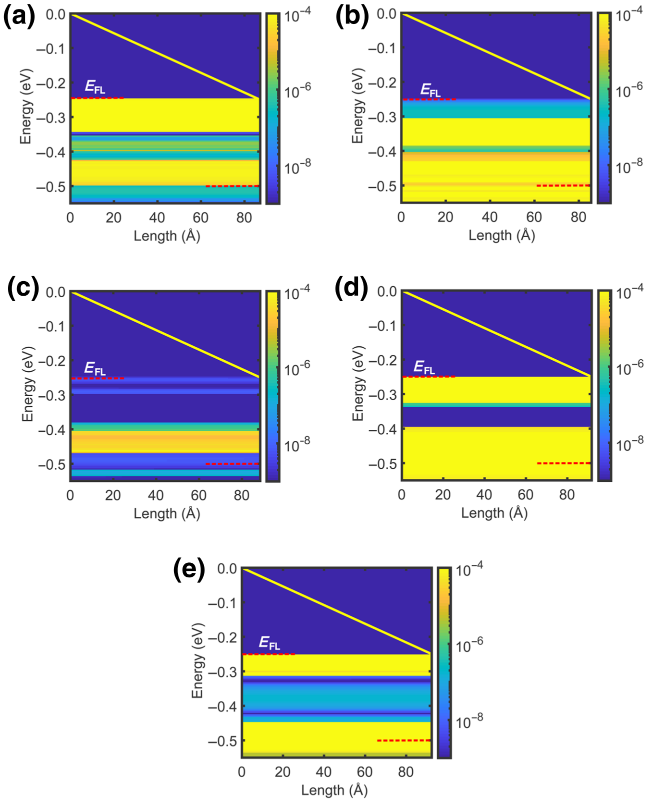

FIG. 10. Energy-resolved current spectral density in the valence band for (a) (In,Ga)As, (b) (Al,Ga)As, (c) (In,Al)As, (d) (Al,In)(As,Sb), and (e) Al(As,Sb). The bias for the simulation is set to $V = 0.25$ V and the total period length is 30 monolayers.

Besides tunneling processes, it is possible for carriers to jump across energy gaps through inelastic scattering. In APDs, the dominant scattering mechanism is intervalley optical phonon scattering. Using the Boltzmann transport equation (BTE) model described in Sec. II C, the effect of phonon scattering in digital alloys is studied. The low-field carrier mobilities and optical phonon energies of the binary constituents of the alloys used in the BTE simulations are listed in Table II. The low-field mobilities are generally valid for electric fields up to of 1–100 kV/cm depending on the material. The effective mobility at high fields is proportional to the low-field mobility [58]. Experimental data for the high field mobilities or scattering lifetimes of the digital alloys are not available. Thus, it is reasonable to use the

TABLE II. Electron/hole mobilities and optical phonon energies of binary compounds that form the digital alloys [56,57].

Material	μ_h (cm^2/Vs)	E_{opt} (meV)
InAs	500	30
AlAs	200	50
GaAs	400	35
AlSb	400	42

low-field mobilities because the relative difference in the scattering lifetimes of the different digital alloys and their effects are preserved even at high electric fields. An effective scattering strength $H_{\vec{p},\vec{p}'}$ is obtained from the mobility values, as described in Sec. II C. For our BTE simulations, we use the heavy-hole effective masses outlined in Table I. We compute the carrier density distribution in the valence band under a high electric field of 1 MV/cm, by solving the three-dimensional Boltzmann equation with the entire set of tight-binding energy bands within the Brillouin zone of the digital alloy. The optical phonon energy and mobilities of each alloy are taken to be the average of the binary constituent optical phonon energies and their mobilities. The energy-resolved carrier density distribution of the valence band for all the alloys is shown in Fig. 11. The valence band plot in Fig. 11 has contributions from different valence bands like heavy-hole, light-hole, and split-off bands. It shows that the occupation probabilities for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) are lower than those of the other two alloys at high energies. The optical phonon energies of these alloys are not sufficiently large to overcome their minigaps and thus prevent holes from ramping their kinetic energies up to E_{th}.

The top few valence bands of (In,Ga)As are shown on the left side of Fig. 12(a) and the valence band carrier density distribution is projected onto the bottom. The bands are inverted for a better view. For clearer understanding, the (In,Ga)As carrier density distribution contour is also shown on the right. The valence band carrier distributions for the (Al,Ga)As, (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) alloys are respectively shown in Figs. 12(b), 12(c), 12(d), and 12(e). By studying the contours of each material, we observe that the densities for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) are more localized compared to those of (Al,Ga)As and (In,Ga)As. This is once again consistent with the lower hole impact ionizations of (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb).

For (In,Ga)As and (Al,Ga)As, the bandwidths are large enough to allow both holes and electrons to reach E_{th} easily. The resulting values of k for these materials are quite high. Correspondingly, these two alloys have higher excess noise. In contrast, for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb), it is easy for electrons to reach the threshold energy, but the holes are confined close to the valence band edge. This results in asymmetric ionization coefficients that give a low k, leading in turn to low excess noise.

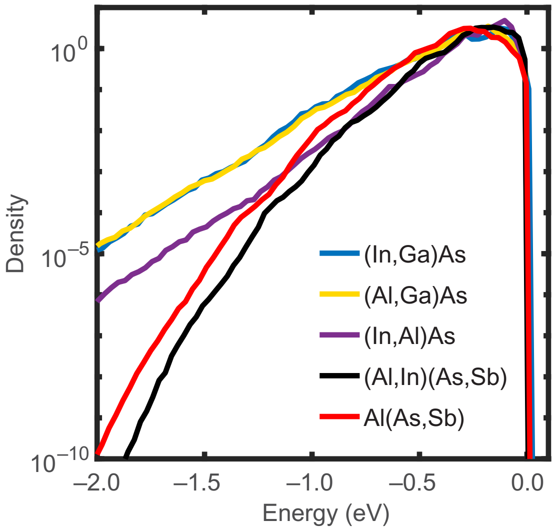

FIG. 11. Carrier density distribution versus energy for the valence band in the presence of optical phonon scattering computed using BTE simulation. (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) have lower occupation probabilities compared to (In,Ga)As and (Al,Ga)As. This prevents holes from gaining the ionization threshold energy.

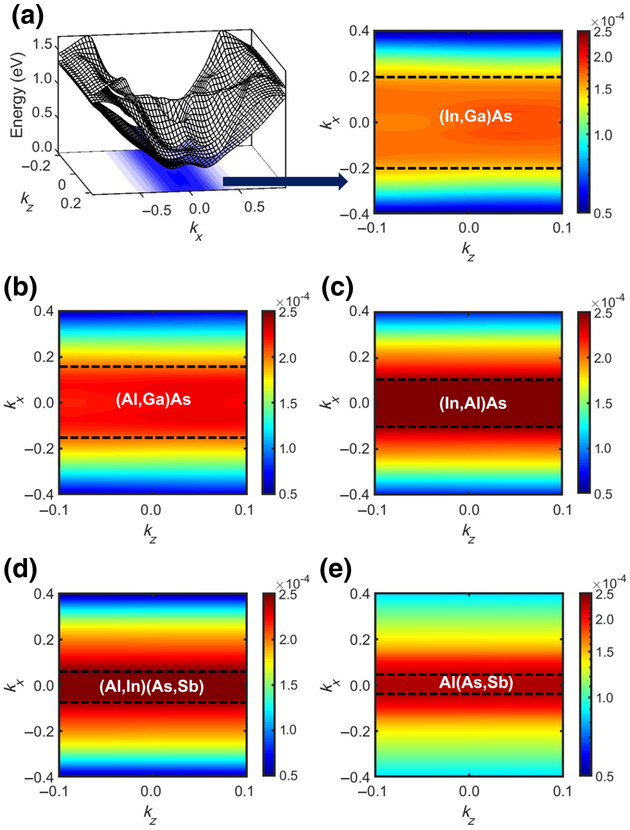

FIG. 12. Carrier density distribution for (a) (In,Ga)As, (b) (Al,Ga)As, (c) (In,Al)As, (d) (Al,In)(As,Sb), and (e) Al(As,Sb).

Armed with these results, we attempt to paint a clearer picture of how the minigaps and band splitting can reduce the excess noise in APDs. Specifically, we propose a set of empirical inequalities that can used to judge the excess noise performance of a digital alloy.

IV. EMPIRICAL INEQUALITIES

Based on our experimental results and theoretical calculations, five inequalities are proposed that use only material parameters like effective mass and minigap size obtained from our material bandstructures as inputs. In this paper, the transport is in the [001] direction. Since the minigaps considered lie in the LH band, we use the unfolded LH effective mass value in the $\Gamma - [001]$ direction for the inequalities. The masses are obtained using the relationship $\hbar^2 k^2/2m^* = E(1+\alpha E)$, where $\alpha = [(1 - m^*/m_0)^2]/E_G$ [59]. In reality, the effective masses are complicated tensors that cannot be included in these empirical inequalities, but are captured by the NEGF simulations described in Sec. II B. A digital alloy material should favor low noise if it satisfies the majority of these inequalities. The four main inequalities are as follows.

Inequality (1): $\Delta E_b/E_{\text{th}} \ll 1.$
Inequality (2): $E_{\text{opt}}/\Delta E_m \ll 1.$

Inequality (3): $\exp\left(-\dfrac{4\sqrt{2m_l}\Delta E_m^{3/2}}{3q\hbar F}\right) \ll 1.$

Inequality (4): $\exp\left(-\dfrac{4\sqrt{2m_l}\Delta E_{\text{LS}}^{3/2}}{3q\hbar F}\right) \ll 1.$

Here, ΔE_b represents the energy difference between the VB maximum and the first minigap edge in VB, E_{opt} is the optical phonon energy, and ΔE_m gives the size of the minigap. The longitudinal effective mass of the band in which the minigap exists is represented by m_l. The energy difference between the LH and SO bands at the Γ point is denoted by ΔE_{LS}. A pictorial view of the different energy differences and inequalities mentioned above is shown in Fig. 13.

The first inequality, inequality (1), states that the energy bandwidth ΔE_b must be less than the ionization threshold energy E_{th}. This means that a carrier cannot gain sufficient kinetic energy to impact ionize before reaching the minigap. When a carrier reaches a minigap, it faces a barrier (Fig. 13), which it can overcome by phonon scattering or quantum tunneling. Within the parabolic band approximation, $E_{\text{th}} = [(2\mu + 1)/(\mu + 1)]E_G$, where for holes, μ is the ratio between the valence band effective mass and conduction band mass. According to this equation, E_{th} varies between E_G and $2E_G$ depending on the value of μ. For the minigaps to be effective, they should ideally be located within an E_G value away from the valence band edge. Inequality (2) sets the condition for phonon scattering across the minigap. If the E_{opt} of the material is less than ΔE_m then the phonon scattering of the carriers across the minigap is inhibited because carriers cannot gain sufficient energy to jump across the gap. It is possible for the carrier to still overcome the minigap by tunneling, and the condition for that is given in inequality (3), in terms of the tunneling probability across the minigap under the influence of an electric field. To compute the tunneling probability, we consider a triangular barrier in the minigap region and use the well-known Fowler-Nordheim equation. Together inequalities (2) and (3) give the effectiveness of the minigap in limiting hole ionization in digital alloys.

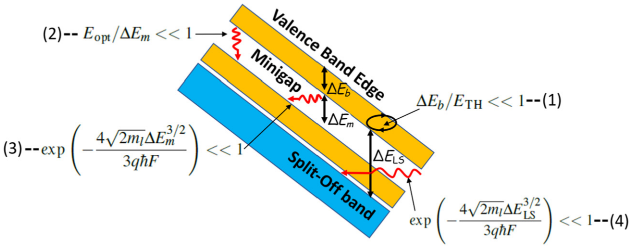

FIG. 13. Criteria for designing low-noise digital alloy APDs. Inequality (1) states that the bandwidth to the first minigap is lower than the ionization threshold energy. Inequality (2) asserts that the optical phonon energy has to be less than the minigap size. The tunneling probability for holes to jump across the minigap or from the light-hole band to the split-off band must be low. These are described by inequalities (3) and (4).

Electron-injected digital alloys can in fact achieve low noise even in the absence of minigaps, for instance in a material with a large separation ΔE_{LS} between the LH and SO bands, like Al(As,Sb). Holes within the heavy-hole (HH) and LH bands are limited near the valence band edge by thermalization (hole-phonon scattering) due to the heavy effective masses in these bands, preventing them from reaching the ionization threshold energy within the band. An alternate pathway to ionization involves the split-off band. Since the split-off band has a low effective mass, holes require much smaller momentum to reach higher energies in this band, so that holes entering this band from HH or LH can quickly gain their ionization threshold energy. The separation between HH or LH and SO bands is controlled by spin-orbit coupling, as shown in Fig. 14. Strong spin-orbit coupling due to the inclusion of heavy elements, like antimony or bismuth, can increase the separation ΔE_{LS}, as shown in Fig. 14(b). When ΔE_{LS} is large, it becomes very difficult for holes to reach the threshold energy. Inequality (4) is accordingly important for APDs in which electron impact ionization is the dominant process, and is a measure of hole tunneling from the light-hole to the split-off band.

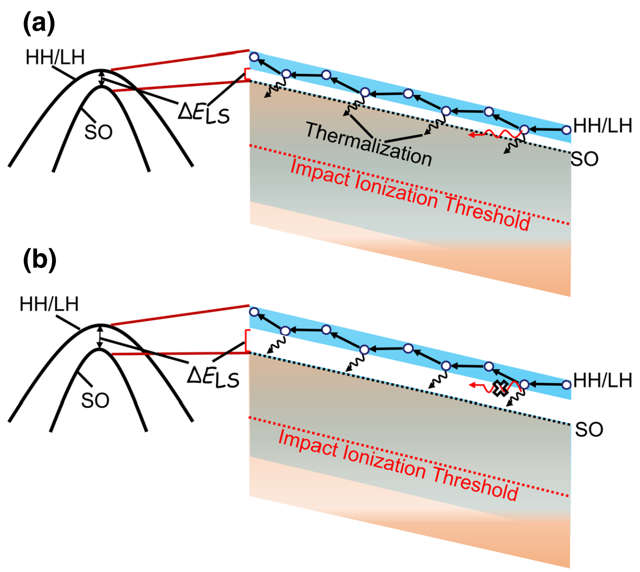

FIG. 14. Effect of spin-orbit coupling on LH/SO separation. (a) Weak coupling results in small ΔE_{LS} and (b) strong coupling results in large ΔE_{LS}.

An inherent fifth inequality, satisfied by these five alloys, is

$$E_{\text{SC}} < E_{\text{th}}. \tag{12}$$

Here E_{SC} is the energy gained by a hole between successive phonon scattering events, expressed as $E_{\text{SC}} = \lambda_{\text{mfp}}/F$. The z-directed mean free path, $\lambda_{\text{mfp}} = v_{\text{sat}}\tau_{\text{SC}}/2$, where v_{sat} is the saturation velocity and τ_{SC} is the scattering lifetime. This inequality states that phonon scattering events reduce the carrier energy and prevent them from directly gaining the ionization threshold energy between two successive scattering events. This makes it more difficult for the carriers to impact ionize. As a result, the carriers need to traverse several mean free paths to gain sufficient energy for ionization. The E_{SC} values of the five alloys at electric fields of 100 and 500 kV/cm are given in Table III. We extract τ_{SC} for an alloy by assuming a virtual crystal approximation of the component binary alloy scattering times. The τ_{SC} values for InAs, GaAs, AlAs, and AlSb are 0.08, 0.09, 0.08, and 0.11 ps, respectively [62]. A similar average is done for the ternary alloy saturation velocities. Because of unavailability of AlSb v_{sat}, InAs v_{sat} is used for (Al,In)(As,Sb) and AlAs v_{sat} for Al(As,Sb). InAs, GaAs, and AlAs v_{sat} values used are 5×10^4, 9×10^4, and 8×10^4 m/s, respectively [63].

TABLE III. The E_{SC} values at $F = 100$ kV/cm and $F = 500$ kV/cm, and E_{th} of the five alloys. For a material with equal conduction and valence band effective masses, considering parabolic bands, the threshold energy $E_{\text{th}} = 1.5E_G$ [60]. The same assumption is made here for the fifth inequality, as this is standard practice in the APD literature.

Material	E_{SC} (eV) at 100 kV/cm	E_{SC} (eV) at 500 kV/cm	E_{th} (eV)
(In,Ga)As	0.029	0.149	0.95
(Al,Ga)As	0.036	0.181	3.91
(In,Al)As	0.028	0.138	1.85
(Al,In)(As,Sb)	0.024	0.119	1.79
Al(As,Sb)	0.038	0.19	2.4

In order to validate these inequalities as design criteria, we apply them to the set of digital alloys mentioned in this paper. We consider a high electric field of 1 MV/cm for inequalities (3) and (4). The values of the left sides of the inequalities for the five alloys—(In,Ga)As, (Al,Ga)As, (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb)—are given in the first four columns of Table IV, while the measured k is provided as reference in column 6. The table cells are colored green or red. Green cells aid in noise suppression (left sides of the inequalities are relatively small) and red is detrimental to reducing noise (left sides larger and corresponding inequalities not satisfied). Additionally, the color intensities highlight the strength of that inequality (how far the left side is from equality with the right side). A lighter shade represents a smaller impact, while a darker shade means that the condition has a greater effect on the impact ionization noise. For example, in the case of (In,Ga)As, inequality (1) is shaded light green, which means that

TABLE IV. Suitability of digital alloys for attaining low noise is judged using the proposed inequalities. Here, the color green means beneficial for low noise and red indicates it is detrimental. The impact of the inequality in determining the experimentally determined ionization coefficient ratio k of the material is depicted by the color shades. A darker shade indicates that the inequality has a greater impact on the value of k. The experimental random alloy k values of the five alloys are given in the last column.

Material	Inequality 1	Inequality 2	Inequality 3	Inequality 4	k (digital alloy)	k (random alloy)
(In,Ga)As	0.38	1.08	0.88	0.006	0.3[18]	0.5[18]
(Al,Ga)As	1	∞	1	7.2×10^{-4}	0.1[19]	0.2[19]
(In,Al)As	0.16	0.33	0.17	5.6×10^{-4}	0.05[15]	0.2[15]
(Al,In)(As,Sb)	0.17	0.59	0.53	7.9×10^{-7}	0.01[16]	0.018[23]
Al(As,Sb)	0.22	0.45	0.3	3.4×10^{-7}	0.005[17]	0.05[61]

it does not effect noise performance significantly. However, the remaining inequalities for (In,Ga)As are shaded dark red, indicating their key role in the high noise and hence high k of (In,Ga)As. The inequalities for (Al,Ga)As, which has a slightly lower k, have a lighter shade of red. There are no minigaps for (Al,Ga)As in the light-hole band. There is a minigap in the SO band of (Al,Ga)As that is very deep in the valence band and there are other available states at that energy. Thus, holes can gain sufficient momentum to jump to other bands and bypass the minigap. So, we consider $\Delta E_m = 0$ for it. We accordingly expect that (Al,Ga)As has a lower noise. However, since the LH effective mass for (Al,Ga)As is greater than (In,Ga)As, it has lower hole impact ionization and thus lower noise compared to (In,Ga)As. The remaining alloys have significantly lower noise compared to these two.

The cells for (In,Al)As, (Al,In)(As,Sb), and Al(As,Sb) are all green. This means that these three alloys are quite favorable for attaining low excess performance. (In,Al)As has a minigap size $\Delta E_m = 0.12$ eV, which is larger than its optical phonon energy. It also has a large LH effective mass that prevents quantum tunneling across the minigap, as well as the LH-SO separation ΔE_{LS} that is comparable to that of (Al,Ga)As and (In,Ga)As. (Al,In)(As,Sb) has a low value for inequality (1), so that cell is shaded dark green. However, for inequalities (2) and (3), the values for (Al,In)(As,Sb) are higher than those of (In,Al)As and are thus shaded in a lighter color. (Al,In)(As,Sb) has a larger LH-SO separation than (In,Al)As and hence its inequality (4) has a darker shade. For In Al(As,Sb), the values for inequalities (1)–(3) have medium shades as they lie between the maximum and minimum values in each of these columns for the corresponding inequalities. However, Al(As,Sb) has a large $\Delta E_{\mathrm{LS}} = 0.54$ eV, so its inequality-(4) cell is shaded dark green. The (Al,In)(As,Sb) and Al(As,Sb) minigaps are larger than the optical phonon energies and have favorable locations away from the valence band edge. Thus, these gaps have secondary contributions in limiting hole impact ionization in these Sb-containing materials. Based on the inequality values, it would seem that (In,Al)As would have the lowest noise since it has the darkest shades. However, looking at the inequality-(4) values for these three materials we can infer that the LH-SO separation plays a critical role in reducing noise. Here, Al(As,Sb) has the lowest $k = 0.005$ and also the largest ΔE_{LS}. On the contrary, (In,Al)As has the highest $k = 0.1$ and the smallest ΔE_{LS}. Finally, the fifth inequality, Eq. (12), discussed in the context of split-off states is trivially satisfied by all five studied alloys. While important, it is thus not tabulated here, as it does not alter the status quo.

The random alloy k values are given in final column to provide comparison with the digital alloy k values. In general, the random alloy values are higher than the digital alloy values, but there are some subtleties that should be pointed out. The (Al,In)(As,Sb) random alloy value is very close to the digital alloy value. The small discrepancy might be within the experimental error limits. Moreover, the thickness of the multiplication region in the Al(As,Sb) digital APD is 1550 nm [17] and is 250 nm for the random alloy counterpart [61]. In the thinner Al(As,Sb) random alloy the larger k value can be due to the presence of higher electric fields where α and β start to converge. Data for thick random alloy Al(As,Sb) APD are not available as it is very difficult to grow good quality thick random alloy Al(As,Sb) structures [64]. Also, the period thickness in the Al(As,Sb) digital alloy is very small (about 1.3 nm). This can introduce some randomness in the alloy composition.

In short, the values of the inequalities in Table IV give a fairly good understanding of the excess noise performance of the set of digital alloys considered in this paper. They can potentially serve as empirical design criteria for judging alternative digital alloys in consideration as potential material candidates for digital alloy superlattice APDs.

V. CONCLUSION

In this paper, we have studied the digital alloy valence band carrier transport using NEGF and BTE formalisms. Based on our simulation results, we explain how minigaps and LH/SO offset impede hole impact ionization in APDs and improve their excess noise performance. When these gaps and offsets are sufficiently large, they cannot be bridged across by quantum tunneling or phonon scattering processes. Furthermore, we propose five inequalities as empirical design criteria for digital alloys with low-noise performance capabilities. Material parameters calculated computationally are used as inputs for these. We validate these criteria by explaining the excess noise performance of several experimentally fabricated digital alloy APDs. The design criteria can be used to computationally design alternative digital alloy structures and benchmark them before actually fabricating these.

ACKNOWLEDGMENTS

This work is funded by the National Science Foundation under Grant No. 1936016. The authors thank Dr. John P. David of the University of Sheffield and Dr. Seth R. Bank of the University of Texas-Austin for important discussions and insights. The calculations were carried out using the University of Virginia (Rivanna) High-Performance Computing systems, and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation under Grant No. ACI-1548562.

[1] A. Tosi, N. Calandri, M. Sanzaro, and F. Acerbi, Low-noise, low-jitter, high detection efficiency ingaas/inp

single-photon avalanche diode, IEEE J. Sel. Top. Quantum Electron. **20**, 192 (2014).

[2] J. C. Campbell, in *Optical Fiber Telecommunications V A (Fifth Edition)*, Optics and Photonics, edited by I. P. Kaminow, T. Li, and A. E. Willner (Academic Press, Burlington, 2008), p. 221, fifth edition ed..

[3] N. Bertone and W. Clark, Avalanche photodiode arrays provide versatility in ultrasensitive applications, Laser Focus World **43** (2007).

[4] P. Mitra, J. D. Beck, M. R. Skokan, J. E. Robinson, J. Antoszewski, K. J. Winchester, A. J. Keating, T. Nguyen, K. K. M. B. D. Silva, C. A. Musca, J. M. Dell, and L. Faraone, in *Intelligent Integrated Microsystems*, Vol. 6232, edited by R. A. Athale and J. C. Zolper: International Society for Optics and Photonics (SPIE, 2006), p. 70.

[5] G. M. Williams, Optimization of eyesafe avalanche photodiode lidar for automobile safety and autonomous navigation systems, Opt. Eng. **56**, 1 (2017).

[6] M. Nada, F. Nakajima, T. Yoshimatsu, Y. Nakanishi, S. Tatsumi, Y. Yamada, K. Sano, and H. Matsuzaki, High-speed iii-v based avalanche photodiodes for optical communications-the forefront and expanding applications, Appl. Phys. Lett. **116**, 140502 (2020).

[7] K. Pasquinelli, R. Lussana, S. Tisa, F. Villa, and F. Zappa, Single-photon detectors modeling and selection criteria for high-background lidar, IEEE Sens. J. **20**, 7021 (2020).

[8] S. Li, L. Da Xu, and S. Zhao, 5g internet of things: A survey, J. Ind. Inf. Integr. **10**, 1 (2018).

[9] M. Z. Chowdhury, M. K. Hasan, M. Shahjalal, E. B. Shin, and Y. M. Jang, in *2019 ICAIIC* (IEEE, 2019), p. 004.

[10] A. Y. Liu and J. Bowers, Photonic integration with epitaxial iii–v on silicon, IEEE J. Sel. Top. Quantum Electron. **24**, 1 (2018).

[11] J. C. Campbell, Recent advances in avalanche photodiodes, J. Lightwave Technol. **34**, 278 (2016).

[12] R. McIntyre, Multiplication noise in uniform avalanche diodes, IEEE Trans. Electron Devices **ED-13**, 164 (1966).

[13] M. Teich, K. Matsuo, and B. Saleh, Excess noise factors for conventional and superlattice avalanche photodiodes and photomultiplier tubes, IEEE J. Quantum Electron. **22**, 1184 (1986).

[14] N. Z. Hakim, B. E. A. Saleh, and M. C. Teich, Generalized excess noise factor for avalanche photodiodes of arbitrary structure, IEEE Trans. Electron Devices **37**, 599 (1990).

[15] J. Zheng, Y. Tan, Y. Yuan, A. Ghosh, and J. Campbell, Strain effect on band structure of inalas digital alloy, J. Appl. Phys. **125**, 082514 (2019).

[16] S. R. Bank, J. C. Campbell, S. J. Maddox, A. K. Rockwell, M. E. Woodson, M. Ren, A. Jones, S. March, J. Zheng, and Y. Yuan, in *2018 IEEE RAPID* (IEEE, 2018), p. 1.

[17] X. Yi, S. Xie, B. Liang, L. W. Lim, J. S. Cheong, M. C. Debnath, D. L. Huffaker, C. H. Tan, and J. P. David, Extremely low excess noise and high sensitivity alas$_{0.56}$sb$_{0.44}$ avalanche photodiodes, Nat. Photonics **13**, 683 (2019).

[18] A. Rockwell, M. Ren, M. Woodson, A. Jones, S. March, Y. Tan, Y. Yuan, Y. Sun, R. Hool, and S. Maddox *et al.*, Toward deterministic construction of low noise avalanche photodetector materials, Appl. Phys. Lett. **113**, 102106 (2018).

[19] Y. Yuan, J. Zheng, Y. Tan, Y. Peng, A.-K. Rockwell, S. R. Bank, A. Ghosh, and J. C. Campbell, Temperature dependence of the ionization coefficients of inalas and algaas digital alloys, Photonics Res. **6**, 794 (2018).

[20] J. Zheng, S. Z. Ahmed, Y. Yuan, A. Jones, Y. Tan, A. K. Rockwell, S. D. March, S. R. Bank, A. W. Ghosh, and J. C. Campbell, Full band monte carlo simulation of alinassb digital alloys, InfoMat **2**, 1236 (2020).

[21] J. Zheng, Y. Yuan, Y. Tan, Y. Peng, A. Rockwell, S. R. Bank, A. W. Ghosh, and J. C. Campbell, Simulations for inalas digital alloy avalanche photodiodes, Appl. Phys. Lett. **115**, 171106 (2019).

[22] J. Zheng, Y. Yuan, Y. Tan, Y. Peng, A. K. Rockwell, S. R. Bank, A. W. Ghosh, and J. C. Campbell, Digital alloy inalas avalanche photodiodes, J. Lightwave Technol. **36**, 3580 (2018).

[23] S. H. Kodati, S. Lee, B. Guo, A. H. Jones, M. Schwartz, M. Winslow, N. A. Pfiester, C. H. Grein, T. J. Ronningen, J. C. Campbell, and S. Krishna, Alinassb avalanche photodiodes on inp substrates, Appl. Phys. Lett. **118**, 091101 (2021).

[24] J. Taylor-Mew, V. Shulyak, B. White, C. H. Tan, and J. S. Ng, Low excess noise of al$_0$.85ga$_{0.15}$as$_{0.56}$sb$_{0.44}$ avalanche photodiode from pure electron injection, IEEE Photonics Technol. Lett. **33**, 1155 (2021).

[25] İ. H. Oğuzman, Y. Wang, J. Kolník, and K. F. Brennan, Theoretical study of hole initiated impact ionization in bulk silicon and gaas using a wave-vector-dependent numerical transition rate formulation within an ensemble monte carlo calculation, J. Appl. Phys. **77**, 225 (1995).

[26] F. Osaka, T. Mikawa, and O. Wada, Analysis of impact ionization phenomena in inp by monte carlo simulation, Jpn. J. Appl. Phys. **25**, 394 (1986).

[27] Y. Liu, X. Yi, N. J. Bailey, Z. Zhou, T. B. Rockett, L. W. Lim, C. H. Tan, R. D. Richards, and J. P. David, Valence band engineering of gaasbi for low noise avalanche photodiodes, Nat. Commun. **12**, 1 (2021).

[28] Y. Tan, M. Povolotskyi, T. Kubis, T. B. Boykin, and G. Klimeck, Transferable tight-binding model for strained group iv and iii-v materials and heterostructures, Phys. Rev. B **94**, 045311 (2016).

[29] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. Scott, and N. Wilkins-Diehr, Xsede: Accelerating scientific discovery, Comput. Sci. Eng. **16**, 62 (2014).

[30] Y. P. Tan, M. Povolotskyi, T. Kubis, T. B. Boykin, and G. Klimeck, Tight-binding analysis of si and gaas ultrathin bodies with subatomic wave-function resolution, Phys. Rev. B **92**, 085301 (2015).

[31] D. Kienle, J. I. Cerda, and A. W. Ghosh, Extended hückel theory for band structure, chemistry, and transport. i. carbon nanotubes, J. Appl. Phys. **100**, 043714 (2006).

[32] D. Kienle, K. H. Bevan, G.-C. Liang, L. Siddiqui, J. I. Cerda, and A. W. Ghosh, Extended hückel theory for band structure, chemistry, and transport. ii. silicon, J. Appl. Phys. **100**, 043715 (2006).

[33] J. Heyd, G. E. Scuseria, and M. Ernzerhof, Hybrid functionals based on a screened coulomb potential, J. Chem. Phys. **118**, 8207 (2003).

[34] S. Z. Ahmed, Y. Tan, D. S. Truesdell, B. H. Calhoun, and A. W. Ghosh, Modeling tunnel field effect transistors–from interface chemistry to non- idealities to circuit level performance, J. Appl. Phys. **124**, 154503 (2018).

[35] Y. Tan, F. W. Chen, and A. W. Ghosh, First principles study and empirical parametrization of twisted bilayer mos 2 based on band-unfolding, Appl. Phys. Lett. **109**, 101601 (2016).

[36] T. B. Boykin, N. Kharche, and G. Klimeck, Brillouin-zone unfolding of perfect supercells having nonequivalent primitive cells illustrated with a si/ ge tight-binding parameterization, Phys. Rev. B **76**, 035310 (2007).

[37] T. B. Boykin, N. Kharche, G. Klimeck, and M. Korkusinski, Approximate bandstructures of semiconductor alloys from tight-binding supercell calculations, J. Phys.: Condens. Matter **19**, 036203 (2007).

[38] J. A. Støvneng and P. Lipavský, Multiband tight-binding approach to tunneling in semiconductor heterostructures: Application to Γx transfer in gaas, Phys. Rev. B **49**, 16494 (1994).

[39] B. Stickler, Theory and modeling of spin-transport on the microscopic and the mesoscopic scale (na, 2013).

[40] S. Datta, Nanoscale device modeling: The green's function method, Superlattices Microstruct. **28**, 253 (2000).

[41] C. Lee, R. Logan, R. Batdorf, J. Kleimack, and W. Wiegmann, Ionization rates of holes and electrons in silicon, Phys. Rev. **134**, A761 (1964).

[42] J. Conradi, The distribution of gains in uniformly multiplying avalanche photodiodes: Experimental, IEEE Trans. Electron Devices **19**, 713 (1972).

[43] W. Grant, Electron and hole ionization rates in epitaxial silicon at high electric fields, Solid-State Electron. **16**, 1189 (1973).

[44] T. Kaneda, H. Matsumoto, and T. Yamaoka, A model for reach-through avalanche photodiodes (rapd's), J. Appl. Phys. **47**, 3135 (1976).

[45] A. Marshall, C. Tan, M. Steer, and J. David, Electron dominated impact ionization and avalanche gain characteristics in inas photodiodes, Appl. Phys. Lett. **93**, 111107 (2008).

[46] A. R. Marshall, P. J. Ker, A. Krysa, J. P. David, and C. H. Tan, High speed inas electron avalanche photodiodes overcome the conventional gain-bandwidth product limit, Opt. Express **19**, 23341 (2011).

[47] W. Sun, S. J. Maddox, S. R. Bank, and J. C. Campbell, in *72nd Device Research Conference* (IEEE, 2014), p. 47.

[48] W. Sun, Z. Lu, X. Zheng, J. C. Campbell, S. J. Maddox, H. P. Nair, and S. R. Bank, High-gain inas avalanche photodiodes, IEEE J. Quantum Electron. **49**, 154 (2012).

[49] P. J. Ker, A. R. Marshall, A. B. Krysa, J. P. David, and C. H. Tan, in *2012 17th Opto-Electronics and Communications Conference* (IEEE, 2012), p. 220.

[50] J. D. Beck, C.-F. Wan, M. A. Kinch, and J. E. Robinson, in *Materials for Infrared Detectors*, (International Society for Optics and Photonics, 2001), Vol. 4454, p. 188.

[51] J. D. Beck, C.-F. Wan, M. A. Kinch, J. E. Robinson, P. Mitra, R. E. Scritchfield, F. Ma, and J. C. Campbell, in *Infrared Detector Materials and Devices*, (International Society for Optics and Photonics, 2004), Vol. 5564, p. 44.

[52] Yu Ling Goh, Jo Shien Ng, Chee Hing Tan, W. K. Ng, and J. P. R. David, Excess noise measurement in $\mathrm{in}_{0.53}\mathrm{ga}_{0.47}\mathrm{as}$, IEEE Photonics Technol. Lett. **17**, 2412 (2005).

[53] C. Lenox, P. Yuan, H. Nie, O. Baklenov, C. Hansing, J. Campbell, A. Holmes Jr, and B. Streetman, Thin multiplication region inalas homojunction avalanche photodiodes, Appl. Phys. Lett. **73**, 783 (1998).

[54] S. Z. Ahmed, Y. Tan, J. Zheng, J. C. Campbell, and A. W. Ghosh, in *2018 IEEE Photonics Conference (IPC)* (IEEE, 2018), p. 1.

[55] W. Wang, J. Yao, J. Wang, Z. Deng, Z. Xie, J. Huang, H. Lu, and B. Chen, Characteristics of thin inalas digital alloy avalanche photodiodes, Opt. Lett. **46**, 3841 (2021).

[56] J. Piprek, *Semiconductor Optoelectronic Devices: Introduction to Physics and Simulation* (Elsevier, San Diego, USA, 2013).

[57] M. S. Shur, *Handbook Series on Semiconductor Parameters* (World Scientific, Singapore, 1996), Vol. 1.

[58] D. Caughey and R. Thomas, Carrier mobilities in silicon empirically related to doping and field, Proc. IEEE **55**, 2192 (1967).

[59] W. Fawcett, A. Boardman, and S. Swain, Monte carlo determination of electron transport properties in gallium arsenide, J. Phys. Chem. Solids **31**, 1963 (1970).

[60] B. K. Ridley, *Quantum Processes in Semiconductors* (Oxford university press, Oxford, UK, 2013).

[61] J. Xie, S. Xie, R. C. Tozer, and C. H. Tan, Excess noise characteristics of thin alassb apds, IEEE Trans. Electron Devices **59**, 1475 (2012).

[62] Q. L. Yang, H. X. Deng, S. H. Wei, S. S. Li, and J. W. Luo, Materials design principles towards high hole mobility learning from an abnormally low hole mobility of silicon, arXiv:2011.02262 [cond-mat.mtrl-sci] (2020).

[63] V. Palankovski, Simulation of heterojunction bipolar transistors (2000).

[64] X. Yi, Ph.D. thesis, University of Sheffield, 2020.