Correct implementation of polarization constants in wurtzite materials and impact on III-nitrides

Cyrus E. Dreyer,[1, 2] Anderson Janotti,[1, *] Chris G. Van de Walle,[1] and David Vanderbilt[2]

[1]*Materials Department, University of California, Santa Barbara, CA 93106-5050*

[2]*Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08845-0849*

(Dated: December 22, 2015)

Accurate values for polarization discontinuities between pyroelectric materials are critical for understanding and designing the electronic properties of heterostructures. For wurtzite materials, the zincblende structure has been used in the literature as a reference to determine the effective spontaneous polarization constants. We show that, because the zincblende structure has a nonzero formal polarization, this method results in a spurious contribution to the spontaneous polarization differences between materials. In addition, we address the correct choice of "improper" versus "proper" piezoelectric constants. For the technologically important III-nitride materials GaN, AlN, and InN, we determine polarization discontinuities using a consistent reference based on the layered hexagonal structure and the correct choice of piezoelectric constants, and discuss the results in light of available experimental data.

I. INTRODUCTION

Pyroelectric materials have emerged in a variety of electronic and optoelectronic applications. Because of the symmetry of their crystal structure these materials exhibit spontaneous (SP) and piezoelectric (PZ) dipole moments [1] which manifest themselves as electric fields in heterostructure layers and sheet charges at interfaces. In the technologically important III-nitrides, which have the wurtzite (WZ) structure (space group $P6_3mc$), polarization differences allow for strong carrier confinement and the formation of a two-dimensional electron gas (2DEG) with high density at AlGaN/GaN interfaces, exploited in high electron mobility transistors (HEMTs). The effect of polarization can also be detrimental, for example causing the quantum-confined Stark effect in quantum wells of light-emitting diodes (LEDs), which reduces radiative recombination rates and shifts the emission wavelength. For both HEMTs and LEDs, accurate values of the SP and PZ polarization constants are required for a fundamental understanding as well as for device design.

Since experimental determination of the separate SP and PZ contributions to the total polarization is very difficult, calculated values of SP and PZ polarization constants are widely used in simulations. The PZ polarization constants are, in principle, fairly straightforward to explicitly measure or calculate [2]. However, the reported values exhibit a considerable spread [3]. In addition, the difference between so-called "proper" and "improper" PZ constants [2, 4] is often overlooked, even though it can give rise to significant quantitative changes in the resulting polarization fields. This difference is one issue that is elucidated in the present paper.

The definition of SP polarization constants is even more subtle, and they are typically not amenable to explicit experimental determination. The calculation of SP polarization requires the choice of a reference structure, which in the case of WZ semiconductors, has invariably been chosen to be zincblende (ZB) [5–7]. In this work we will show that though ZB as a reference structure is intuitively appealing, the SP polarization constants that result have been misinterpreted, introducing a source of error into the predicted values for bound sheet charge densities (and polarization fields). We also demonstrate that a proper choice of reference structure can eliminate these problems, and we provide revised values that can be directly inserted in current simulation tools.

While our theoretical considerations are general, we choose the nitride semiconductors because they provide a suitable example to illustrate the derivations, and because our findings have a significant impact on this materials system of high (and still increasing) technological importance. In Sec. II we review the underlying theory. In Sec. III we address the problems with choosing zincblende as a reference structure, and propose a solution. Sec. IV deals with piezoelectric contributions, specifically the issue of proper versus improper constants. In Sec. V we show that our findings have important consequences for nitride device structures and compare with previous implementations and with experiment. Section VI concludes the paper.

II. CALCULATING POLARIZATION CONSTANTS IN WURTZITE

For WZ films grown in the [0001] direction (*i.e.*, the $+c$ direction), the polarization component P_3 is given by the sum of the SP polarization at the wurtzite material's own lattice parameters, P_{SP}, and the z component of the

———

* Current address: Materials Science and Engineering, University of Delaware, Newark, Delaware 19716-1501, USA

PZ polarization [1]. That is, for material m,

$$P_3^m = P_{\mathrm{SP}}^m + (\varepsilon_1 + \varepsilon_2)e_{31}^m + \varepsilon_3 e_{33}^m, \qquad (1)$$

where (in Voigt notation) ε_i (i=1,2,3) is the strain in the i direction and e_{3i} are the corresponding piezoelectric constants (specifically, the "improper" ones; see Section IV). Henceforth we drop the subscript "3" from P for simplicity; all unbolded quantities pertaining to wurtzite are assumed to be in the c direction.

A. The modern theory of polarization

Direct calculation of the polarization constants in Eq. (1) by first-principles electronic-structure methods was enabled by the formulation of a rigorous theory of bulk polarization, known as the modern theory of polarization (MTP) [8, 9]. For a given structure λ, the MTP allows calculation of the so-called "formal" polarization [8]:

$$\mathbf{P}_{\mathrm{f}} = \mathbf{P}_{\mathrm{ion}} + \mathbf{P}_{\mathrm{el}} \qquad (2)$$
$$= \frac{e}{\Omega} \sum_s Z_s^{\mathrm{ion}} \mathbf{R}_s^{(\lambda)} + \frac{ief}{8\pi^3} \sum_j^{\mathrm{occ}} \int_{\mathrm{BZ}} d\mathbf{k} \langle u_{j,\mathbf{k}}^{(\lambda)} | \nabla_{\mathbf{k}} | u_{j,\mathbf{k}}^{(\lambda)} \rangle,$$

where Ω is the cell volume, Z_s^{ion} is the charge of the ion s and $\mathbf{R}_s^{(\lambda)}$ is its position in the λ structure, f is the spin degeneracy of the bands, the sum j runs over occupied bands, and $u_{j,\mathbf{k}}^{(\lambda)}$ are the cell periodic parts of the Bloch wave functions. The formal polarization is defined only modulo the "quantum of polarization" $e\mathbf{R}/\Omega$, where $\mathbf{R}$ is any lattice constant and e is the electron charge [8, 9].

In the MTP, only *differences* between formal polarizations of appropriate structures, $\lambda = 0$ and $\lambda = 1$, are well defined:

$$\Delta \mathbf{P} = \mathbf{P}_{\mathrm{f}}^{(\lambda=1)} - \mathbf{P}_{\mathrm{f}}^{(\lambda=0)}. \qquad (3)$$

The choice of the "appropriate" structures $\lambda = 0$ and $\lambda = 1$ rests on one of two possible considerations to ensure that physical conclusions can be drawn from their formal polarization differences. First, if the two structures are connected by an adiabatic, gap-preserving deformation path [8, 9], then their difference in polarization [$\Delta \mathbf{P}$ in Eq. (3)] is given by the expression

$$\Delta \mathbf{P} = \int_0^1 d\lambda \frac{\partial \mathbf{P}}{\partial \lambda}, \qquad (4)$$

and corresponds to the zero-field adiabatic displacement current. This quantity can, in principle, be determined experimentally. An obvious application is the calculation of piezoelectric constants, which involves polarization differences between structures with slightly different lattice constants and/or internal structural parameters.

B. Interface theorem

The second consideration, as shown by Vanderbilt and King-Smith [10], is that *if* an insulating interface can be constructed between two structures, the difference in formal polarization gives the bound charge, σ_{b}, that builds up at the interface as a result of the continuity of the displacement field over an interface with no free charge:

$$\sigma_{\mathrm{b}} = (\mathbf{P}_{\mathrm{f}}^{\lambda=1} - \mathbf{P}_{\mathrm{f}}^{\lambda=0}) \cdot \hat{\mathbf{n}}. \qquad (5)$$

This is often referred to as the "interface theorem." Since there is no adiabatic path necessary between the two structures in this consideration, $\lambda = 0$ and $\lambda = 1$ can be different polymorphs of the same material (such as WZ and ZB structures of GaN) or different materials altogether (such as GaN and AlN); as long as they form an insulating interface, Eq. (5) will give the bound charge accumulation at the interface.

From the interface theorem [Eq. (5)] and Eq. (1), the bound polarization charge at the interface between different III-nitride materials (m and n) is

$$\sigma_{\mathrm{b}} = [P_{\mathrm{SP}}^m + e_{31}^m(\varepsilon_1^m + \varepsilon_2^m) + e_{33}^m \varepsilon_3^m]$$
$$- [P_{\mathrm{SP}}^n + e_{31}^n(\varepsilon_1^n + \varepsilon_2^n) + e_{33}^n \varepsilon_3^n]. \qquad (6)$$

As an example, we will take a realistic situation that occurs in heterostructures, by assuming that material n is strained coherently to m ($\epsilon_1^m = \epsilon_2^m = \epsilon_3^m = 0$), *i.e.*, under plane stress ($\epsilon_1^n = \epsilon_2^n, \epsilon_3^n = -2C_{13}^n/C_{33}^n\epsilon_1^n$, where C_{ij} are the elastic constants). Therefore we have

$$\sigma_{\mathrm{b}} = (P_{\mathrm{SP}}^m - P_{\mathrm{SP}}^n) - 2\epsilon_1^n(e_{31}^n - e_{33}^n C_{13}^n/C_{33}^n)$$
$$= \Delta P_{\mathrm{SP}}^{\mathrm{int}} - 2\epsilon_1^n(e_{31}^n - e_{33}^n C_{13}^n/C_{33}^n). \qquad (7)$$

Note that σ_{b} is the charge density of electrons at an interface for which material n has been grown on top of material m in the $+c$ direction.

III. REFERENCE STRUCTURE FOR SPONTANEOUS POLARIZATION

A. Effective spontaneous polarization constants

We will first address the difference in spontaneous polarization in Eq. (7), $\Delta P_{\mathrm{SP}}^{\mathrm{int}}$. Strain effects will be taken into account separately in the PZ part, so that $\Delta P_{\mathrm{SP}}^{\mathrm{int}}$ is simply the difference of formal polarizations of the respective zero-strain structures,

$$\Delta P_{\mathrm{SP}}^{\mathrm{int}} = P_{\mathrm{f}}^m\big|_{\varepsilon=0} - P_{\mathrm{f}}^n\big|_{\varepsilon=0}. \qquad (8)$$

For purposes of Eq. (1), we would like to define a SP polarization constant that is a property of a single material. Simply taking P_{f}^m of Eq. (8) as P_{SP}^m is problematic, since formal polarization is multivalued, being only well-defined modulo a quantum of polarization $e\mathbf{R}/\Omega$. Therefore, in every situation in which Eq. (1) is applied to

determine σ_b at an interface, it must be confirmed that formal polarizations of the two materials are taken on the same "branch" of $e\mathbf{R}/\Omega$. A better approach is to take P_{SP}^m in Eq. (1) as a so-called "effective" SP polarization, $\mathbf{P}_{\mathrm{eff}}$, defined by Resta and Vanderbilt [11] to be the $\Delta\mathbf{P}$ in Eq. (3) that results as the system is taken from a high-symmetry "reference" structure ($\lambda = 0$) to the structure of interest ($\lambda = 1$). That is,

$$\mathbf{P}_{\mathrm{eff}} = \mathbf{P}_{\mathrm{f}}^{(\lambda=1)} - \mathbf{P}_{\mathrm{f}}^{(\lambda=0)} = \mathbf{P}_{\mathrm{f}} - \mathbf{P}_{\mathrm{f}}^{\mathrm{ref}}. \qquad (9)$$

Using $\mathbf{P}_{\mathrm{eff}}$ to define the SP polarization of the material removes the indeterminacy inherent to the formal polarization.

The reference structure is often chosen to be centrosymmetric, but it is important to recognize that the formal polarization of centrosymmetric crystals is not necessarily zero. This is because, as stated above, $\mathbf{P}_{\mathrm{f}}$ is a multivalued vector field, so it is possible for a nonzero formal polarization to be unchanged (modulo $e\mathbf{R}/\Omega$) under the inversion operator. Nevertheless, high symmetry puts restrictions on the possible values of $\mathbf{P}_{\mathrm{f}}^{(\lambda=0)}$ [10].

While in principle effective polarization constants are still *differences* in formal polarization between $\lambda = 1$ and $\lambda = 0$ (reference) structures, in practice they can be used to compare spontaneous polarizations of different materials to obtain ΔP_{SP} if such materials share a reference structure with the same formal polarization. Such a comparison then correctly yields the interface charge density according to the interface theorem of Ref. 10. In such cases, ΔP_{SP} is just given by the difference in *effective* SP polarization of the materials,

$$\Delta\widetilde{P}_{\mathrm{SP}}^{\mathrm{int}} = P_{\mathrm{eff}}^m - P_{\mathrm{eff}}^n. \qquad (10)$$

In the more general case that the reference formal polarizations do not match, the correct change in SP polarization following from Eqs. (8) and (9) is

$$\Delta P_{\mathrm{SP}}^{\mathrm{int}} = \Delta\widetilde{P}_{\mathrm{SP}}^{\mathrm{int}} + (P_{\mathrm{f}}^{m,\mathrm{ref}} - P_{\mathrm{f}}^{n,\mathrm{ref}}). \qquad (11)$$

That is, a correction term of the form

$$\Delta P_{\mathrm{corr}}^{\mathrm{ref}} \equiv P_{\mathrm{f}}^{m,\mathrm{ref}} - P_{\mathrm{f}}^{n,\mathrm{ref}} \qquad (12)$$

has to be added to Eq. (10). Unfortunately, this correction term is not typically implemented in device simulation packages (*e.g.*, Ref. 12) or used in the interpretation of experimental data (*e.g.*, Ref. 13, which is considered a standard reference in the field).

When the PZ terms are included as well, the total interface charge given by Eq. (7) becomes

$$\sigma_{\mathrm{b}} = \Delta\widetilde{P}_{\mathrm{SP}}^{\mathrm{int}} + \Delta P_{\mathrm{corr}}^{\mathrm{ref}} - 2\epsilon_1^n(e_{31}^n - e_{33}^n C_{13}^n/C_{33}^n). \qquad (13)$$

Equation (13) is a central result of the present work.

B. Correction term for the effective spontaneous polarization with the zincblende reference structure

As mentioned before, previous studies [5–7] have exclusively used ZB (space group $F\bar{4}3m$) as a reference structure for calculating the SP polarization of the WZ. This structure is not centrosymmetric, although it has sufficient symmetry to preclude any SP polarization [1]. The fact that an insulating (111) interface can be constructed between the WZ and ZB polytypes [7] makes it an appropriate reference structure.

However, there is a subtlety with using ZB as a reference structure: it has a nonzero formal polarization in the [111] direction, $P_{\mathrm{f}}^{\mathrm{ZB}}$ (modulo $e\mathbf{R}/\Omega$). Again, this is consistent with the symmetry considerations because P_{f} is a multivalued vector quantity, and can be nonzero while still remaining unchanged (modulo $e\mathbf{R}/\Omega$) under the $F\bar{4}3m$ symmetry operations. These symmetry operations dictate the possible values of $P_{\mathrm{f}}^{\mathrm{ZB}}$, and therefore the resulting value depends only on the lattice constant, not on the chemical species of the atoms [10]. The ZB reference structures for the reported effective SP polarization values for the III-nitrides were those with lattice constants equal to the in-plane lattice constant of the corresponding wurtzite material [5–7] (as confirmed by our calculations), so $P_{\mathrm{f}}^{\mathrm{ZB}}$ will be different for each material. Therefore, for the effective SP polarization constants with the ZB reference to be implemented in Eq. (7), the correction term of Eq. (12) is required, as in Eq. (13).

Consider the example of the interface charge between InN and GaN. Although, as mentioned above, the formal polarization of zincblende does not necessarily vanish, the symmetry of the structure severely limits the possible values. Specifically there are two possible values of the formal polarization in the [111] direction that are consistent with the symmetry: either P_{f} vanishes, or it is equal to $e\sqrt{3}a_n/2\Omega_n$ (both modulo ea_n/Ω_n), where a_n is the lattice constant of material n and Ω_n is the volume [10]. For the III-nitrides our calculations result in the latter, giving a correction term for GaN/InN:

$$\begin{aligned}
\Delta P_{\mathrm{corr}}^{(\mathrm{ZB\ ref})} &= P_{\mathrm{f}}^{\mathrm{GaN,ZB}} - P_{\mathrm{f}}^{\mathrm{InN,ZB}} \\
&= \frac{2e\sqrt{3}}{4}\left(\frac{1}{(a_{\mathrm{GaN}})^2} - \frac{1}{(a_{\mathrm{InN}})^2}\right) \qquad (14) \\
&= 0.28 \ \mathrm{C/m}^2.
\end{aligned}$$

When considering the SP polarization differences between WZ nitrides, this represents a significant correction. In fact, as we will show in Section III C, the correction is an order of magnitude larger than the effective polarizations when they are calculated with the zincblende reference [5, 6]. As we shall see later in Section V E, this error is substantially reduced in practice by an approximate error cancellation that occurs in connection with the treatment of the PZ response.

There is nothing intrinsically wrong with using ZB as the reference structure, as long as the $\Delta P_{\mathrm{corr}}^{\mathrm{ref}}$ term

[Eq. (12), or Eq. (14) for the example of GaN/InN] is explicitly included in expressions such as Eq. (11) or Eq. (13) . To our knowledge, however, this has not been properly implemented in the numerous previous evaluations of SP polarization for nitride interfaces, and it would require changes in the software for the many simulation tools that include modeling of polarization fields in heterostructures.

C. $P6_3/mmc$ hexagonal layered structure as an alternative reference

In order to avoid extensive changes in the simulation software, and to enhance physical insight, we advocate another approach, namely to determine effective SP polarization constants with respect to a reference structure for which the formal polarization is explicitly zero (so that $\Delta P_{\text{SP}}^{\text{int}} = \Delta \widetilde{P}_{\text{SP}}^{\text{int}}$). A straightforward choice for this reference structure is the layered hexagonal (H) structure (space group $P6_3/mmc$), as was used for hexagonal $P6_3mc\ ABC$ materials [14]. This structure is centrosymmetric, and we will show below with explicit first-principles calculations that it remains insulating and its formal polarization vanishes. The layered hexagonal structure can be obtained by an adiabatic (gap preserving) increase of the internal structural u parameter from $u \approx 0.37 - 0.38$ of the WZ structure to $u = 0.5$. All that is required to avoid correction terms like Eq. (14) is to replace the effective SP polarization constants currently used in the field (the ones referenced to ZB [5, 6]) with those referenced to the H reference structure. We have explicitly verified that this leads to expressions that are identical to those that would be obtained for the ZB reference, provided the second term in Eq. (11) or Eq. (13) is included.

The first-principles calculations of P_{f} for the H, WZ, and ZB structures of the III-nitrides were performed using density functional theory with the screened hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE) [15] as implemented in the VASP code [16]. Hartree-Fock mixing parameters of 31 % for AlN and GaN, and 25 % for InN were used to correctly describe the band gaps and structural parameters of each material. Conventional functionals based on the local density approximation (LDA) or generalized gradient approximation (GGA) predict InN to be a metal, precluding the calculation of the polarization constants if the Γ point is included in the k-point mesh (which is required in VASP). Projector augmented wave potentials (PAW) [17], with the In and Ga d electrons frozen in the core, were used. All calculations were performed on bulk primitive cells, with a $6 \times 6 \times 8$ Monkhorst-Pack [18] k-point mesh to sample the Brillouin zone, and a large energy cutoff of 600 eV for the plane-wave basis set, chosen to ensure convergence of the internal structural parameter u. The calculated lattice parameters and band gaps, listed in Section S1 of the supplemental material (SM) [19], show good agreement
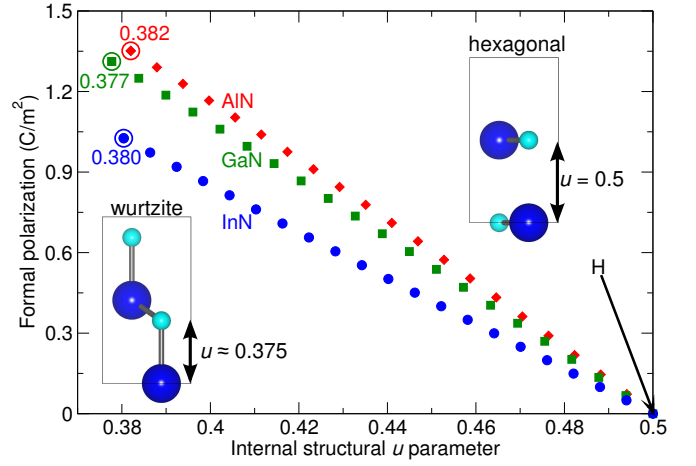

FIG. 1. Formal polarization of InN, GaN, and AlN for structures as a function of the internal structural parameter u, varying between fully relaxed WZ (circled symbols, labelled with relaxed u value) and H ($u = 0.5$), as shown schematically by ball and stick models where smaller balls represent N atoms and larger ones represent the cations. All other lattice parameters were fixed at their relaxed WZ values.

with experimental data.

We have calculated the electronic structure for structures with increasing u, ranging from $u \approx 0.37$ to $u = 0.5$ (Fig. 1), and confirmed that this path between WZ and H is gap preserving by checking the electronic structure for each calculation. These calculations also show that the formal polarization of the H structure is zero (modulo $e\mathbf{R}/\Omega$) for the III-nitrides (Fig. 1). We remind the reader that this was not guaranteed, since $\mathbf{P}_{\text{f}}$ can be nonzero and still consistent with inversion symmetry, if the inversion operator changes $\mathbf{P}_{\text{f}}$ by a multiple of $e\mathbf{R}/\Omega$. We have therefore verified that the hexagonal phase is a reference structure for which there is no spurious term in Eq. (11).

In addition, by correcting for any discontinuities (in the amount of a multiple of $e\mathbf{R}/\Omega$) that may occur in the calculations of formal polarizations along the path between WZ and H, we have insured that we are comparing formal polarizations of WZ GaN, AlN, InN on the same branch of $e\mathbf{R}/\Omega$ [11].

The calculated spontaneous polarization coefficients for the WZ structure using either H or ZB as a reference are given in Table I. The results obtained by Bernardini *et al.* [5, 6] are listed for comparison. The GGA functional used in that work provides results that are very close to those we obtained with HSE; the discrepancy is the largest for InN, which is probably related to the fact, mentioned above, that GGA predicts InN to be a metal. Table I also shows, however, that the choice of reference structure makes a significant difference. The magnitudes of the coefficients are much larger, and their signs are different when H is used as the reference. We observe that it is not just the absolute values, but also the relative differences between the calculated polarization constants of the three materials that differ from the

TABLE I. Effective spontaneous polarization constants in units of C/m^2 of wurtzite (WZ) GaN, AlN, and InN calculated using either the hexagonal (H, space group $P6_3/mmc$) or zincblende (ZB, space group $F\bar{4}3m$) reference structures. The lattice constant of the ZB structure is chosen to match the in-plane lattice constant of the WZ structure for the same material. Results from previous calculations [6] that used the ZB reference are listed for comparison.

	$P_{\text{eff}}^{(\text{H ref})}$	$P_{\text{eff}}^{(\text{ZB ref})}$	$P_{\text{eff}}^{(\text{ZB ref})}$, prev.[a]
GaN	1.312	−0.035	−0.034
AlN	1.351	−0.090	−0.090
InN	1.026	−0.053	−0.042

[a] From Ref. 6.

previously reported values [5, 6].

The difference in sign of $P_{\text{eff}}^{(\text{H ref})}$ compared to $P_{\text{eff}}^{(\text{ZB ref})}$ demonstrates that the conventional wisdom that the SP polarization in WZ points in the -c direction is misleading. The formal polarization of WZ has no definite sign as this would depend on the chosen branch. The effective SP depends on the polarization difference, and therefore the sign will depend on the sign and magnitude of the formal polarization of the reference structure.

IV. IMPROPER VERSUS PROPER PIEZOELECTRIC CONSTANTS

We now address the specifics of the PZ terms in Eqs. (7) and (13). A complication that must be addressed is the choice between improper and proper e_{31} (e_{33} has no such complication) [2, 4].

As we have done above, consider a thin layer of a WZ material grown in the c direction. If the layer is strained perpendicular to the c direction, the total bound charge on the $+c$ and $-c$ surfaces will change as a result of the polarization current, or redistribution of charge, in the layer. If metallic contacts on the $+c$ and $-c$ surfaces are short-circuited when the strain occurs, the current flow can be measured directly and will give the proper PZ constant, denoted e_{31}^{prop} [2, 4].

If the $+c$ and $-c$ faces are in open-circuit boundary conditions, the layer will have a field across it due to the SP polarization, which will be modified by the strain via two mechanisms. The first is the same as in the proper case, as the strain will cause a flow of polarization current. But in addition, since the field depends on the charge *density*, the change in the area of the c-plane as a result of ε_1 will dilute or concentrate the pre-strain bound charge. For small strains the latter is given by the zero-strain formal polarization [2, 4]. Taking both of these mechanisms into account gives the improper PZ constant, e_{31}^{imp}.

In the case of, *e.g.*, Eq. (13), the PZ constants correspond to the improper case, since their role in the equation is to take into account the change in formal polar-

ization of material n with strain, so that σ_b corresponds to the bound charge at the coherent interface with the in-plane lattice constant of material m. The change in formal polarization with strain is an alternative definition of the improper PZ constants [4].

From Refs. 2 and 4, the improper PZ constant $e_{31}^{n,\text{imp}}$ is related to the proper constant by

$$e_{31}^{n,\text{imp}} = e_{31}^{n,\text{prop}} - P_{\text{f}}^n\big|_{\varepsilon=0}, \qquad (15)$$

where $P_{\text{f}}^n\big|_{\varepsilon=0}$ is the zero-strain formal polarization of material n. There is no change to the e_{33} PZ constant. The proper PZ constant is a well-defined bulk quantity, as it is related to the polarization current; however the improper PZ constant is branch dependent [4]. Here also, defining polarization with respect to the H reference proves useful. Since the formal polarization of the H structure vanishes (Fig. 1), $P_{\text{eff}}^{n,(\text{H ref})} = P_{\text{f}}^{n,\text{WZ}}\big|_{\varepsilon=0}$; this also ensures that improper PZ constants for the different materials are taken on the same branch, in the same way as this is confirmed for the SP polarization constants. Therefore, consistent use of the H reference structure allows us to write Eq. (7) as

$$\sigma_{\text{b}} = \Delta\widetilde{P}_{\text{SP}}^{\text{int},(\text{H ref})}$$
$$- 2\epsilon_1^n\left(e_{31}^{n,\text{prop}} - P_{\text{eff}}^{n,(\text{H ref})} - e_{33}^{n,\text{prop}}C_{13}^n/C_{33}^n\right), \qquad (16)$$

where $P_{\text{eff}}^{n,(\text{H ref})}$ can be taken from Table I.

TABLE II. Calculated piezoelectric polarization constants in units of C/m^2 compared with reported values from the literature.

		proper	improper	prev. reported[a]
GaN	e_{31}	−0.551	−1.863	−0.22 to −0.55
	e_{33}	1.020	1.020	0.43 to 1.12
AlN	e_{31}	−0.676	−2.027	−0.38 to −0.81
	e_{33}	1.569	1.569	1.29 to 1.94
InN	e_{31}	−0.604	−1.63	−0.23 to −0.59
	e_{33}	1.238	1.238	0.39 to 1.09

[a] From Ref. 3 and references therein.

Calculated proper PZ constants are given in the "proper" column of Table II. Since the HSE hybrid functional was used (and therefore density functional perturbation theory was not implemented), finite differences were used to calculate the derivatives with strain, following the procedure outlined in Eqs. (4)-(6) in Ref. 5. Specifically, improper PZ constants were calculated, and converted to proper constants by adding $P_{\text{f}}^n\big|_{\varepsilon=0}$ [see Eq. (15)] as determined in the calculation. This removes any dependence on the branch choice used in finite-difference calculations [4]. We then convert *back* to improper constants using $P_{\text{eff}}^{n,(\text{H ref})}$ as discussed above, in order to ensure that the constants are reported for the same branch for each material ("improper" column in Table II).

It is important to comment on the PZ constants reported in the literature [3]. When PZ polarization constants have been implemented in simulations (*e.g.* Refs. 3, 12, and 13) it has never been specified *which* PZ constants are used for WZ III-nitrides. However, by comparing our calculations of proper and improper PZ constants ("proper" and "improper" columns of Table II) with the reported PZ constants in the literature ("prev. reported" of Table II) we have found that the reported constants are more likely to be the proper PZ constants.

In previous work [6], $P_{\text{eff}}^{n,(\text{ZB ref})}$ was used instead of $P_{\text{f}}^{n}|_{\varepsilon=0}$ in Eq. (15) to convert improper to proper e_{31} PZ constants (cf. Table VI and V of Ref. 6). Because of the nonvanishing formal polarization of the ZB reference structure, $P_{\text{eff}}^{n,(\text{ZB ref})} \neq P_{\text{f}}^{n,\text{WZ}}|_{\varepsilon=0}$; instead, we see from the discussion resulting in Eq. (14) that Eq. (15) can be expressed as

$$e_{31}^{n,\text{imp}} = e_{31}^{n,\text{prop}} - \left(P_{\text{eff}}^{n,\text{ZB ref}} + \frac{e\sqrt{3}}{2a_n^2} \right), \qquad (17)$$

where a_n is the equilibrium lattice constant of material n. To our knowledge the inclusion of the last term in Eq. (17) has not been discussed in the literature.

Because of the small magnitude of $P_{\text{eff}}^{n,(\text{ZB ref})}$, neglecting the last term in Eq. (17) led to the conclusion in Ref. 6 that the difference between the proper and improper PZ constants is small, seemingly rendering the distinction of no consequence. Instead, because of the large magnitude of $P_{\text{eff}}^{n,(\text{H ref})}$ [and $e\sqrt{3}/2a_n^2$ in Eq. (17)], the distinction between proper and improper PZ constants is actually very significant.

V. COMPARISON WITH REPORTED EXPERIMENTAL RESULTS

A. Correct expressions for total polarization for wurtzite materials

Before discussing specific quantitative results for nitride semiconductors, we briefly summarize the main points of the previous sections and rigorously express the polarization of a given WZ material [Eq. (1)]. Spontaneous polarization constants must be defined with respect to a reference structure, and this choice of reference structure must be taken into account when evaluating polarization discontinuities at interfaces. We determined a correction term [Eq. (12)] that is necessary when effective SP polarization constants are used to determine the SP polarization difference between materials at an interface. This correction term is significant when the ZB reference structure is used [*e.g.*, Eq. (14)], but is zero for the H reference. Using H as a reference is therefore more straightforward and is the approach we advocate, with the SP constants $P_{\text{eff}}^{(\text{H ref})}$ listed in Table II.

In addition, the improper PZ constants should be used to determine interface bound charge and fields in heterostructure layers. These can be obtained from the proper constant e_{31}^{prop} by subtracting $P_{\text{eff}}^{(\text{H ref})}$ (Table II). Therefore, in the notation of this paper, Eq. (1) is written rigorously as

$$P = P_{\text{eff}}^{(\text{H ref})} + (\varepsilon_1 + \varepsilon_2)\left(e_{31}^{\text{prop}} - P_{\text{eff}}^{(\text{H ref})} \right) + \varepsilon_3 e_{33}^{\text{prop}}. \quad (18)$$

B. Calculation of sheet charges for III-nitrides

Because of the important impact of polarization on device performance and design, a plethora of experimental studies have been aimed at determining the effects of polarization at GaN/InGaN and GaN/AlGaN heterostructures. We have plotted these reported results in Fig. 2, expressed as the magnitude of polarization bound charge at the interface, as a function of alloy content (a full list of references is provided in Section S2 of the SM [19]).

For GaN grown in the $+c$ direction with the InGaN (AlGaN) grown on top, the sign of the bound charge at the interface will be negative (positive) [20].

In Fig. 2, the black dashed curves correspond to the current practice in the field: sheet charges are predicted based on (i) SP constants referenced to the ZB structure ($P_{\text{eff}}^{(\text{ZB ref})}$ in Table I) and Eq. (13) *without* the correction term $\Delta P_{\text{corr}}^{(\text{ZB ref})}$; and (ii) proper PZ constants ("proper" column in Table II). Quantities for alloys were obtained using linear interpolation (Vegard's law). For an explicit expression in terms of alloy content, see Eq. (1) in Section S3 of the SM [19]. Elastic constants were taken from Ref. 21.

The red solid line in Fig. 2 corresponds to the implementation recommended in this work, *i.e.,* using the H reference structure and the improper PZ constants, as in Eqs. (16) and (18) [and Eq. (3) in Section S3 of the SM [19]]. In view of the arguments given above, it may seem surprising that the dashed black and solid red curves agree as well as they do; we will return to this point in Section V E.

C. InGaN/GaN interfaces

For the InGaN/GaN system, most experimental studies have applied optical techniques to determine the polarization fields in GaN/InGaN/GaN quantum wells (QWs). This field can be probed by varying QW width [22] or external biases [23] and measuring the change in the optical properties of the QW (labeled "optical" in Fig. 2). In addition, there have also been studies using time-resolved PL to measure shifts due to screening of the polarization field by photoexcited or electrically injected carriers [24]. Other studies have been based on electron holography [25], where cross-sectional transmission electron microscopy is conducted on InGaN/GaN
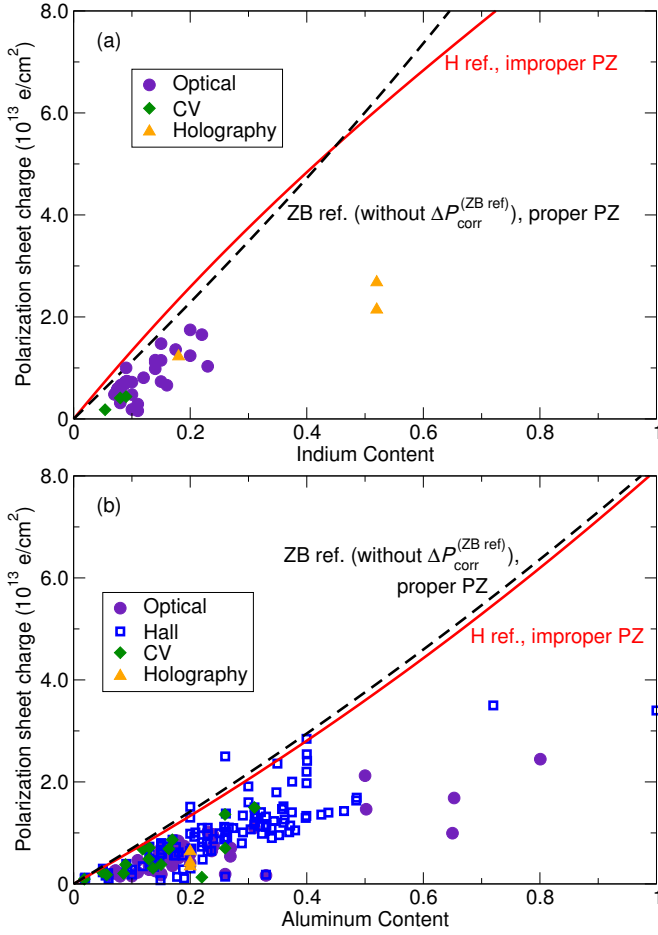
FIG. 2. Absolute values for polarization sheet charges at the (a) InGaN/GaN and (b) AlGaN/GaN interface as a function of alloy content predicted from the spontaneous polarization constants calculated using either the zincblende (ZB) reference structure [without correction term, Eq. (14)] and the proper piezoelectric constants (black dashed curve), or the hexagonal (H) reference structure and improper piezoelectric constants (red solid curve). Points are experimental values from the literature (see section S2 of the SM [19] for references and values).

heterostructures to determine the depth-resolved electrostatic potential in the growth direction, and capacitance-voltage (CV) profiling of the fields [26]. When fields are reported, we convert to bound charge for the purposes of Fig. 2(a), assuming GaN/InGaN/GaN quantum well (with thick barriers such that the electric field in the barriers is presumed zero) using a simple parallel-plate capacitor model ($E = \sigma/\epsilon_0 \epsilon_r$, using a relative dielectric constant for GaN of 10 [27] and for InN of 15 [28] and a linear interpolation for the dielectric constant of InGaN). For specific values of the points in Fig. 2, see Section S2 of the SM [19].

The red curve in Fig. 2(a) is indeed in reasonable agreement with the experimental observations, appearing to be an upper bound of the data. The optical experiments usually rely on Schrödinger-Poisson simulations to de-

termine the field magnitude from the measured optical properties. Uncertainties in input parameters to these models such as well widths, compositions, and composition profiles [29, 30] can result in quantitative differences. Deviations from an ideal QW structure are expected to be significant for InGaN/GaN because of the large lattice mismatch and the large difference in optimal growth temperatures for GaN and InGaN.

D. AlGaN/GaN interfaces

For the AlGaN/GaN system, there are two basic strategies for experimentally determining polarization effects. The first is to directly measure the polarization field in an AlGaN/GaN/AlGaN (QW) structure with the same methods as used in the InGaN/GaN case [31, 32]. For the purposes of Fig. 2(b) we have converted these fields to bound sheet charge densities in an AlGaN/GaN/AlGaN quantum well (using a relative dielectric constant for GaN of 10 [27]).

The other strategy is to measure the density of the 2DEG at the AlGaN/GaN interface in a HEMT structure (GaN channel, AlGaN barrier); from this, the bound interface charge, σ_b, can be derived [33]. The 2DEG density can be determined either by Hall effect [13] or CV [33] measurements.

Again, the predicted sheet charges are in reasonable agreement with experimental observations. The significant scatter in the experimental data in Fig. 2(b) may have several origins. There are experimental uncertainties that can influence fields, such as incomplete strain relaxation in buffer layers [34], and differences in background doping. For optical measurements, Schrödinger-Poisson modeling is again typically used to interpret the measured properties, and the same uncertainties arise as discussed in the case of InGaN/GaN [35, 36].

E. Comparison of theoretical implementations

The degree of agreement between results obtained based on the current practice in the field (ZB reference, no correction term, proper PZ) and our revised implementation (H reference, improper PZ constants) merits some discussion. The success of current practice in the field has hinged on a fortuitous cancellation of errors between the spurious term from the use of the ZB reference for the SP polarization and the proper PZ constants instead of the improper ones. Comparing the revised implementation [Eq. (3) of the SM [19]] to the current practice in the field [Eq. (1) of the SM [19]] gives the difference between the solid red and dashed black curves for a given x in Fig. 2 as $x\Delta P_{\text{corr}}^{(\text{ZB ref})} + 2\varepsilon_1(x)P_{\text{eff}}^{n,(\text{H ref})}(x)$ (see Section S4 of the SM [19] for a detailed derivation). For both AlGaN/GaN and InGaN/GaN, $\Delta P_{\text{corr}}^{(\text{ZB ref})}$ and $2\varepsilon_1 P_{\text{eff}}^{(\text{H ref})}$ have opposite signs and a tendency to cancel.

To understand why, note that for small lattice-constant differences, the correction term can be approximated as $x\Delta P_{\text{corr}}^{(\text{ZB ref})} \simeq -2\epsilon_1(x)P_{\text{f}}^{m,\text{ZB}}$ [see Eq. (6) of the SM [19]]. Therefore, the difference between the implementations depends on the in-plane strain and the difference between $P_{\text{f}}^{m,\text{ZB}}$ and $P_{\text{eff}}^{n,(\text{H ref})}$ [see Eq. (7) of the SM [19]]. We see from Table I that $P_{\text{eff}}^{\text{ZB ref}} = P_{\text{f}}^{\text{WZ}} - P_{\text{f}}^{\text{ZB}}$ is one to two orders of magnitude less than $P_{\text{eff}}^{(\text{H ref})} = P_{\text{f}}^{\text{WZ}}$, therefore $P_{\text{eff}}^{(\text{H ref})} \sim P_{\text{f}}^{\text{ZB}}$.

The similarity between P_{f}^{ZB} and P_{f}^{WZ} is not unexpected. Although ZB has sufficient symmetry to preclude SP polarization, in the [111] direction the structure only differs from the WZ c direction by the stacking of the cation/anion planes.

For AlGaN/GaN, the relatively modest difference in lattice constants between AlN and GaN (and therefore modest ε_1 values for coherently strained alloy layers), and an almost exact cancellation between $P_{\text{f}}^{\text{GaN,ZB}}$ and $P_{\text{eff}}^{\text{AlGaN,(H ref)}}$ means that the difference between implementations is small over the whole composition range (cf. Section S4 of the SM [19]). For InGaN/GaN, the large lattice mismatch of InN and GaN and a less complete cancellation of $P_{\text{f}}^{\text{GaN,ZB}}$ and $P_{\text{eff}}^{\text{InGaN,(H ref)}}$ results in a significant deviation at higher In content; this will be important for the prediction of polarization fields in applications such as tunnel field-effect transistors based on thin, high In-content interlayers [37].

The two implementations differ significantly in the *relative* contributions of SP and PZ polarization. An effect of this is illustrated by the case where there is strain relaxation in the alloy layer. In Fig. 3 the predicted polarization bound charges for $\text{In}_{0.2}\text{Ga}_{0.4}\text{N}/\text{GaN}$ (blue curves) and $\text{Al}_{0.2}\text{Ga}_{0.4}\text{N}/\text{GaN}$ (green curves) are shown as a function of strain relaxation of the layer, modeled by simply scaling ϵ_1. For both InGaN/GaN and AlGaN/GaN, the revised implementation of this work predicts a much faster decrease in bound charge at the interface than the current practice in the field. Of course, strain relaxation is associated with the presence of edge dislocations at the interface, which may themselves influence the interface bound charge; this effect has not been taken into account in either version of the implementation.

VI. CONCLUSIONS

We have derived a rigorously correct implementation of polarization constants in wurtzite materials, focusing on the example of the III-nitrides. Our derivation has demonstrated the impact of the choice of reference structure when calculating spontaneous polarization constants using the modern theory of polarization. Insufficient care in using the values can result in spurious contributions to the polarization discontinuities at heterostructure in-terfaces. We have provided new values calculated with a consistent hexagonal (rather than zincblende) reference
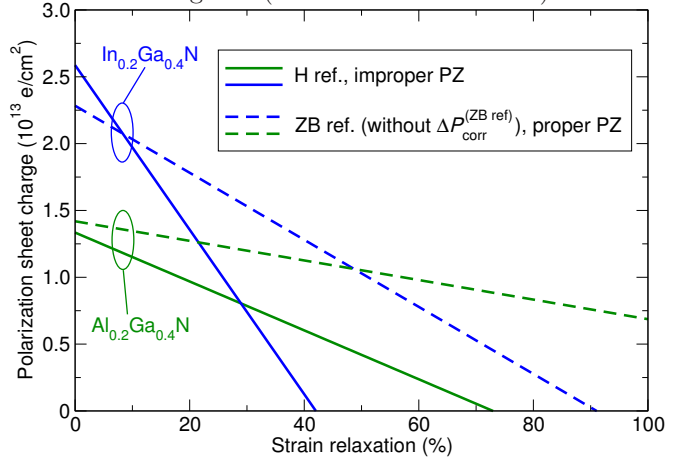

FIG. 3. Absolute values for polarization sheet charges at the $\text{In}_{0.2}\text{Ga}_{0.4}\text{N}/\text{GaN}$ (blue) and $\text{Al}_{0.2}\text{Ga}_{0.4}\text{N}/\text{GaN}$ (green) interface as a function of percent strain relaxation. 0 % relaxation corresponds to perfectly strained layers, 100 % relaxation to an unstrained overlayer at its bulk lattice constant. Solid curves are the correct implementation described in this work (H reference for SP and improper PZ constants); dashed curves are the current practice in the field (ZB reference for SP, without the correction term, and proper PZ constants).

structure. In addition, we have demonstrated the importance of choosing the correct piezoelectric constants (improper), and provided values for these improper constants. These revised values of the spontaneous and piezoelectric constants can be directly used in simulations and to interpret experimental observations. The revised implementation predicts a more rapid decrease of polarization charge with strain relaxation for an alloy layer on GaN.

[1] J. F. Nye (1957). *Physical Properties of Crystals* (pp. 300). London: Oxford University Press.

[2] R. M. Martin, Phys. Rev. B **5**, 1607 (1972).

[3] M. Feneberg and K. Thonke, J. Phys.: Condens. Matter **19**, 403201 (2007).

[4] D. Vanderbilt, J. Phys. Chem. Sol. **61**, 147 (2000).

[5] F. Bernardini, V. Fiorentini, D. Vanderbilt, Phys. Rev. B **56**, R10024 (1997).

[6] F. Bernardini, V. Fiorentini, D. Vanderbilt, Phys. Rev. B **63**, 193201 (2001).

[7] F. Bechstedt, U. Grossner, and J. Furthmüller, Phys. Rev. B **62**, 8003 (2000).

[8] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).

[9] R. Resta, Rev. Mod. Phys. **66**, 899 (1994).

[10] D. Vanderbilt and R. D. King-Smith, Phys. Rev. B **48**, 4442 (1993).

[11] R. Resta and D. Vanderbilt, in *Physics of Ferroelectrics: A Modern Perspective*, edited by K. Rabe, Ch. H. Ahn, and J.-M. Triscone (Springer-Verlag, Berlin, 2007).

[12] SiLENSe Physics Summary v5.2 (2011).

[13] O. Ambacher, B. Foutz, J. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Murphy, R. Dimitrov, A. Mitchell, and M. Stutzmann, J. Appl. Phys. **87**, 334 (2000).

[14] J. W. Bennett, K. F. Garrity, K. M. Rabe, and D. Vanderbilt, Phys. Rev. Lett. **109**, 167602 (2012).

[15] J. Heyd, G. E. Scuseria, M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003); *ibid.* **124**, 219906 (2006).

[16] G. Kresse and J. Furthmüller, Phys. Rev. B. **54**, 11169 (1996).

[17] P. E. Blöchl, Phys. Rev. B 50, 17953 (1994).

[18] H.J. Monkhorst, J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[19] See Supplemental Material at [URL will be inserted by publisher] for a table of structural properties and band gaps, a list of experimental references, equations used to generate curves in Fig. 2, and equations demonstrating the difference between implementations.

[20] O. Ambacher, R. Dimitrov, M. Stutzmann, B. E. Foutz, M. J. Murphy, J. A. Smart, J. R. Shealy, N. G. Weimann, K. Chu, M. Chumbes, B. Green, A. J. Sierakowski, W. J. Schaff, and L. F. Eastman, Phys. status solidi (b) **216**, 381 (1999).

[21] I. Vurgaftman and J. R. Meyer, J. Appl. Phys. **94**, 3675 (2003).

[22] A. Hangleiter, F. Hitzel, S. Lahmann, and U. Rossow, Appl. Phys. Lett. **83**, 1169 (2003).

[23] F. Renner, P. Kiesel, G. H. Döhler, M. Kneissl, C. G. Van de Walle, and N. M. Johnson, Appl. Phys. Lett. **81**, 490 (2002).

[24] D. Turchinovich, P. Uhd Jepsen, B. S. Monozon, M. Koch, S. Lahmann, U. Rossow, and A. Hangleiter, Phys. Rev. B **68**, 241307 (2003).

[25] D. Cherns, J. Barnard, and F. A. Ponce, Solid State Commun. **111**, 281 (1999).

[26] H. Zhang, E. J. Miller, E. T. Yu, C. Poblenz, and J. S. Speck, J. Vac. Sci. Technol. B **22**, 2169 (2004).

[27] A. S. Barker, M. Ilegems, Phys. Rev. B **7**, 743 (1973).

[28] O. Madelung, *Semiconductors: Data Handbook* (Springer Berlin Heidelberg, 2004).

[29] Y.-R. Wu, R. Shivaraman, K.-C. Wang, and J. S. Speck, Appl. Phys. Lett. **101**, 083505 (2012).

[30] P. M. McBride, Q. Yan, and C. G. Van de Walle, Appl. Phys. Lett. **105**, 083507 (2014).

[31] N. Grandjean, B. Damilano, S. Dalmasso, M. Leroux, M. La ugt, and J. Massies, J. Appl. Phys. **86**, 3714 (1999).

[32] M. Leroux, N. Grandjean, J. Massies, B. Gil, P. Lefebvre, and P. Bigenwald, Phys. Rev. B **60**, 1496 (1999).

[33] E. T. Yu, G. J. Sullivan, P. M. Asbeck, C. D. Wang, D. Qiao, and S. S. Lau, Appl. Phys. Lett. **71**,2794 (1997).

[34] Y. Zhou, B. Shen, T. Someya, H. Yu, J. Liu, H. Zhou, R. Zhang, Y. Shi, Y. Zheng, and Y. Arakawa, Jpn. J. Appl. Phys. **41**, 2531 (2002).

[35] S. Arulkumaran, T. Egawa, H. Ishikawa, and T. Jimbo, J. of Vac. Sci. Technol. B **21**, 888 (2003).

[36] K. A. Mkhoyan, J. Silcox, Z. Yu, W. J. Schaff, L. F. Eastman, J. Appl. Phys. **95**, 1843 (2004).

[37] W. Li, S. Sharmin, H. Ilatikhameneh, R. Rahman, Y. Lu, J. Wang, X. Yan, A. Seabaugh, G. Klimeck, D. Jena, and P. Fay, IEEE J. Explor. Solid-State Comput. Devices Circuits **1**, 28 (2015).