Revisiting semiconductor band gaps through structural motifs: An Ising model perspective

Robert A. Makin ●, Krystal York, and Steven M. Durbin ●[*]

Department of Electrical and Computer Engineering, Western Michigan University, Kalamazoo, Michigan 49008, USA

Roger J. Reeves

School of Physical and Chemical Sciences, University of Canterbury, Christchurch 8140, New Zealand

We present an alternative perspective on semiconductor band-gap energies in terms of structural motifs, viewed through the lens of an Ising model as a means of quantifying the corresponding degree of lattice ordering. The validity of the model is demonstrated experimentally first through $ZnSnN_2$ as an archetype ternary heterovalent semiconductor, in which variation of cation disorder enables the band gap to be tuned from above its equilibrium phase value, through zero, to negative values which correspond to inverted bands. The model is then applied to example binary compounds InN, GaN, and ZnO where anion-cation antisite defects form the basis for structural motifs, and we present experimental evidence that the same range of band-gap tuning is also possible for such materials. The case of alloys is treated by applying a three-spin Potts model to $In_xGa_{1-x}N$. The Ising model also applies to elemental semiconductors, and is used to explain the wide range of reported values for silicon and nanoporous graphene in the context of vacancy-based structural motifs.

I. INTRODUCTION

Semiconductor materials are characterized by multiple properties that must be considered when contemplating the fabrication of an electronic device, but arguably the most fundamental of them is the band-gap energy. Often described as arising from the periodic potential describing the lattice in which a particular material crystallizes [1], the intrinsic gap of a semiconductor is defined as the energy difference between the valence band the uppermost band containing electrons which participate in chemical bonding and the band immediately above it, referred to as the conduction band. Disorder is acknowledged as having an impact on the band gap, although largely categorized as undesirable, except perhaps in the context of what is termed an amorphous semiconductor. Engineering of the band-gap energy dates to the early days of light-emitting diodes [2], and can be accomplished by alloying one or more compounds in order to achieve a desired value. It is also well known that strain can also be used to engineer the energy gap, by exploiting heteroepitaxial structures potentially in tandem with alloying. However, the traditional perspective masks what may be best described as a hidden variable, requiring a variety of phenomena to be invoked including many-body effects, degenerate carrier populations, and impurity doping to explain a wide range of experimentally determined numerical values for the band-gap energy. While generally successful at predicting trends, a fundamental limitation of the conventional approach is a lack of recognition for the full range of achievable band-gap values, which we demonstrate can only be discerned through a quantitative description of disorder cast in terms of structural motif composition.

The concept of structural motifs provides the basis for describing a wide range of constructs, spanning such diverse fields as biology [3–6], music [7,8], and art [9]. The term has its origin as a means of understanding RNA conformations and corresponding secondary structures [3,4], and structural similarities among different kinase enzymes performing the same function [5]; in such a context, a motif can be defined as a component of a structure with identifiable but variable order, where the resulting physical modification leads to altered properties or functionality. In the case of semiconductor compounds, a structural motif describes any of the possible variants to the nearest-neighbor bonding environment of a given atom. In an ideal crystal of a given compound with fixed stoichiometry, the number of structural motifs is limited to a single type; the occurrence of variations of this base motif depends directly upon the degree of disorder and stoichiometry characterizing the corresponding lattice. Cluster expansion theory also describes disorder, but does so from a viewpoint of lattices and sublattices, building up clusters that can occur within a given lattice from single points, to bonds (pairs) to triangles and higher-order clusters. While cluster expansion theory can include the complete set of structural motifs, such a lattice-based approach can mask the fundamental importance of the variety of structural motifs.

In an otherwise interstitial-free lattice, the set of structural motifs arising from disorder is defined through the range of possible antisite defects. Since such defects can have large energies of formation, it is usually argued that they are unlikely to form in lattices. A disordered lattice, however, has a higher entropy than a perfectly ordered lattice, and as such the driving force behind the formation of disordered lattices and their

*steven.durbin@wmich.edu

corresponding numerical distribution of the complete set of possible structural motifs is, as for any endothermic process, entropy. Thus, by designing processes to be entropy driven instead of energy driven, it is possible to obtain a wide range of structural motif distributions. For example, by varying the conditions during crystal growth, different degrees of ordering can be achieved [10]; the two-step process of annealing and quenching a sample is another common technique employed to obtain a desired degree of ordering [11,12]. Additionally, nonequilibrium crystal growth techniques, which can easily access metastable structures, are frequently used to grow disordered lattices. In reality, even so-called equilibrium growth conditions will only be approximately so, and hence will yield nonzero distributions of nonequilibrium structural motifs.

Initial efforts directed toward quantifying the effects of varying degrees of order on semiconductor properties focused on alloys, where exchanging atoms with the same number of valence electrons does not violate the octet rule. Wei, Laks, and Zunger first proposed constructing an Ising model to describe cation ordering in common-anion $(A_{1-x}B_xC)$ semiconductor alloys: In their formulation, A cations were assigned a $+1$ spin and B cations were assigned a spin of -1 [13,14]. To quantify the degree of ordering characterizing the cation sublattice, they applied the Bragg-Williams order parameter [15–18] S, which takes on values between 1 and 0 for ordered and completely disordered lattices, respectively. Specifically, they defined $S = r_A + r_B - 1$, where r_A is the fraction of A cations on A-cation sites and r_B is the fraction of B cations on B-cation sites (as defined in terms of the perfectly ordered structure). Within the Wei-Laks-Zunger formulation, then, any property P characterizing a binary alloy sample of mole fraction x in a given state of cation sublattice ordering S is expressed as

$$P(x, S) = P(x, 0) + S^2[P(0.5, 1) - P(0.5, 0)] \quad (1)$$

so long as P is dominated by pair (in this case, cation atom) interactions. Wei *et al.* note that while their model is accurate for the case of $x = 0.5$, it becomes inaccurate as the stoichiometry moves away from equal cation composition. Recently, we extended the concept of a cation ordering Ising model to the case of two ternary heterovalent, environmentally friendly, sustainably sourced, earth-abundant element semiconductors, $ZnSnN_2$ and $MgSnN_2$, which have the ability to span the visible spectrum not through traditional alloying, but rather through the controlled introduction of varying degrees of cation ordering [10]. We further demonstrated that the same model accurately predicts the reported experimental measured band-gap values for similar compounds, including $ZnGeN_2$, $ZnSnP_2$, $CuGaS_2$, and $CuInS_2$. The Bragg-Williams order parameter can be extracted through any number of experimental techniques, including Raman spectroscopy, x-ray diffraction, reflection high-energy electron diffraction (RHEED), and, as we demonstrate in this paper, scanning electron microscopy (SEM) with a high degree of reproducibility. All data regarding band gaps and measurements of S for every sample used in this paper can be found in the Supplemental Material [19].

Our previous study focused exclusively on stoichiometric material [10]. Since the two different cations have opposite effects on the relative energy of the valence band, deviations

from stoichiometry should also lead to variation in the band-gap energy, which can be combined with variable ordering for additional tuning ability. Here, we demonstrate that this is indeed the case, and enables the band-gap energy to not only exceed the value corresponding to the stoichiometric equilibrium structure, but also provides the means to tune the band gap, through zero, and into negative values (corresponding to band inversion). This ability can be understood in terms of the direct numerical relationship between S and the distribution of different structural motifs, which has a stoichiometry dependence; the Ising model enables us to predict the full range of accessible band-gap values. We subsequently invoke the Potts model to properly describe alloys, which can have any degree of sublattice ordering.

However, the fundamental method we describe here also applies to binary compounds such as GaN, and to elemental semiconductors such as silicon and nanoporous graphene. As we illustrate, it is then possible to accurately account for the range of reported band-gap energies for these and other materials, without invoking phenomena such as the Burstein-Moss effect [20,21] or band-gap renormalization [22].

II. GENERALIZED ISING MODEL FOR ARBITRARY CATION STOICHIOMETRY

Wei, Laks, and Zunger express any lattice property P as a series expansion using the orthonormal basis of correlation functions of all possible class of figures (groupings of $k = 1, 2, 3 \ldots$ lattice sites) [13,14]. Following their framework of limiting the cluster expansion to just single and pair figures, we express a property of a lattice (that is dominated by single and pair interactions) with a given degree of ordering and stoichiometry as

$$P(x, S) = p_p \Pi_p(x, S), \quad (2)$$

where p_p is the contribution of pair interactions to the lattice property and $\Pi_p = \langle \sigma_0 \sigma_r \rangle$ is the spin-pair correlation function. Departing from the derivation of Wei *et al.*, we equate S^2 with the reduced spin-pair correlation function $\Gamma(r) = [\langle \sigma_0 \sigma_r \rangle - \langle \sigma_0 \rangle^2]/[\langle \sigma_0^2 \rangle - \langle \sigma_0 \rangle^2]$, and rearranging we can find Π_p in terms of S^2 and spins

$$\Pi_p = S^2[\langle \sigma_0^2 \rangle - \langle \sigma_0 \rangle^2] + \langle \sigma_0 \rangle^2. \quad (3)$$

In the completely disordered state ($S = 0$), no spins are correlated, and the spin-pair correlation function reduces to $\langle \sigma_0 \rangle^2$; in the completely ordered lattice ($S = 1$), the spins are completely correlated and thus the spin-pair correlation function has unity value. Thus,

$$\Pi_p(x, S) = S^2[\Pi_p(0.5, 1) - \Pi_p(x, 0)] + \Pi_p(x, 0) \quad (4)$$

which yields

$$P(x, S) = S^2[P(0.5, 1) - P(x, 0)] + P(x, 0). \quad (5)$$

Consequently, we may express the expected value of the band gap as

$$E_g(x, S) = S^2[E_g(0.5, 1) - E_g(x, 0)] + E_g(x, 0). \quad (6)$$

Depending on the band gap in the fully disordered state for a particular stoichiometry, as we will demonstrate, Eq. (4) allows for the possibility to completely close the band gap,
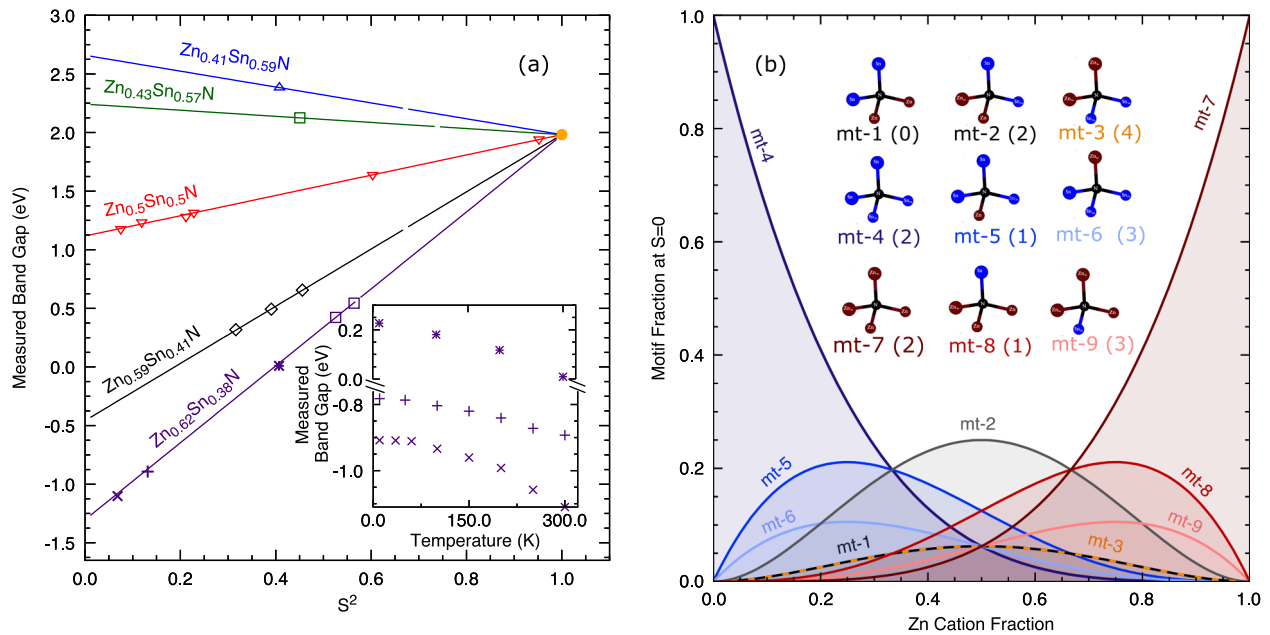
FIG. 1. (a) Measured band-gap energy of a series of $ZnSnN_2$ thin films with a range of cation stoichiometries as a function of Bragg-Williams order parameter S (plotted as S^2). Symbols represent experimental values, and linear fits include the perfectly ordered ($S = 1$) stoichiometric point determined in Ref. [10]. The yellow dot indicates the extrapolated end point. The dashed lines indicate physically inaccessible values of S. Inset: Band-gap energy as a function of temperature for the two $Zn_{0.62}Sn_{0.38}N$ samples with inverted bands, plotted as a negative band-gap energy, and one sample having a nearly zero ($\sim$10 meV) room-temperature gap. The same symbols are used as those in the corresponding Ising model line. No features consistent with secondary phases (e.g., Zn_3N_2) were detected by electron diffraction during growth. (b) Plot of the motif composition (normalized to unity) corresponding to a fully randomized cation sublattice ($S = 0$) as a function of the cation stoichiometry ($Zn + Sn = 1$). The inset depicts the structural motif corresponding to each curve, with the central atom representing nitrogen, the blue atoms representing tin, and the red atoms representing zinc. The numbers in parentheses indicate the corresponding number of cation antisite defects. The color scheme is preserved in the plots of motif composition, with motifs having equal numbers of Zn and Sn atoms indicated with a black dashed line (labeled mt-1 and mt-3) or a gray line (mt-2). Details pertaining to calculation of the motif distributions are provided in the Supplemental Material [19].

and even force it to a negative value (corresponding to inversion of the valence and conduction bands) by tuning S. It is worth noting, however, that a perfectly ordered structure ($S = 1$) is only achievable under balanced conditions, when the number of A cations equals the number of B cations, and none of them lie on anion sites (or vice versa). More generally, the maximum achievable value of S is $2x$ for $x \leqslant 0.5$, and $2(1 - x)$ for $x \geqslant 0.5$.

In order to evaluate the predictive ability of the stoichiometry-dependent Ising model as formulated, a series of $ZnSnN_2$ thin films were grown with different Zn:Sn ratios using plasma-assisted molecular beam epitaxy (PAMBE). The experimentally determined band gaps are plotted in Fig. 1(a) as a function of S^2. As expected, all samples with the same stoichiometry exhibit a linear relationship between band gap and S^2, and share the $S = 1$ band-gap end point, confirming the validity of this cation ordering Ising model for ternary heterovalent materials. Furthermore, two films intentionally grown Zn rich and with low-S values exhibit clear evidence of inverted bands, indicated in Fig. 1(a) as negative band-gap values and confirmed through the temperature dependence of the magnitude of the band-gap energy [23]. A third film with a nearly zero room-temperature band gap ($\sim$10 meV) was also successfully grown as part of the same series. Two films grown with slightly Sn-rich stoichiometries are included in

Fig. 1(a) to demonstrate that the equilibrium phase stoichiometric end point is not in fact the maximum band-gap energy for the material. Consequently, it is possible to tune the band-gap energy of this semiconductor from a value significantly in excess of 2 eV, through zero, to comparable separations of inverted bands.

The underlying mechanism responsible for the trends evident in Fig. 1(a) is the direct correlation of the Bragg-Williams order parameter to cation motif distributions as influenced by stoichiometry. For $ZnSnN_2$ there are nine types of such motifs, each of which is based upon a central N atom surrounded by four cation atoms, as illustrated in Fig. 1(b). For the two extreme stoichiometries (i.e., ZnN, SnN), only one type of motif will exist. At intermediate stoichiometries, however, a mixture of motifs occurs for $S < 1$, the precise distribution of which is dictated by the stoichiometry in conjunction with the value of the Bragg-Williams order parameter. In the perfectly ordered ($S = 1$) $ZnSnN_2$ lattice which, by definition, must also be stoichiometric, only one type of motif exists the zero antisite defect Zn_2Sn_2 motif (Supplemental Material Fig. 2 [19]). However, as S is decreased, other motifs appear in directly correlated amounts. The Zn_2Sn_2 motifs are neutral since the average charge of the cations is balanced by the nitrogen atom. The other types of motifs all have unequal numbers of Zn and Sn atoms, and so the cation charges are not locally
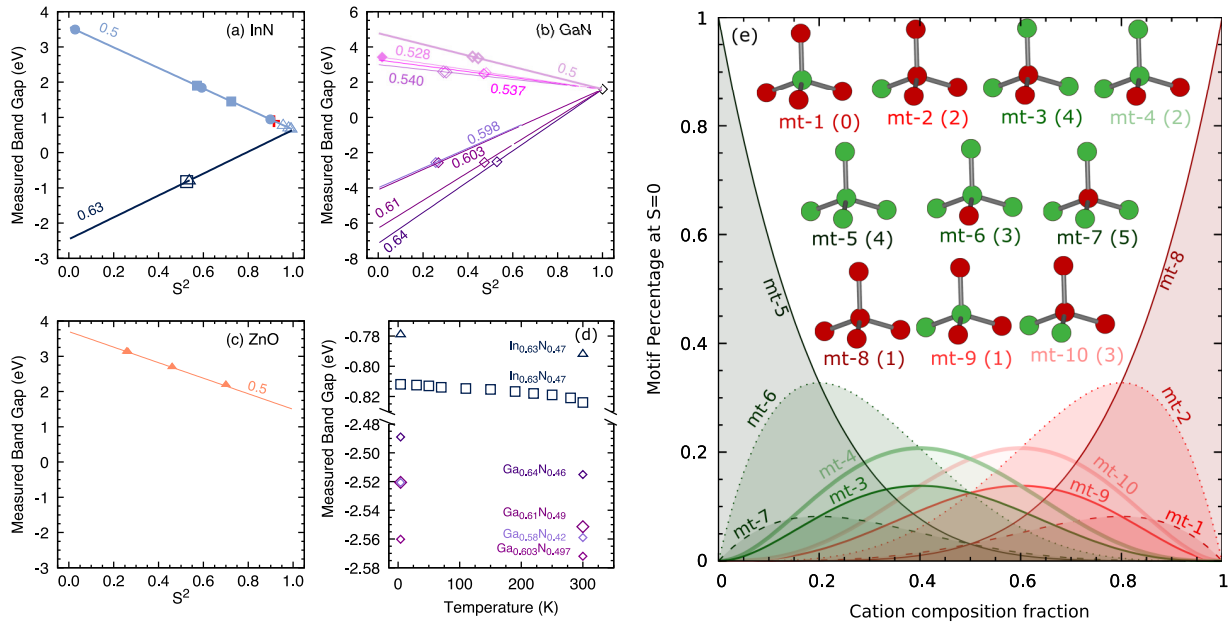
FIG. 2. The Ising model applied to example compound semiconductors. (a) Measured band-gap energy of InN as a function of order parameter. Experimental points are identified by symbols. Upper curve: nominally stoichiometric material. Light blue solid circles from [29]. Light blue solid squares from [30]. Red cross: Data extracted from [31]. Lower curve: an indium-rich sample which exhibits inverted bands (represented as a negative band gap). (b) Experimentally derived measured band-gap energy of GaN as a function of order parameter, including a range of gallium-rich samples which are characterized by inverted bands. Open symbols: this work. Closed diamond: data from [32]. (c) Measured band-gap energy as a function of ordering for stoichiometric ZnO (data extracted from [33]). (d) Temperature dependence of InN and GaN samples (preserving the same symbols) in (a) and (b) characterized by inverted bands, plotted as a negative gap. (e) Plot of the motif-type composition (normalized to unity) corresponding to the fully randomized case ($S = 0$) as a function of the stoichiometry. The inset depicts the structural motif corresponding to each distribution curve, with the green atoms representing anions and the red atoms representing cations. The numbers in parentheses indicate the number of corresponding antisite defects.

balanced by the nitrogen atom. Motifs consisting of more Sn than Zn atoms will therefore be positively charged, while Zn-rich motifs will be negatively charged. In $ZnSnN_2$ samples with excess Sn concentration, Sn-rich motifs occur in higher quantities than Zn-rich motifs and thus dominate the band-gap behavior. Specifically, the nitrogen p level (and hence the valence band) is pushed down by the Coulombic attraction of the positively charged Sn to its counterpart anion [10,24]; thus, as more Sn-rich motifs appear with decreasing cation ordering, the band gap will correspondingly increase. We note that in the case of excess Sn, the corresponding weakening of p-d repulsion ($2p$ nitrogen states and $3d$ Zn states) would also be expected to have a similar effect on the valence band. For $ZnSnN_2$ samples with excess Zn concentration, Zn-rich motifs, which as result of their negative charge push up the nitrogen p levels, will occur in larger quantities than Sn-rich motifs, and therefore will dominate the band-gap behavior. Consequently, the band gap will shrink, and ultimately the valence and conduction bands will invert, as the imbalance between Zn-rich and Sn-rich motifs pushes the nitrogen p level (and hence the valence band) upward.

III. APPLICATION TO BINARY SEMICONDUCTORS

The Ising model framework can also be applied to binary semiconductors, if we assign one spin to cations and the other spin to anions. As examples, we have extracted data

for InN, GaN, and ZnO from the literature as well as from samples previously grown using the same PAMBE technique employed for the $ZnSnN_2$ films. The order parameters were calculated from either Raman spectra, RHEED, x-ray diffraction, or SEM. As can be seen in Fig. 2, each of the three materials follows the Ising model trend predicted by Eq. (4), including the effects of off stoichiometry. Significantly, the range of reported band-gap values for InN, the subject of considerable controversy and discussion in the literature [25], is accounted for through variable ordering (i.e., a distribution of different motifs consisting of antisite defects), without the need for invoking the effects of degenerate electron populations [26] or oxygen contamination [27]. Further, as shown in Fig. 2(a), there is experimental evidence that the band gap of this material can also be tuned through zero and into negative values corresponding to band inversion, in the same fashion as $ZnSnN_2$. There is a considerably larger body of literature for GaN, making it straightforward to identify a series of samples within one study in which separate measures of both the band gap and the order parameter are provided. As seen in Fig. 2(b), the same trends exist for GaN as they do for InN (and $ZnSnN_2$). If we extrapolate the plots to the perfectly ordered stoichiometric case, we obtain the value of the $S = 1$ band gap of the material in this case, 1.60 eV, in contrast with the generally accepted value of 3.4 eV; a similar situation exists for ZnO [Fig. 2(c)], with the Ising model predicting an $S = 1$ gap of 1.55 eV, as opposed to the accepted value of
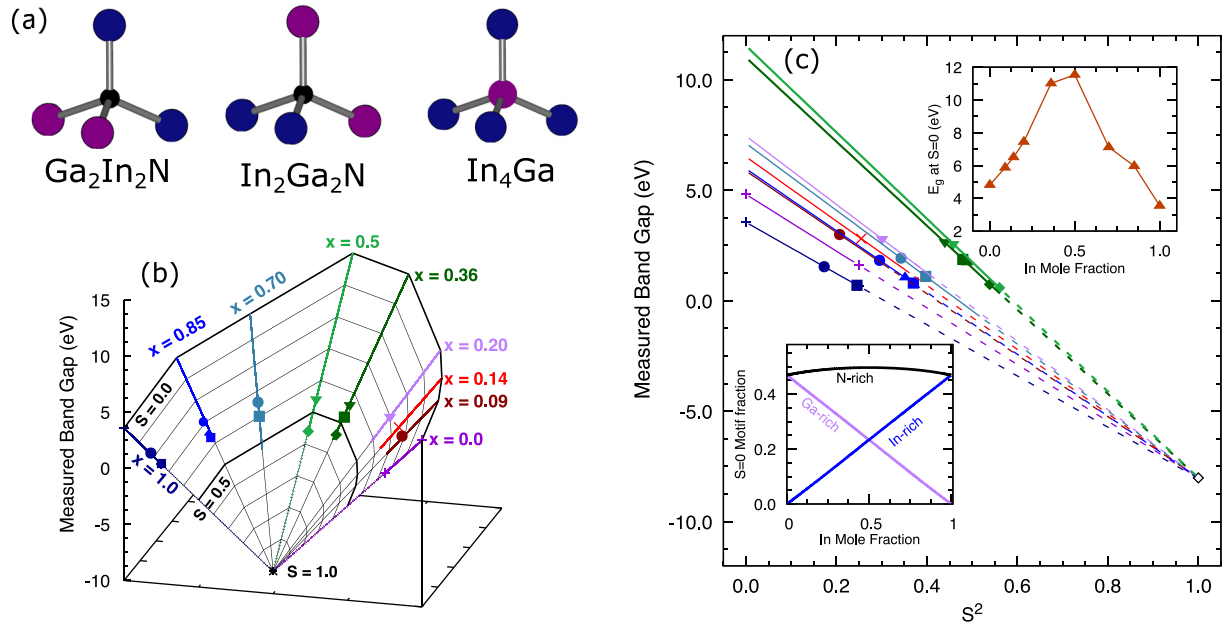
FIG. 3. Employing the Potts model to correctly represent a ternary alloy. (a) The three fundamental motifs for perfectly ordered, stoichiometric InGaN. (b) Measured band-gap energy for $In_xGa_{1-x}N$ as a function of mole fraction x and order parameter S, illustrating the wide range of accessible band-gap values. (c) The plot in (a) recast in Cartesian coordinates, with experimental points identified. Squares: [35]. Circles: [36]. Up triangles: [37]. Down triangles: [38]. Diamonds: [39]. Note that there is a maximum value of S corresponding to each mole fraction value, with physically unattainable values represented by dashed lines. Top inset: trend of $S = 0$ band gap as a function of indium mole fraction x. Bottom inset: trend of nitrogen-rich, In-rich, and Ga-rich motifs at $S = 0$ for $In_xGa_{1-x}N$ as a function of the indium mole fraction. The motif equations for InGaN are provided in the Supplemental Material [19].

3.37 eV [28]. In Fig. 2(d), we plot the temperature dependence of the band gap for InN and GaN samples characterized by inverted bands.

As in the case of $ZnSnN_2$, the trends which become apparent for binary compounds within the context of the Ising model can be explained by the motif distribution dictated by the stoichiometry and order parameter. There are 10 types of motifs possible in fourfold-coordinated binary materials as illustrated in Fig. 2(e); the cation and anion charges are balanced within the AB_4 motifs, but this is not the case in cation- or anion-rich motifs. The cation-rich motifs, all of which have a central atom and are positively charged, will push down the cation s levels and thus lower the conduction band maximum. In contrast, the anion-rich motifs will push up the anion p levels due to Coulombic repulsion, however, the anion p levels are pushed down due to strong anion bonds that form between the anions in the anion-rich motifs. For stoichiometric partially ordered lattices, the highest percentage motifs are anion rich, and thus increasingly dominate the band-gap behavior as the order parameter decreases. For cation-rich binaries, the cation-rich motifs dominate the band-gap behavior, which causes the band gap to decrease and even invert for large cation concentrations, as shown for InN in Fig. 2. Similarly, for anion-rich binaries, the anion-rich motifs dominate the motif distribution, and this results in an increase of the band gap. We can also understand the quantitative differences in behavior among the compounds in Fig. 2 from the context of structural motifs. In ZnO, the O-O bonds in antisite motifs are not as strong as the N-N bonds in GaN antisite motifs, and thus the variation in the band gap for the stoichiometric case is less for ZnO than it is for GaN. Because Ga is slightly

more electronegative than In, the Ga cation-centered motifs in stoichiometric GaN will push down the cation s levels less than In will for stoichiometric InN, resulting in a slightly larger ($\sim$280 meV) increase in the band gap with decreasing S^2 for GaN than for InN.

IV. ALLOYS AND GENERALIZATION TO THREE OR MORE SPIN STATES

The spin-based disorder model we have presented, which leads to Eqs. (5) and (6), can also be applied to materials with more than two elements that may exchange sites (e.g., a common-anion or common-cation alloy) by generalizing from the Ising model to the multispin Potts model [34]. The Potts model assigns each of the q total spins a complex vector $\exp(i2\pi\sigma_n/q)$ where $\sigma_n = 1, 2, 3, \ldots q$, and reduces to the Ising model when $q = 2$. The derivation in Eqs. (1)–(4) makes no assumption about the number of spin states, and therefore applies directly to the more general Potts model. Thus, even with multiple elements that can exchange sites in a lattice, any lattice property (such as band-gap energy) which is dominated by pair interactions will follow Eq. (3). As an illustration, we consider the alloy $In_xGa_{1-x}N$. We begin by defining the order parameter as $S = \frac{(ra+rb+rc)}{2} - \frac{1}{2}$, where r_a is the fraction of In atoms on In sites, r_b is the fraction of Ga atoms on Ga sites, and r_c is the fraction of N atoms on N sites. The terminal case of $S = 1$, corresponding to the zero-field state in the Potts model, is comprised of equal numbers of In, Ga, and N atoms, not the charge-balanced stoichiometric composition $In_xGa_{1-x}N$. This necessitates that a minimum of three motifs, shown in Fig. 3(a), must be used to describe the
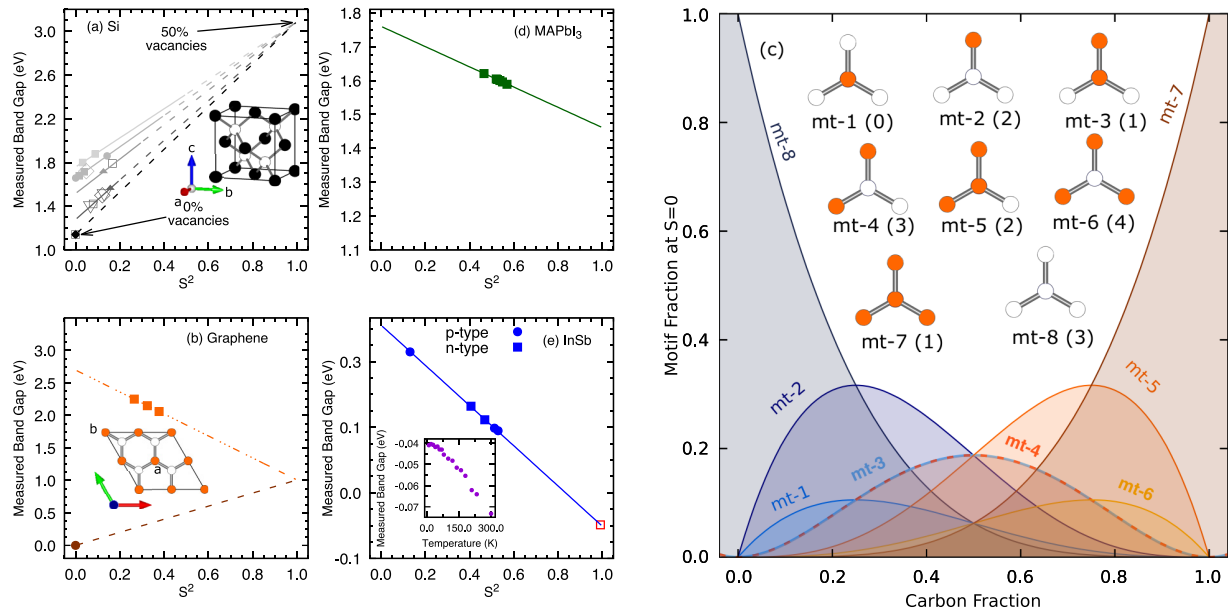
FIG. 4. (a) Motif-driven Ising model analysis applied to silicon, with experimental points extracted from the literature. The inset illustrates how silicon atoms and vacancies represent the two spin states within the Ising model. Hence, $S = 1$ corresponds to a lattice with equal numbers of silicon atoms and vacancies in a perfectly ordered structure, with a predicted band gap of 3.1 eV. Data extracted from [40] (solid upward triangle), [41] (solid square), [42] (solid circle), [43] open diamond, [44] (open square), and [45] (open downward triangle). (b) A similar analysis applied to nanoporous graphene sheets (data extracted from [46]). The solid circle represents the $S = 0$ vacancy-free case [47]. (c) Plot of the motif distribution for graphene with a range of vacancy concentrations corresponding to the fully randomized ($S = 0$) case. The inset defines the eight possible structural motifs, with vacancies represented by white circles. The extreme represented by zero carbon fraction would be a nonexistent layer, and the other extreme is a graphene ribbon having no vacancies. (d) Ising model applied to the hybrid perovskite MAPbI$_3$. Data extracted from [48]. (e) Band gap of InSb (n type, $\sim6 \times 10^{15}$ cm^{-3} [49] and p type, $\sim10^{18}$ cm^{-3} [50]) as a function of S^2. The additional point (red square) is a commercially obtained bulk sample with $n = 1.9 \times 10^{16}$ cm^{-3}. The inset shows the band gap for the commercial sample as a function of temperature.

alloy since any number less than three will not result in the correct composition for the $S = 1$ state.

Figure 3 shows the results of plotting the band gap of InGaN as a function of S^2 for a range of mole fractions; values were extracted from the literature, as well as from films previously grown using the same PAMBE system later employed for ZnSnN$_2$. Consistent with what is expected from applying the Ising model to two-atom disordered lattices, the band gap of InGaN across a range of mole fractions follows a linear trend with S^2 as predicted by Eq. (4), and hence is strongly influenced by the distribution of structural motifs dictated by specific combinations of S and x. We note that the extrapolated value of the perfectly ordered, stoichiometric, $S = 1$ structure corresponds to a negative band gap, caused by the high concentration of the InGa motif (characterized by an overall positive charge), which pushes down the conduction band. All of the compositions in Figs. 3(b) and 3(c) have a 50% nitrogen composition, and so relative to the ordered $S = 1$ state (in the Potts model this corresponds to 1:1:1 In:Ga:N) they are nitrogen rich and hence dominated by nitrogen-rich motifs. As in the binary materials, the nitrogen-rich motifs will push nitrogen p levels lower due to the stronger N-N bonds, and cause the band gap to increase with increasing disorder.

The top inset of Fig. 3(c) shows the variation of the $S = 0$ E_g as a function of the mole fraction x; the $S = 0$ E_g increases with x initially, until it peaks at $x = 0.5$, and then decreases until $x = 1$. This variation can be understood by considering the changing motif percentages at $S = 0$ as x

changes. As shown in the bottom inset of Fig. 3(c), at $x = 0$ (GaN) there are equal percentages of Ga-rich and N-rich motifs, and as x increases the fraction of nitrogen-rich motifs increases, which causes the band gap to increase. The maximum fraction of nitrogen-rich motifs occurs at $x = 0.5$, which corresponds to the largest $S = 0$ E_g band gap. As x increases to 1, the fraction of nitrogen-rich motifs decreases causing the band gap to decrease.

Consequently, this alloy (and any such alloy, for that matter) has a profoundly larger range of band-gap energies than is presently appreciated, accessible through combined tuning of cation ordering and alloy composition. This also provides an alternative explanation for the wide variation in reported emission wavelengths in InGaN for ostensibly similar mole fractions.

V. ELEMENTAL AND HYBRID TERNARY SEMICONDUCTORS

Returning to the Ising model, the basic framework can also be applied to elemental semiconductors if one spin is assigned to atoms of the element and the other spin to vacancies. Thus, for silicon in the $S = 1$ case, one of the atoms in the two-atom basis of the diamond lattice is replaced by a vacancy [Fig. 4(a), inset]. The motifs in the vacancy model framework for silicon then map to the same motifs as the binary materials, with the cation replaced by a Si atom and the anion by a vacancy, and thus exhibit the same band-gap dependencies

as a function of S and stoichiometry [Fig. 4(a)]. For single-layer graphene in the $S = 1$ case, one of the atoms in the two-atom basis of the planar hexagonal unit cell is replaced by a vacancy [Fig. 4(b), inset]; the carbon atoms are threefold coordinated and thus there are only eight possible motif types, as illustrated in Fig. 4(c). In both materials, vacancies cause the band gap to increase, and so as seen in Fig. 4 the extrapolated $S = 1$ cases for silicon and graphene have a larger band gap than the vacancy-free $S = 0$ end points. Consequently, both silicon and graphene (specifically nanoporous graphene) follow the predictions of Eq. (4) when vacancies are properly accounted for.

It should be noted that the solid triangles and open squares in Fig. 4(a) correspond to samples identified as amorphous silicon (a-Si) and all other points (solid squares, solid circles, open diamonds, open downward triangle) in Fig. 4(a) correspond to samples identified as porous silicon (p-Si). While these types of silicon have been previously treated separately, they both fit within the framework of the vacancy-based Ising model. In fact, there are reports of porous silicon, which has been regarded as having a higher degree of order than amorphous Si, exhibiting signatures of amorphous silicon in Raman spectra [51,52]. Additionally, while a-Si shows no discernible order, signs of short-range tetrahedral ordering of the diamond lattice have been observed in a-Si, and a-Si is generally regarded to have a latticelike structure [53]. Furthermore, many features of a-Si point to the presence of vacancies: a-Si is less dense than crystalline silicon; the average atomic coordination in amorphous silicon is less than 4 [54], and evidence points to the main defect in amorphous silicon being a threefold-coordinated Si atom with a dangling bond [53,54]. Consequently, viewed from the perspective of structural motifs which incorporate vacancies, there is no physical difference between amorphous and porous silicon in general, as indicated by the Ising model relationship. It also suggests that a more overt pathway to tuning the band-gap energy of silicon (and nanoporous graphene, as well as other elemental semiconductors) exists by the controlled introduction of disorder in terms of vacancies or potentially impurity atoms, at either doping or alloy levels.

The Ising model spins can also represent whole molecules on a lattice site, such as in hybrid perovskites, which are the focus of considerable recent interest for potential applications in photovoltaics. Figure 4(d) shows the results of such an analysis for methylammonium lead iodide (MAPbI$_3$) using order-parameter and band-gap values extracted from the literature [48]. In this case, one spin represents the methylammonium ($CH_3NH_3^+$, or MA) molecule and the second spin represents the Pb atom. As can be seen in Fig. 4(d), this material also follows the linear trend predicted by Eq. (4). Consequently, an Ising model of disorder can be invaluable in predicting the band gap of such materials for purposes of tuning the property to the precise value desired for a particular device application, and assist in the determination of which crystal processing parameters might be responsible for any observed sample-to-sample variation.

A final note is in order regarding the fitting of experimentally determined band-gap energies. As discussed, there are typically a wide range of reported values for any given material. This was noted by Burstein [20] and separately by Moss [21] for the particular case of InSb, with the variation attributed to significant band-filling effects arising from degenerate doping and a very light electron effective mass. However, the Burstein-Moss effect leads exclusively to an apparent increase in the band gap as the carrier concentration increases, and thus is insufficient to properly account for the full range of reported values. Subsequently, band-gap renormalization was proposed [22], which has the effect of reducing the electronic band gap through a rigid shift of either the conduction band, the valence band, or both [55]. Consequently, the two effects, arising from the same phenomenon (degenerate carrier populations), shift the band gap in opposite directions, and a mixture of the two can be used to match an experimentally observed trend of band-gap variation with carrier concentration. Both effects have been the subject of numerous investigations [56–59]. In this work, neither effect is necessary to match the experimentally observed band-gap values, including InSb [Fig. 4(e)].

VI. CONCLUSIONS

We have presented a spin-based model perspective on the physical mechanism which controls the value of band-gap energy in any semiconductor material, specifically, the variety and numerical distribution of structural motifs which define a particular lattice. The validity of the approach has been verified using experimentally determined values of band-gap energy and stoichiometry-dependent Bragg-Williams order parameter for a variety of different semiconductors, including the ternary heterovalent compound ZnSnN$_2$, binary compounds InN, GaN, and ZnO, the representative alloy InGaN, Si, graphene and the hybrid perovskite MAPbI$_3$. Consequently, focusing on exerting control over structural motif composition enables the realization of a previously (unappreciated) range of accessible band-gap energies, and an alternative route to engineering the critical parameter of a semiconductor for electronic device applications.

[1] R. D. L. Kronig, W. G. Penney, and R. H. Fowler, Quantum mechanics of electrons in crystal lattices, Proc. R. Soc. London, Ser. A **130**, 499 (1931).

[2] N. Holonyak and S. F. Bevacqua, Coherent (visible) light emission from $Ga(As_{1-x}P_x)$ junctions, Appl. Phys. Lett. **1**, 82 (1962).

[3] R. S. Morgan, Reciprocal ribose interactions: A possible structural motif in and between RNA'S, Biosystems **5**, 95 (1973).

[4] S. Bleckley and S. J. Schroeder, Incorporating global features of RNA motifs in predictions for an ensemble of secondary structures for encapsidated MS2 bacteriophage RNA, RNA **18**, 1309 (2012).

[5] C. Anderson, F. Zucker, and T. Steitz, Space-filling models of kinase clefts and conformation changes, Science **204**, 375 (1979).

[6] R. A. Makin and S. M. Durbin, Structural motifs, disorder, and the efficacy of viral vaccines, bioRxiv (2020), doi: 10.1101/2020.06.08.139907.

[7] K. P. Aichele, Paul Klee's "Rhythmisches": A recapitulation of the Bauhaus years, Z. Kunstgeschichte **57**, 75 (1994).

[8] B. Janssen, J. A. Burgoyne, and H. Honing, Predicting variation of folk songs: A corpus analysis study on the memorability of melodies, Front. Psychol. **8**, 621 (2017).

[9] C. Cray and G. Rowley, Chinese and western composition, College Art J. **15**, 6 (1955).

[10] R. A. Makin, K. York, S. M. Durbin, N. Senabulya, J. Mathis, R. Clarke, N. Feldberg, P. Miska, C. M. Jones, Z. Deng, L. Williams, E. Kioupakis, and R. J. Reeves, Alloy-Free Band Gap Tuning across the Visible Spectrum, Phys. Rev. Lett. **122**, 256403 (2019).

[11] J. M. Cowley, X-ray measurement of order in single crystals of Cu_3Au, J. Appl. Phys. **21**, 24 (1950).

[12] S. Nakatsuka and Y. Nose, Order-disorder phenomena and their effects on bandgap in $ZnSnP_2$, J. Phys. Chem. C **121**, 1040 (2017).

[13] D. B. Laks, S.-H. Wei, and A. Zunger, Evolution of Alloy Properties with Long-Range Order, Phys. Rev. Lett. **69**, 3766 (1992).

[14] S. Wei, D. B. Laks, and A. Zunger, Dependence of the optical properties of semiconductor alloys on the degree of long-range order, Appl. Phys. Lett. **62**, 1937 (1993).

[15] W. L. Bragg and E. J. Williams, The effect of thermal agitation on atomic arrangement in alloys, Proc. R. Soc. London, Ser. A **145**, 699 (1934).

[16] W. L. Bragg and E. J. Williams, The effect of thermal agitaion on atomic arrangement in alloys, Proc. R. Soc. London, Ser. A **151**, 540 (1935).

[17] E. J. Williams, The effect of thermal agitation on atomic arrangement in alloys-III, Proc. R. Soc. London, Ser. A **152**, 231 (1935).

[18] B. D. Cullity, *Elements of X-ray Diffraction* (Addison-Wesley, Boston, 1978).

[19] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.102.115202 for motif probability equations. The complete set of data and analysis for S^2 measurements are available at https://doi.org/10.5281/zenodo.3904743. For band gaps measured by our group the data (Tauc plot or PL) are available in the repository, for samples from the literature the band gap reported in the paper along with relevant table or figures in the manuscript are provided along with the

manuscript doi. Note: That error bars for S^2 values and band-gap measurements (where available or performed by our group) are within the size of points on all graphs.

[20] E. Burstein, Anomalous optical absorption limit in InSb, Phys. Rev. **93**, 632 (1954).

[21] T. S. Moss, The interpretation of the properties of indium antimonide, Proc. Phys. Soc., Sect. B **67**, 775 (1954).

[22] J. Slotboom and H. de Graaff, Measurements of bandgap narrowing in Si bipolar transistors, Solid-State Electron. **19**, 857 (1976).

[23] J. R. Dixon and R. F. Bis, Band inversion and the electrical properties of $Pb_xSn_{1-x}Te$, Phys. Rev. **176**, 942 (1968).

[24] J. Ma, H.-X. Deng, J.-W. Luo, and S.-H. Wei, Origin of the failed ensemble average rule for the band gaps of disordered nonisovalent semiconductor alloys, Phys. Rev. B **90**, 115201 (2014).

[25] T. D. Veal, C. F. McConville, and W. J. Schaff, *Indium Nitride and Related Alloys* (CRC Press, Boca Raton, FL, 2011).

[26] W. Walukiewicz, S. Li, J. Wu, K. Yu, J. Ager, E. Haller, H. Lu, and W. J. Schaff, Optical properties and electronic structure of InN and In-rich group III-nitride alloys, J. Cryst. Growth **269**, 119 (2004).

[27] A. G. Bhuiyan, K. Sugita, K. Kasashima, A. Hashimoto, A. Yamamoto, and V. Y. Davydov, Single-crystalline InN films with an absorption edge between 0.7 and 2 eV grown using different techniques and evidence of the actual band gap energy, Appl. Phys. Lett. **83**, 4788 (2003).

[28] A. Mang, K. Reimann, and S. Rübenacke, Band gaps, crystal-field splitting, spin-orbit coupling, and exciton binding energies in ZnO under hydrostatic pressure, Solid State Commun. **94**, 251 (1995).

[29] E. Kurimoto, M. Hangyo, H. Harima, M. Yoshimoto, T. Yamaguchi, T. Araki, Y. Nanishi, and K. Kisoda, Spectroscopic observation of oxidation process in InN, Appl. Phys. Lett. **84**, 212 (2004).

[30] D. B. Haddad, J. S. Thakur, V. M. Naik, G. W. Auner, R. Naik, and L. E. Wenger, Optical band gap measurements of InN films in the strong degeneracy limit, MRS Proc. **743**, L11.22 (2002).

[31] T. Inushima, V. Mamutin, V. Vekshin, S. Ivanov, T. Sakon, M. Motokawa, and S. Ohoya, Physical properties of InN with the band gap energy of 1.1 ev, J. Cryst. Growth **227-228**, 481 (2001).

[32] D. E. Lacklison, J. W. Orton, I. Harrison, T. S. Cheng, L. C. Jenkins, C. T. Foxon, and S. E. Hooper, Band gap of GaN films grown by molecular-beam epitaxy on GaAs and GaP substrates, J. Appl. Phys. **78**, 1838 (1995).

[33] S.-H. Cho, Effects of growth temperature on the properties of ZnO thin films grown by radio-frequency magnetron sputtering, Trans. Elect. Electron. Mater. **10**, 185 (2009).

[34] R. B. Potts, Some generalized order-disorder transformations, Math. Proc. Cambridge Philos. Soc. **48**, 106 (1952).

[35] J. W. Ager, W. Walukiewicz, W. Shan, K. M. Yu, S. X. Li, E. E. Haller, H. Lu, and W. J. Schaff, Multiphonon resonance Raman scattering in $In_xGa_{1-x}N$, Phys. Rev. B **72**, 155204 (2005).

[36] Q. Guo, Y. Kusunoki, Y. Ding, T. Tanaka, and M. Nishio, Properties of InGaN films grown by reactive sputtering, Jpn. J. Appl. Phys. **49**, 081203 (2010).

[37] V. Y. Davydov, A. A. Klochikhin, V. V. Emtsev, A. N. Smirnov, I. N. Goncharuk, A. V. Sakharov, D. A. Kurdyukov, M. V. Baidakova, V. A. Vekshin, S. V. Ivanov, J. Aderhold, J. Graul,

A. Hashimoto, and A. Yamamoto, Photoluminescence and Raman study of hexagonal InN and In-rich InGaN alloys, Phys. Status Solidi B **240**, 425 (2003).

[38] S. R. Meher, K. P. Biju, and M. K. Jain, Raman spectroscopic investigation of phase separation and compositional fluctuations in nanocrystalline In$_x$Ga$_{1-x}$N thin films prepared by modified activated reactive evaporation, Phys. Status Solidi A **208**, 2655 (2011).

[39] L. Robins, A. Paul, C. Parker, J. Roberts, S. Bedair, E. Piner, and N. El-Masry, Optical absorption, Raman, and photoluminescence excitation spectroscopy of inhomogeneous InGaN films, MRS Proc. **537**, G3.22 (1998).

[40] A. P. Sokolov, A. P. Shebanin, O. A. Golikova, and M. M. Mezdrogina, Structural disorder and optical gap fluctuations in amorphous silicon, J. Phys.: Condens. Matter **3**, 9887 (1991).

[41] A. K. Sood, K. Jayaram, and D. V. S. Muthu, Raman and high-pressure photoluminescence studies on porous silicon, J. Appl. Phys. **72**, 4963 (1992).

[42] R. Tsu, H. Shen, and M. Dutta, Correlation of Raman and photoluminescence spectra of porous silicon, Appl. Phys. Lett. **60**, 112 (1992).

[43] S. M. Prokes and O. J. Glembocki, Role of interfacial oxide-related defects in the red-light emission in porous silicon, Phys. Rev. B **49**, 2238 (1994).

[44] B. Fogal, S. O'Leary, D. Lockwood, J.-M. Baribeau, M. Noël, and J. Zwinkels, Disorder and the optical properties of amorphous silicon grown by molecular beam epitaxy, Solid State Commun. **120**, 429 (2001).

[45] D. J. Lockwood and A. G. Wang, Quantum confinement induced photoluminescence in porous silicon, Solid State Commun. **94**, 905 (1995).

[46] T. V. Cuong, V. H. Pham, Q. T. Tran, S. H. Hahn, J. S. Chung, E. W. Shin, and E. J. Kim, Photoluminescence and Raman studies of graphene thin films prepared by reduction of graphene oxide, Mater. Lett. **64**, 399 (2010).

[47] A. K. Geim and K. S. Novoselov, The rise of graphene, Nat. Mater. **6**, 183 (2007).

[48] V. D'Innocenzo, A. R. Srimath Kandada, M. De Bastiani, M. Gandini, and A. Petrozza, Tuning the light emission properties by band gap engineering in hybrid lead halide perovskite, J. Am. Chem. Soc. **136**, 17730 (2014).

[49] V. Dixit, B. Rodrigues, H. Bhat, R. Venkataraghavan, K. Chandrasekaran, and B. Arora, Growth of InSb epitaxial layers on GaAs (001) substrates by LPE and their characterizations, J. Cryst. Growth **235**, 154 (2002).

[50] V. Senthilkumar, S. Venkatachalam, C. Viswanathan, S. Gopal, S. K. Narayandass, D. Mangalaraj, K. C. Wilson, and K. P. Vijayakumar, Influence of substrate temperature on the properties of vacuum evaporated InSb films, Cryst. Res. Technol. **40**, 573 (2005).

[51] J. M. Perez, J. Villalobos, P. McNeill, J. Prasad, R. Cheek, J. Kelber, J. P. Estrera, P. D. Stevens, and R. Glosser, Direct evidence for the amorphous silicon phase in visible photoluminescent porous silicon, Appl. Phys. Lett. **61**, 563 (1992).

[52] H. Tanino, A. Kuprin, H. Deai, and N. Koshida, Raman study of free-standing porous silicon, Phys. Rev. B **53**, 1937 (1996).

[53] M. H. Brodsky and R. S. Title, Electron Spin Resonance in Amorphous Silicon, Germanium, and Silicon Carbide, Phys. Rev. Lett. **23**, 581 (1969).

[54] K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J. L. Robertson, J. Wang, and S. C. Moss, High Resolution Radial Distribution Function of Pure Amorphous Silicon, Phys. Rev. Lett. **82**, 3460 (1999).

[55] B. Arnaudov, T. Paskova, P. P. Paskov, B. Magnusson, E. Valcheva, B. Monemar, H. Lu, W. J. Schaff, H. Amano, and I. Akasaki, Energy position of near-band-edge emission spectra of InN epitaxial layers with different doping levels, Phys. Rev. B **69**, 115216 (2004).

[56] J. S. Manser and P. V. Kamat, Band filling with free charge carriers in organometal halide perovskites, Nat. Photonics **8**, 737 (2014).

[57] K. F. Berggren and B. E. Sernelius, Band-gap narrowing in heavily doped many-valley semiconductors, Phys. Rev. B **24**, 1971 (1981).

[58] H. Kalt and M. Rinker, Band-gap renormalization in semiconductors with multiple inequivalent valleys, Phys. Rev. B **45**, 1139 (1992).

[59] L. Kong, G. Liu, J. Gong, Q. Hu, R. D. Schaller, P. Dera, D. Zhang, Z. Liu, W. Yang, K. Zhu, Y. Tang, C. Wang, S.-H. Wei, T. Xu, and H.-k. Mao, Simultaneous band-gap narrowing and carrier-lifetime prolongation of organic–inorganic trihalide perovskites, Proc. Natl. Acad. Sci. USA **113**, 8910 (2016).