Ordering in GaAs co-doped with Bi and N

T.J.F. Verstijnen,[1,*] D. Tjeertes,[1] A. D. Rice,[2] K. Alberi,[2] and P. M. Koenraad[1]

[1]*Department of Applied Physics and Science Education, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

[2]*National Renewable Energy Laboratory, Golden, CO 80401, United States of America*

(Dated: April 17, 2024)

Introducing only a few atomic percent of Bi or N in GaAs has a large effect on the band gap of the material. Specifically Bi doped GaAs shows potential for local band gap engineering in optoelectronic applications. The incorporation of Bi and N into GaAs is difficult due to strain effects. In this work we study the ordering of these dopants at the atomic scale in order to get a better understanding of the behavior of these dopants in the host lattice. Cross-sectional scanning tunneling microscopy (X-STM) is used to find the exact position of Bi and N dopants in the GaAs matrix, allowing us to study both their nearest neighbor pair occurrences and pair correlation functions. An attractive interaction between Bi dopants at short ranges (1-2 nm) is found and a similar effect is observed between N dopants. We find a repulsive interaction with a similar length scale between Bi and N dopants. A similar repulsion is found in the Bi-N nearest neighbor pairs. Density functional theory (DFT) is used to calculate the different nearest neighbor pair energies and test these results to the experimental pair occurrences. It is concluded from the experimental and theoretical results that the growth conditions and N inclusion greatly affects the Bi distribution in GaAs.

I. INTRODUCTION

Introducing dilute amounts of Bi in GaAs has drastic effects on the band structure. The band gap is reduced by 400 meV when 5% of Bi is introduced, due to the valence band maximum (VBM) shifting to higher energy[1]. At the same time the split-off band is pushed down, increasing the split-off energy with increasing Bi content[2]. This results in a cross-over point at a Bi concentration around 10-11 %, where the split-off energy becomes larger than the band gap[3]. Reaching this concentration is an important milestone for the development of efficient (near)-infrared lasers, since Auger recombination will be suppressed, theoretically making lasers much more energy efficient. The band structure reordering of GaAs with the introduction of Bi is described by the valence band anti-crossing (VBAC) model[4]. Unfortunately, (homogeneous) Bi incorporation in GaAs is difficult to achieve and requires very specific growth conditions in MBE[5]. Sub-optimal growth conditions can lead to formation of Bi droplets on the surface and inhomogeneous incorporation[5–8], drastically reducing the optical quality of the material. Inhomogeneous incorporation can happen at the atomic level, through the formation of nearest neighbor Bi pairs, or at a larger scale, with the formation of Bi-rich and Bi-poor regions. An important aspect of the incorporation of Bi in GaAs is the size difference between Bi and the As it replaces. Because of this, the Bi atom induces a local compressive strain in the lattice. The alloy created in this way is part of the highly mismatched alloys (HMA's).

Introduction of N atoms in GaAs also reduces the band gap of GaAs drastically. A small amount of N can cause a band gap reduction of 600 meV[9–11], even though the band gap of GaN is 3.4 eV, much larger than the 1.48 eV band gap of GaAs. Contrary to Bi in GaAs, N causes the conduction band minimum (CBM) to move downwards. This effect is also described by the band anti-crossing model[12]. N atoms also induce a local strain, but because the N atom is much smaller than the As atom, the strain is tensile. Combining Bi and N doping in GaAs probably enables individual control of the VBM and CBM. In this way the band gap can be engineered to reach desired parameters. Possible applications for GaAs with Bi and N co-doping lie in the domain of near- and mid-infrared devices[13]. For these purposes a homogeneous doping distribution is required, since Bi or N localization results in low optical quality. Due to the opposite strain behavior of Bi and N it is also possible to grown strain free on GaAs with certain ratios of Bi and N. A previous study with cross-sectional scanning tunneling microscopy (X-STM) on Bi in InP showed a enhanced presence of first and second nearest neighbor Bi pairs and a local clustering range of about 2 nm[14]. For N atoms in GaAs a strongly reduced occurrence of first and fourth nearest neighbor pairs has been observed with X-STM[15]. Adding an atomic Bi flux during the growth of Bi and N doped GaAs has been shown to enhance the N incorporation[13] and atomic scale ordering of Bi and N in GaAs has been observed by transmission electron microscopy (TEM)[16]. In this work we study GaAs co-doped with Bi and N with X-STM and density functional theory (DFT). This allows us to directly observe Bi and N atoms in the GaAs (110) surface and to calculate relative pair energies. We analyze the distribution of these atoms, both on a nearest neighbor level and longer ranges up to 10 nm.

II. METHODS

A. X-STM measurements

Several Bi and N doped GaAs samples were grown by molecular beam epitaxy (MBE). All samples were grown on n^+-doped GaAs (001) substrates by and contained multiple layers with Bi and/or N concentrations varying from 0.3% to 2%. N and Bi growth conditions were determined from samples containing single layers and lattice constants obtained via X-ray diffraction. The data analyzed in this work is collected from two samples referred to as sample 1 and sample 2, respectively. The X-STM measurements were performed using

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

a commercial omicron LT-STM system. The measurements were performed at a temperature of 77K. Additional contacts were deposited on the samples made from Sn, Au and Ge to improve electronic contact while measuring. The samples were brought into the system and annealed at a temperature of >150 °C to ensure vacuum quality. Then the samples were brought into the ultra high vacuum STM chamber, $<10^{-11}$ mbar, and cleaved along the (110) natural cleavage direction. This exposes a clean, atomically flat cross-section of the samples on which the X-STM measurements were conducted. The X-STM tips were made by electro-chemical etching of poly crystalline W wire. After loading the tips into the STM system they are also annealed at <170 °C and sputtered with an Ar ion gun to improve tip quality.

B. DFT calculations

The DFT calculations performed in this work are made using the Vienna Ab Initio Simulation Package (VASP)[17]. Perdew-Burke-Ernzerho (PBE) functionals[18] are used, a type of Generalized Gradient Approximation (GGA) pseudopotentials. The electron-ion interaction is described using the Plane Augmented Wave (PAW) method[19]. In this work we consider 512 atom bulk GaAs supercells consisting of 4x4 conventional unit cells. This size is chosen to exclude inter supercell crosstalk. Using the aforementioned pseudopotentials and supercells, we find a GaAs lattice constant of 5.75. This is an overestimation of the lattice constant which is expected for PAW-PBE pseudopotentials[20]. A plane wave cutoff energy of 400 eV was found to be sufficient for these cells after convergence testing. The sampling of the Brillouin Zone was done using a Monkhorst-Pack grid[21] of 2x2x2. Different dopant atom pairs are introduced into these GaAs cells and the cells are left to relax until the force on the individual atoms is less than 0.05 eV/Å. The total energies of these different cells can then be compared in order to predict which dopant pairs are likely to appear in our X-STM images.

C. Notation and theory

Single Bi and N atoms in different layers in the surface will be indicated by the element followed by a hyphen and the number of the layer that they are located in, for example Bi-1 for a Bi atom in the first layer below the surface. Nearest neighbor pairs will be marked by the two elements involved (BiBi, BiN or NN) followed by a subscript number indicating the nearest neighbor order of the pair, for example a 2nd nearest neighbor pair of a Bi atom and an N atom will be indicated by BiN_2. To avoid confusion with a nitrogen nitrogen pair, which is noted as NN_n, nearest neighbor will not be abbreviated to NN, instead it is written in full.

To analyze the distribution on a longer scale, we use a pair correlation study. The pair correlation function (PCF) or radial distribution function gives the probability of finding an atom at a certain distance from another atom, as given by the following equation:

$$g(r) = \frac{dn(r)}{\rho * 2\pi r dr} \tag{1}$$

where $dn(r)$ is a function that describes the total number of particles found at a distance r from another particle, ρ is the average density and dr is the bin size. The PCF will be zero up to the minimum distance between to atoms in the lattice since no atoms can be present. For a completely random distribution of atoms the PCF will be 1 after the nearest neighbor distance. Deviations from 1 will indicate a higher (>1) or lower (<1) occurrence of other atoms at that specific distance.

D. Analysis

During measurements large scale, typically 60×80 nm, STM images are taken of the different layers in the sample. To obtain proper statistics on the nearest neighbor pairs, large (2×10^4 nm^2) areas have to be imaged, depending on the concentration of doping atoms in the layer. These images are stitched together and analyzed by hand, marking the location of Bi and N atoms and nearest neighbor pairs. Using the total area of the STM images and the number of Bi and N atoms counted, the concentration of these atoms is calculated and compared to the value estimated by XRD. The concentration is used to calculate the expected amount of nearest neighbor pairs by considering the [110] surface as a lattice gas. In a limiting case, where no interaction is present between the particles the probability that two neighboring sites are occupied is given by the square of the concentration. Scaling this to the total number of lattice sites gives the total amount of expected nearest neighbor pairs. This result has been confirmed through the generation of random lattices and extracting the amount of nearest neighbor pairs therein. By dividing the observed number of pairs with expected number in a random distribution, the relative occurrence of this specific pair is calculated. A number higher than 1 means the specific pair occurs more frequently than in a random distribution, lower than 1 means it occurs less frequently.

A typical measurement set of a single layer spans 400 to 800 nm with a width of 40 nm. All the atoms in the image are detected and given a coordinate. Since the data set is finite, edge effects need to be considered before calculating the PCF. An atom near the edge will have fewer neighbors since no atoms are present beyond the edge of the data. To account for this, atoms within the PCF range (typically 10 nm) with respect to the edge of the data set are excluded from the set of origins for the PCF calculations. The distances between the origin atoms to all other atoms in the data set are calculated. These distances are then binned for a certain dr and the PCF is calculated for each bin, resulting in a PCF plot. Image processing and data analysis are performed using NumPy[22] and OpenCV[23] in Python.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

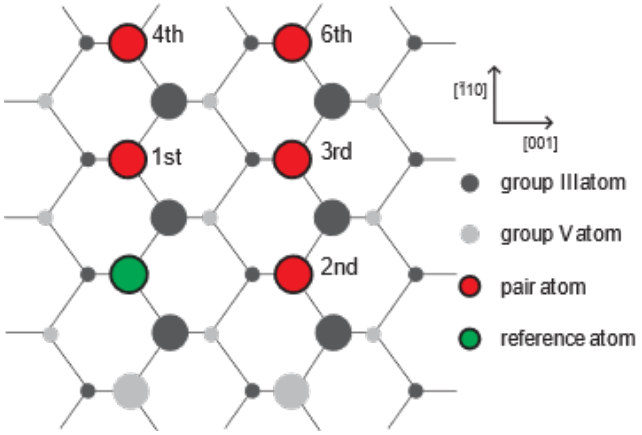

FIG. 1. Schematic depiction of the (110) surface of GaAs, with group III and group V atom in dark grey and light grey, respectively. With respect to a reference group V atom (green) different nearest neighbor group V atoms are indicated.

E. Nearest neighbor pairs

We will consider the distribution of dopant atoms in the topmost surface layer in this work. The different nearest neighbor pairs in the (110) surface of III-V semiconductor are schematically depicted in Fig 1. A group V reference atom, from which the different pairs are defined, is colored green. The first nearest neighbor group V atom is located in the $[\bar{1}10]$ direction at a distance of $a/\sqrt{2}$. A different first nearest neighbor atom is located in the $[1\bar{1}0]$ direction, so each group V atom has two first nearest neighbor group V atoms in the (110) surface. The second nearest neighbor is located in the $[001]/[00\bar{1}]$ direction, at a distance of a. The third and sixth nearest neighbor pairs both have four equivalent locations and will occur twice as often in a random distribution as the first, second and fourth nearest neighbor pairs. The fifth nearest neighbor location is located two atomic layers below the second nearest neighbor location, which makes it difficult to observe in experiments.

III. RESULTS

Fig. 2 displays a typical filled state STM image of a Bi and N doped GaAs layer of sample 2. Atomic corrugation lines can be seen running along the $[\bar{1}10]$ direction. Bi atoms are identified as bright protrusions of the surface at a single atomic site. This is caused by the fact that Bi atoms are significantly larger than the As atoms they replace. The brightest Bi atoms are located in the surface layer and are indicated by Bi-0[14,24], Bi atoms two layers below the surface, as marked by Bi-2, show a much weaker contrast. N atoms, instead, are imaged as dark depressions of a single atomic site of the surface, since they are smaller than the As atom that is replaced. For N atoms, the darkest ones are located in the surface and indicated by N-0, the less deep ones are located two layers below the surface[25–27] and an example is marked with N-2. The electronic contrast related to these dopants is suppressed at high negative biases. This contrast is related to the shape of the wave function of Bi and N dopants, is more extended and would make the identification of the exact position of these dopants more challenging. This more extended electronic contrast is present both at positive bias and at smaller negative bias[2829]. Because of this, the X-STM images considered in this work are made with gap voltages between -3V and -3.7V. An example of a $BiBi_2$ pair is marked in Fig. 2. A typical surface defect observed in these samples is marked with the letter D. This is a row of vacancies in the $[\bar{1}10]$ direction, created during cleaving. In layers with higher Bi or N concentration these defects occur more often and extend for longer distances. This is most probably caused by the higher strain in the layer affecting the cleaving process. As a result, it is unfortunately not possible to obtain reliable statistics on these high-concentration layers. The following sections will show and discuss the results of the nearest neighbor pair analysis.

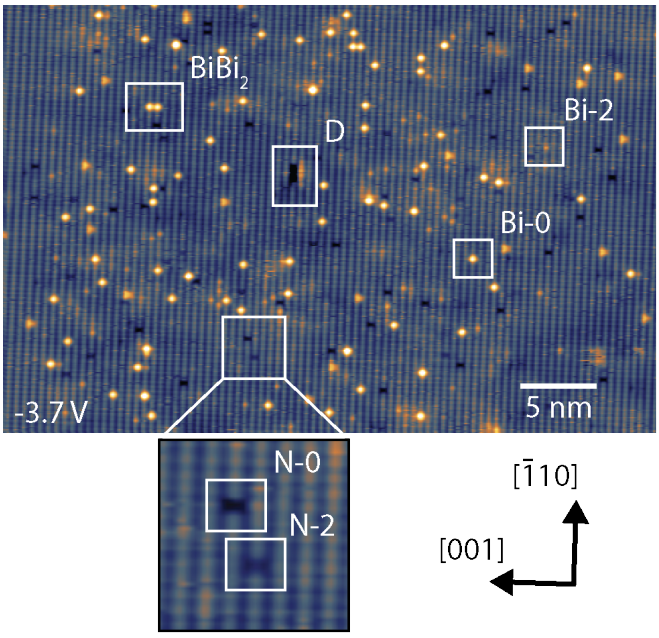

FIG. 2. Filled state STM image obtained at −3.7 V and 120 pA. A surface Bi atom is marked with Bi-0 and a Bi atom two layers below the surface by Bi-2. A 2nd nearest neighbor Bi pair is marked with $BiBi_2$. The white rectangle marks the area that is shown zoomed in underneath the main image. Here an N atom in the surface and two layers below the surface are marked with N-0 and N-2, respectively. A cleaving induced defect is marked with D.

A. $BiBi_n$ pairs in GaAsBi layers

A nearest neighbor pair analysis was performed on a multitude of layers of both sample 1 and sample 2. Fig. 3(a) shows the nearest neighbor analysis of a Bi-doped layer in sample 1 with a Bi concentration of 0.8% in orange. The red line indicates the expected pairs for a random distribution. In sample 1 $BiBi_1$ occurs more than twice as often compared to a ran-

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

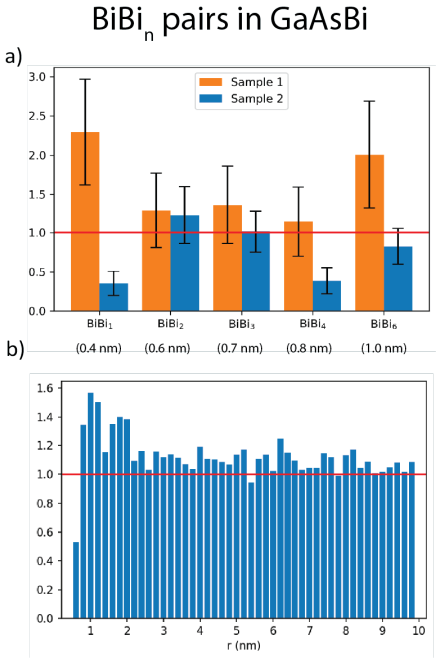

FIG. 3. (a) Relative occurrence of nearest neighbor BiBi pairs in Bi-doped layers in sample 1 and 2. (b) PCF of Bi atoms in a Bi-doped layer of sample 2. The red line in (a) and (b) indicates the amount expected for a completely random distribution.

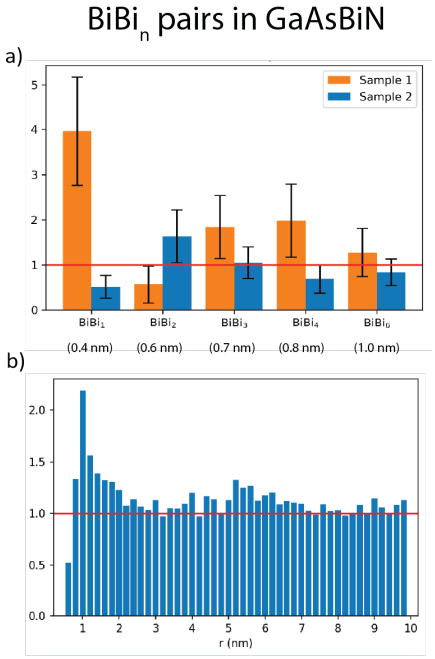

FIG. 4. (a) Relative occurrence of nearest neighbor BiBi pairs in Bi- and N-doped layers in sample 1 and 2. (b) PCF of Bi atoms in a Bi- and N-doped layer of sample 2. The red line in (a) and (b) indicates the amount expected for a completely random distribution.

dom distribution, showing a clear tendency for the formation of these pairs during the growth. $BiBi_2$, $BiBi_3$ and $BiBi_4$ all occur slightly more, but looking at the uncertainty, this could be due to the limited statistics. The uncertainty in the relative occurrence is dominated by the limited number of pairs observed. Even in a large area of 3.6×10^4 nm^2, that takes multiple hours to measure, around 50 $BiBi_3$ pairs are observed. Assuming the observation is uncorrelated the distribution of pairs is Poissonian and the uncertainty in the pairs scales with $\sqrt{N}$, where N is the amount of pairs observed. $BiBi_6$ shows a clear enhancement, occurring almost twice as often. Overall, the Bi atoms in this layer of sample 1 show an enhanced tendency to form nearest neighbor pairs. The analysis of a Bi-doped layer of sample 2 is displayed in blue, the layer contains 1.9% Bi. Here, a very different distribution is observed. $BiBi_1$ and $BiBi_4$ are strongly suppressed, occurring less than half as often as expected for a random distribution. The other pairs occur in numbers close to their expected values.

There is a large difference in the observed $BiBi_n$ pairs in sample 1 and sample 2, especially for $BiBi_1$, $BiBi_4$ and $BiBi_6$. Even though both graphs are corrected for the total concentration, the difference in Bi concentration could be one of the causes. Unfortunately, the two samples did not have a GaAsBi layer with the same Bi concentration in order to make a direct comparison.

The PCF of Bi atoms in a GaAsBi layer of sample 2 is displayed in Fig. 3(b), where the bin size of the PCF is 0.2 nm. The first bin is empty since the 1st nearest neighbor distance is $a/\sqrt{2}$, about 0.4 nm for GaAs. Between 0.8 and 2 nm a clear enhancement of the PCF is observed above the random distribution. This is quite similar to previous observations for Bi in InP[14]. At distances larger than 2.0 nm, the distribution quickly comes close to a random distribution, although it is consistently slightly higher. An observation like this points to the formation of clusters with a higher Bi concentration, with a radius of about 2 nm. The PCF at very small distances cannot be compared to the relative occurrences, since multiple pair types can fall into one bin and are averaged over the full size of the bin. Since pair distances are discreet at this scale, the relative occurrence of the different pairs as shown in Fig. 3(a) is the correct way to analyze this.

B. $BiBi_n$ pairs in GaAsBiN layers

The relative occurrence of $BiBi_n$ in the layers co-doped with both Bi and N of sample 1 and 2 is displayed in Figure 4(a). In sample 1 the layer contains 0.5% Bi and 0.1% N, in sample 2 this is 0.7% Bi and 0.5%N. Again, a large enhancement of the $BiBi_1$ pairs is observed for sample 1. For sample 2 a reduction of these pairs is observed instead. All other pairs show slight enhancements and reductions but all of them are in range of a random distribution when taking into account the uncertainty range.

Comparing these results with the ones obtained for $BiBi_n$ pairs in Bi-doped layers shows very little difference beyond statistical variation. The main difference between the $BiBi_n$ pairs in the GaAsBi and GaAsBiN layers are the amount of $BiBi_6$ pairs observed in sample 1, but this can almost completely be attributed to statistical variation, as can be seen

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

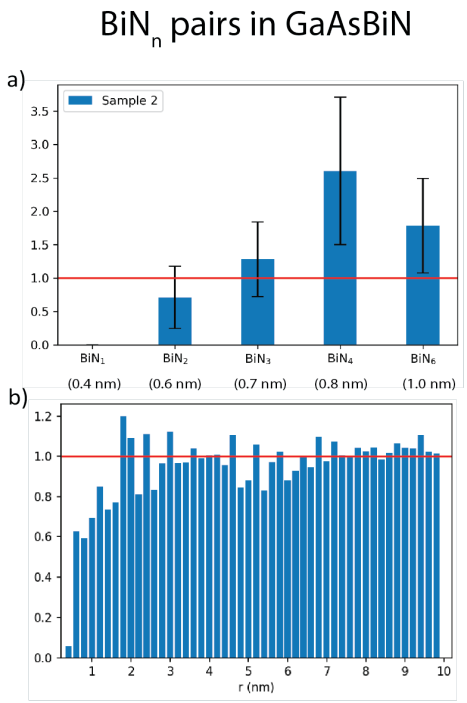

FIG. 5. (a) Relative occurrence of nearest neighbor BiN pairs in Bi- and N-doped layers in sample 2. (b) PCF of BiN atoms in a Bi- and N-doped layer of sample 2. The red line in (a) and (b) indicates the amount expected for a completely random distribution. The Bi concentration in the layer is 0.7% and the N concentration is 0.5%.

from the error bars.

The PCF results for the $BiBi_n$ are displayed in Fig. 4(b). Again, we see low correlation at the length scales below the first nearest neighbor distance. A clear enhancement is seen between 0.8 and 2.0 nm, similar to $BiBi_n$ in GaAsBi. At longer length scales, between 4 and 7 nm a second enhancement can be seen, which peaks around 5.5 nm. Beyond this the distribution follows a random distribution. Again, the enhancement between 0.8 to 2.0 nm points to small clusters with higher Bi concentration. The peak around 5.5 might indicate a weak secondary ordering at longer length scales. This peak can be interpreted as the typical distance between Bi clusters. Such a spread of clusters with a typical inter-cluster distance of 5.5 nm would result in a second peak in the PCF as seen in Fig 4(a). Regardless of the exact interpretation, the comparison between 3(b) and 4(b) shows that the Bi incorporation at longer scales is different in Bi-doped layer compared to Bi and N co-doped layers.

C. BiN_n pairs in GaAsBiN layers

The maximum N concentration in sample 1 is 0.1%, which results in an extremely low amount of BiN_n pairs observed. The analysis of these results contains very large uncertainty ranges, making it not significant enough to be included. In sample 2 the N concentration in the layers ranged from 0.3% to 0.6% providing enough pairs to make a statistically significant analysis.

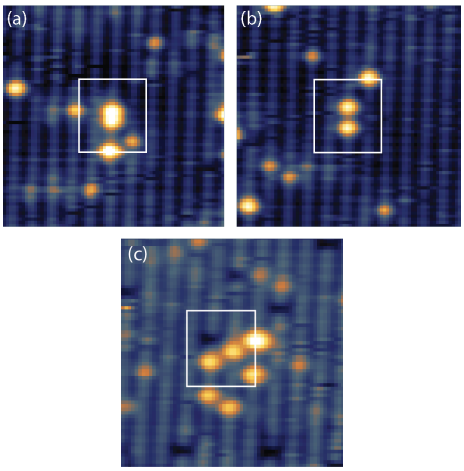

FIG. 6. 5×5 nm^2 STM images of GaAsBi and GaAsBiN. In (a), GaAsBi containing a $BiBi_1$ pair and in (b), GaAsBi containing a $BiBi_4$ pair (b). In (c), GaAsBiN containing both a $BiBi_3$ and $BiBi_4$ pair. The white rectangle in each image indicates the exact location of the pair(s).

In Fig. 5(a) the relative occurrence of BiN_n pairs is displayed. One detail that is immediately visible is that no BiN_1 pairs are observed. $BiBi_2$ and $BiBi_3$ occur about as much as expected for a random distribution. BiN_4, instead, occur more than twice as often and the observation of BiN_6 is also enhanced. Multiple aspects need to be considered in regard to the fact that no BiN_1 pairs are observed and BiN_4 is strongly enhanced.

First of all, the classification of N atoms is more difficult than of the Bi atoms. The difference between a surface N and an N atom two layers below the surface is small and is strongly dependent on the contrast of the STM image. In very high quality images, they are easy to distinguish, but if the image quality is slightly lower this becomes quite difficult. As a result, some N atoms in layer 2 could have been classified as N atoms in the surface. This would affect the statistics, increasing the number of observed pairs.

Secondly, the orientation of BiN_1 and BiN_4 is the same, both lie perpendicular to the growth direction, as explained in Fig. 1. The difference being the spacing between the two atoms, for BiN_4 one group V site lies in between the Bi and N atom. A possible explanation for the amount of BiN_1 and BiN_4 pairs observed, is that the manual classification was incorrect and all BiN_1 have been marked as BiN_4 instead. We argue that this is unlikely, since in our measurements individual atomic sites are visible. As a result, the difference between a BiN_1 and BiN_4 pair should be clearly observable. This is confirmed by comparing $BiBi_1$ and $BiBi_4$ pairs as displayed in Fig. 6. In Fig. 6(a) the $BiBi_1$ can be observed as two Bi atoms with no separation, since they are so close together they can almost not be distinguished as two separate atoms anymore. The $BiBi_4$ pair in Fig. 6(b) has a clear separation between the two Bi atoms. This shows that distinguishing between $BiBi_1$ and $BiBi_4$ can be reliably done and as an extension also be-

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

tween BiN_1 and BiN_4. Since no BiN_1 pairs were observed, such a direct comparison cannot be made between BiN_1 and BiN_4 pairs. In Fig. 6(c), a BiN_4 pair is shown as well as a BiN_3 pair that shares the same N atom. These two pairs are an indication of what a BiN_1 pair would look like compared to a BiN_4 pair since the pairing Bi atom in the BiN_3 pair has equal spacing to the N atom along the $[1\bar{1}0]$ direction as the BiN_1 pair would have.

In Fig. 5(b) the PCF is plotted for the BiN_n pairs in a GaAsBiN layer of sample 2. At distances below 2 nm a strong reduction of the PCF is observed, around 3 and 4 nm the distribution is close to random and around 5 and 6 nm again a reduction is observed. At larger distances the PCF follows a random distribution. The reduction at distances below 2 nm points to an unexpected repulsion between Bi and N atoms, forming separated Bi and N rich areas. The reduction around 5 nm correlates with the peak observed in the PCF of $BiBi_n$ pairs in GaAsBiN of Fig. 4(b). If clusters of higher Bi concentration are present at this distance, less N atoms will be present, resulting in the dip observed in Fig. 5(b).

D. NN_n pairs in GaAsBiN layers

The concentration of N atoms in the layers of sample 1 and 2 is about 0.6% at most. This results in a low number of observed pairs with a large uncertainty range. Combined with the fact that N atoms in the surface and layer 2 are harder to distinguish, makes the analysis not significant enough to be considered. From the results that we could obtain we observed no clear trends, with large variations of the NN_n pair statistics between the layers. For the BiN_n pairs the low concentration of N atoms was compensated by the high Bi concentration, making it possible to do an analysis on those. A future experiment on a sample with higher N concentration would help to make the analysis more reliable. We have found that the layers with high Bi and especially high N concentration can cleave poorly making these layers difficult to measure by X-STM.

E. Growth effects

We expect several of the observations made in these measurement to be related to the growth of the samples. First, the fact that the nearest neighbor statistics vary strongly between sample 1 and 2. Especially, the $BiBi_1$ pairs, that show strong enhancement in sample 1 and a strong reduction in sample 2. These sample were grown in different batches with maintenance on the MBE in between. Although calibration growths are performed before each proper growth batch, ideal growth conditions might have changed. The reflection high-energy electron diffraction (RHEED) system was also not functioning properly during the growth, which made observations of the reconstruction of the growth surface very difficult. Regardless it is very well possible that the two samples have been grown with a different surface reconstruction, which is known to affect the Bi incorporation and strongly depends on the exact growth conditions[13]. Work from Masnadi-Shirazi *et al.* shows that for an As_2:Ga flux of 0.5-1.0, there are at least four different surface reconstructions possible in a 100 °C range[30]. A second observation is the similarities and differences in the Bi distributions for the layers with and without N atoms. As can be seen when comparing Fig. 3(a) and 4(a), the nearest neighbor distribution does not change much between layer with and without N atoms, both in sample 1 and sample 2. But when looking at the PCFs in Fig. 3(b) and 4(b) a difference is observed at longer length scales. This enhancement of the PCF around 5.5 nm is attributed to the formation of small Bi rich areas with an inter-area distance of typically 5.5 nm. The formation of these Bi rich areas does not affect the nearest neighbor pair statistics. A possible cause for this is the strain effect that Bi and N incorporation has on the GaAs host. Bi increases the lattice constant of the host when it is incorporated. As a result, an area with more Bi in it might prove as a easier place for more Bi atoms to incorporate, since the lattice constant is already increased at the local scale. For N atoms it could be harder to incorporate in these areas, since they decrease the local lattice constant. Local alloy fluctuations like this have been observed before by X-STM[31]. Lastly, the first and fourth nearest neighbor pairs are expected to be influenced the most by the growth conditions since both are located on top of the (001) growth surface. This is an additional explanation for the behavior of the BiN_1 and BiN_4 pairs shown in Fig. 5(a). The interaction between the Bi and N atom in the growth surface can have a great effect on the formation of the two pairs. If the interaction between the Bi and N atom is repulsive at the very local scale, it could result in Bi or N atoms incorporating as BiN_4 pairs instead of BiN_1 pairs, doubling the observed BiN_4 pairs.

F. DFT calculations

DFT calculations of the various different dopant pairs discussed in this work were also performed. This is done to compare the experimental results to a theoretical bulk case and to obtain the relative energies of the various dopant pairs. The supercells described in section II B are used to calculate the energies of bulk GaAs supercells with different $BiBi_n$, NN_n and BiN_n pairs present. These calculations are performed in the bulk configuration since the Bi and N dopants are not expected to be displaced in the (110) plane during the cleaving process. We compare the total energy of the supercells with the various different pairs to predict which nearest neighbor pairs would occur the most often and see if these agree with our experimental results. Since our experimental analysis considers the first, second, third, fourth and sixth nearest neighbor pairs, the same is done in our DFT analysis. The total energy of the supercells containing the dopant atom pairs after relaxation relative to the energy of the 1^{st} nearest neighbor cell is given in Fig. 7

We can observe some general trends in figure 7. The BiBi and NN pairs behave qualitatively similar, with the first and fourth nearest neighbor pairs being the highest in energy. The BiN pairs show the opposite behavior, with the first and fourth nearest neighbor pair being the lowest in energy. The first and

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

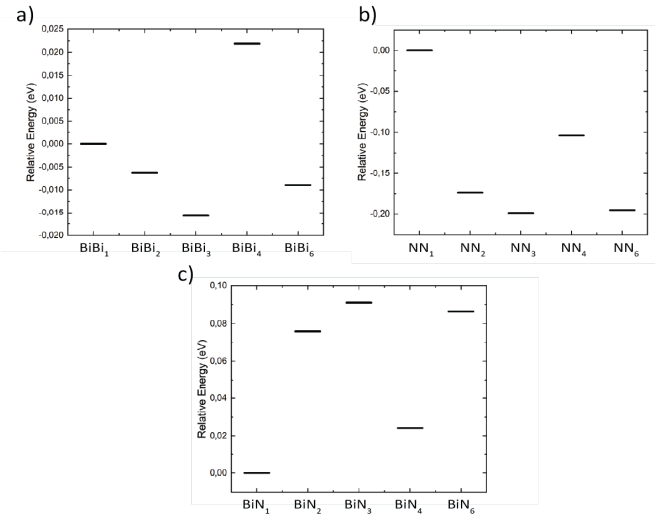

FIG. 7. Energies of the relaxed supercells as calculated with DFT for BiBi, NN and BiN pairs in a), b) and c) respectively. The energies of the supercells are given relative to the energy of a supercell with a first nearest neighbor pair of the same two dopants. This means that the energy of the $BiBi_2$ cell in a) is the energy of the cell relative to the $BiBi_1$ cell.

fourth nearest neighbor pair are both oriented along the $[\bar{1}10]$ direction like discussed in section III E. These DFT results indicate that this is a direction along which it is unfavorable to have two similar dopants. This is most likely related to the very different size of these dopants compared to the As atom they replace. The strain introduced into the lattice through these dopants seems higher along the $[\bar{1}10]$ direction. If we introduce atoms with opposite strain, like for BiN pairs, this behavior is mirrored and now the $[\bar{1}10]$ direction becomes the most favorable for dopant pairs.

We now confront the DFT results with the observed nearest neighbor occurrences shown in Fig. 3 and 5. The results of sample 2 shown in Fig. 3 agree very well with our simulated energies. The $BiBi_1$ and $BiBi_4$ pairs occur less often and we calculate those to have the highest energies. For sample 1, the $BiBi_1$, $BiBi_4$ and $BiBi_6$ do show a different behavior. This could be caused by growth conditions and their effect on the growth surface as discussed in section III E. It it challenging to simulate growth conditions with DFT since the growth is strongly temperature and III/V flux dependent, which cannot be simulated. The BiN pairs shown in Fig. 5 agree partially with the simulations. The high amount of BiN_4 pairs are consistent with the lower pair energy shown in Fig. 7. The fact that no BiN_1 pairs are observed is not consistent with the calculated energy though. As previously discussed, BiN_1 pairs are difficult to observe in X-STM images. Nevertheless, a higher amount of BiN_1 pairs would be expected with this low pair energy. This discrepancy could also be caused by the local alloy fluctuations discussed in section III E. The fact that Bi rich or N rich areas can form during the growth of the material which locally change the lattice constant is not taken into account in the DFT calculations. The DFT calculations all assume a homogeneous lattice constant of GaAs. These local alloy fluctuations would be an additional effect driving Bi and N atoms further from each other. The NN_n pair energies cannot be compared to experimental results from this work but it is possible to compare to previous experimental work from Plantenga et al. on N doped GaAs[15]. They found a low occurrence of NN_1 and NN_4 pairs present in N doped GaAs. This is again consistent with the high simulated energy of these pairs.

IV. CONCLUSION

We performed X-STM measurements and DFT simulations on Bi and BiN doped GaAs layers in two different samples. We have analyzed our experiments by observing the nearest neighbor pair distribution and the longer-range pair correlation function to see the distribution of Bi and N atoms throughout these layers. In all samples and layers a short range (1-2 nm) attractive interaction between Bi atoms is observed in the PCFs. A large difference in the nearest neighbor pair distribution between sample 1 and 2 is most likely related to different growth conditions for the two samples, possibly resulting in a different growth surface reconstruction. From the PCFs of $BiBi_n$ pairs in GaAsBiN we observe the formation of Bi rich and N rich areas within GaAsBiN. This repulsive interaction of the Bi and N atoms is also observed in the nearest neighbor pair distributions, where no BiN_1 pairs are observed. We simulated the energies of the different dopant pairs with DFT. These simulation agreed well with our experiments in most cases. Especially in the case of the second sample studied, the $BiBi_n$ pairs agreed very well with the simulations. The BiN_1 pairs were an outlier, which could be caused by the difficulty in correctly identifying these pairs experimentally. In this research we found indications that the growth conditions and N inclusion can strongly affect the Bi distribution, but further research is required to make a definite statement on this.

ACKNOWLEDGMENTS

We thank the Nederlandse Organisatie voor Wetenschappelijk onderzoek (NWO) for funding this research through Zwaartekracht project on Integrated Nanophotonics with Project No. [10018478] and for also providing the computational resources needed for this study. NWO granted us access to the Snellius supercomputer in Amsterdam through project EINF-3528 which was used to carry out the DFT calculations. This work was authored in part by the Alliance for Sustainable Energy, LLC, then Manger and Operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. K.A. and M.A.S contributed equally to the development of the model, generation analysis of the results and composition of the manuscript. Samples were grown with funding from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences as part of the "Light-Stimulated Epitaxy" program. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript.
The published version of the article is available from the relevant publisher.

U.S. Government retains and the publisher, by accepting the article for publication, acknowledge that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Author T.J.F. Verstijnen and D. Tjeertes have contributed equally to the work presented here.

* t.j.f.verstijnen@tue.nl

[1] S. Tixier, M. Adamcyk, T. Tiedje, S. Francoeur, A. Mascarenhas, P. Wei, and F. Schiettekatte, Applied Physics Letters **82**, 2245 (2003), publisher: American Institute of Physics.

[2] B. Fluegel, S. Francoeur, A. Mascarenhas, S. Tixier, E. C. Young, and T. Tiedje, Physical Review Letters **97**, 067205 (2006), publisher: American Physical Society.

[3] S. J. Sweeney and S. R. Jin, Journal of Applied Physics **113** (2013), 10.1063/1.4789624.

[4] K. Alberi, O. D. Dubon, W. Walukiewicz, K. M. Yu, K. Bertulis, and A. Krotkus, Applied Physics Letters **91**, 89 (2007).

[5] D. A. Beaton, A. Mascarenhas, and K. Alberi, Journal of Applied Physics **118** (2015), 10.1063/1.4937574.

[6] X. Lu, D. A. Beaton, R. B. Lewis, T. Tiedje, and Y. Zhang, Applied Physics Letters **95**, 2007 (2009).

[7] S. Imhof, A. Thränhardt, A. Chernikov, M. Koch, N. S. Köster, K. Kolata, S. Chatterjee, S. W. Koch, X. Lu, S. R. Johnson, D. A. Beaton, T. Tiedje, and O. Rubel, Applied Physics Letters **96**, 131115 (2010).

[8] R. B. Lewis, M. Masnadi-Shirazi, and T. Tiedje, Applied Physics Letters **101**, 082112 (2012), publisher: American Institute of Physics.

[9] M. Weyers, M. Sato, and H. Ando, Japanese Journal of Applied Physics **31**, L853 (1992), iSBN: 0021-4922.

[10] W. Shan, W. Walukiewicz, J. W. Ager, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz, Physical Review Letters **82**, 1221 (1999), iSBN: 1521-3951.

[11] W. Shan, W. Walukiewicz, J. W. Ager, E. E. Haller, J. F. Geisz, D. J. Friedman, J. M. Olson, and S. R. Kurtz, Journal of Applied Physics **86**, 2349 (1999), iSBN: 0021-8979.

[12] J. Wu, W. Shan, and W. Walukiewicz, Semiconductor Science and Technology **17**, 860 (2002).

[13] J. Occena, T. Jen, E. E. Rizzi, T. M. Johnson, J. Horwath, Y. Q. Wang, and R. S. Goldman, Applied Physics Letters **110** (2017), 10.1063/1.4984227.

[14] C. M. Krammel, M. Roy, F. J. Tilley, P. A. Maksym, L. Y. Zhang, P. Wang, K. Wang, Y. Y. Li, S. M. Wang, and P. M. Koenraad, Phys. Rev. Mater. **1**, 034606 (2017).

[15] R. C. Plantenga, *Probing impurities at the atomic scale in semiconductor structures*, PhD Thesis, Technische Universiteit Eindhoven, Eindhoven (2019), iSBN: 9789038649023.

[16] J. Occena, T. Jen, H. Lu, B. A. Carter, T. S. Jimson, A. G. Norman, and R. S. Goldman, Applied Physics Letters **113**, 109 (2018).

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

[18] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

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

[20] J. G. Lee, *Computational materials science: an introduction* (CRC press, 2016).

[21] J. D. Pack and H. J. Monkhorst, Phys. Rev. B **16**, 1748 (1977).

[22] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, Nature **585**, 357 (2020), number: 7825 Publisher: Nature Publishing Group.

[23] G. Bradski, Dr. Dobb's Journal: Software Tools for the Professional Programmer **25** (2000).

[24] F. J. Tilley, M. Roy, P. A. Maksym, P. M. Koenraad, C. M. Krammel, and J. M. Ulloa, Phys. Rev. B **93**, 035313 (2016).

[25] H. A. McKay, R. M. Feenstra, T. Schmidtling, and U. W. Pohl, Applied Physics Letters **78**, 82 (2001).

[26] H. A. McKay, R. M. Feenstra, T. Schmidtling, U. W. Pohl, and J. F. Geisz, Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures **19**, 1644 (2001).

[27] J. M. Ulloa, P. M. Koenraad, and M. Hopkinson, Applied Physics Letters **93**, 1 (2008).

[28] C. M. Krammel, A. R. da Cruz, M. E. Flatté, M. Roy, P. A. Maksym, L. Y. Zhang, K. Wang, Y. Y. Li, S. M. Wang, and P. M. Koenraad, Phys. Rev. B **101**, 024113 (2020).

[29] R. C. Plantenga, V. R. Kortan, T. Kaizu, Y. Harada, T. Kita, M. E. Flatté, and P. M. Koenraad, Physical Review B **96**, 155210 (2017), arXiv: 1605.09006 Publisher: IOP Publishing ISBN: 0021-8979.

[30] M. Masnadi-Shirazi, D. A. Beaton, R. B. Lewis, X. Lu, and T. Tiedje, Journal of Crystal Growth **338**, 80 (2012).

[31] P. Offermans, P. M. Koenraad, J. H. Wolter, J. D. Song, J. M. Kim, S. J. Bae, and Y. T. Lee, Applied Physics Letters **82**, 1191 (2003), publisher: American Institute of Physics.

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.