Computational study of positron annihilation parameters for cation mono-vacancies and vacancy complexes in nitride semiconductor alloys

Shoji Ishibashi¹,², Akira Uedono³, Hiori Kino², Takashi Miyake¹,² and Kiyoyuki Terakura¹

¹ Research Center for Computational Design of Advanced Functional Materials (CD-FMat), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8568, Japan
² Center for Materials Research by Information Integration (CMII), Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki 305-0047, Japan
³ Division of Applied Physics, Faculty of Pure and Applied Science, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

E-mail: shoji.ishibashi@aist.go.jp

Received 2 April 2019, revised 11 July 2019
Accepted for publication 25 July 2019
Published 20 August 2019

Abstract
We calculate positron annihilation parameters, namely the $S$ and $W$ parameters from the Doppler broadening spectroscopy and the positron lifetime $\tau$, for defect-free states as well as cation mono-vacancies and vacancy complexes in nitride semiconductor alloys Al$_{0.5}$Ga$_{0.5}$N, In$_{0.5}$Ga$_{0.5}$N and Al$_{0.5}$In$_{0.5}$N. The obtained distributions of these parameters differ from compound to compound. Especially, the $S$–$W$ relation for In$_{0.5}$Ga$_{0.5}$N is very different from that for Al$_{0.5}$Ga$_{0.5}$N. For the cation mono-vacancies, introducing local structural parameters, their correlations with $S$, $W$ and $\tau$ are investigated. The $S$ and $\tau$ variations are well described with the size distributions of the vacancies while the $W$ variation is related to the presence of localized d electrons. For the vacancy complexes as well as the cation mono-vacancies, multiple-linear-regression models to describe $S$, $W$ and $\tau$ are successfully constructed using the local structural parameters as descriptors. The $S$–$W$ and $S$–$\tau$ relations are also compared with those for AlN, GaN and InN.

Keywords: positron annihilation, vacancy-type defect, nitride semiconductor alloy

(Some figures may appear in colour only in the online journal)
MOVE and MBE, the major defect species can be identified as vacancy cluster such as $V_{\text{Si}}$–$nV_{\text{Si}}$ ($n = 1–3$), while, for LEO, vacancy-oxygen complexes such as $V_{\text{Al}}$–$nO_{\text{N}}$ ($n = 3$, 4) seem to have significant contributions. In Mg-doped GaN, $V_{\text{Ga}}$–$nV_{\text{N}}$ ($n = 2, 3$) were observed [4] and they were identified as the major non-radiative recombination centers [5]. In the above description, $V_{\text{A}}$ represents a vacancy at the A site. In practical applications of nitride semiconductors, the alloying technique is used to control the band gap. We have experimentally investigated (Al,Ga)N and (In,Ga)N systems and observed a significant concentrations of vacancy-type defects in high-Al content samples for the former [6], and in samples close to the In$_{0.5}$Ga$_{0.5}$N composition for the latter [7].

Theoretical predictions of positron states and annihilation parameters have often been found to be crucial in enabling specific defect species to be identified from the results of experiments. In consequence, we have been calculating positron lifetime and positron-electron momentum distributions for various types of defects in nitride semiconductors [8]. For alloy systems, in general, each of the nominally similar atomic sites are in fact not equivalent due to the random distribution of the two (or more) cations on the neighbor sites. The calculated positron annihilation parameters show finite distributions which differ from compound to compound. The differences among the three compounds are finite distributions which differ from compound to compound. The differences among the three compounds are.

In the present study, we calculate positron annihilation Doppler broadening spectra and positron lifetime ($\tau$) values for defect-free (DF) states as well as for cation mono-vacancies and vacancy complexes $V_{\text{N}}$ in Al$_{0.5}$Ga$_{0.5}$N and In$_{0.5}$Ga$_{0.5}$N and Al$_{0.5}$In$_{0.5}$N. The Doppler broadening spectra are analyzed by evaluating the positron lifetime $\tau$ values for the vacancy-type defects $V_{\text{N}}$ in nitride semiconductors ‘disordered’ alloys Al$_{0.5}$Ga$_{0.5}$N, In$_{0.5}$Ga$_{0.5}$N and Al$_{0.5}$In$_{0.5}$N. The Doppler broadening spectra are analyzed by evaluating $S$ and $W$ parameters. For the vacancy-type defects $V_{\text{N}}$, there are a variety of configurations reflecting the disordered nature, and the obtained $S$, $W$ and $\tau$ values show finite distributions which differ from compound to compound. We use several data-science methods to analyze the calculated results. The differences among the three compounds are explained in terms of local structural properties of the vacancy by evaluating linear correlation coefficients between the positron annihilation and local structural parameters. The characteristic behaviors of $S$, $W$ and $\tau$ are successfully modeled by the multiple linear regression using the local structural parameters. Furthermore, to figure out overall relationships among the parameters, we perform principal component analyses. For each compound and $n$, the averaged values of $S$, $W$ and $\tau$ are evaluated and compared with the values for the DF states as well as those for AlN, GaN and InN.

2. Computational methods

For each calculation in the present study, an orthorhombic supercell equivalent to $4 \times 4 \times 2$ hexagonal unit cells of the wurtzite structure is used. There are 64 cations and 64 N atoms in the supercell if there is no vacancy. The alloy structures for Al$_{0.5}$Ga$_{0.5}$N, In$_{0.5}$Ga$_{0.5}$N and Al$_{0.5}$In$_{0.5}$N are generated using the special quasirandom structure (SQS) approach [12]. The numbers of the possible vacancy complexes $V_{\text{N}}$–$nV_{\text{N}}$ in the above-mentioned supercell are 64, 256, 384 and 256 for $n = 0, 1, 2$ and 3, respectively. The lattice parameters for the alloy supercells are evaluated by averaging the experimental lattice parameters of the appropriate pair of wurtzite binary structures (AIN, GaN and InN).

Electron and positron states are calculated with our computational code QMAS [13], which adopts the projector augmented-wave (PAW) method [14] and the plane-wave basis. The generalized gradient approximation for the electronic exchange-correlation energy is used [15]. Internal atomic positions are optimized with evaluating the Hellmann–Feynman force. It has been shown previously that the effects of the positron localization on the atomic configurations and positron annihilation parameters [16–18] are insignificant for AlN, GaN and InN [9, 19], hence they are not considered in the present study. The vacancy-type defects $V_{\text{N}}$–$nV_{\text{N}}$ are thought to be in different charge states depending on the Fermi level [22]. For example, in GaN, the charge state of $V_{\text{Ga}}$ varies from 0 to 3 – according to density-functional-theory calculations by Neugebauer and Van de Walle [23]. Recent hybrid functional calculations by Diallo and Demchenko [24] suggest $+, 0, -2$ – and 3 – charge states for $V_{\text{Ga}}$ while $3+, +, 0, -2$ – states for $V_{\text{Ga}}$. To investigate the effect of the charge state on the annihilation parameters, we have calculated positron annihilation parameters for 64 types of $V_{\text{N}}$ in Al$_{0.5}$Ga$_{0.5}$N using supercells with the charge states of 0 and 3 –. For the 3 – cases, additional three electrons are distributed on N atoms in the vicinity of $V_{\text{N}}$ similarly to the previous calculations for GaN [23, 24]. Although the obtained $S$, $W$ and $\tau$ for each $V_{\text{N}}$ are slightly different between the two charge states, their overall distributions show rather simple parallel shifts. We therefore use neutral supercells throughout the present study. First we obtain the electronic structure with the atomic positions optimized, and then, fixing the electronic structure and atomic positions, the positron state is calculated. The electron–positron correlation energy is included using the local-density-approximation expression obtained for a positron in a homogeneous electron gas [16]. The plane-wave energy cutoff and the force convergence criterion are set to be 20 Ha and $5 \times 10^{-5}$ Ha/bohr, respectively. Only the Γ point is used for the Brillouin zone sampling. We also calculate positron annihilation parameters based on electronic structures obtained using $2 \times 2 \times 2 k$ points for 32 cases of $V_{\text{N}}$ in Al$_{0.5}$Ga$_{0.5}$N. The results are very similar to those obtained with the Γ-point sampling. Positron states are calculated with the positron effective potential constructed using core electron wavefunctions from isolated atom calculations and valence electron wavefunctions described in the PAW formalism [20]. With the obtained positron and electron wavefunctions, positron annihilation Doppler broadening spectra $\rho(\mathbf{p})$, which corresponds to the one-dimensional projection of the electron–positron momentum density $\rho(\mathbf{p})$, and positron lifetimes $\tau$ are calculated as follows:
\[ \rho(p) \propto \sum_j \left| \int \psi_+(r) \psi_j(r) \sqrt{\gamma(n_-(r))} \exp(-ip \cdot r) \, dr \right|^2. \]

\[ \rho_{1D}(p_z) = \int \rho(p) \, dx \, dy, \]

\[ \frac{1}{\tau} = \pi r_0^2 \int n_+(r)n_-(r)\gamma(n_-(r)) \, dr, \]

where \( \psi_+, \psi_j, n_+ \), and \( n_- \) are the positron wavefunction, the \( j \)th electron wavefunction, the positron density and the electron density, respectively. The direction of \( \gamma \)-ray emission corresponds to the \( z \) direction in Cartesian coordinates, \( r_0 \) represents the classical electron radius and \( c \) is the speed of light, and \( \gamma \) is the enhancement factor. In the present study of semiconductors, we use the expression

\[ \gamma_{BN}(r_s) = 1 + 1.23r_s + 0.8295r_s\sqrt{r_s} - 1.26r_s^2 + 0.3286r_s^3\sqrt{r_s} + r_s^4/6. \]

The electron–positron correlation energy is correspondingly modified by scaling [21]. As for \( \varepsilon_\infty \), averages of the experimental values for AlN, GaN and InN are used: 5.10 for Al\(_{0.5}\)Ga\(_{0.5}\)N, 6.95 for In\(_{0.5}\)Ga\(_{0.5}\)N and 6.55 for Al\(_{0.5}\)In\(_{0.5}\)N.

The Doppler broadening spectra \( \rho_{1D} \) are calculated using spherically averaged momentum densities \( \rho(p) \), since directionally averaged analyses are applied in the following. For spherical averaging, first, interpolated values of \( \rho(p) \) are obtained with an interval of 0.02 a.u. along 16 directions using the tricubic interpolation method by Lekien and Marsden [25], and then averaged. The 16 directions derive from the 32
icosahedral-group points for numerical integration on a sphere [26]. The obtained spectra are convoluted with a Gaussian whose full width at half maximum (FWHM) is 1.0 keV to mimic the experimental resolution. The resultant spectra are hereinafter referred to as $\rho^*_{1D}(p_z)$. The positron annihilation $S$ parameter is defined as the ratio of the center area in the momentum region between $\pm 0.419$ a.u. to the total spectrum area while the $W$ parameter is that for the higher momentum areas of $\pm (1.828 \sim 3.688)$ a.u.:

$$S = \int_{-0.419 \text{a.u.}}^{0.419 \text{a.u.}} \rho^*_{1D}(p_z) dp_z / \int_{-\infty}^{\infty} \rho^*_{1D}(p_z) dp_z,$$

(6)

$$W = \left( \int_{-5.688 \text{a.u.}}^{-3.688 \text{a.u.}} \rho^*_{1D}(p_z) dp_z + \int_{3.688 \text{a.u.}}^{5.688 \text{a.u.}} \rho^*_{1D}(p_z) dp_z \right) / \int_{-\infty}^{\infty} \rho^*_{1D}(p_z) dp_z.$$

(7)

For pairs of data $(x_i; y_i), i = 1, \ldots, N$, the linear correlation coefficient $r$ is given as
\[ r = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2}}. \]

where \( \bar{x} \) and \( \bar{y} \) are the mean values of \( x_i \)’s and \( y_i \)’s, respectively. The value \( r \) lies between \(-1 \) and \( 1 \). If \( r = 1 \), the data points \((x_i; y_i)\) line on a straight line with a positive slope while, for \( r = -1 \), on that with a negative slope. To evaluate \( r \) values, we use the \( \text{cor()} \) function in the R software package \([28]\), which is an integrated software package to perform data-scientific analyses. For multiple-linear-regression modeling, we use the \( \text{lm()} \) function in R. The regression results are investigated in terms of the coefficient of determination \( R^2 \) (adjusted \( R^2 \)), which is given as

\[ R^2_{adj} = 1 - \frac{\sum_{i=1}^{N} (y_i - f_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \frac{N - 1}{N - p - 1}. \]

where \( y_i \)’s and \( f_i \)’s are observed (in the present case, theoretically calculated) values and predicted (by multiple linear regression) values. \( N \) and \( p \) are the numbers of \( (y_i, f_i) \) pairs and explanatory variables in the regression model, respectively. \( R^2_{adj} = 1 \) indicates that the model is perfect. In a summary of application of the \( \text{lm()} \) function in R, the \( R^2_{adj} \) value is given. Principal component analysis is a procedure to construct a set of linearly uncorrelated variables called principal components from a possibly correlated observed variables. The first principal component is determined to maximize the variance of observed variables. Then, the first principal component reflects their variability as much as possible. The second principal component is determined to have the largest variance under the constraint that it is orthogonal to the first principal component. Similarly, subsequent components are determined. The principal components are linear combinations of the original observed variables. The \( \text{prcomp}() \) function in R is used to perform principal component analyses in the present study.

3. Results and discussion

3.1. Defect-free states

In figure 1, the calculated \( S \), \( W \) and \( \tau \) for the DF states in \( \text{Al}_{0.5}\text{Ga}_{0.5}\text{N}, \text{In}_{0.5}\text{Ga}_{0.5}\text{N} \) and \( \text{Al}_{0.5}\text{In}_{0.5}\text{N} \) are shown. The values for \( \text{AlN}, \text{GaN} \) and \( \text{InN} \) at their experimental lattice constants as well as those at the lattice constants of the alloys are also plotted for comparison. For \( \text{AlN}, \text{GaN} \) and \( \text{InN} \), it is clearly shown that the \( S, W \) and \( \tau \) values change significantly depending on the lattice constants. For example, the \( W \) values are in order of \( \text{GaN} \gg \text{InN} > \text{AlN} \) at the experimental lattice constants while the \( W \) values for \( \text{GaN} \) and \( \text{InN} \) are close to each other at the \( \text{In}_{0.5}\text{Ga}_{0.5}\text{N} \) lattice constants. The actual local environment in an alloy system is thought to be in between the two extremes. Anyway the \( S, W \) and \( \tau \) values for an alloy system can be well approximated by the average of the values for their constituent binary systems at the experimental lattice constants or by the average of those at the alloy lattice constants. The present results also indicate that the epitaxial strain can affect the positron annihilation parameters significantly.
3.2. Cation mono-vacancies

Figure 2(a) represents the atomic configuration around a cation mono-vacancy \( V_c \) surrounded by four N atoms and twelve cations. In figure 2(b), a positron density distribution trapped at a cation mono-vacancy in In\(_{0.5}Ga_{0.5}\)N is shown as an example. The positron density is well localized at the vacancy site. We calculate Doppler broadening spectra and positron lifetime values for all such states (64 sites \( \times \) 3 compounds). As mentioned above, the cation sites are not equivalent to each other because of the randomness. To show variations in the Doppler broadening spectra, two lines connecting the maximum or minimum points among the 64 spectra at respective momentum values are plotted in each panel of figure 3. To observe the distributed results more clearly, the \( S \) and \( W \) parameters are evaluated using the energy windows described above and in figure 3.

The obtained \( S \), \( W \) and \( \tau \) values are shown as \( S \)-\( W \) and \( S \)-\( \tau \) plots together with their frequency distributions in figure 4. The relations between \( S \) and \( W \) differ from compound to compound while those between \( S \) and \( \tau \) look rather simple. The \( S \)-\( W \) relation is almost linear for Al\(_{0.5}Ga_{0.5}\)N while it is obscure for In\(_{0.5}Ga_{0.5}\)N. The situation for Al\(_{0.5}In_{0.5}\)N looks in between them. The distributions for \( S \) and \( \tau \) are wide in order of Al\(_{0.5}In_{0.5}\)N (4.99 \( \times \) 10\(^{-3} \)), In\(_{0.5}Ga_{0.5}\)N (3.41 \( \times \) 10\(^{-3} \)), and Al\(_{0.5}Ga_{0.5}\)N (1.88 \( \times \) 10\(^{-3} \)), 2.13 ps) while the \( W \) distribution is in order of Al\(_{0.5}Ga_{0.5}\)N (6.95 \( \times \) 10\(^{-4} \)) \( \gg \) Al\(_{0.5}In_{0.5}\)N (5.44 \( \times \) 10\(^{-4} \)) \( \gg \) In\(_{0.5}Ga_{0.5}\)N (2.90 \( \times \) 10\(^{-4} \)). In parentheses, the standard deviations are shown. We try to understand the reasons of these differences in terms of local structures of the cation mono-vacancies. As shown in figure 2(a), a cation mono-vacancy \( V_c \) is surrounded by four N atoms and twelve cations. Three structural parameters that characterize the local environment of the vacancy defect in the A\(_{0.5}B_{0.5}\)N alloy are introduced: (1) the fraction of the 12 neighbor sites occupied by element B, \( x_B = N_B/12 \), where \( N_B \) is the number of B atoms on the sites, (2) the average distance of the four N atoms \( d_N \) from the center of the tetrahedron they form and (3) the average distance of the twelve cations \( d_c \) from the center of the tetrahedron. Correlations between the positron annihilation parameters and local structural parameters \( S \), \( W \), \( \tau \), \( x_B \), \( d_N \) and \( d_c \) are investigated. In table 1, the obtained linear correlation coefficients (\( \tau \)) are listed.

There are several common characteristics in the \( r \) values among the three compounds. The pairs \((S, \tau)\), \((\tau, d_N)\), and \((\tau, d_c)\) show strong correlations \((r > 0.9)\) although the \( S \) and \( \tau \) distributions differ from compound to compound. Here, we define an additional parameter \( d_N^{DF} \), which is defined as the average distance between the cation and four N atoms in the DF systems. There is one-to-one correspondence between \( d_N^{DF} \) and \( d_N \) since they are evaluated before and after introduction of a vacancy at a certain site and optimization of atomic positions. Relations between \( d_N^{DF} \) and the corresponding \( d_N \) are plotted together with their frequency distributions in figure 5. It is shown that the distribution profiles of \( S \), \( \tau \) and \( d_N \) for each compound have some similarity to one another, corresponding to the large \( r \) values between these parameters listed in table 1. Thus, the distribution characteristics of \( S \) and \( \tau \) reflect the size distribution of the N tetrahedrons surrounding the cation mono-vacancies. The distributions of \( d_N^{DF} \) (figure 5) show two peaks reflecting the difference in the cation size. According to Shannon [30], the effective ionic radii of four-coordinated trivalent Al, Ga and In are 0.39, 0.47 and 0.62 Å, respectively. It is clearly shown that the width of the \( d_N \) distribution is closely related to the separation between the two peaks in the \( d_N^{DF} \) distribution. It seems that the \( d_N \) value does not depend on whether A or B is removed to create the vacancy.
in Al0.5Ga0.5N. We cannot tell whether \( V_c \) is \( V_A \) or \( V_B \) from its local configuration although it can be practically determined by counting the number of \( A \) or \( B \) atoms in the supercell. The center of \( d_N \) distribution is larger than that of \( d_{\text{DF}} \) as shown in figure 5. For cation mono-vacancies in the binary compounds AlN, GaN and InN, computational studies report the outward relaxation of the neighboring N atoms [9]. For In0.5Ga0.5N, the \( r \) values for the pairs \((S, x_B)\), \((\tau, x_B)\) and \((d_N, x_B)\) are opposite in sign to those for the Al0.5Ga0.5N and Al0.5In0.5N cases. As for the cation size, \( Al < Ga, In < Ga \) and \( Al < In \). Only for the In0.5Ga0.5N case, \( B \) is smaller. Since it is thought that \( S, \tau \) and \( d_N \) are closely related to the defect size, the difference of \( r \) sign can be attributed to the size relation between the cations.

The correlation between \( W \) and \( x_B \) is also strong although \( r = 0.865 \) (\(<0.9\)) for In0.5Ga0.5N. As mentioned above, the width of \( W \) distribution is in order of \( Al < Ga, In < Ga \) and \( Al < In \). We can interpret these results as follows. Al has no d electrons while Ga and In have localized d electrons. The 3d orbitals of Ga are more localized than the 4d orbitals of In. The more localized electronic state, the wider distribution in the momentum space, which contributes to an increase of \( W \). Al0.5Ga0.5N consists of Al atoms with no d electrons and Ga with the most localized 3d electrons. Depending on the cation site, the \( x_B \) value is different. These reasons explain why the distribution of \( W \) is the widest (figure 4) and why there is a strong correlation between \( W \) and \( x_B \) for Al0.5Ga0.5N. As shown in figure 6, in terms of \( W \), the difference between \( Al \) and \( Ga \) is the most drastic and that between \( Al \) and \( In \) is the next. The difference between \( Ga \) and \( In \) is small since both elements have d electrons. Since the relations of the valence-electron localization are \( Al < Ga, In < Ga \) and \( Al < In \), the \( r \) values have the same sign among the three compounds.

### Figure 8.
Factor loadings from principal component analyses of \( S, W, \tau, x_B, d'_N \) and \( d'_c \). FL1 and FL2 represent the factor loadings concerning the first and second principal components, respectively.

### Figure 9.
\( S \)--\( W \) plot of the averages for each \( n \) of \( V_c-nV_N \). The values for the DF states and those in AlN, GaN and InN are also plotted for comparison. Filled symbols correspond to the DF state while open symbols to the vacancy complexes. With \( n \) increasing from 0 to 3, the symbol size increases. Since there are two types of vacancy configurations for the case of \( n = 1\)–\(3 \) in AlN, GaN and InN, two values are plotted for each case of them.

### 3.3. Vacancy complexes
We expand our calculations to the vacancy complexes \( V_c-nV_N \) (\( n \leq 3 \)). Figure 6 represents the relations between the calculated \( S \) and \( W \) values for the vacancy complexes (the results for the cation mono-vacancies are also shown). For each compound, the distributions are similar to that for the
mono-vacancies although their positions shift toward the right (a higher $S$). This is quite reasonable since the defect size increases with $n$. Figure 7 represents the relation between $S$ and $\tau$ for vacancy complexes. Good correlations are observed even for In$_{0.5}$Ga$_{0.5}$N. With $n$ increasing, the distributions shift toward higher values of $S$ and $\tau$. Al$_{0.5}$In$_{0.5}$N shows the largest distribution for each $n$. This is reasonable in light of the ionic-radius difference.

Similarly to the mono-vacancy case, we try to analyze the obtained positron annihilation parameters $S$, $W$ and $\tau$ in relation to local structural parameters. To deal with vacancy complexes $V_{-n}V_N$ with various $n$ on equal terms, we introduce new local structural parameters $d'_N$ and $d'_c$, which are measured from the redefined defect center, namely the center of twelve cations. For each case of the cation mono-vacancies, the difference between $d_N$ and $d'_N$ and that between $d_c$ and $d'_c$ are marginal. Correlations between $S$, $W$, $\tau$, $x_B$, $d'_N$ and $d'_c$ are investigated and the obtained $r$ values are listed in appendix. The $r$ values concerning $d'_N$ and $d'_c$ for $n = 0$ are very similar to the corresponding values for $d_N$ and $d_c$ in table 1. With $n$ increasing, the correlation between $S$ and $\tau$ remains strong while the correlations for the $(S, d'_N)$ and $(\tau, d'_N)$ pairs weaken. Since the number of N atoms decrease with $n$ increasing, the observed variations are thought to be reasonable.

For each compound and $n$ value, we construct multiple-linear-regression models based on the following equations using all the sets of parameters (64, 256, 384 and 256 sets for $n = 0, 1, 2$ and 3, respectively):

$$ S = a^S + b^S x_B + c^S d'_N + d^S d'_c, $$  

(10)

$$ W = a^W + b^W x_B + c^W d'_N + d^W d'_c, $$  

(11)

$$ \tau = a^\tau + b^\tau x_B + c^\tau d'_N + d^\tau d'_c. $$  

(12)

Although, as might be expected from $r$ values, it is revealed that not all the three parameters $x_B$, $d'_N$ and $d'_c$ are significant in several regression models by checking the t-statistic and corresponding $p$-value, all of them are included in the regression models to keep the same footing. In table 2, the resultant adjusted $R^2$ values are listed. The closer value to unity, the better regression model. For the cation mono-vacancy ($n = 0$) case, the obtained multiple-linear-regression models are excellent. For the other cases also, the regression models are rather good although, as mentioned above, the correlations between $S$ and $d'_N$ and between $\tau$ and $d'_N$ weaken with $n$ increasing. It is thought that the three descriptors complement one another.

In order to strengthen the above discussion, we perform principal component analyses of the parameters $S$, $W$, $\tau$, $x_B$, $d'_N$ and $d'_c$ with respect to each compound and $n$. As mentioned in the previous section, the first principal component reflects the variability of these parameters as much as possible. The second principal component reflects the remaining variability as much as possible under the orthogonality constraint. Factor loadings, which correspond to correlation values between the original parameters and the principal components, are crucial values to understand the nature of a specific component. In the present study, the cumulative proportion values up to the second component, which describe how much information of the parameters is included in the first and second principal components, run from 83% through 99%. We therefore evaluated factor loadings concerning the first and second principal components and plot in figure 8. The linear correlation coefficients in tables 1, A1–A3 represent the quantitative relationship between each pair of the parameters. With the plots of the factor loadings, we can comprehend the relations among all the parameters visually at a time. It should be noted that the sign of each factor loading is meaningful only in a relative sense. Even if the signs of all the values are inverted concerning a certain principal component axis, the interpretation must be the same. If directions from the origin concerning two parameters are parallel or antiparallel to each other, the correlation between them is strong. Figure 8 enables statistically significant correlation between the objective variables $S$, $W$ and $\tau$ and the descriptors $x_B$, $d'_N$ and $d'_c$ to be identified. For example, in the plot for Al$_{0.5}$Ga$_{0.5}$N and $n = 0$, it is shown that $x_B$ is an excellent descriptor for $W$ while $d'_N$ is an excellent descriptor for $\tau$ and a good descriptor for $S$ also. In the plot for In$_{0.5}$Ga$_{0.5}$N and $n = 2$, it looks that each of $x_B$, $d'_N$ and $d'_c$ alone is not a good descriptor for $S$, $W$, or $\tau$. By (linearly) combining them, good descriptors are expected to be obtained. The plots for Al$_{0.5}$In$_{0.5}$N are roughly similar to one another. The directions of $S$, $W$ and $\tau$ for Al$_{0.5}$Ga$_{0.5}$N and Al$_{0.5}$In$_{0.5}$N are nearly parallel or antiparallel to the first principal component axis while those for In$_{0.5}$Ga$_{0.5}$N are not. This can be also related to the distinct feature in the $S$–$W$ plot for In$_{0.5}$Ga$_{0.5}$N.

In actual experiments, except for the case that spectral decomposition is possible in positron lifetime measurements, only average values of the positron annihilation parameters for various annihilation sites are obtained for each sample. Nonetheless, the results shown in figures 6 and 7 are meaningful since there is a possibility that vacancy complexes show preferences for surrounding atoms. Here, we assume that such a preference does not exist and take the average for each compound and $n$. In figure 9, the $S$–$W$ relation is plotted together with the results for the DF states. For comparison, the results for AlN, GaN and InN also are plotted. Since there two types of vacancy configurations for $n = 1$–3 in the wurtzite structure,
Table A1. Linear correlation coefficients ($r_{xy}$) between positron-annihilation and structural parameters for vacancy complexes $V_c-nV_N$ ($n = 0–3$) in $\text{Al}_{0.5}\text{Ga}_{0.5}\text{N}$.

| $x \backslash y$ | $S$ | $W$ | $\tau$ | $x_B$ | $d'_N$ | $d'_c$ |
|-----------------|-----|-----|--------|-------|--------|--------|
| $S$             |     |     |        |       |        |        |
| $W$             |     |     |        |       |        |        |
| $n = 0$         |     |     |        |       |        |        |
| $W$             | $-0.962$ | 1 | $-0.923$ | $-0.797$ | 1 |       |
| $x_B$           | $-0.952$ | 0.995 | $-0.779$ | 1 |       |
| $d'_N$          | 0.889 | $-0.744$ | 0.991 | $-0.729$ | 1 |       |
| $d'_c$          | $-0.583$ | 0.759 | $-0.255$ | 0.788 | $-0.205$ | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.897$ | 1 | $-0.917$ | $-0.825$ | 1 |       |
| $x_B$           | $-0.852$ | 0.944 | $-0.677$ | 1 |       |
| $d'_N$          | 0.788 | $-0.492$ | 0.800 | $-0.445$ | 1 |       |
| $d'_c$          | 0.683 | $-0.716$ | 0.887 | $-0.497$ | 0.524 | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.984$ | 1 | $0.931$ | $-0.899$ | 1 |       |
| $x_B$           | $-0.970$ | 0.981 | $-0.882$ | 1 |       |
| $d'_N$          | 0.348 | $-0.297$ | 0.532 | $-0.393$ | 1 |       |
| $d'_c$          | 0.199 | $-0.168$ | 0.339 | $-0.054$ | $-0.191$ | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.989$ | 1 | $0.952$ | $-0.947$ | 1 |       |
| $x_B$           | $-0.966$ | 0.983 | $-0.917$ | 1 |       |
| $d'_N$          | 0.342 | $-0.344$ | 0.487 | $-0.389$ | 1 |       |
| $d'_c$          | 0.178 | $-0.171$ | 0.251 | $-0.150$ | $-0.267$ | 1 |

Table A2. Linear correlation coefficients ($r_{xy}$) between positron-annihilation and structural parameters for vacancy complexes $V_c-nV_N$ ($n = 0–3$) in $\text{In}_{0.5}\text{Ga}_{0.5}\text{N}$.

| $x \backslash y$ | $S$ | $W$ | $\tau$ | $x_B$ | $d'_N$ | $d'_c$ |
|-----------------|-----|-----|--------|-------|--------|--------|
| $S$             |     |     |        |       |        |        |
| $W$             |     |     |        |       |        |        |
| $n = 0$         |     |     |        |       |        |        |
| $W$             | $0.014$ | 1 | $0.989$ | 0.006 | 1 |       |
| $x_B$           | $0.450$ | 0.860 | 0.442 | 1 |       |
| $d'_N$          | 0.972 | 0.172 | 0.972 | 0.546 | 1 |       |
| $d'_c$          | 0.089 | $-0.962$ | 0.097 | $-0.829$ | $-0.043$ | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.108$ | 1 | $0.946$ | $-0.202$ | 1 |       |
| $x_B$           | $0.589$ | 0.443 | 0.598 | 1 |       |
| $d'_N$          | 0.778 | 0.339 | 0.762 | 0.544 | 1 |       |
| $d'_c$          | 0.270 | $-0.896$ | 0.409 | $-0.176$ | $-0.143$ | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.314$ | 1 | $0.957$ | $-0.333$ | 1 |       |
| $x_B$           | $0.697$ | 0.205 | 0.684 | 1 |       |
| $d'_N$          | 0.396 | 0.541 | 0.444 | 0.411 | 1 |       |
| $d'_c$          | 0.447 | $-0.912$ | 0.497 | $-0.024$ | $-0.350$ | 1 |
| $S$             |     |     |        |       |        |        |
| $W$             | $-0.030$ | 1 | $0.920$ | 0.025 | 1 |       |
| $x_B$           | $0.680$ | 0.540 | 0.733 | 1 |       |
| $d'_N$          | 0.212 | 0.639 | 0.332 | 0.356 | 1 |       |
| $d'_c$          | 0.047 | $-0.846$ | 0.062 | $-0.516$ | $-0.307$ | 1 |
Table A3. Linear correlation coefficients ($r_{xy}$) between positron-annihilation and structural parameters for vacancy complexes $V_c-nV_N$ ($n = 0–3$) in Al$_{0.5}$In$_{0.5}$N.

| $x$ | $y$ | $S$ | $W$ | $\tau$ | $x_B$ | $d_{N}^d$ | $d_{C}^d$ | $n$ |
|-----|-----|-----|-----|--------|-------|--------|--------|-----|
| $V_c$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $W$ | -0.870 | -0.778 | 1 | 1 | 1 | 1 |
| $x_B$ | -0.708 | 0.910 | -0.574 | 1 | 1 | 1 |
| $d_{N}^d$ | 0.947 | -0.687 | 0.969 | -0.535 | 1 | 1 |
| $d_{C}^d$ | -0.184 | 0.529 | -0.020 | 0.796 | -0.013 | 1 |

| $V_c$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $W$ | -0.867 | -0.819 | 1 | 1 | 1 | 1 |
| $x_B$ | -0.717 | 0.900 | -0.614 | 1 | 1 | 1 |
| $d_{N}^d$ | 0.847 | -0.556 | 0.837 | -0.460 | 1 | 1 |
| $d_{C}^d$ | 0.218 | -0.150 | 0.409 | 0.162 | 0.072 | 1 |

| $V_c$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $W$ | -0.902 | -0.841 | 1 | 1 | 1 | 1 |
| $x_B$ | -0.778 | 0.905 | -0.668 | 1 | 1 | 1 |
| $d_{N}^d$ | 0.627 | -0.371 | 0.663 | -0.366 | 1 | 1 |
| $d_{C}^d$ | 0.269 | -0.133 | 0.386 | 0.160 | -0.071 | 1 |

| $V_c$ | 1 | 1 | 1 | 1 | 1 | 1 |
| $W$ | -0.938 | -0.879 | 1 | 1 | 1 | 1 |
| $x_B$ | -0.815 | 0.891 | -0.735 | 1 | 1 | 1 |
| $d_{N}^d$ | 0.410 | -0.238 | 0.526 | -0.302 | 1 | 1 |
| $d_{C}^d$ | -0.101 | 0.271 | 0.045 | 0.453 | -0.077 | 1 |

Figure B1. (a) Frequency distribution of $d_{N}^d$, (b) frequency distribution of $x_B'$, (c) $S$–$W$ plot and (d) $S$–$\tau$. As for abbreviations CT1, CT2, SG1 and SG2, see text.
two values are plotted for each of them. Depending on the system, different behaviors are observed with n varying. As already reported in our previous studies [3, 4, 8, 27], with n increasing, the (S, W) values shift toward the right in the S–W plot with a negative slope for AlN or with a positive slope for GaN and InN. The shift directions for the alloy systems are in between the corresponding two of them. In figure 10, the (S, τ) values are plotted with the results for AlN, GaN and InN. In contrast with the S–W relation, there are small variations among the systems. A strong correlation between S and τ is clearly shown.

4. Conclusions

For the DF states as well as the cation mono-vacancy and vacancy complexes V_c–nV_N (n = 0–3) in Al_{0.5}Ga_{0.5}N, In_{0.5}Ga_{0.5}N and Al_{0.5}In_{0.5}N, the positron annihilation parameters S, W and τ have been calculated for all the possible configurations in the supercells, which contain 128 atoms if there is no vacancy. For the mono-vacancy, the relations between S and W as well as S and τ are presented together with the distributions of S, W and τ. They differ from compound to compound. For example, the S–W plots for Al_{0.5}Ga_{0.5}N and In_{0.5}Ga_{0.5}N are very different from each other. We introduce the local structural parameters x_b, d_N and d'_c and investigate correlations between the positron annihilation parameters and these local structural parameters. S and τ show strong correlations with d_N while W has a strong correlation with x_b. The former is understood in terms of the vacancy size while the latter is thought to be related to the localization property of valence electrons. The S–W and S–τ relations are also presented for the vacancy complexes. For each compound, the distributions are similar to that for the mono-vacancies although their positions shift toward a higher S (and also toward a higher τ for S–τ) with n increasing. Using the local structural parameters x_b, d_N and d'_c as the descriptors, multiple-linear-regression models are successfully constructed to describe S, W and τ. By performing principal component analyses, the factor loadings for S, W, τ, x_b, d_N and d'_c are obtained. The results are consistent with the other analyses in the present study. The average values of S, W and τ are evaluated and compared with the values for the DF states and those for AlN, GaN and InN in the S–W and S–τ plots, which are expected to be useful in interpreting experimental results.

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant No. JP16H06424 and by ‘Materials research by Information Integration’ Initiative (MI²I) project of the Support Program for Starting Up Innovation Hub from JST.

Appendix A. Linear correlation coefficients for vacancy complexes V_c–nV_N (n = 0–3)

Appendix B. Additional analyses for the Al_{0.5}Ga_{0.5}N and n = 1 case

As shown in figure 7, for the Al_{0.5}Ga_{0.5}N and n = 1 case, it appears that the data-point distributions in the S–τ plots split into two groups. In the wurtzite structure, there are two types of configurations for V_c–V_N. The configuration type where V_c and V_N are aligned along the c direction is referred to as CT1 and the other configuration type as CT2 (figure B1 inset). In figure B1(a), the frequency distributions of d_c' are shown for CT1 and CT2 separately. For each configuration, the distribution splits into two subgroups. The subgroup with lower d_c' values is referred to as SG1 and that with higher d_c' values as SG2. Here, we introduce an additional local parameter x'_Ga, which is the Ga concentration on the three cation sites around V_N. In figure B1(b), the frequency distributions of x'_Ga are shown for each configuration and subgroup. For each configuration type CT1 or CT2, the x'_Ga distribution is significantly different between SG1 and SG2. In figures B1(c) and (d), the S–W and S–τ plots are represented with the configuration types and the subgroups distinguished. It is clearly shown that the upper group in the S–τ plot corresponds to the subgroup CT2 SG2. In the S–W plot, depending on the subgroup, the data distributions are slightly different.

ORCID iDs
Shoji Ishibashi https://orcid.org/0000-0002-4896-3530

References

[1] Krause-Rehberg R and Leipner H S 1999 Positron Annihilation in Semiconductors (Berlin: Springer)
[2] Tuomisto F and Makkonen I 2013 Rev. Mod. Phys. 85 1583
[3] Uedono A et al 2009 J. Appl. Phys. 105 054501
[4] Uedono A, Ishibashi S, Tenjinbayashi K, Tsutsui T, Nakahara K, Takamizu D and Chichibu S F 2012 J. Appl. Phys. 111 014508
[5] Chichibu S F, Shima K, Kojima K, Takashima S, Edo M, Ueno K, Ishibashi S and Uedono A 2018 Appl. Phys. Lett. 112 211901
[6] Uedono A, Tenjinbayashi K, Tsutsui T, Shimahara Y, Miyake H, Hiramatsu K, Oshima N, Suzuki R and Ishibashi S 2012 J. Appl. Phys. 111 013512
[7] Uedono A, Ishibashi S, Watanabe T, Wang X Q, Liu S T, Chen G, Sang L W, Sumiya M and Shen B 2012 J. Appl. Phys. 112 014507
[8] Ishibashi S and Uedono A 2014 J. Phys.: Conf. Ser. 505 012010
[9] Ishibashi S and Uedono A 2016 J. Phys.: Conf. Ser. 674 012020
[10] Ishibashi S, Kino H, Uedono A, Miyake T and Terakura K 2017 J. Phys.: Conf. Ser. 791 012023
[11] Tuomisto F, Normman V and Makkonen I 2014 J. Phys.: Conf. Ser. 505 2042
[12] Zunger A, Wei S-H, Ferreira L G and Bernard J E 1990 Phys. Rev. Lett. 65 353
[13] Ishibashi S, Tamura T, Tanaka S, Kohyama M and Terakura K 2007 Phys. Rev. B 76 153310
[14] Blöchl P E 1994 Phys. Rev. B 50 17953
[15] Perdew J P, Burke K and Ernzerhof M 1996 Phys. Rev. Lett. 77 3865
[16] Borotti G and Nieminen R M 1986 Phys. Rev. B 34 3820
[17] Puska M J, Seitsonen A P and Nieminen R M 1995 Phys. Rev. B 52 10947
[18] Ishibashi S 2015 J. Phys. Soc. Japan 84 083703
[19] Ishibashi S 2017 Acta Phys. Pol. A 132 1602
[20] Ishibashi S 2004 Mater. Sci. Forum 445–6 401
[21] Puska M J, Miiikinen S, Manninen M and Nieminen R M 1989 Phys. Rev. B 39 7666
[22] Van de Walle C G and Neugebauer J 2004 J. Appl. Phys. 95 3851
[23] Neugebauer J and Van de Walle C G 1994 Phys. Rev. B 50 8067
[24] Diallo I C and Demchenko D O 2016 Phys. Rev. Appl. 6 064002
[25] Lekien F and Marsden J 2005 Int. J. Numer. Methods Eng. 63 455
[26] McLaren A D 1963 Math. Comput. 17 361
[27] Uedono A, Nakamori H, Narita K, Suzuki J, Wang X, Che S-B, Ishitani Y, Yoshikawa A and Ishibashi S 2009 J. Appl. Phys. 105 054507
[28] R Development Core Team 2015 R: a Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing)
[29] Momma K and Izumi F 2011 J. Appl. Crystallogr. 44 1272
[30] Shannon R D 1976 Acta Crystallogr. A 32 751